Deep Learning-based diagnostic system for malignant liver detection

A dissertation submitted to the

College of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

in the Department of Mechanical Engineering

University of Saskatchewan

Saskatoon

By

Rayyan Azam Khan

Permission to Use

In presenting this dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my dissertation work was done. It is understood that any copying or publication or use of this dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my dissertation.

Requests for permission to copy or to make other uses of materials in this dissertation in whole or part should be addressed to:

Head of the Department of Mechanical Engineering
University of Saskatchewan
Engineering Building, 57 Campus Drive
Saskatoon, Saskatchewan S7N 5A9 Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9 Canada

Disclaimer

Reference in this dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Abstract

Cancer is the second most common cause of death of human beings, whereas liver cancer is the fifth most common cause of mortality. The prevention of deadly diseases in living beings requires timely, independent, accurate, and robust detection of ailment by a computer-aided diagnostic (CAD) system. Executing such intelligent CAD requires some preliminary steps, including preprocessing, attribute analysis, and identification.

In recent studies, conventional techniques have been used to develop computer-aided diagnosis algorithms. However, such traditional methods could immensely affect the structural properties of processed images with inconsistent performance due to variable shape and size of region-of-interest. Moreover, the unavailability of sufficient datasets makes the performance of the proposed methods doubtful for commercial use.

To address these limitations, I propose novel methodologies in this dissertation. First, I modified a generative adversarial network to perform deblurring and contrast adjustment on computed tomography (CT) scans. Second, I designed a deep neural network with a novel loss function for fully automatic precise segmentation of liver and lesions from CT scans. Third, I developed a multi-modal deep neural network to integrate pathological data with imaging data to perform computer-aided diagnosis for malignant liver detection.

The dissertation starts with background information that discusses the proposed study objectives and the workflow. Afterward, Chapter 2 reviews a general schematic for developing a computer-aided algorithm, including image acquisition techniques, preprocessing steps, feature extraction approaches, and machine learning-based prediction methods.

The first study proposed in Chapter 3 discusses blurred images and their possible effects on classification. A novel multi-scale GAN network with residual image learning is proposed to deblur images. The second method in Chapter 4 addresses the issue of low-contrast CT scan images. A multi-level GAN is utilized to enhance images with well-contrast regions. Thus, the enhanced images improve the cancer diagnosis performance. Chapter 5 proposes a deep neural network for the segmentation of liver and lesions from abdominal CT scan images. A modified Unet with a novel loss function can precisely segment minute lesions. Similarly, Chapter 6 introduces a multi-modal approach for liver cancer variants diagnosis. The pathological data are integrated with CT scan images to diagnose liver cancer variants.

In summary, this dissertation presents novel algorithms for preprocessing and disease detection. Furthermore, the comparative analysis validates the effectiveness of proposed methods in computer-aided diagnosis.

Acknowledgements

First, I would like to express my gratitude to my supervisor Prof. Fang-Xiang Wu, for allowing me to pursue my Ph.D. under his guidance and supervision. He encouraged me to achieve my objectives by providing novel scientific ideas in my research that broadened my horizon. Moreover, he stood with me and pushed me through my most difficult times of research, for which I am indebted to him. I feel lucky to have an experienced supervisor that made it possible for my dreams to come true. I am also thankful to my co-supervisor, Dr. Yigang Luo, for being readily available anytime for discussion and giving me valuable input from a medical perspective. He also played a vital role in obtaining ethical approvals for this research.

I am also grateful to my advisory committee members, Prof. Chris Zhang, Prof. James Johnston, and Prof. Francis Bui, for their valuable comments and suggestions during my Ph.D. program.

Furthermore, I am thankful to all my lab mates, especially Fei Wang and Ali Akbar Hamali Beyrami, for providing a friendly and supportive environment in my research.

Foremost, I admire and acknowledge my wife's unconditional love and efforts to stand with me through thick and thin. Also, I am obliged to thank my master's supervisor, Prof. Noman Naseer, and friends particularly Dr. Noori FM for their moral support and motivation.

Finally, I am gratified to the Natural Science and Engineering Research Council of Canada (NSERC) and the College of Graduate and Postdoctoral Studies (CGPS) for providing financial support throughout my program.

I would like to dedicate this thesis to my parents, Shaheen Azam Khan, Muhammad Azam Khan, and grandmother Razia Begum who passed away during the Covid-19 pandemic.

Contents

List of Tables

List of Figures

List of Abbreviations

AC	Active Contour
AMSE	Absolute Mean Sharpness Error
ASSD	Asymmetric Surface Distance
AUC	Area Under Curve
BE	Benign
BKE	Backward Elimination
CAD	Computer-aided Diagnosis
CC	Cholangiocarcinoma
CCT	Contourlet Coefficient Texture
CH	Chan-vese
CL	Clustering
CLD	Chronic Liver Disease
CNN	Convolution Neural Network
CR	Cirrhosis
CT	Computed Tomography
CY	Cyst
DBN	Deep Belief Network
DCNN	Deep Convolution Neural Network
DFT	Discrete Fourier Transform
DNN	Deep Neural Network
DRB	Dense Residual Network Block
DSC	Dice Coefficient
DT	Decision Tree
DWT	Discrete Wavelet Transform
ELM	Extreme Learning Machine
FB	Fibrosis
FDCT	Fast Discrete Curvelet Transform
FFT	Fast Fourier Transform
FISCR	Fisher Criterion
FL	Fatty Liver
FLD	Fatty Liver Disease
FNH	Focal Nodular Hyperplasia
FOS	First Order Statistics

FT	Fourier Transform
GAN	Generative Adversarial Network
GC	Graph Cut
GE	Geometric
GL	Gray Level
GLCM	Gray-level Co-occurrence Matrix
GLDM	Gray-level Difference Matrix
GLRLM	Gray-level Run Length Matrix
GR	Gradient
GWO	Gray Wolf Optimization
HCC	Hepatocellular Carcinoma
HEM	Hemangioma
HEP	Hepatomegaly
ICA	Independent Component Analysis
IM	Hepatic Metastasis
KC	Kinetic Curve
k-NN	k-nearest Neighbor
LDA	Linear Discriminant Analysis
LE	Lesion
LFT	Liver Function Test
LOOCV	Leave One Out Cross-validation
LR	Logistic Regression
LSTM	Long Short-term Memory
MA	Malignancy
MAE	Mean Absolute Error
MAP	Maximum a-Posterior
mCRC	Metastatic Colorectal Cancer
MET	Metastasis
MI	Moment Invariant
ML	Machine Learning
MLP	Multi-layer Perceptron
MO	Multifocal HCC
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
MSSD	Maximum Surface Symmetric Distance
NN	Neural Network

NOR	Normal
OvR	One versus Rest
PATH	Pathology
PCA	Principal Component Analysis
PIQE	Perception-based Image Quality Evaluator
PL	Primary Liver Cancer
PNN	Probabilistic Neural Network
PSF	Point Spread Function
PSNR	Peak Signal to Noise Ratio
PSO	Particle Swarm Optimization
ReLU	Rectified Linear Unit
RBF	Radial Base Function
REG	Regression
RF	Random Forest
RFE	Recursive Feature Elimination
RIL	Residual Image Learning
RK	Ranklet
RL	Richardson-Lucy
RNN	Recursive Neural Network
ROC	Receiver Operating Curve
ROI	Region of Interest
RVD	Relative Volume Difference
RMSSD	Root Mean Square Surface Distance
SBE	Sequential Backward Elimination
SFS	Sequential Forward Selection
SL	Secondary Liver Cancer
SSEQ	Spatial and Spectral Entropies
SSIM	Structural Similarity Index
SSM	Statistical Shape Model
SSR	Single Scale Retinex
SVM	Support Vector Machine
SVR	Support Vector Regression
TEX	Texture
TH	Threshold
TV	Tikhonov Regularization
US	Ultrasonography

VOE	Volumetric Overlap Error
WAT	Watershed
WGAN	Wasserstein Loss
WHO	World Health Organization
WT	Wavelet Transform

1 Introduction

1.1 Background

Chronic liver diseases are progressive and potentially fatal due to asymptotic behavior. According to the American Institute of Cancer Research Report 2018, liver cancer is the second most common cause of death in human beings. The percentage has been growing steadily in past years as it was ranked worldwide the sixth most common cancer in 2012 and fifth in 2018. In 2012, there were $782,000$ cases, while in 2018, $840,000$ cases were reported of liver cancer, more common in men than in women [15]. Since liver cancer does not have any symptoms in the early stage, the mean age of survival for five years was around 12 percent only during 2000 to 2007. Similarly, in most cases, liver cancer was diagnosed over the age of 75. Based on given statistics, a liver cancer diagnosis is currently a focus of interest for biomedical researchers [16, 17, 18, 19, 20].

Liver diseases can be categorized as focal and diffuse. The former type may include benign such as cysts or malignant lesions related to primary or secondary cancer. In contrast, the latter covers virus hepatitis, fibrosis, cirrhosis, fatty liver, etc. Liver malignancy is divided into primary- and secondary cancer and has high mortality with an incidence rate of 0.95.

Multiple clinical tests are performed to diagnose liver malignancy, including biopsies, laparoscopy, and imaging scans. Unfortunately, such tests are painful, time-consuming, laborious, and cumbersome, while the results are often not completely reliable. Moreover, specimens obtained for histopathology require biopsy techniques that are associated with potential patient morbidity or mortality. In contrast, non-invasive methods such as medical imaging may be used to develop reliable diagnosis algorithms to identify different clinical outcomes. Here, computer-aided algorithms can assist experts, such as assessing and categorizing liver lesions.

Medical imaging of liver disease may utilize three different imaging modalities: ultrasound (US); magnetic resonance imaging (MRI); and computed tomography (CT). CT is more common for abdomen image acquisition due to its operational independence, user-friendly interface, and better resolution than US [21, 22, 23, 24, 25, 26, 27].

An intelligent computer-aided diagnosis (CAD) requires preliminary steps, including preprocessing, attribute analysis, and classification [20]. In recent studies, several methodologies utilize conventional preprocessing techniques that affect the structural properties of bio-medical images. Also, often the extraction of region-of-interests are performed manually. Such methods can be tiresome to perform with high precision due to variable shape and size. After preprocessing, the features are selected based on some state-of-the-art methods.

In contrast, selected features can be classified with pre-defined machine learning algorithms [20]. The classified results are utilized to identify ailment regions for further operations such as radiation therapy, hepatectomy, cryosurgery, [28]. The unavailability of sufficient datasets, however, makes the performance of the proposed methods doubtful for commercial or clinical use. Similarly, from a medical perspective, pathological data correlate with imaging data. However, none of the recent diagnostic studies integrates clinic data with imaging data for a reliable prognosis [29, 30, 31, 32, 33, 34, 35, 36].

In my research, I first propose two models of a modified generative adversarial network that could deblur and enhance the biomedical images with minimal structural changes [11, 12]. Afterwards, I designed an improved Unet with a novel loss function to segment the liver and lesions [13]. Lastly, I developed a multi-modal deep neural network to diagnose liver cancer variants based on pathological and imaging data.

1.2 Motivations and Objectives

This thesis comprises the following objectives:

Multiple studies have been proposed for computer-aided diagnosis. However, these methodologies have various limitations in several manners that include data constraint, automatic segmentation, image processing, and diagnosis techniques. Hence, I proposed a systematic review with a critical analysis of the best-performed studies in my first objective given as,

1. To present a systematic review of machine learning-based liver disease diagnosis.

Most methodologies utilize iterative image processing techniques to process non-homogeneous artifacts and noises. Applying such traditional methods can affect the structural properties of bio-medical images by losing spatial detail of pixels. Therefore, the second objective of this thesis is to develop deblurring techniques as follows,

2. To develop a multi-scale GAN with residual image learning for removing heterogeneous blur.

Previously, manual contrast adjustment has been performed for a series of images. However, such conventional intensity adjustment techniques result in nonuniform contrast due to variable exposure or depth of field during one complete scan. Hence, low contrast CT scans require a global intensity adjustment method that has developed in my third objective as,

3. To propose multi-level GAN based enhanced CT scans for liver cancer diagnosis.

Manual and semi-automatic methods have been commonly proposed for image segmentation. The semi-automatic models require user input to select the region of interest. Thus, it requires more computation time with inaccurate outcomes for medical applications. However, precise automatic segmentation with maximum accuracy is required to achieve the best diagnosis. This gives us our next objective as follows,

4. To design a residual multi-scale UNet for liver and lesion segmentation.

Image modality is frequently utilized to perform computer-aided diagnosis while clinical data is highly correlated with image data that can help diagnose. Therefore, we integrated multi-modal data to diagnose liver cancer variants. Thereafter segmentation, the last objective became,

5. To develop a multi-modal deep neural network for multi-class liver cancer diagnosis.

1.3 Organization of the Dissertation

This dissertation is formatted in a manuscript style such that most content presented here is published or submitted papers that I have written during my Ph.D. program. A brief introduction at the beginning of each chapter describes a connection of the manuscript to the dissertation. All manuscripts have been reformatted with unified bibliography to maintain consistency.

The remainder of this dissertation is organized as follows. Chapter 2 gives a systematic review of machine learning-based liver disease diagnosis studies. The preprocessing, feature extraction, and machine learning-based algorithms are discussed in detail. Chapter 3 proposes a multi-scale GAN with residual image learning to perform deblurring. Chapter 4 suggests a multi-level GAN for the enhancement of CT scans. Chapter 5 proposes an improved Unet with a novel cost function to segment liver and lesions. Similarly, Chapter 6 introduces a multi-modal network that integrates clinical data with image data to diagnose liver cancer variants.

Finally, Chapter 7 summarizes the work presented in this dissertation and discusses the future directions of this research. The list of published research articles are enlisted in Appendix B, while the copyright permissions are included in Appendix C.

2 Machine Learning based Liver Disease Diagnosis: A Systematic Review

This chapter presents a systematic review of machine learning models for liver disease diagnosis. First, the advancement in the preliminary steps, including image acquisition, preprocessing, attribute analysis, and machine learning models, have been discussed in detail with a comparative analysis of best-performed studies. Afterward, future works are proposed to overcome the shortcomings.

Abstract

The computer-based approach is required for the non-invasive detection of chronic liver diseases that are asymptomatic, progressive, and potentially fatal in nature. In this study, we review the computer-aided diagnosis of hepatic lesions in view of diffuse- and focal liver disorders. This survey mainly focuses on three image acquisition modalities: ultrasonography, computed tomography, and magnetic resonance imaging. We present the insightful analysis with pros and cons for each preliminary step, particularly preprocessing, attribute analysis, and classification techniques to accomplish clinical diagnostic tasks. In preprocessing, we explore and compare commonly used denoising, deblurring and segmentation methods. Denoising is mainly performed with nonlinear models. In contrast, deep neural networks are frequently applied for deblurring and automatic segmentation of region-of-interest. In attribute analysis, the most common approach comprises texture properties. For classification, the support vector machine is mainly utilized across three image acquisition modalities. However, comparative analysis shows the best performance is obtained by deep learning-based convolutional neural networks. Considering biopsy samples or pathological factors such as overall stage, margin, and differentiation can be helpful for improving the prediction performance. In addition, technique breakthrough is expected soon with advances in machine learning models to address data limitation problems and improve the prediction performance.

2.1 Introduction

According to World Health Organization (WHO) report 2018, cancer is considered the second most common cause of death in human beings. Whereas liver cancer is the fifth most common cancer in men and ninth in women. In 2012, $782,000$ while in 2018, $840,000$ cases have been reported of liver cancer with most of the diagnosis over the age of 75. Out of these total incident cases, 7.5% are men, thus more common than in women. More specifically, among the standardized age incidence cases per $100,000$ male population, 9.3 are in North America, 11.1 in Europe and 31.9 in Eastern Asia [15]. Additionally, the liver cancer prognosis is very poor due to the fact that the overall ratio of mortality to incidence is only 0.95. Therefore, the mean age of survival for 5 years is only around 12% from 2000 to 2007 [37].

Liver is formally divided into eight functionally independent segments, while its cancer disease is mainly categorized into primary and secondary [38]. Primary liver cancer begins in the cells of the liver. These cancerous cells can accumulate to make a single lump or can start at multiple sites of a liver. Patients with severe liver damage are more prone to several growth sites. Primary liver cancer is further categorized into Hepatocellular carcinoma (HCC), Cholangiocarcinoma (CC), Angiosarcoma and Hepatoblastoma. HCC accumulates up to 75% of primary liver cases, while CC accounts for approximately 10 to 20% of all liver cancers. HCC most likely occurs due to metabolic syndromes, hepatitis virus or alcohol abuse. CC is commonly known as the bile duct cancer, where duct carries bile to the gallbladder to assist with the digestion. The cancer can begin out- or inside section of the liver duct, thus named as extra- or intra-hepatic bile duct cancer [39]. Secondary liver cancer is also named as metastasis, occuring when cancer begins in other primary organs of the body such as lungs, stomach, breast, pancreatic etc., while the cancerous cells break down to travel through the bloodstream or lymphatic system to reside somewhere inside the liver. Due to the possible multiple source sites, metastasis is more common than the primary liver cancer [40, 41].

Besides several possible symptoms of the liver cancer such as fever, vomiting, fatigue, etc., multiple diagnostic steps are performed to confirm any possible treatment. Diagnosis may involve one or various possible tests that includes blood tests, imaging scans, biopsy, laparoscopy, and some others. Imaging scans assist in structural analysis of the region of interest (ROI) and thus help the overall treatment. Medical imaging can be performed through various types of modalities. However, abdominal cancer treatment involves one or more of the three commonly used modalities, namely, Ultrasonography (US), Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) [21, 22, 23, 24, 25, 26, 27]. These modalities can operate on multiple protocols to acquire structural images with various contrasts. Considering independent, timely and adequate detection, MRI and CT are common in practice [26]. Still there is a room of improvement as researchers have been working for accurate robust detection and classification of the liver tumor or lesion through US, MRI or CT input images [42, 43, 44, 45, 46, 47]. The discernment of such computer-aided diagnosis (CAD) system requires some preceding steps that includes preprocessing, attribute analysis and classification [48, 49, 50, 28].

Images can be blurry due to subjects' involuntary movements and may also contain environmental and instrumental noises. Such noises and artifacts can affect performance while analysing images through state-of the art algorithms [51, 52]. Likewise, the liver size varies with the body shape, gender and age, also the malevolent tissue often has low contrast with the normal tissue, thus making malignant tissue detection difficult. Hence, for adequate structural analysis; for instance, after artifact removal, denoising, deblurring and registration, the accurate localization and lesion detection are performed [53, 54, 55]. Consequently, pre-processing images reduces the false negative rate. Afterwards, distinct and meaningful features are required to differentiate between normal and malignant tissues [50]. After extraction and selection, these features are utilized to perform classifications between desired outcomes. Various algorithms have been reported to perform such diagnosis in the best possible ways [42, 56]. The overall schematic of the CAD system is shown in Fig.2.1. The classified outcomes are employed to identify the ailment regions for further treatment such as radiation therapy, hepatectomy, cryosurgery, etc., [28].

To identify the relevant articles, we extensively searched Web of Science, EMBASE, PubMed and Google scholar databases till April 2021. The search includes several renown studies that consist of following key-words: Liver, Liver mass, Metastasis, Malignancy, Computer-aided diagnosis, Hepatocellular carcinoma, Primary liver cancer, Computed tomography, Magnetic resonance imaging, Ultrasound, Cyst, Benign and Fibrosis. Moreover, to increase the search yield, we also review the references of the primary survey articles.

Multiple preliminary studies [41, 57, 58, 59, 60, 61] discusses such CAD algorithms with evaluation criteria to formally target unresolved problems. To the best of our knowledge, so far malignant liver CAD literature lacks in various perspectives that are formally addressed in this review, which is thus the first study to summarize:

- systematic survey of image preprocessing and attribute analysis for malignant liver diagnosis;
- comparative overview of state-of-the-art liver CAD systems based on multiple modalities (US, MRI and CT);
- the scope of artificial intelligence in the identification of multi-class liver cancer;
- comprehensive analysis on the diagnosis of primary- versus secondary liver cancer, fibrosis versus cirrhosis, hepatocellular- versus cholangio-carcinoma, etc.

The remainder of this study is arranged as follows. Section 2.2 briefly discusses preliminary steps such as image-acquisition, preprocessing, segmentation and feature extraction techniques required for the development of intelligent systems in view of developed algorithms. Liver CAD systems are discussed in section 2.2.5. Subsection 2.2.5 and 2.2.5 critically discusses these proposed algorithms. Finally, prospects and conclusion are summarized in section 2.2.3 and 2.4.
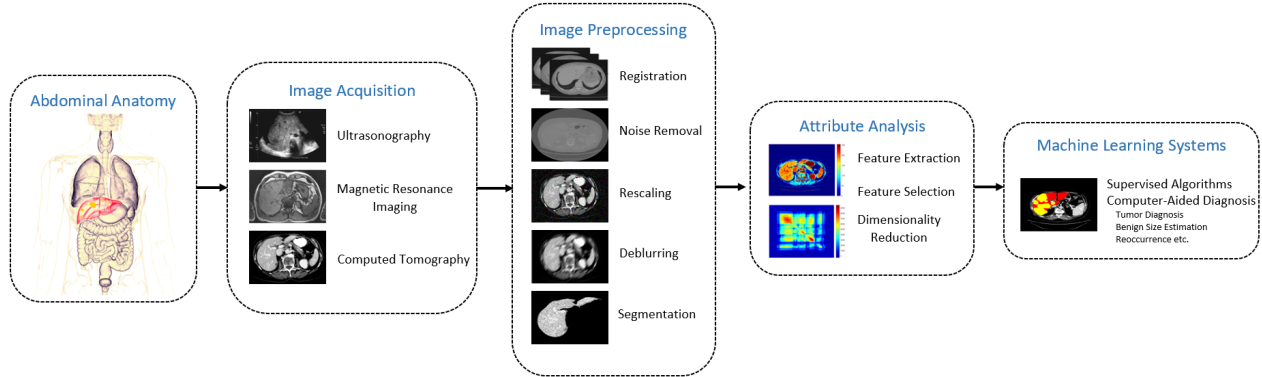
Figure 2.1: The schematic flow of malignant liver diagnostic system

2.2 Method Preview

2.2.1 Image Acquisition

Patients experiencing chronic liver diseases are more prone to complications with liver failure. In the diagnosis and therapeutic assessment of such perennial diseases, it requires to estimate the degree of hepatic fibrosis stage as substantial predictive factor for the occurrence of HCC. LI-RADS [62] and Barcelona Staging System [63] are well-known to predict cancer stage. However, for multi-class diagnosis such as cirrhosis versus benign, hepatocellular- versus cholangio-carcinoma etc., we require intelligent computer-aided diagnosis systems.

The liver biopsy serves as the reference test for staging the fibrosis and is still considered as gold-standard for its grading [64]. However, it is an invasive procedure with potential side effects; for instance, pain in 30 to 40% cases [65, 66, 67]. To decrease potential use of such painful biopsies, non-invasive methods such as US, MRI and CT are utilized to perform screening and radiological follow-ups of potentially risky patients so that early detection can reduce the mortality rate. These modalities can perform at various protocols to distinguish malignant hepatic tumors from benign. For optimized examinations, we discuss these techniques with commonly used protocols as follows [26].

Ultrasonography (US)

US is the first commonly used imaging modality for abdominal parenchymal organs, due to the fact of non-invasiveness, low-cost and wide availability. However, due to the variable diagnosis efficacy for hepatic fibrosis, it is operated by well experienced and qualified personnel. It utilizes high-frequency sound waves to get the inside view of the body. A transducer is directly placed on the skin with a gel in between to increase the conductive medium for the sound waves. Due to its real-time characteristics, it allows dynamic scanning. Contrary to its multiple advantages, US has a comparatively low-contrast resolution [26].

Magnetic Resonance Imaging (MRI)

MRI is used for detailed anatomical image analysis. It has some advantages over CT; for instance, non-ionizing radiations, high-contrast resolution and multi-planar imaging capabilities. MRI uses magnetic waves to align excited photons with the field. Thus, the synchronized radio-frequency current pulses with magnetic field are utilized to analyze various type of tissue molecules. Abdominal MRI can be performed commonly with contrast enhanced axial spin-echo T1-weighted and fast spin-echo T2-weighted imaging. T2-weighted imaging is commonly preferred for abdominal scans [26].

Computed Tomography (CT)

CT utilizes X-rays to acquire images with sequential or spiral techniques. Due to spiral- and sub-second scanning, CT is preferred for the detection of hepatic masses. Axial reconstruction of images with optimal intervals reduces averaging of lesions partial volume. Likewise, it allows the scanning of hepatic enhancement during various phases named as arterial, portal and delayed. Moreover, registration issues due to respiration are minimized due to single hold breath. Lastly, spiral CT scans provide three-dimensional imaging datasets. Routine abdominal CT scans are performed after intravenous examination of contrast agents with variable thickness, pitch and reconstruction intervals. Sometimes, sequential scans are also performed for further analysis [26].

2.2.2 Image Preprocessing

The biomedical images obtained using MRI or CT are mostly structural images that may contain artifacts or blurriness, environmental- and instrumental noises due to subjects' unintentional movements. Such aspects can adversely affects image quality and eventually disease detection accuracies while processing or analysing the images through state-of-the-art algorithms [57, 68, 69]. More frequently, nonlinear bilateral- and median filters are used for noise removal [70, 71]. However, some adaptive multi-dimensional and wavelet-based techniques have been proposed to minimize motion artifacts and noises [72, 73]. Similarly, deep learning has also been widely used to discover and remove various image distortions [74, 75, 76, 77, 78]. Some commonly used techniques are explained as follows.

Artifacts Removal

Streak effects, also known as metal artifacts, appear due to the presence of high-attenuation objects such as dental fillings, hip prosthesis, surgical clips, etc.,[79]. Linear prediction [80], non-weighted and weighted non-missing neighbor projection [81], and polynomial interpolation [79], are some of the techniques utilized to remove such artifacts. Motion artifacts appear due to the patient involuntary movements, such as the cardiac motion, or the voluntary movement of the chest in expiration and inspiration. Scans through spiral CT are short enough to capture an image during a breath hold. Whereas the involuntary movement problem is solved

by some newly developed scanners that can synchronize with the heart rhythmic beat thus minimizing the motion [82]. Pixel-wise filter back-propagation methods are used to further minimize these effects [83, 84]. Lastly, spiral artifacts are produced during the interpolation process of continuous projection of individual slices. These artifacts can be minimized by adjusting the interpolation algorithm and scanning pitch [85].

With the advancement in deep learning algorithms, the trained convolution neural networks (CNNs) can classify any possible motion-perturbed and motion-free coronary cross-sectional images with the level of artifacts [74, 75]. Lossau etal. [74] develop a motion artifact model for supervised learning by producing synthetic data with distortions having scale between 0 and 10. Fantini etal. [75] use CNNs to remove fine grained artifacts. The low level layers of deep convolution neural network (DCNN) are used to extract useful information, which is then used to classify with the binary output. Similarly, several studies [86, 87, 88, 89] utilize supervised deep learning to identify artifacts. While unique evaluation metrics are used for the quantification of artifacts [90]. Due to image distortions, un-supervised reference-free deep learning models are in demand to process images.

Noise Removal

The most common type of noises in an image is a white Gaussian noise and an impulse (salt and pepper) with a normal distribution of zero-mean. These noises can be added independently to image pixels either during transmission or acquisition. The noisy image can be expressed as

$$I_n = I_\circ \eta_m + \eta_\alpha \tag{2.1}$$

where I_n is the degraded (noisy) image, $I_\circ$ is the original (noiseless) image, η_m and $\eta_\alpha \sim N(0, \sigma_N^2)$ are the pixel based multiplicative and additive noises respectively. Filtering techniques are mainly divided into spatial and variational methods. The former is categorized into nonlinear- and linear filters. Linear filter, such as Mean can reduce noises by applying weighted average technique. However, it over-smooths the image [91]. On the other hand, nonlinear filters such as Wiener [92], Median [93], and Bilateral [51], preserves structural properties with noise reduction. Other hybrid filtering techniques include diffusion- [94], adaptive- [72], non-local means- [95], Gaussian- [96] and spectral based [97, 98, 99, 100], which could retain structural properties with more clarity and less rippling artifacts.

Additionally, acquired data may contain speckle noise or artifacts due to the variable contrast dye resolution. Such noises and distortions can impinge diagnostic performance by making image details drastically hazy, unclear, thus demean image features [69, 101]. Hence, deep learning based models have been used with end-to-end training to minimize such distortions [102, 103, 104]. Jiao etal. [102] explore the structural residual information to restore distorted images. The residual image learning assist to recover fine details lost due to denoising. Yuan etal. [104] investigate spatial- and spectral multi-scale and multi-level features with deep neural networks (DNNs) to obtain denoised images. Moreover, generative adversarial networks (GANs) [105] have been used to produce high resolution synthetic images with minimal noise and artifacts

[77, 76, 78]. Acquired images may contain random noises and distortions with unknown distributions. Thus, un-supervised DNNs are in demand to restore images without loss of high-resolution information.

Deblurring

The deep learning- and traditional model-based techniques are used to perform deblurring. Deep learning-based methods consider DNNs with end-to-end [4] or adversarial [14] training scheme to produce deblurred images. Traditional model-based methods uses parameter estimation schemes based on regularization [106, 107, 108], maximum likelihood [109, 110, 111], singular value decomposition [112], etc., to process images. The basic blurring model is given by

$$B = I \otimes K \tag{2.2}$$

where $\otimes$ is a convolution operator, B is a blurred image, I is the latent or true image and K is a blur kernel or invariant point spread function (PSF) [113, 114].

Based on available information, deblurring is further categorized into blind- and non-blind deblurring problems. Several studies have been proposed to address an image deblurring with blur kernel estimation. These studies include blind deconvolution based on patch priors [115], residual layers [116, 117, 118], spectral irregularities [119], random transform [120] and CNN [121, 122]. In Xu etal. [106], a unified framework is proposed for uniform and non-uniform motion deblurring by using novel regularization technique with gradient based regularization function for blind deblurring. Fergus etal.[108] assumes a camera blur to be uniform, therefore Bayesian based maximum a-posterior model is proposed for the blur kernel estimation by probabilistic distribution. Levin etal. [107] proposes a non-blind prior model-based deconvolution with residual deconvolution enhancement technique to reduce blur, artifacts and noise with statistical blur kernel estimation.

Deep learning based neural networks have also showed impressive results for blind deblurring applications [123, 4, 124, 125, 74]. Kupyn etal. [14] apply a GAN to deblur natural images. Similarly, Tao etal. [126] consider muti-level neural networks (NNs), whereas Ye etal. [127] apply upscale-iterative networks to enhance deblurring performance.

Blind deblurring is common in practise, yet with some disadvantages for instance, artifacts in deblurred images, high computation cost due to ill-posed problem while non-blind deblurring can provide more stable and accurate solutions [117]. To this end, optical motion flow can provide some initial estimate of latent images [128, 129]. In addition, deep learning- or traditional model-based techniques can be utilized to further refine blur kernels and latent images with a reduced number of iterations [110]. Moreover, as discussed in studies [127, 126], a DNN with multi-channel and multi-scale image input can be used to improve performance.

Registration

US, MRI or CT modality is used to acquire biomedical images mostly in longitudinal directions. Such images consist of consecutive series of frames across each patient, each frame containing multi-parametric

information of tissues at higher temporal and spatial resolution [130, 131, 132, 133, 134, 135]. Meanwhile, the patients unintentional movements; heartbeat, respiration and other factors can misalign these images. Formally, image registration is divided into dichotomies as deformable vs. rigid, retrospective vs. prospective, intra- vs. inter-subject, inter- vs. intra-modality, inter- vs. intra-protocols and intensity- vs geometry-based image registration [54]. Registration can be defined as mapping between the two or more images, a challenging task in terms of intermix geometries due to complex spatial transformations and similarities. Thus, registration requires optimized and complex numerical schemes to properly overlay two images with necessary transformations. Commonly, affine transformation is applied to handle inter-images variations [84, 136, 137, 138, 139]. Affine transformation includes rotation, translation, shear and rescale parameters to compare and align two or more images.

Recently, deep learning models have been developed to minimize registration time and to improve accuracies, for example, deformable [140, 141, 142, 143], deformable and unsupervised [144], model-to-image [145] and 3-D pairwise medical image registration [146]. de Vos etal. [144] employ CNNs to perform the intensity-based registration by exploiting image similarity in end-to-end training. Once the network is trained it can be used to perform affine and deformable registration. Similarly, Balakrishnan etal. [146] apply an un-supervised learning model for deformable medical image registration. Once the network is trained, it can be used to perform registration without any supervised information. Hence, a semi-supervised well-trained CNN can speed up medical image analysis by performing registration more precisely and accurately.

Segmentation

Biomedical images acquired using US, MRI or CT are gray scale images with homogeneous background and ambiguous appearance. Therefore, for improving diagnostic performance, CAD algorithms demand quantification and localization of ROI [53, 147, 148]. Segmentation is mainly categorized into automatic, semi-automatic and manual segmentation. Semi-automatic techniques comprise of numerical energy-based optimization algorithms, require some initial user input for segmentation, methods include active-contour [149], graph-cut [150], intensity-based [148, 151, 152], and some other iterative techniques [49]. Automatic algorithms are user-independent approaches, perform segmentation based on supervised learning or numerical formulation. However, supervised learning requires enough training data for the generalized performance [153, 154, 155, 156]. Due to improved computation capabilities in recent years, neural networks have been able to give promising results [157, 158]. Table 2.1 illustrates automatic and semi-automatic liver- and lesion segmentation studies with the best outcomes highlighted across open access datasets. Likewise, MRI is rarely used for the lower abdomen image acquisition due to the large scanning time, metal implant constraint and some other factors. However, MRI has better spatial resolution compared to CT [65]. Literature [24, 157, 158, 159] has been reported for MRI liver segmentation. Furthermore, Table 2.3 illustrates, mostly manual segmentation preferred for US images. This is due to the fact of low contrast, ambiguous background and heterogeneous appearances of US images (further explained in Section 2.2.5).

11

Currently, liver and lesion segmentation algorithms are commonly evaluated on publicly available MRI or CT scan datasets, including Dircadb [7], LiTS [9], Silver07 [8], Chaos [160], and others. It is evident from Table 2.1 that study [161] performs consistently better for liver and lesion segmentation. Li etal. [161] explore intra- and inter-slice features with a hybrid feature fusion layer to improve the automatic segmentation. Similarly, Chlebus etal. [162] first apply 2-D CNN, and conduct the postprocessing with random forest to perform coarse to fine segmentation. In comparison to similar studies for LiTS dataset, this study outperforms in terms of relative volume difference (RVD), asymmetric surface distance (ASSD) and root mean square surface distance (RMSSD). Qin etal. [163] perform multinomial classification to classify pixels into liver: inner-region, boundary and background, and thus can explicitly identify boundary region. The method [164] works better for automatic lesion segmentation on Dircadb dataset, which utilizes multi-channel CNN to segment liver tumors for CT images. Network is trained for each phase while features are combined. Similarly, semi-automatic studies [165, 166] perform satisfactory. Zhou etal.[165] consider level set- and shape prior-based approach to segment liver. Wu etal. [166] implement fuzzy C-mean and graph cut technique to segment lesions in CT images.

Comparatively, semi-automatic studies are more prone to change in intensity or image distortions. Consequently, such methodologies perform over- or under-segmentation with hazy backgrounds. Therefore, model-based methods are considered favourable for small datasets with localized performance only.

The liver and lesion segmentation studies show various state-of-the-art methods [49], while neural network techniques prove to be more reliable for the segmentation of liver with variable size and shape [55]. However, deep learning-based segmentation requires large training datasets. To the best of our knowledge, so far proposed CAD systems lack precise and robust automatic segmentation. Disadvantages may include high computation cost with strong graphic processing units for automatic segmentation. While, semi-automatic methods such as: thresholding, region-growing, watershed etc., are fast and easy to implement with homogeneous images. However, they are user-dependent and highly susceptible to image noises. On the other hand, manual segmentation is a repetitive, time consuming and tiring task for medical application. Due to required trade off between precision, computation time and hardware, algorithms are in demand that could provide accurate un-supervised segmentation with reduced computational cost [179].

Multiple metrics are computed to evaluate segmentation performance; that includes, Dice- and Jaccard coefficient, RVD, Volumetric overlap error (VOE), ASSD, Maximum surface symmetric distance (MSSD), RMSSD, etc., [180].

2.2.3 Attribute Analysis

Feature Extraction

Development of CAD systems requires machine learning algorithms to train with discriminating features so that a system is capable enough to identify any abnormality with its optimal performance. To this end, the

Table 2.1: Metrics outcome for liver and lesion segmentation public datasets

Model	Dataset	Year	DICE(%)	VOE(%)	RVD(%)	ASSD(mm)	RMSSD(mm)
			Liver Segmentation				
Christ etal. [167]*	Dircadb	2016	94.3	10.7	-1.4	1.5	-
Lu etal.[168]	Dircadb	2017	95.09	9.36±3.34	0.97±3.26	1.89±1.08	4.15±3.16
Han[155]	Dircadb	2017	93.8±0.02	11.65±4.06	-0.03±0.06	3.91±3.95	8.11±9.68
Li etal. [161]	Dircadb	2018	94.7±0.01	10.02±3.44	**-0.01±0.05**	4.06±3.85	9.63±10.41
Ahmad etal.[169]	Dircadb	2019	91.83	**6.09**	5.59	-	-
Chen etal.[170]	Dircadb	2019	94.6±0.03	9.52±4.65	-0.02±0.07	8.43±9.37	14.21±5.71
Zhou etal. [165]*	Dircadb	2019	**96.53**	6.7±2.3	0.6±3.2	1.5±0.6	**2.4±0.8**
Budak etal. [171]	Dircadb	2020	95.22	9.05	7.03	**1.43**	-
Li etal. [161]	LiTS	2018	96.5	7.4	-1.8	1.45	3.15
Qin etal. [163]	LiTS	2018	**97.31±0.36**	**5.24±0.69**	1.97±1.7	1.77±0.49	-
Chlebus etal. [162]	LiTS	2018	96.50	7.7	**-0.4**	**1.150**	**2.42**
Chen etal.[172]	LiTS	2019	96.7	6.3	2	1.32	-
Hu etal. [173]	Sliver07	2016	**97.25±0.65**	5.35±1.23	**-0.17±1.34**	0.84±0.25	1.78±0.56
Dou etal.[174, 175]	Sliver07	2017	97.24	5.37±0.73	1.32±1.35	**0.67±0.12**	**1.48±0.21**
Lu etal.[168]	Sliver07	2017	96.96	5.9	2.7	0.91	1.88
Ahmad etal.[169]	Sliver07	2019	94.8	**4.31**	1.28	-	-
Shvets etal.[176]	Chaos	2018	**90.01**	**18.16**	-	-	-
Pham etal.[177]	Chaos	2019	73.45±5.93	41.95	-	-	-
			Lesion Segmentation				
Wu etal. [166]*	Dircadb	2017	83±0.06	29.04±8.16	2.2±15.88	**0.72±0.33**	**1.10±0.49**
Han [155]	Dircadb	2017	60±0.12	56.47±13.62	-0.41±0.21	6.36±3.77	11.69±7.60
Sun etal. [164]	Dircadb	2017	**91.54**	**15.6±4.33**	5.8±3.5	2.0±0.9	2.9±1.5
Li etal. [161]	Dircadb	2018	65±0.02	49.72±5.2	-0.33±0.10	5.293±6.15	11.11±29.14
Chen etal. [170]	Dircadb	2019	66.0	41.54±4.32	**0.16±5.03**	2.04±4.32	2.12±5.52
Budak etal. [171]	Dircadb	2020	64.3±34.6	-	-	-	-
Vorontsov etal. [178]	LiTS	2017	77.3	**35.7**	12.12	**1.075**	1.596
Li etal. [161]	LiTS	2018	**82.4**	36.6	**4.272**	1.102	**1.595**
Chlebus etal. [162]	LiTS	2018	79.6	38.3	46.4	1.143	1.728
Chen etal. [172]	LiTS	2019	68.4	43.8	11	25	-

Note: * indicates the semi-automatic approach; - indicates the corresponding metric is not reported.

feature extraction is formally categorized into morphological and hierarchical features [65, 181, 182, 183, 184].

Morphological Features

Morphological features are handcrafted and computed based on transform domain, signal processing and statistical property of an image [3]. Table 2.2 illustrates various models common in practise for morphological feature extraction. These models are explained as follows.

Statistical-based Models Statistical methods are primarily categorized into the first- and second order features. The first order features provide various statistical properties related to gray-level distribution with in an image, ignoring neighboring pixel relations or spatial properties. These parameters include mean, peak, slope, kurtosis, variance, standard deviation, skewness, and others [185, 186]. Alternatively, the second order features rely on co-occurrence properties within neighboring pixels. These features characterize spectral qualities related to intensity of an image, for instance, gray-level co-occurrence matrix (GLCM) and geometrical properties based texture features, etc. [27, 34, 186, 29]. Haralick etal. [187] propose texture features for the very first time and explain the inextricable relationship between texture and tone of an

image. They explore spatially dependent features at 0, 45, 90 and 135 degree of a symmetrical image. The 0- and 90-degree angle considers neighboring pixels, horizontally and vertically with N_x rows and N_y columns, providing $2N_y(N_x - 1)$ number of pixels for 0 degree and $2N_x(N_y - 1)$ symmetrical for 90 degree angle. The 45- and 135-degree angle provides right- and left-diagonal neighboring pixels making total number of $2(N_y-1)(N_x-1)$ pixels for feature extraction. These spatial dependence matrices are then extracted in terms of homogeneity, local variation and linear dependency present between pixels within an image. The texture features are widely explored with multi-contrast and multi-model input data that includes GLCM, moment invariant (MI), gradient (GR), gray-level difference matrix (GLDM) etc., [47, 59, 65, 182, 188, 189, 190, 191]. Table 2.3 illustrates GLCM based texture features widely explored for multi-class predictions.

Similarly, few studies [70, 192] explore geometrical properties based texture features that provides information regarding object profile and physical structure. Mostly, the benign tumors have smooth profiles, while the malignant lesions often contain irregular shapes [193]. Therefore, spatial features, for example, compactness, solidity, ROI area, radial length, circularity, rectangularity etc., provide valuable distinct information for lesion classification.

Processing-based Models   Various sizes of masks are convolved with an image to quantify texture information. Laws' etal. [194] derive such features for the first time by convolving 5×5 size mask with an image. Since then Laws' features have been used with various sizes to explore valuable features [1, 195, 196]. Table 2.2 illustrates various sizes of masks used for Laws' texture feature extraction.

Frequency-based Models   Several transform models have been proposed that are based on frequency-domain features [197]. These are claimed to be effective than spatial-based features due to the fact that noise is reduced while retaining image quality [198]. Some common models include Gabor filter, Wavelet-, Ranklet-, and Contourlet transform, etc., [199, 200, 184].

Hierarchical Features

Recent studies [32, 211, 212, 213] have utilized various DNN models to extract meaningful features. NNs uses back-propagation technique to adaptively learn best combination of spatial features for desired number of outcomes. In the back-propagation, the weights are fine-tuned by minimizing the errors with multiple runs based on defined loss function. Since the obtained features are optimal, no further selection criteria is required. However, such algorithms demand large input data to generalize performance and to avoid overfitting.

Feature Selection

Selection criteria is performed to choose input features that have strong statistical relationship with the output. Multiple studies [21, 182, 214, 215, 209] have utilized statistical tests, recursive eliminations, genetic algorithms and some other methods, to obtain the set of valuable features.

Table 2.2: Developed models for hand-crafted features extraction [1, 2, 3]

Method	Extracted Features
Statistical-based Models	
Gray-level Co-occurrence Matrix (GLCM) [34, 187, 201, 202]; Gray-level Difference Matrix (GLDM) [203, 204]	Correlation information, Angular second moment, Inverse difference moment, Sum of square, Gradient, Sum of variance, Sum of average, Sum of entropy, Contrast, Difference entropy, Difference variance etc.
Gray-level Run Length Matrix (GLRLM) [27, 73, 205]	Gray-level (GL) run length, GL run number, Run emphasis, GL non-uniformity,High-, low-, short- and long GL run emphasis etc.
First Order Statistics (FOS) [34, 29, 201, 206, 202]	Mean, Peak, Slope, Kurtosis, Variance, Standard deviation, Skewness etc.
Geometric features [27, 207]	Coarseness, Periodicity, Roughness, Eccentricity, Regional area, Compactness, Margin, Circularity, Rectangularity, Periphery, Border, Width, Depth etc.
Processing-based Models	
Laws' Mask Analysis [194, 1, 208, 185, 195, 196]	Laws' mask of length range [3-9]
Frequency-based Models	
Transform domain [73, 209, 189, 210]	Fourier power spectrum, Wavelet transform, Gabor wavelet transform, Contourlet transform, Shearlet transform, Ranklet transform etc.

Statistical tests are used to select features based on intrinsic relationship between each input and output variable. These tests are considered valid with some prior assumptions regarding data distribution. Additionally, the features are selected based on independent mean and low variances [216, 217].

Recursive feature elimination (RFE) [182, 188] selects a subset of features based on model performance. The weakest features are eliminated by recursively running the model and evaluating on cross-validation scores. Due to the iterative process, RFE demands high computation cost for larger datasets. Therefore, highly correlated features should be eliminated before applying RFE method for the feature selection.

Genetic algorithm [218] is another feature selection method that performs crossover and mutation to find the optimal chromosomes. Every feature is considered as a chromosome, while each gene contains distinct information such that combination of selective features according to the required feature space. An individual genotype is selected based on a very criterion of survival of the fittest. Multiple crossovers are performed to recombine the best chromosomes and mutation for new structures from various genome.

Moreover, some other techniques, for instance, Grasshopper or Swarm particle optimization have also been used for the optimal feature selection [34, 211, 209, 219].

Dimensionality Reduction

Dimensionality reduction is sometimes required to transform high dimensional input data into a lower dimensional space due to the fact that the number of features should be less than the number of training data

points [69, 71, 34, 200, 193, 204]. High dimensional data can lead towards the overfitting. The principal component analysis (PCA) and linear discriminant analysis (LDA) are used as unsupervised- and supervised data transformation techniques [23, 220, 195]. Additionally, independent component analysis (ICA) technique is used for the data with non-gaussian distribution [221].

2.2.4 Machine Learning Systems

After features extraction and selection, a machine learning model is required to train and classify given inputs into desired outcomes. Previous studies [56, 57, 58, 69, 193, 222, 213] utilize multiple machine learning algorithms that commonly include k-nearest neighbor (k-NN), support vector machine (SVM) and DNN in various forms.

k-Nearest Neighbor

It is evident from the name k-NN, a non-parametric classifier based on k-nearest neighbors. With a pre-specified value of k, the training points are assigned to a class according to the majority of k nearest neighbors. The distance between data points is typically calculated by Euclidean or Mahalanobis distance matrices [204, 215, 223].

Support Vector Machine

Support vector machine (SVM) intends to maximize the distance between the separating hyperplane and support vectors, so that the classes are best separable. In the feature space, the linear separating hyperplane is defined as [224],

$$y = l^T x_i + d \tag{2.3}$$

where x_i is a feature vector of a binary class, d is an intercept and l is the norm vector of the separating hyperplane for the two classes. The cost function of l is given by

$$S(l, k) = \frac{1}{2} \|l\|^2 + C \sum_{i=1}^{E} k_i \tag{2.4}$$

Subject to:

$$y_i \times (l^T x_i + d) \geq (1 - k_i) \qquad k_i \geq 0 \tag{2.5}$$

where C is the regularization parameter, k_i indicates the classification errors, E represents the number of misclassified samples and y_i denotes the class label [224, 225, 226, 227]. Table 2.3 illustrates SVM commonly used to classify various diseases with in a liver.

Neural Networks

Neural Networks (NNs) are biological brain-inspired architectures, comprising of multi-layer neurons to find the optimal hierarchical features based upon back-propagation error minimizing technique. Neurons with

the activation function σ and parameters ξ are used to evaluate output. Parameters $\xi = \{b, w\}$, where b is a bias and w is a weight. An activation function transform a linear combination of the input x with neuron parameters into a neuron output as follows: $\alpha = \sigma(w^T x + b)$, where α is the neuron output. Typical activation functions include sigmoid $1/(1+e^{-x})$, rectified-linear unit $max(0, x)$, hyperbolic tangent $(e^x - e^{-x})/(e^x + e^{-x})$, and others. A multi-layer perceptron is a well-known neural network that contains several layers of these transformations, given as [143]

$$F(x; \xi) = \sigma(W^L \sigma(W^{L-1} \cdots \sigma(W^\circ x + b^\circ) + b^{L-1}) + b^L) \tag{2.6}$$

where W^n is a weight matrix associated with the k activation output, n denotes the number of current layers, here L represents the final layer. Layers in between input- and output layers are known as hidden layers. A neural network with multiple hidden layers is called a DNN. Such architectures have been used in various forms; for instance, AlexNet [228], GoogLeNet [229], ResNet [10], GAN [105] and others [230, 231]. Fig.2.2 shows schematic of feed-forward neural network for x number of input features at input layer with L-hidden layers having k-perceptron in each layer and y-number of outcomes at the output layer. More often, the final layer activation is mapped to a distribution over classes $P(y \mid x; \xi)$ by Softmax function defined as [143]

$$P(y \mid x; \xi) = Softmax(x; \xi) = \frac{e^{(W_i^L)^T x + b_i^L}}{\sum_{k=1}^{K} e^{(W_k^L)^T x + b_k^L}} \tag{2.7}$$

where W_i^L represents the weight vector associated with the class i. Commonly, maximum likelihood with stochastic gradient descent is used to learn parameters ξ to a dataset D. In the stochastic gradient descent, a mini batch is utilized to update the gradient. Maximum likelihood accounts to minimize the negative log-likelihood given as

$$\underset{\xi}{argmin} - \sum_{n=1}^{N} log \left[P(y_n \mid x_n; \xi) \right] \tag{2.8}$$

It results in the multi-class problem by categorical cross-entropy loss or two-class task by binary cross-entropy loss [143]. For information on various DNN models, readers are encouraged to read the literature [232, 233].

NNs have advantage in terms of feature extraction and selection. However, large input data is required for the model to be well trained. Alternatively, transfer learning is performed to address such limitations [54, 234]. Moreover, an entire deep learning model is often uninterpretable. Consequently, proposed DNN models work as a black box for desired classification. An explainable artificial intelligence is required to explain in the case of misclassification or perfect classification. Secondly, image preprocessing such as intensity adjustment or normalization is required in health care applications to make uniformity between multiple subjects data. Finding the correct preprocessing steps is still challenging as it requires a lot of data and longer training to obtain an effective classification model. Lastly, DNNs can be easily fooled by some minor changes to input samples; for example, adding imperceptible noise in an image can cause NN to misinterpret the samples. Thus, machine learning models are susceptible to random manipulations caused by various kind of noises and artifacts [235].
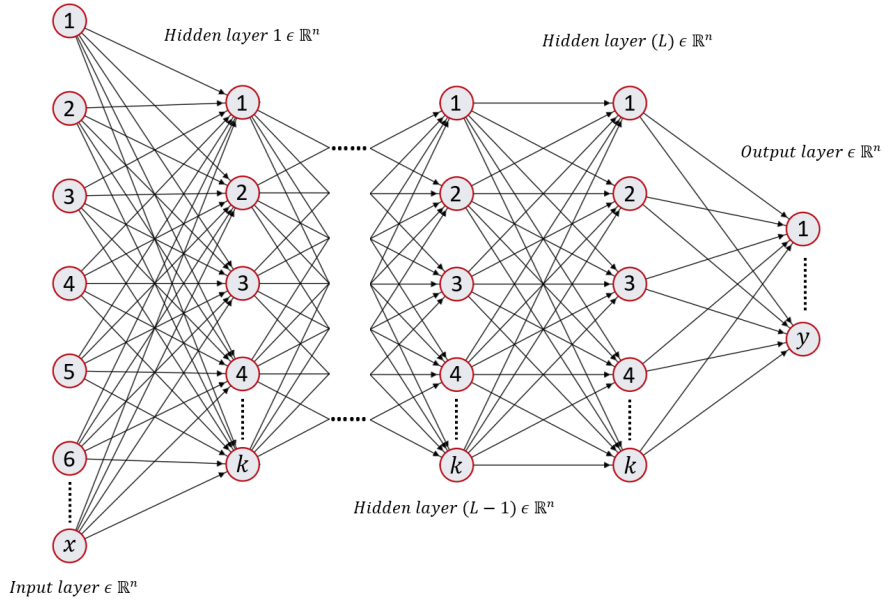
Figure 2.2: Schematic of the feed-forward neural network

Machine learning algorithms are evaluated by some universal quantitative assessment metrics. These metrics are computed mainly based on true positive, true negative, false positive and false negative rate. Such metrics include accuracy, sensitivity, specificity, precision, area under the curve and some others [69].

2.2.5 Computer-Aided Diagnosis (CAD)

Currently, clinical liver cancer diagnosis is performed manually, based on analysis of acquired scans and pathological data. This procedure is tiresome and demands multiple specialties. Therefore, CAD systems are designed to assist in detailed and repetitive diagnosis. The liver disorder can be classified into focal- and diffuse liver diseases. The former type consists of severe liver morbidities that can be further divided as benign (BE), such as hemangioma (HEM), focal nodular hyperplasia, hepatocellular adenoma, cyst (CY), etc., and malignant, such as hepatocellular carcinoma (HCC), metastatic carcinomas (MET), intrahepatic cholangiocarcinoma (CC), etc. In contrast, the latter covers virus hepatitis, fibrosis (FB), cirrhosis (such as alcoholic cirrhosis (CR)), non-cirrhotic chronic liver diseases (such as fatty liver (FL) or steatohepatitis), etc., [3]. The CAD systems are discussed with comparative and critical in-depth analysis as follows.

Primary Analysis

Various studies [70, 73, 204, 32, 219, 31, 208] have utilized machine learning models to classify acquired images into BE, HEM, CY, FB, CR, primary liver cancer (HCC), and secondary liver cancer (MET). Table 2.3 shows some previously developed CAD systems using US, MRI and CT modality. The best performed studies are bold faced for multiple number of disease detection outcomes. Although several studies have proposed novel morphological features for machine learning algorithms training, the overall best performances are observed

for DNN architectures. Commonly, texture features based on GLCM have been proposed. Alternatively, few studies have used CNN based hierarchical features instead of morphological features, to train machine learning algorithms [70, 212, 211]. We have briefly discussed the best studies with respect to each modality and the number of classes in the following paragraphs.

As illustrated in Table 2.3, US is comparatively an old modality, widely used for liver cancer diagnosis. Subramanya etal. [3] propose a binary class CAD system for malignant liver disease detection. The algorithm is designed to classify between diffuse liver diseases such as fatty liver and cirrhosis. Triglyceride fats accumulate to build-up vacuoles in the hepatic cells that is commonly known as fatty liver diseases (FLD). FLD (if not treated) can lead towards fibrosis, while excessive fibrosis can take the condition of cirrhosis. Therefore, they focus on early detection and distinction between FLD and cirrhosis. To this end, radiologists suggest features related to coarse texture, shrinkage of liver size and nodularity for the desired binary outcome. Hence, combination of GLCM, gray-level run-length matrix (GLRLM), first order statistics (FOS), Laws' and gradient based features are used to extract valuable information. Afterwards, k-NN and Nave Bayes are applied with maximum possible combinations and multiple iterations to select a subset of features. The experiment shows FOS and Laws' based features to be optimal for characterization. SVM is then applied to perform binary classification with an accuracy of 99.5%. Considering only the US modality, this methodology outperforms among the discussed studies for the binary classification. Contrary, small dataset of only 29 images is used to train the model. Likewise, feature selection is performed by various machine learning models. Since data type may vary, while feature selection by such a methodology can not guarantee the optimal performance for other models. Hwang etal. [195] comparatively show better performance for the three class problem. The three outcomes, CY, HEM and malignancy (MA) are classified by considering a binary class at a time such as CY vs HEM, CY vs MA and HEM vs MA. GLCM, FOS and Laws' based features are extracted, while the dimensionality reduction is performed using PCA. Thereafter, ANN is used to classify binary class data with total number of 99 focal liver images. The trained model performs with an accuracy of 99.7% for the three-class outcome. Binary classification is performed sequentially to full-fill three-class task. Thus, more computation time is required to solve the problem. Bharti etal. [204] propose the combination of three machine learning models k-NN, SVM and Random forest (RF) to perform voting based classification. Euclidean distance based relevant feature selection and RFE are used to select discriminant features among GLCM, gray-level difference matrix (GLDM) and ranklet transform with total number of 754 images to classify four liver stages. The model performs with the accuracy of 96.6%. Assembling multiple models together highly influence later models performance. Moreover, the disadvantages account for the large computation time. Similarly, Acharya etal. [209] report the best performance for 5-class outcome. Contourlet- and Fourier transform based features are extracted. While ANOVA statistical test is performed to select significant features. Later, Probabilistic Neural Network (PNN) is used to give the prediction with the accuracy of 91.46%. Synthetic data generation is also performed to address the imbalance dataset problems.

MRI has relatively better spatial resolution, but due to high operational cost, longer time consumption,

non-portability etc., its rarely used for lower abdomen disease diagnosis. Xia etal. [70] propose 2-class CAD system based on 3-D and 2-D MRI image data experimentation. The CNN is used to perform binary classification, while SVM is also trained with hierarchical features extracted from the CNN. Four different input methodologies are adopted; for instance, 3-D and 2-D input images in patch and slice form. The best results are obtained by utilizing the 2-D segmented ROIs. 1023 MRI scans are used to train the network for a testing accuracy of 95.5%. 2-D CNN consist of shallow neural network that consist of only four strided convolution layers.

CT is currently on focus due to its rapid scanning and short diagnosis time. Kutlu and Avcı [32] apply the AlexNet neural network to extract features from total number of 112 images. The discrete wavelet transform (DWT) is then applied to reduce vector dimensions. Later, recursive neural network (RNN) based long short-term memory (LSTM) networks are utilized to perform binary classification between benign and malignant liver with an accuracy of 99.1%. Alternatively, the limitation of data is observed that can result in overfitting, however balanced training dataset is preferred to avoid any bias results. Das etal. [219] utilize the watershed segmentation algorithm to segment the ROIs. Afterwards, FOS, GLCM-based texture and geometric features are extracted to train ANN. However, only a three-layer shallow network is proposed for classification. The model performs with an accuracy of 99.39% for the three-class; MET, HCC and HEM. A total number of 225 images are used in the analysis. Likewise, Wang etal. [236] use residual CNN to carry out classification among four-class data: CY, focal nodular hyperplasia (FNH), HEM and HCC. Dataset has total count of 388 images, while the accuracy of the prediction is 91.22%. To overcome data limitation problem, transfer learning with fine tuning is preferred. Balagourouchetty etal. [36] consider six class problem for CAD. The ROI is extracted using region growing segmentation technique. While GoogleNet is considered with the leaky Rectified Linear Unit (ReLU) activation function and some NN changes to enhance performance. Kayaalti etal. [182] apply SVM and k-NN to classify seven possible stages of FB by pair-wise classification. The feature extraction is performed by GLCM, DWT and Fourier transform. After extraction, features are normalized and selected based on sequential forward selection, thus resulting in the accuracy of 90% for 116 CT scans. The class is assigned on basis of the highest vote scheme. In case that more than one class has the same highest votes, the class is assigned based on the highest prior probability. The disadvantage includes bias outcome due to imbalanced dataset. Similarly in [28, 206, 237, 238], various algorithms have been proposed, however due to inefficient segmentation capabilities, increased noise in scans and small training datasets, proposed CAD systems are still non-reliable for clinical applications [60].

Secondary Analysis

Studies [239, 240, 241, 242] have opted variant methodologies to accomplish clinical diagnostic tasks. After preprocessing, liver disease diagnosis requires unique features to classify various diseases. Malignant liver scans sometimes have very small lesions that tend to have homogeneous backgrounds, and thus make it difficult for machine learning models to classify with good precision. Mainly FOS, geometric and GLCM

based texture properties are commonly explored for valuable feature extraction among data from the three various modalities (US, MRI and CT) [70, 208, 186, 187]. However, few studies have also utilize wavelet transform (WT), kinetic curves and some other models to extract novel and distinct features [56, 59, 182, 185, 191, 211, 238, 243, 183, 244, 245]. WT varies spatial resolution to represent suitable scale of textures. Frequency-domain methods give more reliable distinct information than spatial-based methods [22]. Also, NN features have been explored to enhance machine learning algorithms performance. Additionally, in most of the reported cases, liver cancer is diagnosed over the age of 75 with larger count of males [15]. Hence, integrating some general pathological details such as gender, age, overall stage of cancer, size of largest tumor, etc., can also assist in developing more powerful predictive algorithms in terms of mortality and reoccurrence [240].

Feature selection plays a vital role in the development of reliable classification models. Various studies have explored multiple algorithms that includes particle swarm optimization (PSO), recursive forward- and backward selection, statistical tests, genetic algorithm, etc., to select distinct features that have lower inter-class and higher intra-class distances [25, 219, 240, 245]. Thus, multiple selection criteria are encouraged to find optimal combination of features. Moreover, in parallel to the selection criteria, considering importance of features from clinical perspective can assist in development of reliable prognostic models.

Finally, model selection is required to accomplish classification task with optimal performance. Depending on the input data characteristics, for instance, linear or nonlinear classification model is chosen with hyperparameters tuning for training, validation and testing. Table 2.3 illustrates k-NN and SVM machine learning models widely used for predictions. However, advancement in computational machine capabilities, deep learning based neural networks have been widely used in recent years for clinical diagnosis [22, 158, 234].

Due to wide availability, short scanning time and low cost, US imaging is used for general diagnosis. However, due to dependent user input for scanning, low-contrast, low signal-to-noise ratio and low spatial resolution, CT and MRI are more common in practise. Overall best performance for CAD is achieved by NNs in various forms, for example, ANN, CNN and RNN. Such DNN requires large training data to avoid over-fitting with the generalization performance. Yao etal. [240] perform binary classification with comparatively large datasets. Contrary, to the best of our knowledge, most of the studies have utilized small and imbalanced datasets for multi-class problems [246]. Moreover, for multi-class prediction, often studies have either proposed pair-wise classification or assemble-models based voting criteria. The former technique performs sequential binary classification to predict the outcome. Whereas, the later methodology uses cascaded- or multiple fused models with weighted outcome. In such cascade scheme, each model performance is influenced by the prior model output, with exponential increase in computational cost. Likewise, due to small available datasets, studies have used shallow NNs. Such short falls can be avoided by utilizing transfer learning with fine tuning [234, 247, 248]. Furthermore, accuracy metric has been frequently used for performance evaluation in the proposed studies. Whereas, accuracy metric may not be an optimal choice in the case of imbalanced datasets. Therefore, robust diagnostic models with accurate and reliable prediction are required for clinical

diagnosis.

2.3 Future Perspectives

Studies (discussed in Section 2.2.5) have been proposed to classify liver cancer mainly into primary- and secondary cancer. Primarily, data limitation has been observed for development of reliable CAD. Meanwhile, to address such problems, GAN is proposed in various forms to generate synthetic data [249, 250]. Frid etal. [245] utilize GAN to address similar data limitation problem by generating synthetic data for CT scans data augmentation. Likewise, GAN is also used to remove noise and deblur images [14, 251]. However, it has not been explored for biomedical applications. Moreover, human liver is mainly divided into five lobes. These lobes are further sub-divided into eight segments. The identification of malignant segments in the liver scan is time consuming and require experts knowledge. Similarly, lesion size estimation can assist in the cancer stage prediction. It is envisioned that estimating tumor size and segments of the liver are novel tasks that could help to predict the stage of cancer and reoccurrence rate with procurement steps such as chemotherapy, surgery, etc., [38, 27, 207]. Furthermore, multimodal image input data with multi-scale, multi-resolution, multi-contrast and multi-phase can assist in predicting more reliable outcomes [190, 252].

2.4 Conclusion

In this study, we investigate the potential of CAD system for malignant liver detection and treatment. CAD systems can act as a second opinion tool to determine prognostic factors. The existing literature is overviewed up to date in focus with the conventional image acquisition modalities namely; US, MRI, and CT. Acquired images may contain noise, artifacts or blurriness. Accordingly, preprocessing is performed to remove artifacts and further segmentation to select the ROI. Afterwards, features are extracted commonly based on texture, geometry and statistical properties of gray-level co-occurrence matrices. Significant features are selected by various statistical and iterative algorithms. Alternatively, deep learning networks have also been utilized to provide reliable results with transfer learning and fine tuning. Imbalance- and small datasets are some of the pit falls that can be solved by synthetic data generation. Discussed state-of-the-art algorithms utilize various methodologies to fulfil diagnostic task. To the best of our knowledge, there has been no consensus on development of optimal CAD systems. However, we believe that considering patients' history, multiple imaging findings and pathological data can assist to develop reliable diagnostic systems. Moreover, follow-ups can assist to prognosticate stage factors as well in more reliable form.

Table 2.3: Summary of malignant liver computer aided diagnosis systems

Study	Year	# of I/P	Seg	Feature Extract/Select	Classifier	Outcomes	Accuracy%
				Ultrasonography Imaging			
				Focal Liver Disorder			
Horng [253]	2007	I-120	M	GLCM/COR	RBF	NOR, CR, CLD	92.5
Mittal etal.[206]	2011	I-111	M	-	ANN	NOR,HCC,MET,CY,HEM	80.2
Mitrea etal.[192]	2012	I-540	M	GLCM, GR, GE/COR	Adaboost&NN	CR, HCC	81.8
Wu etal.[215]	2012	I-432	M	GLCM,WT/GA	PNN	NOR, CR, HCC	96.6
Streba etal.[254]	2012	P-112	M	GL Intensity/-	ANN	HCC, FL, HEM, MET	94.45
Ribeiro etal.[255]	2012	P-115	M	WT/SFS	k-NN,SVM	2-class (NOR,HCC,CR)	85.59
Virmani etal.[200]	2013	I-56	M	WT/GA	SVM	NOR, HCC, CR	88.8
Lee [256]	2013	I-432	M	WT/-	k-NN& SVM& PNN	NOR, HCC, CR	95.49
Virmani etal.[257]	2013	I-54	M	GLCM, GLRLM, FT/GA	SVM	MET, HCC	91.6
Wu etal.[258]	2014	P-22	M	-	DNN	BE, MA	86.6
Hwang etal.[195]	2015	**I-99**	M	**GLCM, FOS, Laws'/PCA**	**ANN**	**2-Class(CY,HEM,MA)**	**99.7**
Gatos etal.[259]	2015	I-52	M	GL Intensity/-	SVM	BE, MA	90.3
Hassan etal.[71]	2015	I-110	M	GLCM, GLRLM/PCA	SVM	CY, HCC, HEM	96.5
Shan etal.[214]	2016	P-83	M	FOS/Chi-SQ	LR	MA, BE	84.5
Kondo etal.[21]	2017	P-98	M	GL Multi Phase/RFE	SVM	BE, HCC, MA	91.8
Guo etal.[260]	2017	I-93	M	GLCM, FOS/COR, SFS	SVM	BE, MA	90.41
Acharya etal.[209]	**2018**	**I-762**	M	**Contourlet Transform/ANOVA**	**PNN**	**4-Stages Cancer**	**91.46**
Bharti etal.[204]	**2018**	**I-754**	**AC**	**GLCM, GLDM, RK/RFE**	**k-NN& SVM& RF**	**NOR, HCC, CR, CLD**	**96.6**
Xu etal.[188]	2019	I-79	M	GLCM, GLRLM/SFS, SBE	SVM	HCC, CR	88.87
Yamakawa etal.[30]	2019	I-338	M	-	CNN	CY, HCC, HEM, MA	88
Pan etal.[261]	2019	I-93	M	-	3-D CNN	NOR, BE	93.1
				Diffuse Liver Disorder			
Riberio etal[262]	2013	I-148	M	GLCM, WT/REG	SVM& k-NN	NOR, CR	98.67
Singh etal.[263]	2014	I-180	M	GLCM, GLDM, FOS, FT/Z-score	Linear Classifier	NOR, FL	95
Santos etal.[223]	2014	I-162	M	GLCM, GLRLM, FOS/REG	ANN	NOR, FL	79
Kalyan etal.[205]	2014	I-60	AC	GLCM, GLRLM/-	ANN	NOR, FL, CR, HEP	92.5
Acharya etal.[264]	2016	I-100	M	Gabor filter/Wilcoxon, t-test	PNN	NOR, FL	98
Liu etal.[211]	2017	I-91	M	CNN/-	SVM	NOR, CR	86.9
Meng etal.[22]	2017	I-279	M	-	CNN	NOR, FB	93.9
Krishman etal.[73]	2017	**I-364**	**AC**	**WT/-**	**RF**	**10-Class**	**91**
Sharma and Juglan [2]	2018	I-90	M	GLCM, FOS, GE/NHA	SVM	NOR, FL	95.55
Reddy etal.[29]	2018	I-250	M	-	CNN	Four FL grades	93.5
Acharya etal.[265]	2018	I-140	M	WT/ANOVA, PSO	PNN, SVM, k-NN	FL, CR	93.73
Subramanya etal.[3]	**2019**	**I-29**	M	**Hand-craft/RFE**	**SVM**	**FL, CR**	**99.5**
Aggarwal etal.[239]	2020	I-110	M	GL-Intensity/FISCR	SVM	NOR, CR	98.81

(continued on next page)

Table 2.3 *(continued)*

Study	Year	# of I/P	Seg	Feature Extract/Select	Classifier	Outcomes	Accuracy%
				Magnetic Resonance Imaging			
				Focal Liver Disorder			
Meng etal.[266]	2014	I-200	M	Histogram-based/-	SVM	BE, MA	86.67
Ben-Cohen etal.[241]	2017	I-192	M	FOS/-	SVM	PL, MET	90.7
Trivizakis etal.[24]	2019	**I-130**	**M**	-	**2-D patch CNN**	**PL, SL**	**95.50**
Wang etal.[31]	2019	I-60	M	-	CNN	NOR, LE	89.7
				Diffuse Liver Disorder			
Lee etal.[199]	2006	I-146	M	Gabor filter/-	SVM	CY, HCC, HEM	83.95
Yasaka etal.[267]	2018	P-634	-	-	CNN	2-Class (5-FB Stages)	86.33
				Computed Tomography Imaging			
				Focal Liver Disorder			
Kumar etal.[210]	2013	I-300	AC	GLCM, WT, CCT/-	PNN	HCC, HEM	96.7
Adcock etal.[268]	2014	I-132	M	Matching Metric/-	SVM	BE, MA, CY, HEM	80.56
Xia etal.[70]	2016	I-1023	Threshold	GLCM, FOS/-	RF	NOR, HCC	80.71
Ravi etal.[235]	2017	I-62	CL	GLCM, FOS/GWO	SVM	NOR, BE	97
Chang etal.[238]	2017	I-71	M	GLCM, SH, KC/BKE	LR	BE, MA	87.31
Vivanti etal.[269]	2017	I-37	CH	-	CNN	NOR, BE	-
Krishna etal.[270]	2017	I-80	AC	GLCM/-	SVM	NOR, MA	92.5
Yu etal.[23]	2018	I-147	M	GLCM,FOS/PCA	SVM	NOR, HCC	86.36
Ben-Cohen etal.[231]	2018	P-20	CNN	-	FCN	NOR, BE	94.6
Sreeja etal. [271]	2018	I-20	M	GLCM, FOS/-	SVM	NOR, HCC	90
Kuo [272]	2018	I-200	M	GLCM, GLRLM, FOS/PSO	SVM	BE, MA	80
Yu etal.[23]	2018	I-147	M	GLCM, FOS/PCA	SVM	NOR, HCC	87.93
Liang etal.[212]	2018	I-480	GC	CNN/-	SVM	CY,FNH,HEM,HCC	87.58
Wang etal.[236]	2018	**I-388**	**CNN**	-	**CNN**	**CY,FNH,HEM,HCC**	**91.22**
Midya etal.[273]	2018	P-223	TH	-	CNN	HCC, CC	69.72
Yasaka etal.[242]	2018	I-1068	M	-	CNN	HCC, LE, HEM, CY	84
Anter etal.[274]	2018	I-762	TH	FOS, GLCM, Geometric/-	k-NN	NOR, MA, BE	92.6
Frid etal.[245]	2018	GAN	-	-	CNN	MET, HEM, CY	85.7
Nayak etal.[181]	2019	I-726	AC	GL-Multi Phase/-	SVM	NOR, CR, HCC	86.90
Das etal.[219]	**2019**	**I-225**	**WT**	**GLCM, FOS, GE/-**	**ANN**	**MET, HCC, HEM**	**99.38**
Doğantekin etal. [275]	2019	I-200	-	DWT, SVD, CNN/-	SVM, ELM, k-NN	BE, MA	97.3
Kutlu and Avcı [32]	**2019**	**I-112**	**M**	**CNN/-**	**RNN**	**BE, MA**	**99.1**
Chen etal.[276]	2019	I-1091	-	-	CNN	BE, MA	86.90

(continued on next page)

Table 2.3 (continued)

Study	Year	# of I/P	Seg	Feature Extract/Select	Classifier	Outcomes	Accuracy%
Romero etal. [277]	2019	I-230	-	-	CNN	MET, BE	96
Muthuswamy etal. [278]	2019	I-60	TH	GLCM/-	SVM	NOR, MA	95
Das etal.[33]	2019	I-123	CL	DFT/-	DT	HCC, MET	95.8
Remukadevi and Karunakaran [34]	2020	P-583	M	-	DBN	NOR, CR, HCC	98
Yao etal.[240]	2020	I-76914	M	13 LFT PATH/-	DNN	NOR, MA	-
Li etal.[35]	2020	I-165	CNN	-	CNN	3-Class Tumor	98.9
Balagourouchetty etal.[36]	**2020**	**I-634**	**AC**	-	**CNN**	**6-Class**	**89.58**
				Diffuse Liver Disorder			
Mala and Sadasivam [279]	2006	I-100	M	WT/-	PNN	FL, CR	95
Kayaalti etal.[182]	**2012**	**P-116**	**M**	**Hand-craft/SFS**	**SVM**	**5-FB Stages**	**90**
Chen etal.[280]	2013	I-44	M	SSM/COR	SVR, NN	NOR, CR	90

Note: - denotes the missing information; I/P-# denotes the number of images or patients reported in a study; ANN=Artificial neural network; AC= Active contour; BE= Benign; BKE= Backward elimination; CH= Chan-vese; CCT= Contourlet coefficient texture; CLD= Chronic liver disease; CNN= Convolution neural network; CR= Cirrhosis; DT= Decision tree; DFT= Discrete fourier transform; DWT= Discrete wavelet transform; DNN= Deep neural network; DBN= Deep belief network; ELM= Extreme learning machine; FT= Fourier transform; FL= Fatty liver; FDCT= Fast discrete curvelet transform; FNH= Focal nodular hyperplasia; FISCR= Fisher criterion; FOS= First order statistics; GC= Graph-cut; GE= Geometric; GL=Gray-level; GR= Gradient; GLDM= Gray-level difference matrix; GLCM=Gray-level co-occurrence matrix; GLRLM= Gray-level run length matrix; GWO= Gray wolf optimization; FB=Fibrosis; HEP= Hepatomegaly; HCC= Hepatocellular carcinoma; HEM= Hemangiomas; I=Images; CC= Cholangiocarcinoma; KC= Kinetic Curve; LFT= Liver function test; LR= Logistic regression; LE= Lesion; MET= Metastasis; MLP= Multi-layer perceptron; M=Manual; MA=Malignancy; NOR= Normal; P=Participants; PNN= Probabilistic neural network; PATH= Pathology; PL= Primary liver cancer; RK= Ranklet; RFE= Recursive feature elimination; RNN= Recurrent neural network; RBF= Radial base function; RF= Random forest; REG= Regression; SL= Secondary liver cancer; SEG= Segment; SSM= Statistical shape model; SBE= Sequential backward elimination; SFS= Sequential forward selection; SVR= Support vector regression; SVM= support vector machines; TH= Threshold; TEX= Texture; WT= Wavelet transform; WAT= Watershed

3 Multi-scale GAN with Residual Image Learning for Removing Heterogeneous Blur

As discussed in Chapter 2, the CT scans can be acquired with different phases. The contrast enhanced scans are acquired by injecting a contrast agent to enhance tumor pixels before the acquisition of images. However, this enhancement phase and laser light with various focal lengths could result in artifacts and increased noise in images. The deconvolution technique is applied to process such distorted images. However, it may create ripples or rings in an output image. In this chapter, I develop an unsupervised neural network for processing non-homogeneous blur with artifacts. Furthermore, residual image learning is performed to recover some fine details lost in processing the images. The image processing studies mostly compared their performances to natural images. Therefore, I train and validate our proposed method on natural images. This chapter accomplishes Objective 2 of this dissertation.

Abstract

Processing images with heterogeneous blur remains challenging due to multiple degradation aspects that could affect structural properties. In this study, a deep learning-based multi-scaled generative adversarial network (GAN) with residual image learning is proposed to process variant and in-variant blur. Different scaled images concatenated with corresponding gradients are used as multi-channel single input for the proposed GAN. Residual- and dense-networks are combined to explore salient features in the bottleneck section while addressing the vanishing gradient problem. A hybrid content loss function with gradient penalty is designed to minimize the error between generated and ground truth images. Due to structure sparsity, the generated output may lose some information that leads to artifacts. Residual image learning with dilation and end-to-end training is used to resolve this issue by recovering high-resolution anatomical details. Three different datasets: GoPro, Köhler, and Lai, with variant and in-variant blur, are used to train and test the proposed neural network. To validate our proposed method, we performed qualitative and quantitative

analysis. The quantitative analysis is based on eight performance metrics with an equal number of reference and no-reference metrics. To the best of our knowledge, this is the first study to evaluate the proposed deblurring neural network performance extensively for various degradation aspects. Experiments show our proposed method is effective in reducing blur while preserving structural properties compared to multiple preprocessing techniques for image analysis. Moreover, the consistently improved performance over multiple publicly available datasets validates the merits of our proposed method for large data analysis.

3.1 Introduction

The acquired images of real-world scenes or objects can be degraded due to inaccurate acquisition and optical factors that include wrong depth of field, poor focus, camera shaking, objects movement, short exposure, low optical quality, etc. These all factors can lead to produce low contrast, noisy images with artifacts that can immensely affect the image recognition or classification tasks. The general image formation can be described in the spatial domain as follows [114],

$$B(x^{'}, y^{'}) = \iint_{-\infty}^{\infty} k(x^{'} - x, y^{'} - y) I(x, y) dx dy \tag{3.1}$$

where $(x^{'}, y^{'})$ and (x, y) are the image and 3D object in a real-world plane respectively, k is known to be space invariant point-spread function (PSF), B is a blurred image, and I is a latent or true image. Considering additive noises $N(0, \sigma_N^2)$, the image formation model can be represented as:

$$B = k \otimes I + N \tag{3.2}$$

where $\otimes$ is the convolution operator and k is a non-identity matrix with blur information (see further details in the supplementary material A).

Image deblurring methods can be mainly categorized into non-blind and blind deblurring algorithms. In the former algorithms, the blur kernel information is known; thus, the inverse process is carried out to obtain the deblurred image. The latter algorithms require constrained estimation with heuristic parameter tuning to find a best blur kernel. Such methods are considered under conventional model-based techniques that include blind deconvolution based on patch priors [115], spectral irregularities [119], random transform [120], normalized color-line prior [97], and residual-based deconvolution [116]. However, applying deconvolution techniques can produce ripple or ringing artifacts in the estimated image [115, 116, 97]. Various noise and spectral filtering techniques [99, 100, 97, 98] have been proposed to improve images with high contrast and less rippling artifacts. In addition, many studies [108, 107, 281, 282, 106, 283, 284, 285] have focused on optimal blur kernel estimation for preprocessing of images.

Deep learning-based techniques have been widely utilized in image sample estimation [105, 286, 287, 288, 289], style transfer [290, 291], super-resolution [292, 293, 294, 103, 295, 296, 297, 298], and many other applications [299]. Image sample estimation is used to counter data limitation problems by developing

synthetic data with similar data distribution. Similarly, super-resolution neural networks are used to minimize noises and artifacts in an image. Deblurring can be considered as a style-transfer task from blurred to sharp style. Variational auto-encoders and generative adversarial networks (GANs) have been proposed to address data limitation problems in the field of computer vision. However, GANs have proven to be efficient in image style transfer application [105].

In this study, we propose a GAN with multi-scale image and gradient input fused with the residual- and dense-blocks in the bottleneck section to restore non-uniform or uniform blurred images due to various aspects such as camera shake, motion artifacts etc., [300]. Conventionally, the restored images should be equal to reference images. Unfortunately, it is not true due to the fact that while denoising and deblurring, some fine details are lost [116, 301, 117]. To compensate for the lost information in deblurring, we design a neural network with end-to-end training named residual image learning (RIL) in cascade form with the proposed GAN [102]. It notably reduces deconvolution ripples or artifacts by learning the structural and spatial details within an image. Moreover, the RIL block is trained based on the cross-level loss function between the proposed GAN output and the reference image. It improves the semantic visual quality and pixel-wise accuracy with a considerable reduction in the computation cost [103]. Furthermore, to increase the receptive view, stride and dilated convolution layers are adopted with the proposed network.

The proposed method is extensively evaluated qualitatively and quantitatively on three datasets, namely: GoPro [4], Köhler [5], and Lai [6], and is compared with the most recent and standard deblurring methods. Notably, we combine the merits of several deblurring studies that include GAN [14], multi-scale network [126, 4], various cross-level loss functions [300, 14], and residual image learning [117, 102, 103, 296] for style-transfer applications. Our evaluations are in terms of five reference and no-reference performance indexes, namely: Peak signal to noise ratio (PSNR), Structural similarity index (SSIM) [302], Absolute mean sharpness error (AMSE), Perception-based image quality evaluator (PIQE) [303], and No-reference image quality assessment based on spatial and spectral entropies (SSEQ) [304]. The comparative analysis shows better performance of our proposed method in terms of most metrics.

The contributions of this study can be summarized as follows,

- The GAN with the residual- and dense-blocks in the bottleneck layer, which takes multi-scale images and corresponding gradients as input, is designed to restore the spatial properties of variant and non-variant blurred images;

- The various cross-level loss functions are combined to improve pixel-wise properties within an image;

- The generative network is designed to reconstruct a deblurred image while the residual image learning is adopted to improve the high resolution anatomical structural details;

- The extensive experimental comparisons are performed on three different public datasets.

3.2 Related Work and Background

Image restoration is a well-recognized active research area for decades. Techniques for image restoration can be divided into the conventional model- and deep learning-based methods. In this section, beside these techniques, we are also discussing about generative adversarial networks.

3.2.1 Conventional Model-based Techniques

Blind deblurring is an ill-posed inverse problem requiring some optimization techniques to estimate true image. These methods can be categorized into the statistical-, filter- and patch-based techniques [281, 305, 306, 307, 96, 308]. Most existing studies [281, 309, 282, 310, 283, 284, 285] have focused on the model-based selection of adaptive smoothing kernels for optimized measuring protocols.

In statistical-based methods, Fergus et al. [108] assume a camera blur to be uniform; therefore, the Bayesian-based maximum a-posterior (MAP) model is proposed for blur kernel estimation by probabilistic distribution. Yuan et al. [116] considers a pair of noisy and blurred images to estimate deblurred images using non-blind deconvolution. A novel technique of residual deconvolution is also proposed due to the loss of fine details with denoising. Subsequently, Cho and Lee [311] apply non-blind residual deconvolution in a coarse-to-fine manner between an image and blur kernel. Harmeling et al. [312] and Xu and Jia [313] proposes space-variant blind image deblurring techniques with kernel prior cost functions. Levin et al. [314] extend the work of Fergus et al. [108] by considering the blind deblurring model with Gaussian distribution before estimating the blur kernel with MAP approximation.

Similarly, Krishnan et al. [315] propose a novel cost function based on normalized regularization parameter to deblur images. The method is demonstrated on spatially varying and invariant blur images with a considerable reduction in computation cost. In Xu et al. [106] a unified framework is proposed for uniform and non-uniform motion deblurring with gradient-based regularization function for blind deconvolution. Zhang et al. [316] propose a multi-image blind deconvolution model with a novel penalty function to obtain latent images with blur kernel. The algorithm considers multiple blurred images with higher quality image domination to recover latent images with blur kernel. Pan et al. [317] consider regularized intensity with gradient prior to deblur images with the least square estimation. Likewise, Whyte et al. [318] propose RL-based models for saturated pixels that can cause significant artifacts in deblurred images by conventional deconvolution algorithms.

In filter-based methods, Yuan et al. [118] assume the invariant blur kernel for blurred images. Afterwards, they define the correlation between kernel sparseness and translation of a blurred, noisy image. Thus, a good kernel estimation and alignment are performed simultaneously. Shan et al. [319] explore factors of artifacts in a deblurred image and conclude with noise and inaccurate blur kernel estimation due to imperfect distribution assumption. Therefore, a logarithmic image gradient histogram distribution-based function is proposed for

blur kernel estimation. MAP is then utilized for the minimization of energy. Zhong et al. [320] apply directional filters to retain the blur information with inverse radon transform to estimate blur kernel. Lastly, noise-tolerant non-blind deconvolution is used to obtain the latent image. Kim and Lee [321] introduce a deblurring framework based on motion flow and latent image estimation with Tikhonov regularization (TV). Liu et al .[322] use the weighted TV optimization to deblur and denoise the images.

In patch-based methods, Huang et al. [323] perform intensity-based geometric transformation estimation between blurred and non-blurred images. The obtained transformation is then applied to make a non-blurred image close to the blurred image. After that, the method in [118] is applied to obtain blur kernel, and Richardson-Lucy (RL) deconvolution [324] to a deblurred image. Cho and Lee [311] consider gradient input image model with gradient prior to deblur images. Fast Fourier Transform (FFT) is preferred instead of convolution to speed up the iterative process. Whyte et al. [325] explain the blurring process due to camera rotational velocity and propose deblurring model in terms of this parameter. RL deconvolution is further utilized to refine the estimation.

Moreover, Sun et al. [115] models image edge primitives using patch priors. Later, statistical- and synthetic prior are used to refine the blurred image with kernel estimation. Lee et al. [117] consider two consecutive blur and non-blur video frames. After estimating the initial blur kernel, the non-blur frame is synthetically blurred to estimate motion between two consecutive frames. The blurred frame is motion-compensated and deconvolved in an iterative process to refine the blur kernel with a deblurred frame. Michaeli and Irani [326] demonstrate the multi-scale properties for deblurring an image. Moreover, a novel cost function is proposed with multi-scale image dissimilarity and kernel prior. Pan et al. [327] consider threshold-based dark channel prior as a regularization term to measure the sparsity.

3.2.2 Deep Learning-based Techniques

In addition to classification tasks, deep learning-based techniques have been explored in various image processing tasks such as segmentation, enhancement, noise, artifacts removal, deblurring etc., [13, 20]. Xu et al. [328] coin the idea of deconvolution for the very first time using deep neural networks. The blur kernel estimated by statistical estimation is used to perform non-blind deconvolution by a deep neural network. The necessity with a large blur kernel is discussed because a small invariant blur kernel causes artifacts in a latent image. Therefore, a deep neural network can assist in dealing with large heterogeneous blur kernels. Schuler et al. [329] introduce a deep neural network to extract features and to estimate the blur kernel. After that, a deblurred image is estimated to compare with the latent image so that the predicted blur kernel can be further refined. Sun et al. [129] employ convolutional neural network (CNN) for non-uniform motion blur. The patch-based neural network scheme is adopted with Markov random field model to estimate the distribution of smooth motion blur. Afterwards, non-uniform motion is removed using a parametric deblurring model with the patch-level image prior. Nah et al. [4] propose a deep learning-based neural network and a

novel GoPro dataset for blind deblurring. In their study, the kernel free loss training is considered for blur images.

Gong et al. [125] utilize a fully convolutional neural network to estimate motion flow within blur images. Synthetic data is generated for neural network training. Estimated motion flow is then used to deblur images. Li et al. [161] consider CNN binary output as prior to conventional model-based technique to restore blurred images. The MAP estimation with the latent image is used to measure the probability of an image being blurred. Pan et al. [330] consider model- and learning-based techniques to deblur face images. In a model-based algorithm, regularization is used with gradient prior to deblur images. Afterwards, CNN is proposed to remove artifacts and to obtain sharp images. Lu et al. [331] introduce an unsupervised deblurring neural network with unpaired training data. A novel loss function is considered to retain color properties due to unpaired training data. The proposed method is extensively evaluated on natural- and domain-specific images.

Likewise, Wei et al. [297] propose a neural network with local correlation blocks to deblur dynamic scene blurred images. Liu et al. [296] employ a two-stage module to deblur images. The first stage is based on statistical parameter estimation, while the second stage is based on deep neural network restoration to remove artifacts and retain the shape of an image. Three datasets are used to evaluate the proposed algorithm. Jiao et al. [102, 103] introduce residual image learning for deep learning-based deblurring. A content loss function based on pixel, gradient and feature difference is considered for end-to-end training. Shen et al. [332], and Yasarla et al. [333] perform semantic face deblurring by utilizing a multi-scale convolutional neural network and a local structure loss function to regularize the output. Vasu et al. [334] utilize CNN to predict the blur kernel. Large synthetic data is trained to address blur kernel uncertainties. The predicted blur kernel is then used to perform non-blind deblurring.

Nimisha et al. [335] propose an auto-encoder based GAN to process invariant blur. Mean squared and gradient-based loss functions are computed to train the generator. Wu et al. [336] propose an image enhancement framework based on Gaussian-Poisson GAN [105]. A 9-layer autoencoder GAN is used to estimate low-resolution realistic images. Afterwards, Gaussian-Poisson assists in fusing gradient and colour information for constrained optimization. Ramakrishnan et al. [124] consider GAN with skip connections and dense network to deblur the images with four training datasets. Zhang et al. [337] utilize GAN to deblur remote sensing images. Nine residual blocks with stride convolution in the bottleneck section is used in the generative network. Kupyn et al. [14] introduce DeblurGAN to deblur images with a novel loss function. In addition to non-uniform and uniform deblurring, an algorithm is proposed to produce synthetic training data. Liu et al. [338] propose scale-recurrent GAN to deblur images based on extracted spatial and temporal features. Moreover, the progressive loss function is presented to counter the vanishing gradient problem and exploit global and local information of deblurred images. Jiang et al. [77] consider GAN with dense blocks to deblur satellite images. The deblurred outputs contain artifacts. Therefore, a cascaded edge enhancement neural network is used to refine the deblurred output.

Zhang et al. [300] introduce GAN with the residual and dense network to improve deblurring results. The perceptual loss is computed based on ℓ_1 and ℓ_2 loss of VGG-19 pre-trained network output [339]. Zheng et al. [340] propose an edge heuristic multi-scale GAN with hierarchical content loss function to restore non-uniform blur caused by camera shake. Wu et al. [292] utilize the residual neural network to estimate blur kernel with motion flow. The GAN is then fused with the prior information to deblur images. Wu et al. [341] develop a text image deblurring algorithm based on unpaired training data. Two cascaded GANs are used to deblur images by training the prior and the posterior to blur images to minimize the losses for unpaired data. Kupyn et al. [342] improve the performance of DeblurGAN [14] by introducing a feature pyramid network for image restoration and a least-square loss evaluation of local and global patches for the discriminator. Nimisha et al. [335] consider homogeneous blur for deblurring. GAN is fused with an autoencoder to extract variant features. Zhang et al. [251] introduce GAN with res-block in the bottleneck section to exploit blur features. Moreover, the gradient-based content loss function is utilized to improve performance. Cai et al. [343] consider novel threshold-based dark and bright channels with multi-scale GAN to perform deblurring.

Various studies [126, 344, 345, 127] discuss different neural network models that include encoder-decoder, multi-scale, and dilated convolution network. A novel scale-recurrent neural network architecture is proposed to exploit spatial features at variant scales. A sharp latent image generation at each scale is considered input for the subsequent up-scale output until the desired output size is attained.

3.2.3 Generative Adversarial Network

Goodfellow et al. [105] proposed GAN for the first time to address the data limitation problem by generating synthetic data. The network is trained in an adversarial manner, such that the generative network tries to learn the input data distribution so that the discriminator can be fooled by generated synthetic samples close to the real one. While the discriminator network tries to estimate the probability that a generated sample is from the training data. The min-max scheme between generator and discriminator is defined by an objective function as,

$$\min_{g}\max_{d}V(d,g)=\underset{x\sim p_d}{\mathbb{E}}[\log(d(x))]+\underset{z\sim p_g}{\mathbb{E}}[\log(1-d(g(z)))] \tag{3.3}$$

where $\mathbb{E}$ is the expectation operator, x is a sample from the real data distribution p_d, z is a sample from the noise distribution p_z that is a blur image in our case [14, 346], and p_g is the generated model distribution with a mapping of noise space to data space. The discriminator is trained to minimize term $\log(d(x))$ while the generator is trained to minimize term $\log(1-d(g(z)))$.

GAN has been utilized in various forms for image-to-image translation tasks [14]. The conditional GAN is used to learn a mapping of observed image x and random noise vector z to any kind of auxiliary information y, where y could be class label or data from various modalities. Thus, the objective function for the conditional GAN becomes [347],

$$\underset{x\sim p_d}{\mathbb{E}}[\log(d(x|y))]+\underset{z\sim p_g}{\mathbb{E}}[\log(1-d(g(z|y)))] \tag{3.4}$$
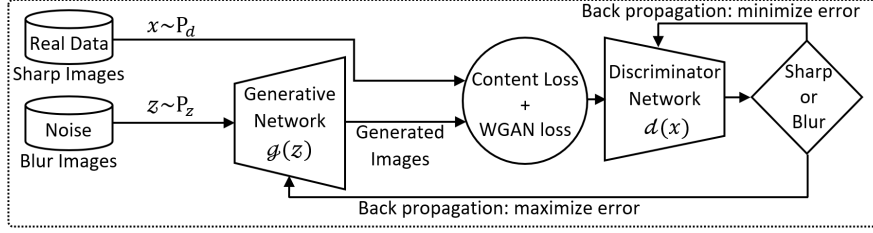
Figure 3.1: Workflow of the proposed GAN training.

Actually the training of original GAN suffers from several problems, such as training instability, vanishing gradients, mode collapse etc., [14, 300]. It is because the original GAN measures the difference between model- and data distribution based on Jensen-Shannon (JS) divergence. Instead, Arjovsky et al. [348] uses the Wasserstein distance to calculate the difference between two distributions. The workflow of the proposed GAN is shown in Figure 3.1. The generative network takes the blurred image as noisy input and generates the estimate of the sharp image. The content- and Wasserstein loss (discussed in section 3.3.4) are used to train the discriminator and generator network. At the test time, only the trained generative network is kept as depicted in Figure 3.2.
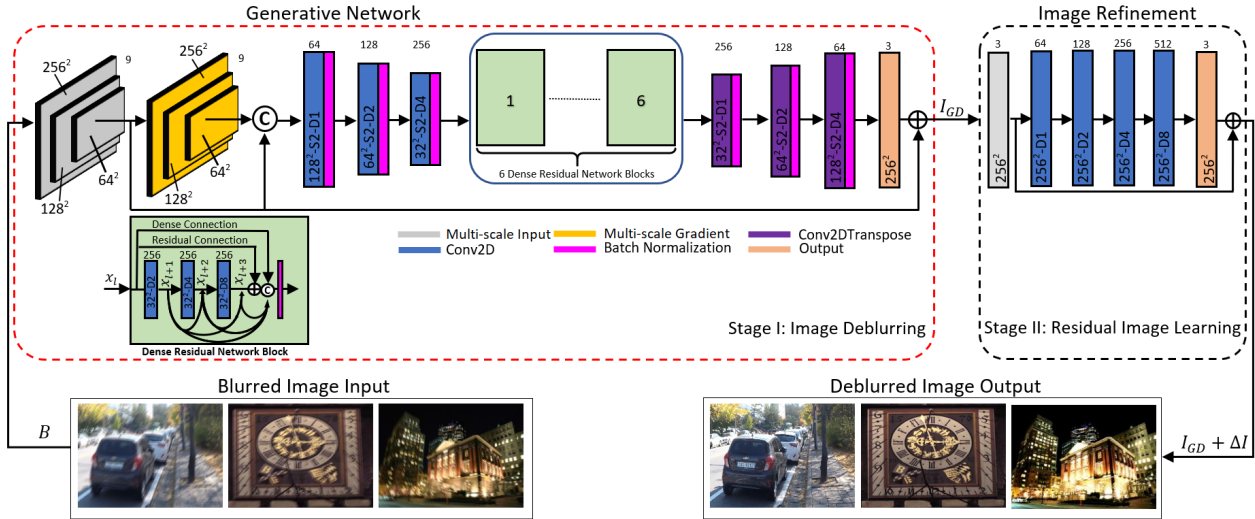
Figure 3.2: Schematic of the proposed deblurring model. $\oplus$ and $\copyright$ denotes addition and concatenation operation.

3.3 Proposed Method

The goal of this study is to reconstruct a high-quality and sharp image I from the blurred image B. To achieve it, we propose a blur-kernel free GAN for blind deblurring. Conventionally, as defined by equation 3.2, the generated deblurred image I_{GD} from GAN should be equal to latent image I but it is not, due to the fact that while performing deblurring to an input image some high-resolution details are lost that causes artifacts in the deblurred output. Thus consider lost information in the conventional model as [116, 301, 117, 102, 103, 296],

$$B = (I_{GD} + \Delta I) \otimes k \tag{3.5}$$

where ΔI is the lost information of deblurred image named as a residual image. The residual image learning network assists in recovering this lost information. Initially, the proposed GAN is trained to estimate I_{GD} corresponding to each blur image B. Then residual image learning is performed to recover lost information ΔI, such that latent image $I = I_{GD} + \Delta I$. The schematic of the proposed deblurring model is shown in Figure 3.2.

3.3.1 The Generator

As explained in section 3.1, stage I, initial deblurring, is based on the generator trained in an adversarial manner. It considers multi-scale image with corresponding gradient concatenated together to form 18 channels input. Three-channel (RGB) image at three variant scales with the resolution of $\{256, 128, 64\}$ and corresponding gradients concatenated together with zero paddings. After that, three 3×3 dilated convolutional layers and three 3×3 transposed dilated convolution layers with a stride of 2 and rectified linear activation function (ReLU), six dense residual network blocks (DRB), a 1×1 convolution layer with hyperbolic tangent activation function and six batch normalization layers have been adopted in the proposed generative network. The batch normalization assists in accelerating the convergence speed of training and in solving the gradient explosion problem. Each layer in the encoder section is downscaled while upscaled in the decoder section. D_x depicts the various dilation rate, and the top index shows a feature map for each layer in Figure 3.2. The blurred image has a high correlation with the corresponding sharp image. Hence, in accordance with the study Kupyn et al. [14], a global residual connection with a 256×256 scale is included to keep track of input with faster and generalize training. The DRB block combines the merits of ResNet [10] and DenseNet [349] neural network to avoid vanishing gradient and over-fitting problems. Each DRB block contains three 3×3 convolution layers with a batch normalization layer. Consider an lth convolution layer output x_l as input of DRB and $r(\cdot)$ be a residual mapping function, then each DRB block can be presented mathematically as [350],

$$
\begin{aligned}
R &= r\left(x_l, W_i\right) + x_l \\
r\left(x_l, W_i\right) &= rW_3\left[rW_2\left(rW_1\left(x_l\right)\right)\right] \\
D &= F\left(\left[x_l, x_{l+1}, x_{l+2}, x_{l+3}, R\right]\right) \\
O &= H\left(D\right)
\end{aligned}
\tag{3.6}
$$

where O is the dense residual network block output, R and D are the outputs of residual and dense connections respectively, $\left(\left[x_l, x_{l+1}, x_{l+2}, x_{l+3}, R\right]\right)$ represents the concatenation of feature maps in layers $(l, \cdots, l+3)$ and residual output R, r is the rectified linear activation function (ReLU), W_i is the weighted dilated convolution layer with a dilation rate of $i = \{2, 4, 8\}$, $F(\cdot)$ is a composite function of ReLU and a 3×3 dilated convolution and H represents a batch normalization function.

3.3.2 The Discriminator

The primary aim of the discriminator is to maximize the error so that the generator is well-trained to generate output close to its reference. Therefore, the discriminator output can be a simplified feature map that can accelerate its training. We consider the discriminator network in accordance with DeblurGAN [14], and PatchGAN [351, 352]. It contains five 4×4 strided convolution layers with the batch normalization, and the LeakyReLU activation function. After that, two fully connected layers are added with the hyperbolic tangent and the sigmoid activation function. The schematic of the discriminator network is shown in Figure 3.3.
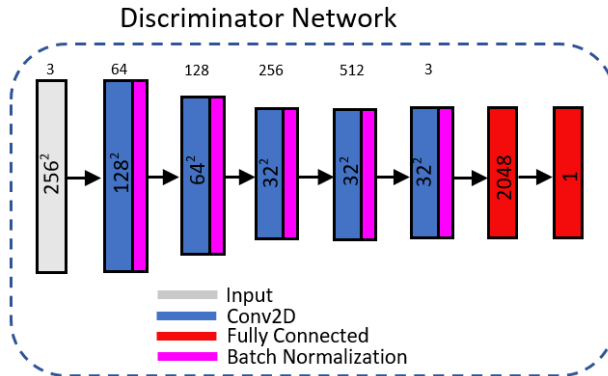
Figure 3.3: Schematic of the discriminator network.

3.3.3 Image Refinement

As discussed in section 3.2, the deblurred output may lack salient-structured regions with high-frequency artifacts or ripples that are irrelevant to input images [77, 296]. Therefore, in accordance with equation 3.5, a residual image learning network is proposed to improve image resolution by reducing image corruption. We consider a heterogeneous blur, which can be processed based on conventional iterative model-based techniques [116, 301, 117] or deep neural networks [102, 103, 77, 296]. However, instead of assuming a random prior for conventional model-based techniques, we propose a deep neural network that can work as a high- and low-pass filter to remove artifacts and estimates highly correlated regions corresponding to latent images. Therefore, a residual image learning network refines I_{GD} to I with four 3×3 dilated convolution layers having ReLU activation function and a 1×1 convolution layer with the hyperbolic tangent activation function. The skip connection is used to keep track of input features. The content and perceptual based cross-level loss functions are fused together to perform the end-to-end training of our proposed residual image learning network (discussed in section 3.3.4). Figure 3.4 shows the residual image ΔI, generated deblurred image I_{GD}, and true image I for the blurred image input.
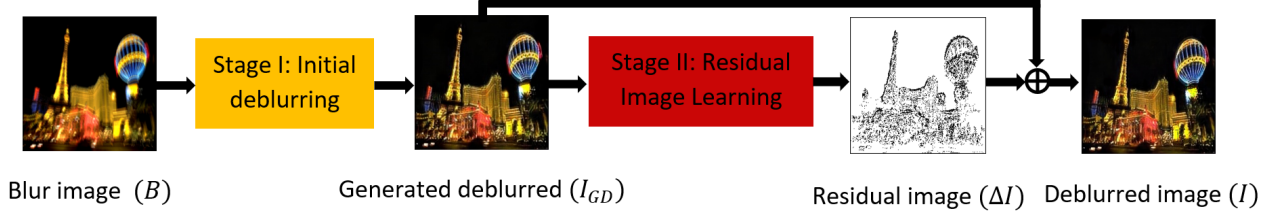
Figure 3.4: The process flow of the blur image input.

3.3.4 Loss Function

We aim to restore a sharp image I given a corresponding blurred image B without any blur kernel information. The combined loss function of the proposed deblurring network is given as,

$$L_{net} = L_{adv} + L_{cont} + L_G \tag{3.7}$$

where L_{adv} indicates the adversarial loss, L_{cont} is the content loss and L_G represents the gradient loss explained as follows.

Adversarial loss

Our proposed GAN takes an image in the form of pairs as blurred B and the corresponding latent or ground truth image I. The blurred image B is processed by the generator to produce I_{GD}. I_{GD} and I are evaluated by the discriminator to compute the loss. We use the same discriminator network loss function as that in Wasserstein GAN [348] study as follows,

$$L_{adv} = \sum_{n=1}^{N} -d_{\theta_d}\left(g_{\theta_g}\left(B(n)\right)\right) \tag{3.8}$$

where g_{θ_g} represents the output of the generative network, d_{θ_d} denotes the trained discriminator with a binary output label, N is the total number of images in a batch, and B is the blurred image.

Hybrid Content loss

The content-based loss is considered as the primary objective function for deblurring. It is computed based on two classical methods: mean absolute error (ℓ_1 norm) or mean squared error (ℓ_2 norm). ℓ_1 loss has a sparse output with a non-smooth gradient at zero. ℓ_2 loss results in detail loss, while it lacks sparseness. Moreover, due to the pixel-wise average computation, these sole optimization functions lead towards blurry and over-smooth targets [14]. To overcome these limitations, we simultaneously consider both losses by leveraging ℓ_1 loss and ℓ_2 loss [300]. Additionally, to avoid blurriness we consider a perceptual loss function. The perceptual loss is a high-level loss computed using CNN feature maps for style-transfer applications. The pre-trained VGG-16 [339] for ImageNet classification is used to extract feature maps as a function ϕ.

Thus, an overall hybrid content loss function becomes,

$$L_{cont} = \alpha \ell_1 + \beta \ell_2$$

$$\ell_1 = \frac{1}{2N} \sum_{n=1}^{N} \left[\frac{1}{W_{i,j} H_{i,j}} \sum_{(x,y)=(1,1)}^{(W_{i,j}, H_{i,j})} \left| \phi_{i,j}(I(n))_{x,y} - \phi_{i,j}(I_\circ(n))_{x,y} \right| \right]$$

$$\ell_2 = \frac{1}{2N} \sum_{n=1}^{N} \left[\frac{1}{W_{i,j} H_{i,j}} \sum_{(x,y)=(1,1)}^{(W_{i,j}, H_{i,j})} \left(\phi_{i,j}(I(n))_{x,y} - \phi_{i,j}(I_\circ(n))_{x,y} \right)^2 \right]$$

(3.9)

where α and β indicates weighted parameters set to 0.25 and 0.5 in this work for the best performance, respectively. $W_{i,j}$ and $H_{i,j}$ are the feature map dimensions, $\phi_{i,j}$ is the feature map obtained by j^{th}-convolution after the ith-max pool layer, $|\cdot|$ represents the magnitude, I is the ground truth image and $I_\circ$ is the deblurred output image. *block1_pool* feature map of VGG-16 is considered because the lower layers of the network provide structural information and higher layers provide spatial information. Therefore, lower layers help better in retaining content information.

Gradient loss

In addition to the perceptual based content loss, we also try to minimize the image gradient difference through the following loss function.

$$L_G = \gamma \left[\frac{1}{2N} \sum_{n=1}^{N} \left| \nabla \left(I(n) \right) - \nabla \left(I_\circ(n) \right) \right| \right]$$

(3.10)

where γ is the weighted parameter set to 0.25, and ∇ denotes the image gradient.

3.4 Experiments

3.4.1 Datasets

We utilize three distinct synthetic and real blurred publicly available datasets to evaluate our proposed deblurring framework. The details of these datasets are as follows.

GoPro Dataset

Nah et al. [4] created this synthetic blur dataset by averaging multiple frames with a total number of 3214 pairs. The dataset consists of natural scenes with 2/3 of the total dataset for training while 1/3 is used for testing.

Köhler Dataset

Köhler et al. [5] create this standard dataset that has been widely used for deblurring algorithms comparison. It consists of a total of 4 images that are blurred by convolving 12 various blur kernels. The blur kernels are

modeled by assuming linear camera response function and 6D real blur camera motion.

Lai dataset

Lai et al. [6] create these images. It contains 100 image pairs from real-world scenes and 100 images generated by convolving multiple degradations and non-uniform blur kernels with 25 sharp images. The blur kernels are generated by 6D camera trajectories.

3.4.2   Training Details

We use K40 NVIDIA GPU with Intel (R) Xenon (R) 2.6Ghz CPU having 30 GB RAM to train and test our proposed deep learning-based deblurring network on Tensorflow 2.0 platform. The proposed generator is adversarially trained to obtain a deblurred image I_{GD}. Afterwards, the image residual learning block is end-to-end trained to obtain image I by removing artifacts from the input image I_{GD}. We performed adaptive training by utilizing the previously trained weights for each epoch. The Adam optimizer with the stochastic gradient descent is used at an initial learning rate of 0.0001. Moreover, we perform data augmentation by utilizing flip, shear, zoom and rotation operations. The size of input images is 256×256 due to computation limitation with a batch size of 5 and the training data portion of 75%.

3.4.3   Evaluation

The degraded images may contain several types of distortion that could immensely affect the analysis. Therefore, the processed images are evaluated based on noise reduction and structural similarity between processed and ground truth images. However, with one given image (distorted), such reference measuring metrics become invalid. Therefore, no-reference metrics are designed to statistically estimate distortions based on various techniques that include data distribution, entropy, and regression models.
No-reference quality assessment algorithms are complex to compute since they use statistical features of input images to evaluate the image quality. We employ five variant metrics to extensively evaluate proposed network performance that are discussed as follows.

Peak signal-to-noise ratio (PSNR)

It is defined as the ratio between maximum pixel intensity and mean square error with respect to reference image. PSNR is expressed in logarithmic decibel scale.

Structure Similarity Index (SSIM)

This index is used to estimate structural changes in an image based on the luminance, contrast, and structure of objects within an image. For comparison, it uses reference image to compare structure with [302].

Absolute Mean Sharpness Error (AMSE)

Various blurred assessment indexes [353, 354, 355, 356, 357, 358, 359, 360] are commonly based on image gradient magnitude due to the fact that sharp images have prominent edges that can be evaluated based on images gradient. However, the blurred images lack this property.

Perception-based Image Quality Evaluator (PIQE)

PIQE is an opinion unaware and unsupervised spatial quality evaluator, which is based on pristine images with blockwise distortion estimation from a local variance of each distorted block [303].

Spatial-Spectral Entropy-based Quality (SSEQ)

SSEQ is a no-reference image quality assessment model that considers spatial and spectral entropy of an image for no-reference quality assessment. It can assess the quality of a distorted image across multiple distortion categories [304].

PSNR, SSIM, PIQE, and SSEQ are implemented with the built-in function of MATLAB while AMSE is computed by calculating mean square error (MSE) between processed ∇G_X and reference image ∇G_Y gradient magnitude. SSIM ranges between $[0, 1]$ with 1 for the best quality while PIQE, and SSEQ ranges between $[0, 100]$ with 0 for the best quality of an image.

3.4.4 Results and Analysis

We perform quantitative and qualitative comparative analysis to evaluate deblur results. Three distinct datasets: GoPro [4], Köhler [5], and Lai [6] with a total number of 3486 training and 523 testing images are used to evaluate and compare various traditional model- and deep learning-based methods with our proposed method. Due to the computation limitation, we adopt the image size of 256×256 for comparative analysis. Hence, a minor difference is observed in results compared to previous studies, considered as the margin of an experimental error [161]. A proposed algorithm averagely takes $0.34s$ to process 256×256 scale image with $0.25s$ for image deblurring and $0.09s$ for image refinement. It can process space-variant and space-invariant blur instead of only invariant blur where the blur kernel is evenly convolved with a latent image to give the blurred output. Therefore, the analysis is performed based on real and synthetically uniform and non-uniform blur, natural and text images.

The codes of aforementioned methods and results can be downloaded from author's websites [5, 6, 361], except for studies [321, 251] whose results have been reported in their previous work. The average PSNR, SSIM, AMSE, PIQE, and SSEQ in multiple studies on Köhler, Lai and GoPro testing datasets are shown in table 3.1. The bold values show the best outcome across each metric. To the best of our knowledge, this is the first study to measure reference and no-reference metrics collectively for deblurred image evaluation. Thus,

deblurred images are extensively evaluated not only for reference noise, structure, and sharpness parameters but also for statistical-based noises and distortions. It is evident from table 3.1 that our proposed method outperforms in terms of most metrics, especially comparing to deep learning methods [4, 14, 126, 127] which are not able to perform well for the Köhler dataset.

Similarly, some conventional model-based techniques [325, 106, 321] perform better with the Köhler dataset but not in terms of SSIM. Our proposed network performs satisfactorily because it takes multi-scale gradients input to keep track of structural information while generating deblurred images. Similarly, we compute no-reference metrics: PIQE, and SSEQ. It is worth noting that PIQE perform favourably best on all the datasets for our proposed network. It is due to the novel hybrid loss function (discussed in section 3.3.4) that counts for multiple loss metrics to attain ground truth results with the multi-scale input image and various dilation rates to probe variant and distinct features.

Likewise, we also compare performance on Lai dataset. Although our proposed methodology performs favourably best, the MSE is large because of the fact that synthetically blurred images consider various non-uniform blur kernels. In contrast, not enough training images are considered to learn unknown blur kernels distribution.

Furthermore, we perform the visual analysis with competing methods. Figures 3.5, 3.6, and 3.7 shows deblur results of various studies for the discussed datasets (more results shown in supplementary material A). It is evident from the visual results that recent CNNs studies can remove considerable blur. However, they are still not pleasing due to ringing artifacts and sometimes remaining blurred edges. For instance, Figure 3.5 shows deblur images for the Köhler dataset. Ye et al. [127] and Tao et al. [126] process images with blur edges remaining, while Nah et al. [4] with the motion artifacts. Moreover, due to structure parsing, they led to more sharp edges. A similar trend can be seen in Figure 3.6 for conventional model-based techniques on the Lai dataset. Local regions tend to over-deblur or under-deblur, specifically eyes, flags, building structures etc., (see supplementary material A). Likewise, multiple studies [14, 344, 127] deblur GoPro images with blurry facial and text regions, as shown in Figure 3.7. Therefore, we consider residual image learning block with end-to-end training to recover fine details that are usually lost while deconvolving the images. However, Cai et al. [343] provide comparable results on the GoPro dataset due to dark and bright channels input as prior constraints. A generalized blind deblurring neural network remains challenging due to training dataset constraints and limitless blur kernels.

Ablation Study

The unique hierarchical features are required for image super-resolution problems. In Zhang et al. [350], a residual dense block is proposed to explore variant features, which validates the effectiveness of the proposed block for image super-resolution tasks. Similarly, in accordance with the studies [350, 251], we consider DRB to explore new hierarchical features at the bottleneck section.

Table 3.1: Performance metrics evaluation on three public datasets

Study	Year	Reference Metrics			Non-reference Metrics	
		PSNR	SSIM	AMSE	PIQE	SSEQ
GoPro Dataset [4]						
Whyte et al. [325]	2012	24.53	0.83	31.31	**26.99**	30.01
Xu et al. [106]	2013	20.29	0.72	43.43	71.14	24.18
Kim and Lee [321]	2014	23.64	0.82	-	-	-
Sun et al. [129]	2015	25.30	0.85	29.39	69.13	24.16
Pan et al. [327]	2016	27.06	0.90	30.97	37.54	**19.76**
Nah et al. [4]	2017	29.02	0.91	21.84	64.24	25.16
Gong et al. [125]	2017	29.49	0.90	27.34	47.14	16.29
Tao et al. [126]	2018	30.21	0.93	19.18	72.89	24.26
Kupyn et al. [14]	2018	27.02	0.89	24.53	65.38	23.67
Zhang et al. [344]	2018	27.07	0.90	24.62	66.52	24.33
Zhang et al. [345]	2019	30.44	0.93	16.50	75.16	40.52
Zhang et al. [251]	2020	28.67	**0.96**	-	-	-
Cai et al. [343]	2020	31.09	0.94	17.38	68.22	24.99
Ye et al. [127]	2020	27.09	0.89	23.50	74.79	36.08
Zhou et al. [362]	2021	24.82	0.82	-	-	-
Feng et al. [363]	2021	29.62	0.89	-	-	-
Wen et al. [364]	2021	26.24	0.86	-	-	-
Wu et al. [365]	2021	30.75	0.91	-	-	-
Pan et al. [366]	2021	31.40	0.94	-	-	-
Proposed	-	**33.10**	0.95	**12.55**	32.85	26.19
Köhler dataset [5]						
Fergus et al. [108]	2006	17.56	0.54	67.96	50.96	11.18
Shan et al. [319]	2008	19.49	0.65	61.68	36.13	8.57
Cho and Lee [311]	2009	18.62	0.62	67.38	42.98	8.45
Harmeling et al. [312]	2010	17.82	0.58	79.04	43.11	9.08
Krishnan et al. [315]	2011	17.81	0.59	74.83	39.56	**7.19**
Whyte et al. [325]	2012	19.73	0.67	61.28	35.93	9.52
Xu et al. [106]	2013	18.48	0.61	67.52	37.52	8.08
Kim and Lee [321]	2014	24.01	0.79	-	-	-
Nah et al. [4]	2017	15.72	0.59	39.50	78.73	36.28
Kupyn et al. [14]	2018	15.80	0.59	39.32	78.90	36.80
Tao et al. [126]	2018	15.79	0.59	39.50	77.72	34.75
Ye et al. [127]	2020	17.32	0.70	50.83	85.38	21.29
Wu et al. [365]	2021	19.83	0.64	-	-	-
Kaufman and Fattal [367]	2021	20.83	0.61	-	-	-
Pan et al. [366]	2021	17.55	-	-	-	-
Proposed	-	**24.16**	**0.81**	**35.92**	**30.25**	10.88
Lai Dataset [6]						
Fergus et al. [108]	2006	15.22	0.57	103.76	46.78	17.19
Cho and Lee [311]	2009	16.71	0.64	85.14	44.49	19.37
Xu and Jia [313]	2010	17.31	0.67	80.58	48.16	19.60
Krishnan et al. [315]	2011	16.43	0.64	87.87	45.90	18.57
Levin et al. [314]	2011	16.02	0.61	89.37	43.35	17.40
Whyte et al. [325]	2012	17.12	0.65	83.08	33.41	20.05
Sun et al. [115]	2013	17.27	0.67	80.77	45.47	20.37
Xu et al. [106]	2013	16.96	0.66	83.31	48.40	19.71
Zhang et al. [316]	2013	16.52	0.63	86.36	43.87	17.59
Zhong et al. [320]	2013	16.60	0.64	85.86	46.97	18.59
Michaeli and Irani [326]	2014	17.45	0.67	80.19	42.29	20.04
Pan et al. [317]	2014	17.28	0.64	83.32	45.02	**14.59**
Perrone and Favaro [310]	2014	16.53	0.65	89.34	48.57	18.90
Zhou et al. [362]	2021	17.82	0.62	-	-	-
Kaufman and Fattal [367]	2021	18.89	0.71	-	-	-
Proposed	-	**19.50**	**0.75**	**57.24**	**31.16**	18.59

Moreover, we perform an ablation analysis on GoPro dataset [4] to quantize the proposed modifications. Hence, we gradually inject the proposed changes, including multi-scale images and corresponding gradients input, DRB block, dilated convolution, and residual image learning. Table 3.2 summarizes the results of each proposed step. It is evident from the performance metrics that each proposed change gradually improves the result. Remarkably, residual image learning shows considerable improvement by recovering high anatomical details in an image. The outputs of Stage I and Stage II are depicted in Figures 3.5, 3.6, and 3.7, respectively.

Table 3.2: An ablation study of proposed GAN on GoPro dataset [4].

Proposed network changes	PSNR	SSIM
Generative network (autoencoder)	28.08	0.897
+Multi-scale images input	29.81	0.912
+Multi-scale gradient input	29.97	0.928
+DRB	31.39	0.935
+Dilated Conv	31.56	0.942
+Residual image learning (proposed)	**33.10**	**0.954**

We also measure computation cost of our proposed network. It consumes 144.63 GFLOPS, which is approximately one-fifth that of deblurGAN [14]. Likewise, it takes around half of a second to produce a deblurred image.
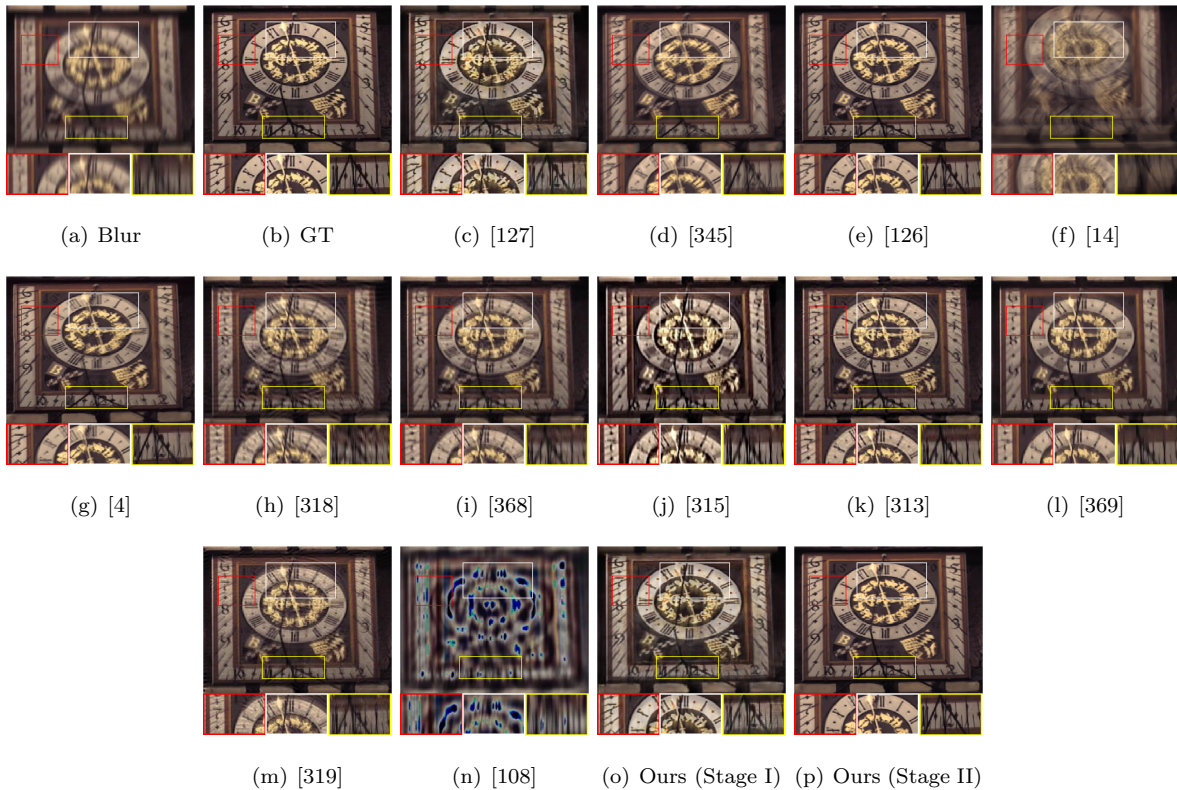
Figure 3.5: Visual deblur results of various studies for Köhler [5] dataset. GT represents ground truth image.
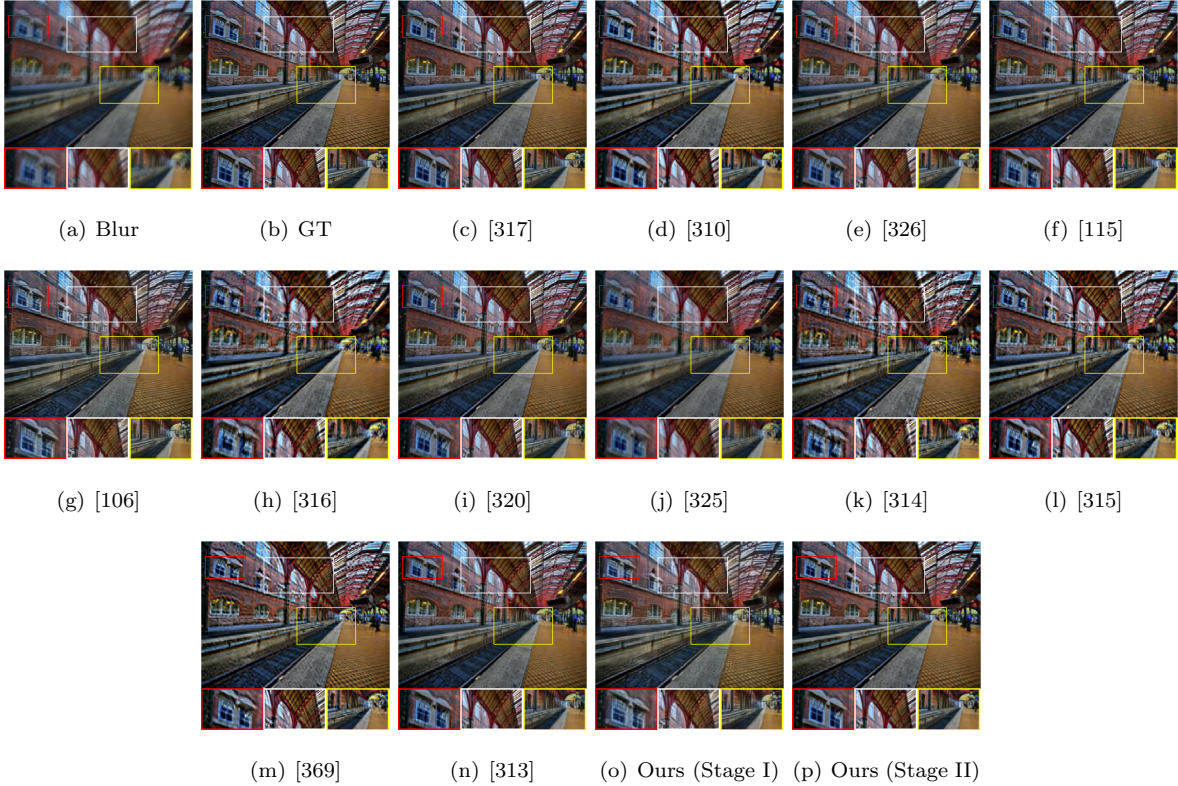
Figure 3.6: Visual deblur results of various studies for Lai et al.[6] dataset. GT represents ground truth image.

3.5 Discussion

Conventional model-based approaches have been commonly used for blind deblurring. Most of these deblurring studies have utilized regularization techniques to estimate blur kernel, and considerable attention has been paid to estimate the regularization term that best describes prior of natural images [370, 371, 372, 373]. However, such optimization techniques are not efficient due to various noises that can drastically affect estimated kernel quality, an iterative framework that requires prior information based on some statistical assumptions, and many more. Contrarily, deep learning-based methods process blur images in an end-to-end manner that effectively minimises errors [328]. Furthermore, generative priors make a more effective breakthrough compared to classical priors [4, 14]. Following, multi-scale and multi-level approaches are introduced with generative priors to improve performance [332, 126, 344, 345, 333, 127]. In a multi-level approach, several auto-encoders are combined blockwise in cascade form to process the same image at various levels. However, utilizing multi-level technique has several demerits such as high dependency on preceding outputs, over-fitting, exponential increase in computation cost, etc. Therefore, we adopt single multi-channel input with multi-scale images concatenated with corresponding gradients to explore salient features in view of deblurred images.
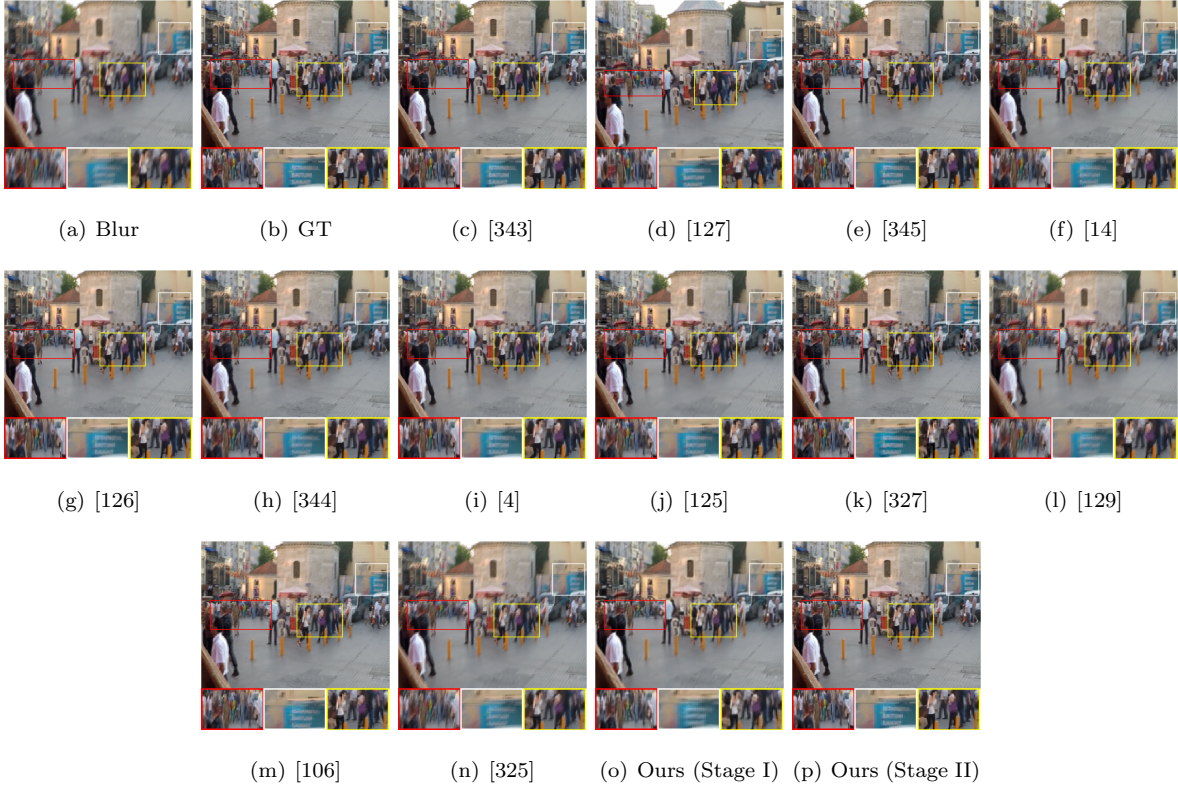
Figure 3.7: Visual deblur results of various studies for GoPro [4] dataset. GT represents ground truth image.

Furthermore, as discussed in section 3.4.4, restoring textures and structures with sparsity result in high-frequency content often referred to as rippling or ringing artifacts. Subsequently, it corrupts hallucinated blur information. Yuan et al. [116] address this issue for the very first time and suggest a model named residual deconvolution. Later, several conventional model-based [311, 117] and deep learning-based [102, 103, 296] studies have utilized this method to minimize artifacts. Nevertheless, we utilize a residual image learning block with dilated convolution and end-to-end training to recover such lost information. Additionally, a hybrid content-loss function with the perceptual loss is proposed to combine the merits of MSE and MAE (discussed in section 3.3.4). Besides that, a loss function with gradient penalty is included to retain sharp boundaries within a generated image. Our proposed neural network is trained to process images based on the intensity of an image. Therefore, the learned network might be less effective when the input image contains any outliers and significant noises. Thus, a denoising filter can be applied before the proposed deblurring process. The limitation of training data always remain problematic for the generalized deep learning model [374]. Consequently, an unsupervised deep neural network could be a choice for deblurring in the future [331, 375].

3.6 Conclusion

The goal of this study is to deblur images with the minimal structural loss. To this end, we develop a multi-scale generative adversarial network (GAN) with residual image learning to process uniform and non-uniform blur. Proposed GAN takes input in the form of multi-scale images concatenated with the corresponding gradients. To avoid ringing artifacts in a deblurred image, we utilize a residual image learning network with a hybrid content loss function and end-to-end training. The deblurred output is quantitatively evaluated with eight reference and no-reference metrics: PSNR, SSIM, AMSE, PIQE, and SSEQ. In contrast, qualitative evaluation consists of comparative analysis with the previous traditional model- and deep learning-based studies. Our analysis shows the proposed method is effective to achieve the required objectives while retaining structural properties. To the best of our knowledge, this is the first study to extensively conduct qualitative and quantitative analysis of deblurred images. One direction of future work could be the development of an unsupervised network to address training data limitations.

4 Multi-level GAN based Enhanced CT Scans for Liver Cancer Diagnosis

As discussed in Chapter 3, after removing artifact and noise with in the acquired CT scans, there comes a problem of uniform intensity levels for various tissues and adjacent organs with vague boundaries that can result in imprecise segmentation and false classification of malignant liver. I propose an unsupervised neural network to enhance low-contrast CT scans, while the conventional techniques result in non-uniform contrast images that could adversely affect diagnostic performance. For fair comparative analysis, previous studies have been evaluated on common CT scans. This chapter covers the Objective 3 of this dissertation.

Abstract

Liver cancer diagnosis requires preprocessing of images with preserved structural details. In this study, a multi-level generative adversarial network (GAN) is proposed to enhance computed tomographic (CT) images. The generated enhanced images are used to perform computer-aided diagnosis between malignant and normal liver. Three publicly available datasets, Ircadb, Sliver07, and LiTS, are used to investigate the performance of the proposed method with qualitative and quantitative analysis, namely performance metrics and computer-aided diagnosis. The mean structure similarity index of 0.45 and peak signal-to-noise ratio of 16.20 dB is achieved for the metric analysis. The AlexNet is adopted to perform binary classification with the testing accuracy of 90.37% and 85.90% for enhanced and non-enhanced images, respectively, which demonstrates the effectiveness of the proposed multi-level GAN in producing enhanced biomedical images with preserved structural details and favourable reduction in artifacts. Moreover, the consistently better performance among three datasets confirms the merits of the proposed multi-level GAN for computer-aided diagnosis.

4.1 Introduction

The prevention of various deadly diseases in human beings requires adequate detection of affected portions. However, the biomedical images obtained using computed tomography (CT) or magnetic resonance imaging (MRI) are primarily structural images that can be low in contrast due to poorly illuminated environments and may contain noise that can inevitably be amplified due to simple brightness adjustment techniques [68]. Also, such distortions affect accuracies while processing or analyzing the images through state-of-the-art algorithms due to intensity-based performance [272, 23, 44, 181, 31, 240]. Similarly, the liver size varies with body shape, sex and age. The malevolent tissue often has low contrast with the normal tissue, thus making malignant tissue detection difficult [164]. Hence, for adequate diagnostic analysis, some pre-classification steps such as enhancement, deblurring, denoising etc., are required for localization and lesion size detection [284, 11].

The enhancement problem can be defined by classic Retinex model decomposition [376],

$$I = R \circ L \tag{4.1}$$

where I is the output image, ($\circ$) is known to be element-wise multiplication, R and L are the reflection and luminance of an image.

Due to the spatial and sequential image acquisition, CT modality is more common in practice for abdomen image analysis. Moreover, the axial formation of CT images with ideal intervals reduces the average of lesions.

Various phases acquire CT scans with a unique intensity range that includes non-contrast, arterial, portal venous, or total body opacification phase [281, 305, 306, 307, 96]. However, the same phase can even produce scans with variable intensities. It may be attributed to variable exposure or depth of field during one complete scan. Figure 4.1 depicts series of portal venous phase CT scans for the same subject. It is evident from Figure 4.1 that series of scans can have variable intensity range that can adversely affect the performance of intensity-based algorithms.

Figure 4.1: The portal venous phase CT scans of a subject, truncated at a best viewable intensity range of [180-255].

The low-contrast image enhancement problem is addressed using the traditional model- or deep learning-based studies [377, 378]. The traditional model-based techniques consider an iterative process with a cost

function optimization. In contrast, deep learning-based techniques perform end-to-end or adversarial training to improve input image illumination. To the best of our knowledge, very few studies have focused on CT image enhancement for the computer-aided diagnosis (CAD) [379, 380, 381, 382, 383, 384]. de Haan et al. [385] give an overview of deep learning-based microscopy image enhancement techniques. Chen et al. [386] utilize a deep convolution neural network for image enhancement with end-to-end training. Georgieva et al. [387] integrate a wavelet decomposition with a homomorphic filter for restoration and gamma correction. Mohammed et al. [388] formulate a conventional model-based stochastic approach to enhance endoscopy images. Munadi et al. [389] enhance X-ray chest images by applying emphasis filter and histogram equalization. Deep neural networks (DNNs) have performed satisfactory for super-resolution [102, 103, 290], image sample estimation [105, 286, 287], and style tansfer [309, 14, 121, 296] applications. Nevertheless, there remains a room of improvement in yielding low-contrast image enhancement techniques.

In this study, we propose a deep learning-based multi-level generative adversarial network (GAN) to enhance low-contrast images. Our proposed generator uses a three-stage simplified auto-encoder structure connected in the cascade form to process images at three variant scales with an adversarial training. Moreover, the proposed auto-encoder contains various dilation rates to explore distinctive features and skip connections to keep track of input. Three publicly available datasets, namely Ircadb [7], Sliver07 [8], and LiTS [9] are used to create the contrast-enhanced image dataset with the help of radiologists. Two evaluation indices, peak signal-to-noise ratio (PSNR) and structure similarity index (SSIM) [302], are used to measure the quantitative performance. In addition, the AlexNet [228] with transfer learning is used to carry out qualitative analysis for liver cancer diagnosis. Finally, the results obtained by two different analyses validates the performance of proposed method.

The contributions of this study are as follows,

- We utilize GAN with auto-encoder blocks at three variant scales in generative network to enhance low-contrast CT scanned images;
- We evaluate the proposed multi-level GAN on three publicly available datasets with both qualitative and quantitative analysis;
- We reproduce several state-of-the-art algorithms for fair comparative analysis.

4.2   Related Work

4.2.1   Traditional Model-based Techniques

Traditional model-based techniques for image enhancement can be categorized into histogram equalization [390, 391, 392, 117, 393, 379], Retinex theory [256, 394, 395, 161], and filters [396].

The histogram equalization methods are used to manipulate the image histogram shape. Such traditional methods include histogram equalization [91] and adaptive histogram equalization [397]. These methods are

widely applied for contrast enhancement. However, they can amplify artifacts and noises by stretching the typical dynamic range for uniform distribution of the target histogram. Similarly, Somal [398] compares local and global histogram equalization techniques for image enhancement.

Retinex theory considers reflection and illumination to form an image [376]. In contrast, some methods remove illumination by considering reflectance as enhanced result. The single-scale Retinex (SSR) [399] and multi-scale Retinex [400] are considered representative algorithms that uses Gaussian filters to discriminate illumination and reflectance. Lee et al. [256] propose a weighted SSR model to eliminate luminance. Li et al. [161] regularize a conventional Retinex model with noise to improve the image quality. The Lagrange multiplier is utilized for the cost function minimization. Ying et al. [401] consider multi-exposed images and fuse them based on a weighted matrix obtained using a camera response model to produce well-exposed synthetic images. Fu et al. [402] develop a maximum-a-posterior (MAP) based technique to estimate reflectance and illumination separately for enhanced images. Guo et al. [260] propose a LIME enhancement model. Three channels, red, green, and blue, are used to estimate illuminate structures using MAP. Afterwards, the illumination map is refined by considering structure prior.

Pătraşcu [403] performs image processing using fuzzification. Fu et al. [404] perform regularization with gradient priors to estimate processed images. Cai et al. [395] formulate a least-square model to investigate intrinsic and extrinsic properties of an image. The shape, illumination, and texture priors are used to outperform the conventional Retinex model. Lin et al. [405] consider a Chi-squared based cost function and a piecewise brightness transform function to enhance microscopic images. Liu et al. [307] propose an enhancement framework that combines the merits of a DNN and MAP optimization. Similarly, few more studies have formulated conventional model-based techniques based on image formation model, illumination estimation, filtration, etc., [406, 345] to address the low exposure problems.

4.2.2 Deep Learning-based Techniques

Lore et al. [407] utilize a deep auto-encoder with end-to-end training to remove noise and dark regions in low-contrast images. Gharbi et al. [408] propose a convolution neural network (CNN) with affine transformations to enhance a low-resolution image. The coefficients of the locally affine model are predicted bilaterally from a low-resolution image. The transformations are then upsampled in the slicing mode and applied to a full resolution image. Anwar et al. [409] perform image enhancement for underwater images. The synthetic images with various wavelength effects are produced to train CNN with mean-squared error (MSE) and SSIM loss functions. Wei et al. [410] develop a deep Retinex-Net with two sub-networks, namely Decom-Net and Enhance-Net, to enhance images. The Decom-Net decomposes an image into reflectance and illumination image. The Enhance-Net uses the illuminated image to improve luminance and reflectance image to reduce noise.

Similarly, Fu et al. [411] develop two sub-models to simultaneously adjust brightness and exposure shift for

low illumination problems. Xu et al. [412] develop a frequency-based decomposition and enhancement model to process the images. The proposed model recovers image objects in low-frequency layers and computes high-frequency detail based on extracted image objects information. Moreover, a low-light image dataset is introduced with real noise. Lv et al. [413] develop a multi-branch CNN with end-to-end training. A synthetic dataset is developed with low light illumination. Afterwards, two attention maps, the first one to explore underexposed regions and the second to distinguish noise from real textures, are learned with a structural- and content loss function. Shen et al. [414] propose a retinal fundus image-enhancing neural network by considering multiple degradation factors and modeling retinal images separately for low quality, retinal structure, and image correction. The processed images for the three unique factors produce enhanced images.

Deng et al. [415] propose an enhanceGAN that takes input in the form of non-paired images and processes to perform piecewise color enhancement. de Stoutz et al. [416] utilize GAN with perceptual loss function to enhance images. Islam et al. [417] utilize a GAN with a Unet based autoencoder network having content and perceptual based loss function. Uplavikar et al. [418] explain the diversity of wavelength dependent underwater images caused by light attenuation due to the variation in depth and quality of ocean water. Ten various types of underwater images are considered for the training of a GAN. Jiang et al. [77] propose an edge-enhanced GAN with a global-local discriminator, unsupervised adversarial learning, and a regularized perceptual loss function to minimize error. Shamsabadi et al. [419] propose an image enhancement filter with structure-aware adversarial learning. Ni et al. [420] propose an unsupervised GAN with unpaired data to perform image enhancement. Fidelity-, identity-, and quality losses are introduced to upgrade the image attributes. Chen et al. [421] utilize Unet structure in generator to explore global features. Adaptive Wasserstein loss (WGAN) is preferred over WGAN with gradient penalty. Moran et al. [422] fuse elliptical, graduated and polynomial filters with a deep neural network to provide a regularization and adjustment for visually pleasing results.

Wang et al. [423] and Guo et al. [424] formulate custom loss functions to perform unsupervised learning of CNNs. The exposure, color, and illumination-based loss functions are computed for no-reference evaluation. Park et al. [425] utilize reference images with distorted color distribution for reinforcement learning of image enhancement. Similarly, various schemes have been proposed to address image enhancement problems using CNN [426, 427, 428, 429, 2, 430, 431, 432, 298, 433, 434, 435, 436, 437, 438] and GAN [439, 294, 440].

4.2.3 Generative Adversarial Network

A classic GAN contains a generator G and discriminator D to solve the image-to-image translation problem [105]. The generator task is to learn the input data distribution so that the generated synthetic data has the same distribution as the real data. In contrast, the discriminator task is to distinguish between real and generated data. In other words, the generator task is to minimize data distribution error, while the discrim-
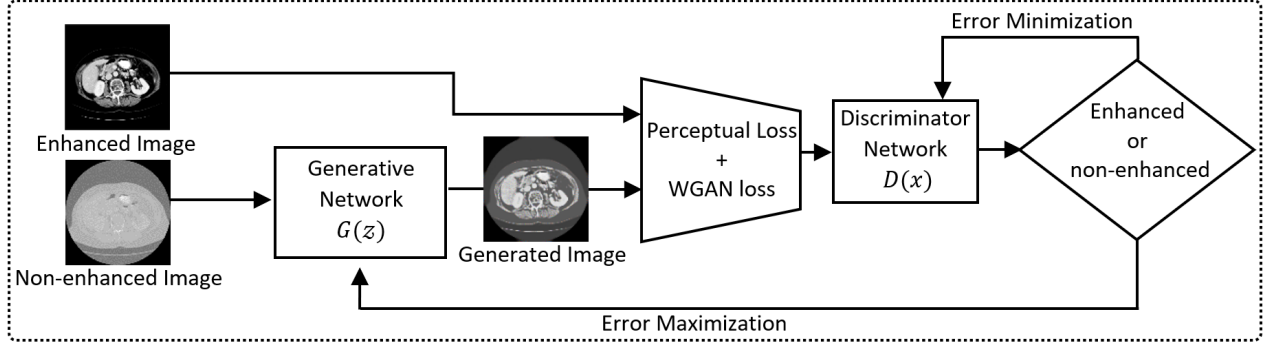
Figure 4.2: The workflow of the proposed GAN.

inator task is to maximize the error. Thus, the generator and discriminator are trained in an adversarial manner.

The training workflow of the proposed GAN is shown in Figure 4.2. The generative network takes the input of a low-contrast image and tries to generate the estimate of the corresponding enhanced image. Therefore, the GAN is a min-max problem that is defined mathematically as,

$$\min_G \max_D \mathcal{V}(D, G) = \mathbb{E}_{I \sim p_R} [\log(D(I))] + \mathbb{E}_{B \sim p_g} [\log(1 - D(G(B)))] \qquad (4.2)$$

where $\mathbb{E}$ is the expectation operator, I is a sample from the reference enhanced images distribution p_R, B is a sample from the low contrast images distribution p_b, and p_g is the generated data distribution with a mapping of low contrast and reference data space [346, 14]. The discriminator minimizes the term $\log(D(I))$ and maximizes the term $\log(D(G(B)))$.

The discriminator and generator networks are trained using perceptual and Wasserstein loss functions (discussed in section 4.3.3). The trained generator is kept only during testing.

4.3 Proposed Method

4.3.1 The Generator

We develop a multi-level generator in accordance with the studies by Nah et al. [4] and Tao et al. [126], and train in an adversarial manner to perform abdominal CT scan enhancement. The schematic of the proposed generator is shown in Figure 4.3. The generator contains three encoder-decoder blocks. Each encoder-decoder block contains symmetric CNN structure that downscale features to explore spatial details and then upscale features to meet the desired output size. Skip connections are added to combine various levels of information with fast convergence and to solve vanishing gradient problem. The three blocks at three variant scales 32, 64, and 128 are combined in a cascade form. Each block contains two 3×3 convolution layers in the encoder section, one 3×3 convolution layer in the bottleneck section, and two 3×3 deconvolution layers in the decoder section with the rectified linear activation function (ReLU). Mathematically, ReLU can be written

as,

$$y = \max(0, x) \tag{4.3}$$

where x and y are the input and output features.

The output consists of a 1×1 convolution layer with a hyperbolic tangent activation function. The batch normalization performs regularization throughout the network. In addition, each layer contains a different dilation rate D_j to explore unique features, where j represents the dilation rate of 1,2 or 4. The low contrast input images are highly correlated to output enhanced images. Therefore, skip connections are added to assist in generalizing training. The proposed encoder-decoder block of the generative network can be written as,

$$y^i = F\left(x_l{}^i, w_j{}^i\right) + x^i \tag{4.4}$$

where $F\left(\cdot\right)$ is the encoder-decoder block, $x_l{}^i$ is the lth convolution output and $w_j{}^i$ is the weighted dilated convolution with j as a dilation rate at the ith scale, x^i and y^i represents the input and output of encoder-decoder block with i^{th} scale index such that $i = 1$ is the coarsest scale. $F\left(\cdot\right)$ in our study can be written as,

$$F\left(x_l{}^i, w_j{}^i\right) = hw_1^i rw_1^i rw_2^i rw_4^i \left[rw_2^i \left(rw_1^i \left(x^i\right)\right)\right] \tag{4.5}$$

where $r\left(\cdot\right)$ represents a composite function of ReLU and batch normalization, h denotes a hyperbolic tangent activation function.
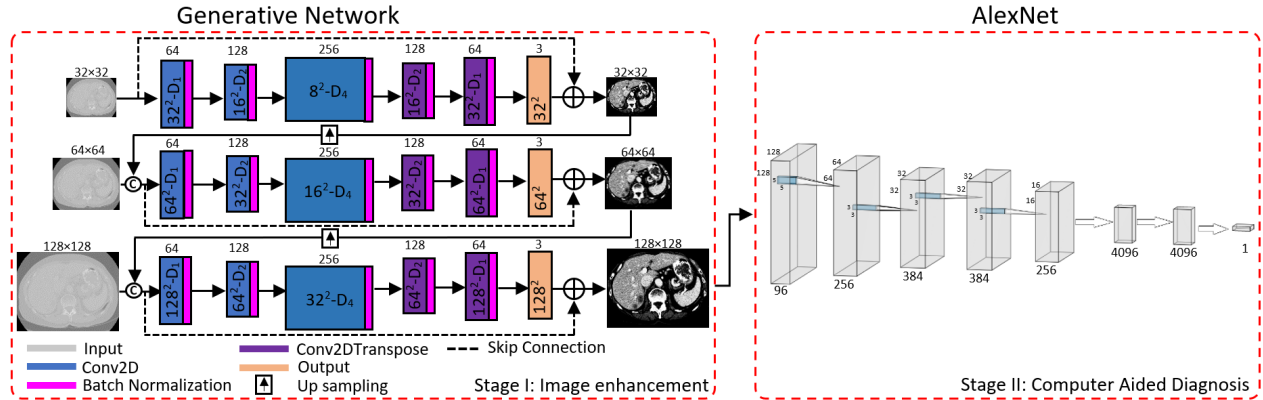
Figure 4.3: The schematic of the proposed image enhancement multi-level generative network with CAD. $\oplus$ and ©️ denotes the addition and concatenation operation.

Coarse-to-Fine Level Network

The proposed multi-level generative network processes images in a coarse-to-fine manner. Due to computation limitations, three input image resolutions of 32, 64, and 128 with 3 channels input and a 0.5 scale difference are considered for a multi-level network. Therefore, we have 6 channels at the next finer level, making feature maps of $(64, 64, 6)$ and $(128, 128, 6)$ at each finer level. In total, the proposed generative network contains

eighteen convolution layers. Several studies [4, 126] process images in a multi-level fashion with the largest scale input at the finest level. Similarly, we consider the smallest scale input for the coarsest level and desired output (largest) scale at the finest level to have a large enough receptive field to explore at the finest level. Likewise, considering the smallest scale at the coarsest level, we can have a large enough receptive field to cover the whole patch of an image.

Moreover, up-convolution [441], reshaping [442] etc., operations have been used to match the output size of the coarsest level to the input of the next finer scale. In contrast, we utilize an up-sampling operation to avoid any checkerboard or ringing artifacts due to stride convolution. The up-sampled features are concatenated with the next finer scale low contrast image input. The proposed multi-level neural network is written as,

$$I^i = N_{ML}\left(B^i, I^{(i-1)\uparrow}; \beta_{ML}\right) \tag{4.6}$$

where i represents the scale index with $i = 1$ for the coarsest level, N_{ML} is our proposed multi-level neural network with training parameters as β_{ML}, I^i and B^i are the enhanced and low contrast images at the ith scale and $(\cdot)\uparrow$ denotes an upscaling operator for the $(i-1)^{th}$ to i^{th} scale.

4.3.2 The Discriminator

In adversarial training, the discriminator maximizes the error between the generated and the reference output. In contrast, the generator tries to fool the discriminator by generating output close to the reference. Therefore, the discriminator can have a simplified binary output. We consider discriminator network according to the study Kupyn et al. [14]. It contains five 4×4 convolution layers with LeakyReLU activation function and the batch normalization. Afterwards, two fully connected layers with the hyperbolic tangent and the sigmoid activation function are added to get the binary output. The schematic of the discriminator network is depicted in Figure 4.4.
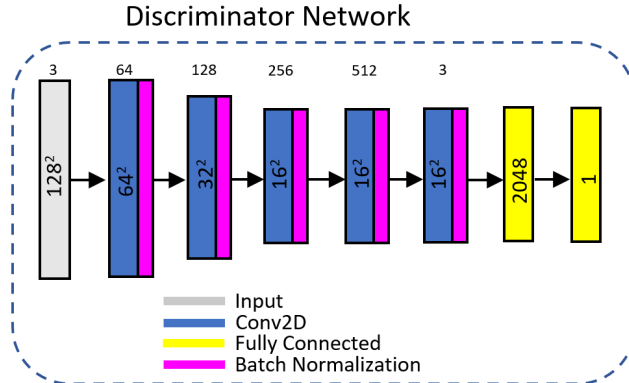
Figure 4.4: The schematic of discriminator network.

4.3.3 Loss Function

The proposed multi-level GAN is trained in an adversarial manner to generate enhanced images. Thus, the generator network weights are updated based on discriminator network output that acts as a critique for the generator network by maximizing error between reference and generated images. Therefore, the net loss of the proposed enhancement network is defined as,

$$\mathcal{L}_{net} = \mathcal{L}_{adv} + \mathcal{L}_{con} \tag{4.7}$$

where $\mathcal{L}_{adv}$ and $\mathcal{L}_{cont}$ are the adversarial and content loss functions explained as follows.

Adversarial Loss

The proposed GAN takes input in pairs as low contrast image B and the corresponding reference image I. The low contrast image B is processed at three variant scales to produce enhanced images (discussed in section 4.3.1). We apply the similar discriminator loss function as in the study Wasserstein GAN [348] given as,

$$\mathcal{L}_{adv} = \sum_{n=1}^{N} -D_{\theta}\left(G_{\theta}\left(B\left(n\right)\right)\right) \tag{4.8}$$

where N represents a number of images in a batch, B is a low contrast input image, G_{θ} denotes the generator network output, and D_{θ} represents the binary label output of the trained discriminator.

Content Loss

Multiple studies [421, 439, 417, 432, 420, 443] have considered perceptual content loss function to enhance low contrast images. For style transfer applications, the perceptual loss is computed using trained CNN feature maps to avoid low contrast regions in an enhanced image. We use the pre-trained VGG-16 [339] neural network on the ImageNet dataset to compute perceptual features for the generated and reference output. Thus, the content loss in terms of perceptual loss function ϕ is defined as,

$$\mathcal{L}_{con} = \frac{1}{2N} \sum_{n=1}^{N} \left[\frac{1}{W_{x,y}H_{x,y}} \sum_{i=1}^{W_{x,y}} \sum_{j=1}^{H_{x,y}} \left(\phi_{x,y}(I_R(n))_{i,j} - \phi_{x,y}(I_G(n))_{i,j} \right)^2 \right] \tag{4.9}$$

where $W_{x,y}$ and $H_{x,y}$ are the output feature map dimensions of the VGG-16 layer, I_R and I_G are the reference and generated output with spatial indices (i, j), respectively. We utilize the "*block1_pool*" output layer feature map of VGG-16 for function ϕ because higher layers provide spatial information while lower layers provide spectral information. Hence, lower layers help better in keeping track of structural information.

4.3.4 Processed Images Evaluation

We perform qualitative and quantitative analysis to validate the proposed GAN performance. The qualitative analysis consists of two performance metrics, namely SSIM [96] and PSNR. In quantitative analysis, we

perform deep learning-based binary classification for liver cancer diagnosis. The performance metrics are explained as follows.

Structure Similarity Index

Image enhancement studies [444, 445, 306, 446] have utilized SSIM to evaluate model performance. The SSIM is assessed on three factors, namely luminance $l(I_R, I_G)$, contrast $c(I_R, I_G)$, and structure $s(I_R, I_G)$. Mathematically it can be written as [302],

$$SSIM(I_R, I_G) = \left[l(I_R, I_G)^{\alpha} \cdot c(I_R, I_G)^{\beta} \cdot s(I_R, I_G)^{\gamma} \right] \tag{4.10}$$

here

$$l(I_R, I_G) = \frac{2\overline{I_R}\,\overline{I_G} + C_1}{\overline{I_R}^2 + \overline{I_G}^2 + C_1}; \quad c(I_R, I_G) = \frac{2\sigma_{I_R}\sigma_{I_G} + C_2}{\sigma_{I_R}^2 + \sigma_{I_G}^2 + C_2}; \quad s(I_R, I_G) = \frac{\sigma_{I_{RG}} + C_3}{\sigma_{I_R}\sigma_{I_G} + C_3}$$

where $\overline{I_R}$ and $\overline{I_G}$ are the local means of reference I_R and generated image I_G, respectively, σ_{I_R} and σ_{I_G} are the corresponding standard deviations, respectively, and $\sigma_{I_{RG}}$ is the covariance of reference and generated image while (C_1, C_2, C_3) are the regularization constants for the three factors, avoiding instability where the standard deviation or local mean is zero. These constants are computed as, $(C_1, C_2, C_3) = \left[(0.01 * R)^2, (0.03 * R)^2, C_2/2 \right]$, where R is the dynamic range of pixels. SSIM with $\alpha = \beta = \gamma = 1$ is adopted in this study and is simplified as,

$$SSIM = \frac{\left(2\overline{I_R}\,\overline{I_G} + C_1\right)\left(2\sigma_{I_{RG}} + C_2\right)}{\left(\overline{I_R}^2 + \overline{I_G}^2 + C_1\right)\left(\sigma_{I_R}^2 + \sigma_{I_G}^2 + C_2\right)} \tag{4.11}$$

Peak Signal-to-Noise Ratio

PSNR is to measure the ratio of a maximum pixel intensity to mean square error of generated image with respect to reference image and is typically expressed in logarithmic decibel scale given as [447, 418, 434],

$$PSNR_{dB} = 10 \log_{10} \left[\frac{[\max(I_G)]^2}{MSE} \right] \tag{4.12}$$

where MSE is the mean squared error between the reference I_R and generated image I_G.

Liver Cancer Diagnosis

The processed high contrast images assist in improving the machine learning based disease diagnosis. We perform qualitative analysis between processed and non-processed images for malignant liver diagnosis. Multiple CNN studies have classified malignant liver between lesions (tumors) versus benign (cyst), secondary versus primary cancer, cholangio versus hepatocellular carcinoma, cirrhosis versus fibrosis, etc. [36, 35, 34, 240]. We sort the Ircadb, Sliver07, and LiTS datasets into a malignant and normal dataset. The former type covers severe liver diseases such as hepatocellular carcinoma, metastatic carcinoma, intrahepatic cholangiocarcinoma,

virus hepatitis, cirrhosis (alcoholic cirrhosis), etc. In contrast, the latter class contains scans related to benign, such as hemangioma, focal nodular hyperplasia, hepatocellular adenoma, cyst, fibrosis, non-cirrhotic chronic liver diseases (such as fatty liver or steatohepatitis), etc. Afterwards, the deep learning-based AlexNet [228] is trained to classify malignant versus normal liver. Although AlexNet is capable of classifying 1000 classes, we consider it for binary class. Therefore, the neural network consists of one 5×5 convolution layer, four 3×3 convolutional layers, and two fully connected layers with ReLU activation function, and one fully connected layer with a sigmoid activation function. The drop-out layers are added with fully connected layers to avoid over-fitting and data augmentation to improve training. The schematic of AlexNet is shown in Figure 4.3.

4.4   Experiments

4.4.1   Datasets

We consider three publicly available datasets, namely Ircadb [7], Sliver07 [8], and LiTS [9], to evaluate our proposed GAN. Moreau et al. [7] develop an image reconstruction for comparison of algorithm database (Ircadb) which contains hepatic tumors of 20 subjects with an equal number of men and women. Similarly, the medical image computing and computer-assisted intervention society (MICCAI) develops a Sliver07 dataset with CT scans from 20 subjects. Likewise, the liver tumor segmentation challenge (LiTS) dataset is produced by CodeLab that contains CT scans of 130 subjects. We develop an enhanced and non-enhanced dataset with the help of a radiologist that contains CT scans of 170 subjects with varying images for each subject that range between 150 to 550. Hence, we consider more than sixty thousand images to produce a multi-modal image enhancement dataset.

4.4.2   Training Details

The proposed GAN is implemented on the Tensorflow 2.0 platform with Intel (R) Xenon (R) having a clock speed of 2.6 Ghz and K40 NVIDIA GPU. The adversarial training is performed with 70% data and Adam optimizer with stochastic gradient descent at a learning rate of 1×10^{-4}. For more generalized training, we perform data augmentation with zoom, flip, rotation, and shear operations. Although, due to computation limitations, we consider a batch size of 5 with an output resolution size of 128×128.

4.4.3   Results and Analysis

We compare performance metrics for several traditional model- and deep learning-based image enhancement methods as illustrated in Table 4.1. These state-of-the-art methods include renown enhancement techniques, namely NPE [394], LIME [260], MBLLEN [447], Retinex-Net [410], etc. Three CT scan datasets are used to compare these studies with the proposed multi-level GAN. A minor performance difference may be observed

due to a different output resolution size of 128×128. We consider this difference as an experimental error. The best-performed metrics for each dataset are highlighted in Table 4.1. To the best of our knowledge, this is the first study to evaluate famous image enhancement methods on common CT scan datasets. Therefore, the enhanced images are extensively evaluated based on noise and structural properties. It is evident from Table 4.1 that Ibrahim and Pik Kong [448] perform second-best among Sliver07 and LiTS datasets. While histogram equalization performs better for the Ircadb dataset. In contrast, the proposed multi-level GAN outperforms among various discussed studies. It is due to the fact that we consider a generator with a multi-level auto-encoder to process images at three variant scales. Processing images at several scales improves the spatial and spectral resolution of an image.

Similarly, studies [430, 386, 422, 427] propose the global enhancement methods. In contrast, the CT scans contain non-uniform image illumination, such that low contrast dark regions with blur boundaries. Therefore, global enhancement results in low-contrast CT scan images with similar intensities for the normal and malignant liver. The visual results of several studies are depicted in Figures 4.5, 4.6, and 4.7. Although methods proposed by Zeng et al. [434] and Ibrahim et al. [448] perform better among discussed studies but still their results are unsatisfactory due to vague region of interest boundaries. Likewise, NPE [394], LIME [260], MBLLEN [447], and Retinex-Net [410] fail to perform enhancement for low-contrast bio-medical images. Comparatively, the histogram equalization performs well as depicted in the visual results. Meanwhile, we propose a local enhancement framework that develops high contrast CT scan images with distinct boundaries of liver lesions.

Furthermore, we perform a CAD to evaluate enhanced images performance. The proposed multi-level GAN output is fed into AlexNet to perform binary classification between malignant and normal liver (discussed in section 4.3.4). AlexNet is trained and evaluated using testing data of the proposed multi-level GAN that is 30% of the whole dataset. Therefore, 70% of nineteen thousand enhanced and non-enhanced images are utilized for training AlexNet. The Adam optimizer is applied with a learning rate of 0.0001 and the early stopping criterion. The early stopping criterion checks at the end of every epoch whether the monitored metric (loss) is improved (decreased) by a minimum change (Δ). If the loss does not improve in the (X) consecutive epochs, then the network stops training [449]. Therefore, the learning process stops if a training loss difference remains less than $\Delta = 0.0001$ for $X = 5$ consecutive epochs and the best weights are restored. This criterion is set to avoid the over-fitting issue and to obtain a better-generalized model [450]. Figure 4.8 shows binary classification accuracy and loss curves for non-enhanced and enhanced CT scans. It can be observed that after exhaustive parameter adjustment and training of AlexNet, stable curves are achieved with 270 epochs with the training accuracy of 97.71% and the testing accuracy of 90.37% for enhanced images.

The non-enhanced low contrast dataset diagnostic accuracy stops improving after 181 epochs with the training accuracy of 91.01% and the testing accuracy of 73.63%. That means loss improves up to 176 epochs and does not improve from 177 to 181 epochs. Hence the training stopped at 181 epoch. However, the validation response of the non-enhanced model still showed gradual improvement at the 181 epoch,

as depicted in Figure 4.8. Therefore, we continue network training (with transfer learning) by changing the stopping criterion to $X = 20$. The loss converges until 196 epochs and stops improving from 197 to 216 epochs, thus giving an improved testing accuracy of 85.90% with 216 total epochs. Besides that, it is worth mentioning that diagnostic performance curves for enhanced dataset are consistent and more stable. Therefore, it can be stated that enhanced images using multi-level GAN are more reliable for the machine learning-based diagnosis.

(a) Raw   (b) Reference   (c) HE   (d) [448]   (e) [394]   (f) [402]

(g) [401]   (h) [406]   (i) [395]   (j) [260]   (k) [161]   (l) [426]

(m) [345]   (n) [447]   (o) [410]   (p) [434]   (q) Ours
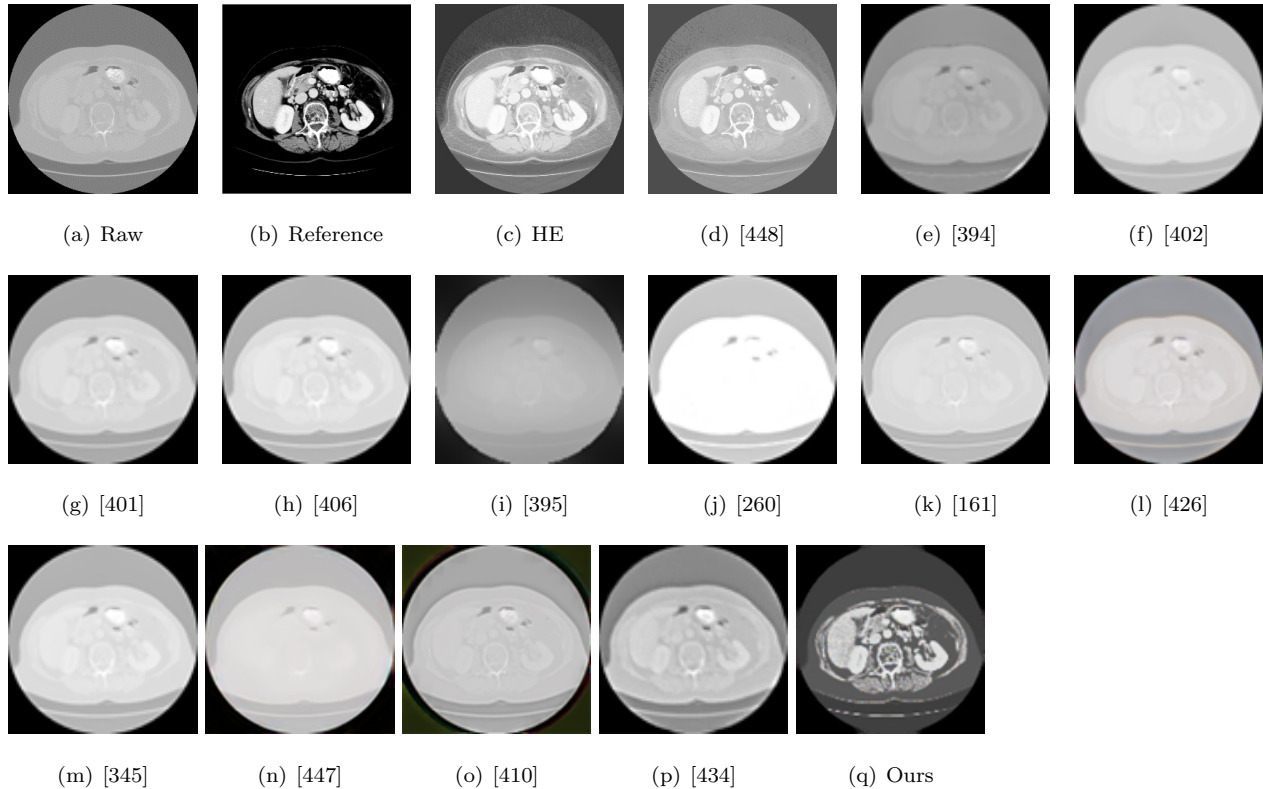
Figure 4.5: Visual results of CT scan enhancement for the Ircadb dataset [7]. HE represents histogram equalization.

4.5   Clinical Application

Clinical cancer diagnosis requires various invasive or non-invasive tests, including biopsy, blood test, medical imaging, etc. Medical imaging assists in structural analysis of a region of interest using one or more commonly used modalities, namely, CT, Ultrasonography (US), and MRI. CT is a well-establish modality used to examine tissues and organs in an abdominal scan, as depicted in Figure 4.1. However, the acquired scans are degraded due to various aspects that include instrumental optical issues such as short exposure, low optical quality, etc., and experimental problems, for instance, expiration and inspiration chest movement while scanning, heartbeat, etc. These factors can combine to affect radiologist analysis by suppressing critical

Table 4.1: Performance metrics evaluation for traditional model and deep learning-based studies on three public datasets at a resolution of (128 × 128).

Study	Year	PSNR(dB)	SSIM
Ircadb Dataset [7]			
Histogram equalization	-	7.804	0.171
Ibrahim and Pik Kong [448]	2007	6.449	0.138
Wang et al. [394]	2013	7.657	0.153
Fu et al. [402]	2015	4.579	0.136
Ying et al. [401]	2017	5.327	0.155
Ying et al. [406]	2017	4.538	0.151
Cai et al. [395]	2017	6.150	0.047
Guo et al. [260]	2017	3.401	0.121
Li et al. [161]	2018	4.544	0.139
Wang et al. [426]	2018	5.508	0.146
Zhang et al. [345]	2019	8.001	0.183
Amsr et al. [447]	2019	4.826	0.068
Wei et al. [410]	2019	5.437	0.073
Zeng et al. [434]	2020	6.030	0.171
Proposed	-	**13.0007**	**0.335**
Sliver07 Dataset [8]			
Histogram equalization	-	8.156	0.234
Ibrahim and Pik Kong [448]	2007	12.186	0.296
Wang et al. [394]	2013	7.919	0.157
Fu et al. [402]	2015	8.951	0.156
Ying et al. [401]	2017	10.119	0.169
Ying et al. [406]	2017	7.135	0.130
Cai et al. [395]	2017	10.461	0.067
Guo et al. [260]	2017	5.463	0.119
Li et al. [161]	2018	8.809	0.167
Wang et al. [426]	2018	7.103	0.118
Zhang et al. [345]	2019	12.132	0.192
Amsr et al. [447]	2019	7.995	0.091
Wei et al. [410]	2019	8.225	0.147
Zeng et al. [434]	2020	10.343	0.357
Proposed	-	**18.013**	**0.54**
LiTS Dataset [9]			
Histogram equalization	-	8.084	0.222
Ibrahim and Pik Kong [448]	2007	11.991	0.265
Wang et al. [394]	2013	7.457	0.122
Fu et al. [402]	2015	8.131	0.100
Ying et al. [401]	2017	9.396	0.124
Ying et al. [406]	2017	6.630	0.104
Cai et al. [395]	2017	10.410	0.057
Guo et al. [260]	2017	5.082	0.111
Li et al. [161]	2018	8.137	0.127
Wang et al. [426]	2018	6.325	0.085
Zhang et al. [345]	2019	11.519	0.155
Amsr et al. [447]	2019	7.582	0.068
Wei et al. [410]	2019	7.614	0.113
Zeng et al. [434]	2020	9.285	0.206
Proposed	-	**17.61**	**0.480**

(a) Raw   (b) Reference   (c) HE   (d) [448]   (e) [394]   (f) [402]

(g) [401]   (h) [406]   (i) [395]   (j) [260]   (k) [161]   (l) [426]

(m) [345]   (n) [447]   (o) [410]   (p) [434]   (q) Ours
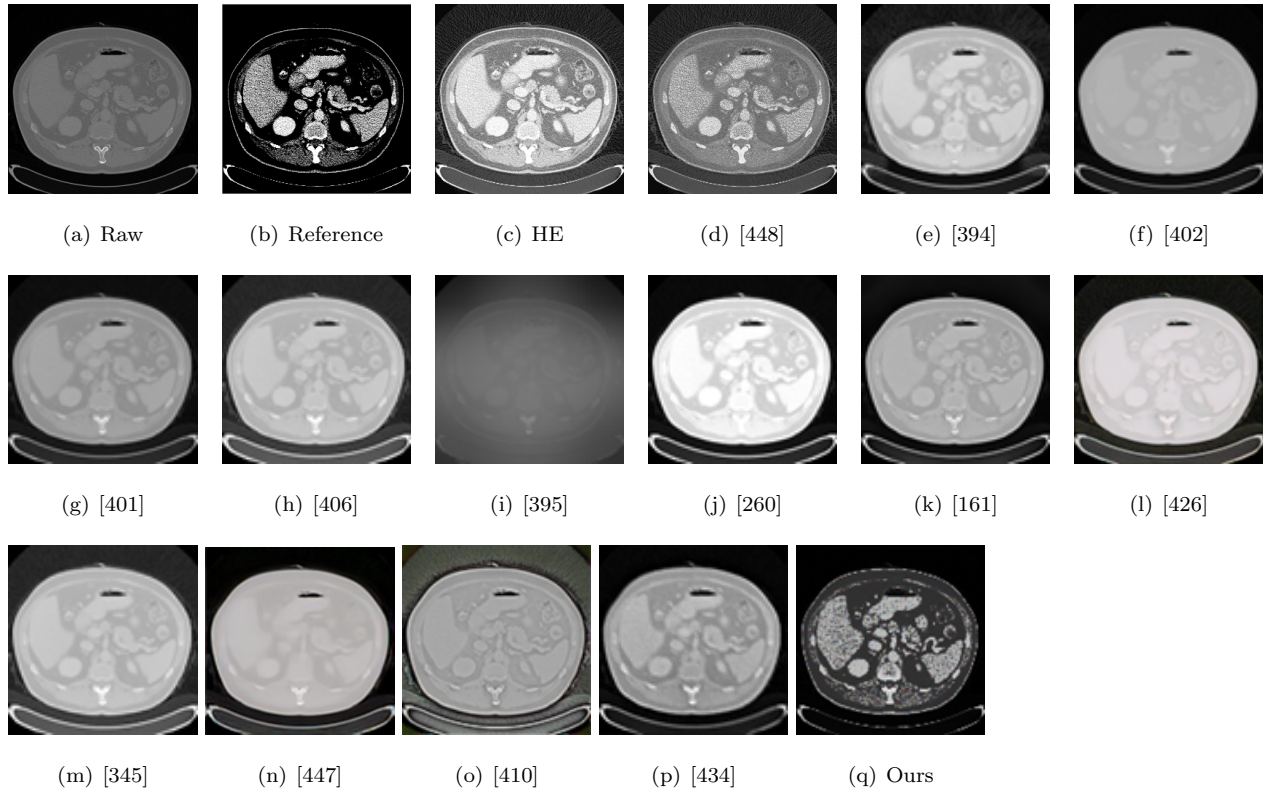
Figure 4.6: Visual results of CT scan enhancement for the Sliver07 dataset [8]. HE represents histogram equalization.

minor details.

Similarly, such degradations can influence CAD performance by impacting each preliminary step required for disease detection through state-of-the-art machine-learning algorithms. Figure 4.9 shows a general process flow for CAD. The preliminary steps mainly include segmentation and feature extraction [13]. In contrast, the low-quality images affect precise segmentation and information extraction for various diseases such as cyst, hemangioma, abscess, hepatocellular carcinoma, metastasis, etc., [245, 36, 451, 20]. Therefore, such techniques are in demand to reduce random effects and distortions in an obtained biomedical image without any structural changes.

Nevertheless, we apply a multi-level GAN that process images with auto-encoder blocks at three variant scales. Processing images with variant scales combined in a cascade form improves illumination details while preserving shape and texture properties. As a result, the output has less artifacts and high contrast regions. Thus, the proposed method can process medical images acquired from any modality to get high-contrast enhanced images with better spatial resolution and in-depth anatomical details. Afterwards, the processed images can be either analyzed by a radiologist or CAD for possible disease detection.

| (a) Raw | (b) Reference | (c) HE | (d) [448] | (e) [394] | (f) [402] |

| (g) [401] | (h) [406] | (i) [395] | (j) [260] | (k) [161] | (l) [426] |

| (m) [345] | (n) [447] | (o) [410] | (p) Ours |
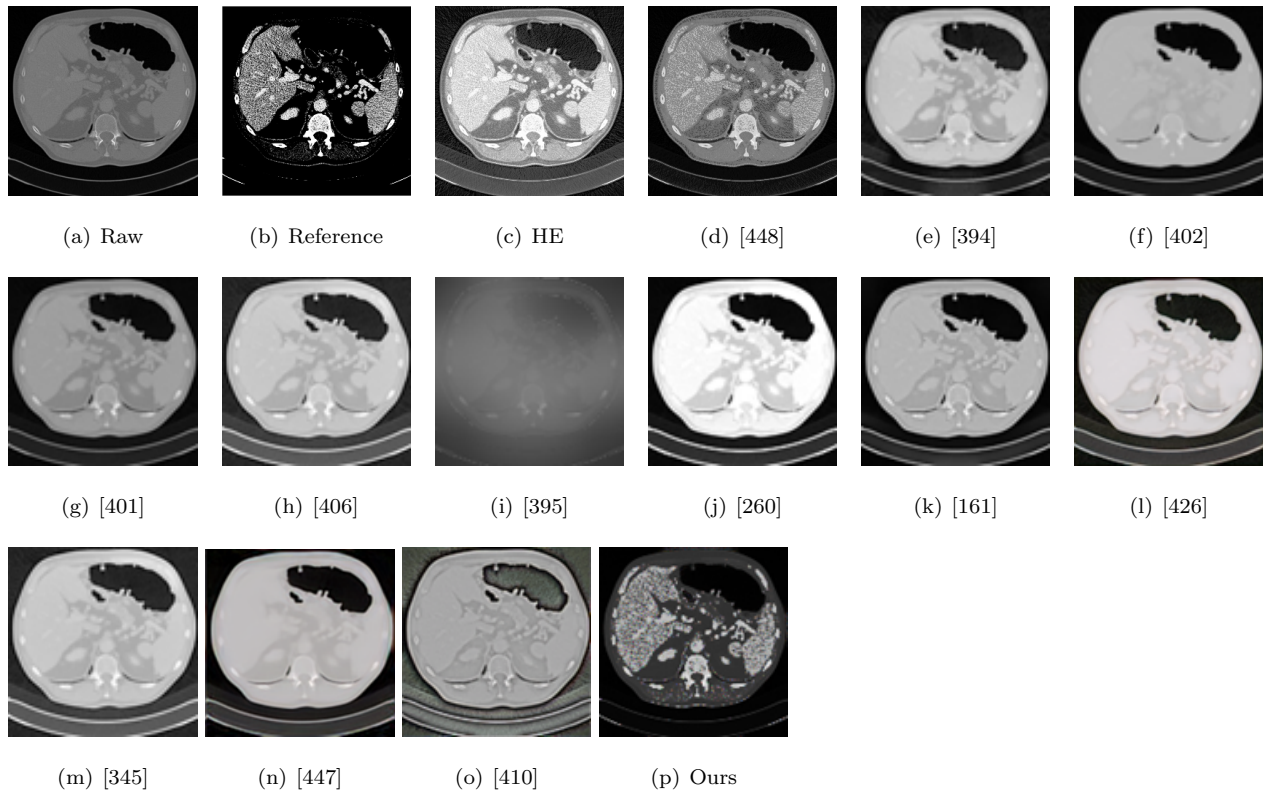
Figure 4.7: Visual results of CT scan enhancement for the LiTS dataset [9]. HE represents histogram equalization.

4.6 Discussion and Limitations

Several studies utilize traditional model-based techniques to process biomedical images. These methods commonly include histogram equalization [452], non-linear polynomial model [52], regularization [453, 454, 455, 117], fuzzy c-means clustering [307], filter [456, 306], voxel adjustment technique [457], etc. Histogram equalization introduces visual deterioration such as unnecessary and non-uniform saturation effects that result in structural loss. In contrast, regularization and fuzzy c-mean clustering produce enhanced lesions with intensity inhomogeneity. In addition, these methods require vary number of iterations for intensity correction of images that make it less applicable for large dataset processing. Moreover, the developed methodologies are good enough to reduce distortions but can cause structural changes that impact disease diagnosis. In comparison, our analyses based three datasets validate the proposed model performance in image processing and CAD.

The proposed multi-level generative network uses three encoder-decoder blocks combine in a cascade form that exponentially increases training parameters. However, to handle low illuminated scans with low contrast and dark regions, larger receptive fields are required, thus resulting in more levels of encoder-decoder blocks for image enhancement.

(a) Non-enhanced (b) Enhanced
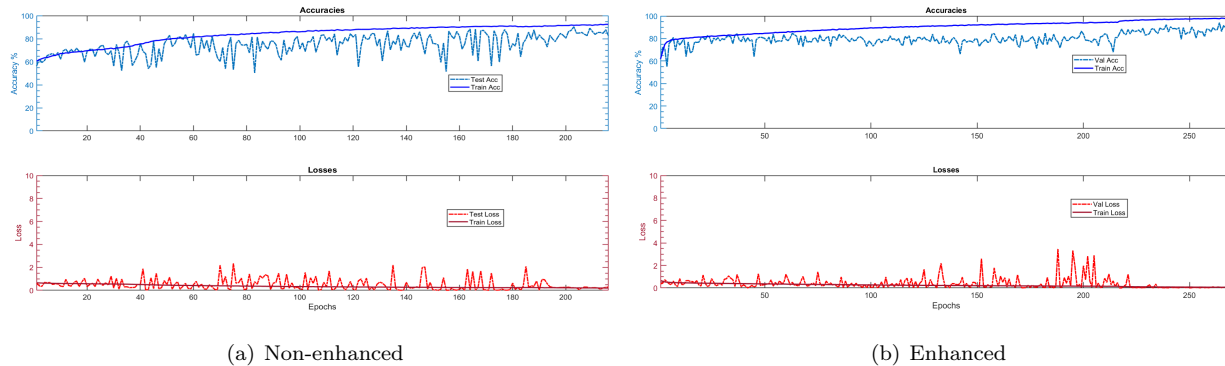
Figure 4.8: Normal versus malignant liver classification performance for (a) non-enhanced and (b) enhanced CT scans dataset.
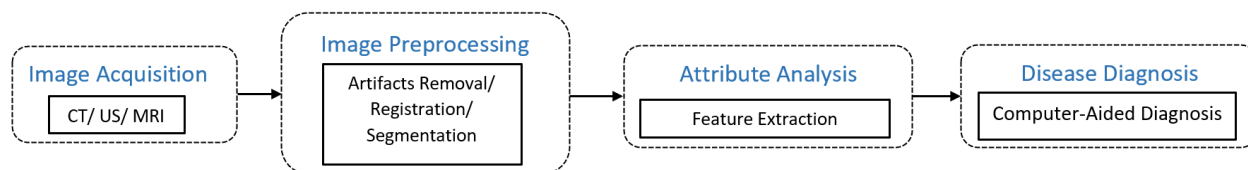
Figure 4.9: The process flow of a typical CAD system.

4.7 Conclusion

This study aims to enhance biomedical images for disease diagnosis. To this end, we develop a multi-level GAN to perform enhancement on CT scans. The images are processed at three variant scales to improve the spatial and spectral resolution of the output. Three public datasets Ircadb, Sliver07, and LiTS, are used to investigate the performance of our proposed method with quantitative and qualitative analysis. The enhanced images are quantitatively evaluated using peak signal-to-noise ratio (PSNR) and structure similarity index (SSIM). The proposed method outperforms with a mean PSNR of 16.20 dB and a mean SSIM of 0.45. In contrast, the qualitative analysis is performed using deep learning-based binary classification for enhanced and non-enhanced images. The AlexNet is used to perform binary classification between normal and malignant liver. The enhanced images surpass a testing accuracy of at least 5% compared to non-enhanced images. Our analysis shows that the proposed method is effective in enhancing images with preserved spatial and spectral resolution. One of the future studies could be developing an unsupervised or semi-supervised enhancement network to counter training data limitations.

5 RMS-UNet: Residual Multi-scale UNet for Liver and Lesion Segmentation

After processing the images for deblurring (chapter 3), and enhancement (chapter 4), segmentation is performed to extract region of interest. The malignant liver detection requires precise and robust segmentation so that minute tumors are not misclassified. A Unet has been proposed for biomedical image segmentation with remarkable performance. I design a Unet with a novel loss function to perform liver and lesion segmentation. The proposed technique can segment very small lesions that are usually ignored with conventional models. The proposed method is evaluated and compared on four publicly available datasets. This chapter accomplishes Objective 4 of this dissertation.

Abstract

Precise segmentation is in demand for hepatocellular carcinoma or metastasis clinical diagnosis due to the heterogeneous appearance and diverse anatomy of the liver on scanned abdominal computed tomography (CT) images. In this study, we present an automatic unified registration-free deep-learning-based model with residual block and dilated convolution for training end-to-end liver and lesion segmentation. A multi-scale approach has also been utilized to explore novel inter-slice features with multi-channel input images. A novel objective function is introduced to deal with fore- and background pixels imbalance based on the joint metric of dice coefficient and absolute volumetric difference. Further, batch normalization is used to improve the learning without any loss of useful information. The proposed methodology is extensively validated and tested on 30% of the publicly available Dircadb, LiTS, Sliver07, and Chaos datasets. A comparative analysis is conducted based on multiple evaluation metrics frequently used in segmentation competitions. The results show substantial improvement, with mean dice scores of 97.31, 97.38, 97.39 and 95.49% for the Dircadb, LiTS, Sliver07, and Chaos liver test sets, and 91.92 and 86.70% for Dircadb and LiTS lesion segmentation. It should be noted that we achieve the best lesion segmentation performance on common datasets. The obtained

qualitative and quantitative results demonstrate that our proposed model outperform other state-of-the-art methods for liver and lesion segmentation, with competitive performance on additional datasets. Henceforth, it is envisaged as being applicable to pertinent medical segmentation applications.

5.1 Introduction

Liver malignancies are a common cause of death. According to World Health Organization report (2018), in Europe, adults diagnosed with liver cancer between 2000 and 2007 had a mean 5-year age-standardized survival rate of 12% and a mortality to incidence of 0.95% [15]. Early detection and diagnosis may certainly help to improve the survival. Therefore, based on these and other statistics, images of liver tumors or lesions by computed tomography (CT) or magnetic resonance imaging (MRI) currently are a focus of interest for researchers in clinical diagnosis [164, 18, 19, 16, 458, 451]. Certainly, the prevention of such deadly outcome in living beings requires the timely, accurate and independent robust diagnostic imaging. However, the process requires preliminary steps, namely preprocessing and segmentation. Various studies [148, 459, 460, 164, 20] have reported tumor segmentation; however, as a means of avoiding false-negative error, the recent literature [461, 462, 463, 464, 465, 466, 467, 468, 147, 53] suggests liver segmentation as the primary image processing step. Segmentation can be divided into three main categories: manual, semi-automatic, and automatic. Semi-automatic methods require a user-defined seed allocation with some energy-based constrained optimization functions to improve segmentation [166, 169, 153]. Moreover, these models are more prone to poor performance, resulting in ambiguous background and heterogeneous appearances of abdominal CT-scan images. Meanwhile, automatic segmentation, which does not require any user interaction, has been proposed, though sufficient training data are required for optimal results [465, 469, 60, 470].

Liver cancer is mainly categorized into primary and metastatic forms. Primary liver cancer begins in the cells of the liver, and can either begin as a solitary focus, spread in the liver, or as multiple foci at the same time within the liver. In addition, after treatment, liver cancer could recur as well. People with known liver damage related to cirrhosis or hepatitis are more likely to develop several liver lesions at different sites simultaneously. Primary liver cancer is further categorized into subtypes of hepatocellular carcinoma (HCC), cholangiocarcinoma, and others. HCC, developing in the hepatocytes, is the most common type of primary liver cancer, accounting for 75% of all cases. It can spread from the liver to other parts of the body, for example the lungs. People with severe liver damage secondary to alcohol abuse, virus hepatitis, or metabolic syndromes are especially prone to HCC. Metastasis or secondary liver cancer originates when cancer cells from a primary organ metastasize to the liver. Contrary to other normal cells in the body, cancer cells can break away from their primary site, which is where the cancer began, and proceed to other regions of the body through the lymphatic system, the bloodstream, or by direct extension. Eventually these migrating cancer cells come to reside and grow in other organs; breast-, colon- or lung-cancer cells, for instance, can metastasize to the liver. Therefore, as a means of heading off metastasis or recurrence, the precise tumor

segmentation is in high demand. Due to the limitation of training data, complex boundaries and diverse anatomy on abdominal CT scans, the proposed architectures [166, 169, 153, 471, 472] are lacking in terms of computational cost and performance due to cascaded multiple models that highly rely on prior ones.

Convolutional neural networks have been widely used in various forms for automatic segmentation; one such example is UNet for the biomedical-image segmentation [441]. UNet is a U-shaped deep-learning network divided into three main sections: down-sampling, bottleneck, and up-sampling. The deep neural network (DNN) is famous for the biomedical image segmentation that consists of 18 3×3 convolutional layers with the rectified linear activation function, 4 2×2 max pool layers, and 4 2×2 up-convolution layers. Various architectures [471, 472, 426, 473, 177], have utilized the basic UNet DNN in their proposed methodologies with variable kernel sizes, 3-D convolution and cascaded multiple models.Unfortunately, increasing network layers with large kernel sizes not only increases computation time but sometimes can result in overfitting. To overcome these drawbacks, dilated convolution in various forms can be utilized to gain large-scale features with a constant kernel size [474, 475, 476, 477].

| (a) | (b) | (c) |
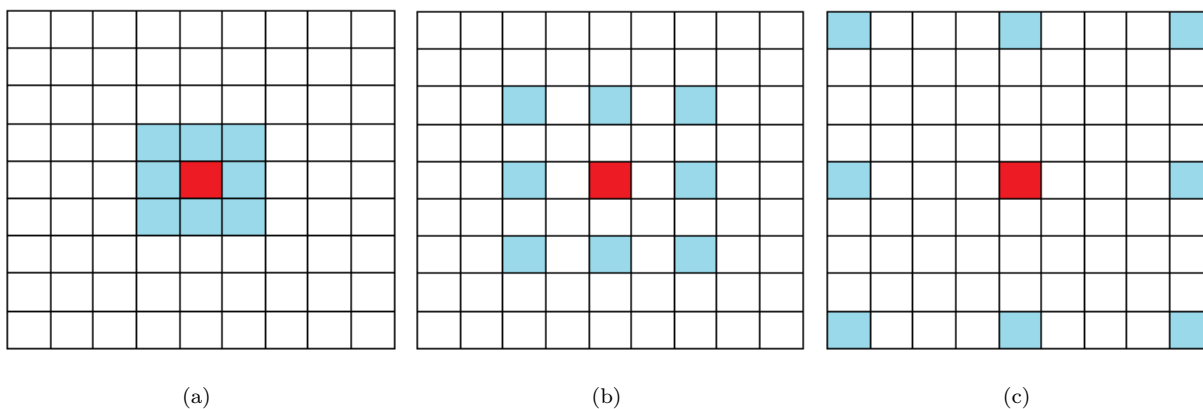
Figure 5.1: A 3×3 kernel with the dilation rate of a) 1, b) 2, and c) 4.

Slice thickness in CT scans varies a lot, while resampling images on equal voxel size may lead to the removal of small lesions [460]. Moreover, cascaded models exponentially increase computational complexity and incur dependency on preceding model performance [478].

In this study, to exponentially increase the receptive view of input images without any change of the kernel size, we introduce the dilated convolution with multiple dilation rates. Figure 5.1 shows three dilation rates 1, 2, and 4 employed throughout the proposed network to increase receptive views without affecting the number of network parameters and computation cost. Similarly, increasing the convolution layers may lead to the training error, residual networks were introduced with the batch normalization to avoid overfitting. The network was initially trained from scratch with the binary cross-entropy loss function. Afterwards, we incorporated a novel objective function with transfer learning based on combined dice coefficient and absolute volumetric difference. In summary, this study makes the following contributions:

- Design of a Residual UNet with the dilated convolution to probe inter-slice generalized features on the large scale for liver and tumor segmentation.

- Proposal of a novel loss-function based on a combination of dice coefficient and absolute volumetric difference that surpasses the performances of dice and cross-entropy-based loss functions.

- Training, validation, and extensive testing of our architecture on four publicly available datasets, namely Dircadb, LiTS, Sliver07, and Chaos.

- For fair comparison, we compare and analyse our model with various methods on similar commonly used open access challenge datasets.

5.1.1 Related Work

Roth et al. [479] develop a cascaded and fully convolutional network using UNet for 3-D segmentation of multiple abdominal organs of the human body such as the liver, spleen, stomach and others, and report a 15%-improved score for their proposed methodology. Drawbacks are observed in terms of computation time and the required hardware due to the cascaded network. In Sun et al. [164], a multi-channel convolutional network is used to segment liver tumors for CT images. The network is trained for each phase while their features were combined. A single-channel consisting of 8 convolutions, 3 subsamplings, 3 deconvolutions and 2 feature-fusion layers and also a multi-channel network is applied to three phase-contrast-enhanced CT images, which improves the algorithm for lesion segmentation of livers. However, the performance evaluation metrics for segmentation are not discussed in their paper. Chlebus et al. [460] implement a 2-D neural network for tumor detection with shape-based post-processing. The model consists of 4 resolution levels for learning local and global features with an accuracy of 77%. A 3-D neural network is not used, due to the fact that CT-scan slice thickness varies a lot, while resampling images to equal voxel size may lead to the removal of small lesions. Drozdzal et al. [480] utilize a fully convolutional network to normalize data while preprocessed data is employed iteratively for precise segmentation by a fully convolutional residual network. The iterative segmentation algorithm is lacking in variant aspects such as a performance metric, the required graphical processing memory and computation time. Li et al. [471] use intra- and inter-slice features with a hybrid feature-fusion layer to enable the improved segmentation. The algorithm yield the best results on the LiTS 2017 database, while, due to training on inter-slice features, the results are comparable to those on Dircadb datasets. Han [155] propose a residual DNN-based UNet segmentation. The proposed model consists of 32 layers in total, and makes use of both the long-range concatenation connections of UNet and short-range residual connections. The model is trained using the 130 LiTS training datasets and achieves an average DSC score of 0.67 for 70 test CT-scan images. Similarly, Jegou et al. [481] apply dense blocks [349] to improve the segmentation without any preprocessing of images. Likewise, two other studies [471, 167] propose DNNs in the cascaded form for segmenting the liver and lesions. Increasing the number of DNN layers with large kernels not only increases computation time but also can result in over-fitting and gradient vanishing problems. To resolve these problems, Kamnitsas et al. [474] and Hai et al. [475] use dilated

convolutional layers while keeping the kernel size constant. Hence, the UNet DNN can be modified with residual or dense blocks and large-scale features that can provide efficient results in terms of accuracy and computation time. Our initial focus is on improving the performance by incorporating dilated convolution layers into the UNet DNN, thereby enabling to grasp significant features while training. Likewise, residual blocks are included in order to minimize training errors with batch normalization layers, thus avoiding over-fitting and gradient vanishing problems. Moreover, several studies [471, 426, 478, 479, 167] have utilized cascaded 2-D and 3-D deep-learning models in various forms to perform coarse-to-fine segmentation. Such overlapping models can greatly influence one anothers performance, but with exponential increases of the computation cost. Hence we propose a single unified residual dilated model to address certain concerns with previously proposed methodologies. Furthermore, our network is trained, validated and tested on multiple datasets to demonstrate improved performance for segmentation. Thereby, the high performance with the minimal loss of segmentation and low computation cost could be achieved. Future potential research may involve analysis of adversarial networks for synthetic data generation with classification [245, 165].

5.2   Methodology

5.2.1   Our Network

To enhance the performance of UNet, we modify its basic deep-learning-based neural network in keras with tensorflow as the backend [441]. Instead of the single-channel input, we utilize a three-channel input with three convolution layers having three distinct dilation rates $(1, 2, 4)$ and the same feature map for each block. The 3×3 kernel size is utilized with the rectified linear unit (ReLU) activation function. Mathematically, this activation function can be written

$$y = \max (0, x) \tag{5.1}$$

where x is the input while y is output. Adding more layers can increase training parameters, though going deeper beyond some point affects the performance by increasing the rate of training error. To minimize such degradation problems, we introduce residual connections to perform identity mapping (see He et al. [10]). Let us suppose that we have x feature map output after utilizing a stacked convolution layer with a dilation rate of 1, as shown in Figure 5.2.

Mathematically, this building block can be written

$$y = F(x, W_i) + x \tag{5.2}$$

where $F(x, W_i)$ is the residual mapping that is to be learned, and x & y are the input and output of the considered layer. In our case, F is equivalent to

$$F(x, W_i) = rW_2 r(W_1 x) \tag{5.3}$$
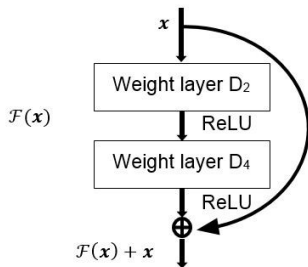
Figure 5.2: Residual block [10].

where r is the rectified linear activation function and W_i represents the dilated convolution layer with weights.

Moreover, the UNet DNN uses valid convolution that reduces the output size at each layer [55]. Consequently, the up-sampling portion is less symmetric than the down-sampling one. We utilize the convolution with the same output feature map as input for integration with the residual connections within the DNN.

Afterwards, a 2×2 max pooling is applied to down the input size. At each down-sampling step, we double the feature channels. This cycle is repeated until we reach the bottleneck section of 32^2 output-layer size. To avoid overfitting and for more generalized training of the DNN, the dropout is applied as in Ronneberger et al. [441]. Dropout acts as a means of regularization; however, it can affect performance by randomly removing parameters during training. Lyu and Simoncelli [482] introduce the concept of the batch normalization that dynamically normalizes the output of a previous activation layer by subtracting the batch mean and dividing it by the batch standard deviation. Thereby, it helps in increasing the learning rate by improving the convergence. Moreover, it provides some regularization effect that works as a dropout for the deep network. Hence, before and after the bottleneck section, we apply the batch normalization in accordance with Han [155] to improve the stability of the neural network by means of faster learning and without any loss in generalization.

In the decoding section, the feature maps are concatenated with the corresponding encoding sections to obtain a more precise location of the output. The long-range connections help produce more fine and precise high-resolution output as the result with corresponding encoding features as input for the decoding section. Each decoding section consists of a 2×2 up-sampling layer concatenated with corresponding skip connections from the encoding section stacked with three convolutional layers at different dilation rates and the same output size with the ReLU activation function. This process is repeated until the output size becomes equal to the input. The schematic of our model is shown in Figure 5.3. D_x shows the dilation rate used for each layer, and the bottom left of the encoding section and the bottom right of the decoding section depict the feature map output size. Likewise, the top index in the encoding section and the bottom index in the decoding section show the output filters of each of the layers.
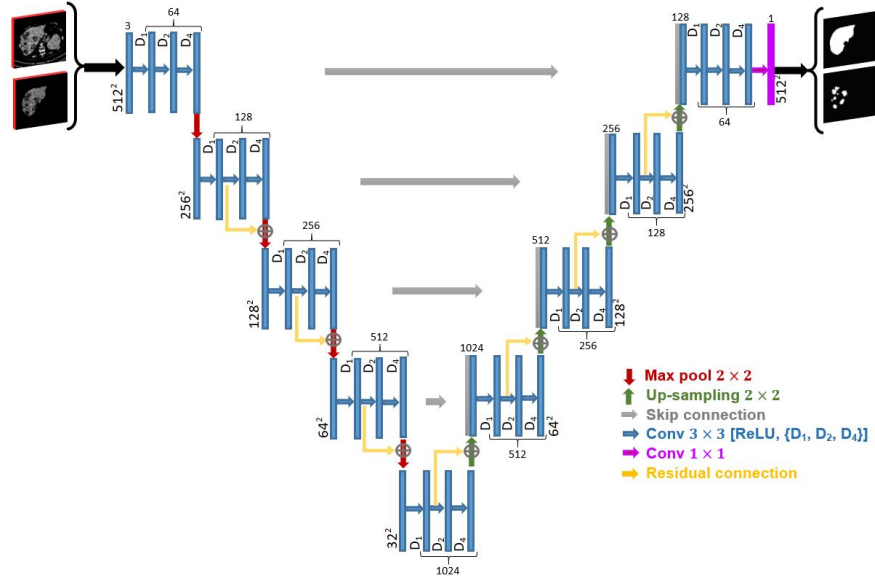
Figure 5.3: Schematic of considered deep neural network.

5.2.2 Training

To maximize the GPU performance, we preferred steps per epoch equal to the number of training images instead of a large batch, thus reducing the batch size to the size of a single image. For supervised learning of corresponding binary masks, the sigmoid energy function $(1/(1 + e^{-z}))$ was used on the final feature map with the binary cross-entropy loss function given as,

$$L_c(y, p) = \frac{-1}{n} \sum_{i=1}^{n} (y_i * \log(p_i) + (1 - y_i) * \log(1 - p_i)) \tag{5.4}$$

where log is a natural logarithm, n is the number of labelled voxels for original $(y_i \epsilon Y)$ binary indicator (0 or 1) if class label **C** is the correct classification for observation **O**, $(p_i \epsilon P)$ is the predicted probability observation **O** of class **C**.

In the medical-images processing, we often come across such situations where the anatomy of intent correspondingly engages a very small region of the scan. As a consequence, the learning process can become trapped in a local minima that leads to the biased background prediction, thus resulting in the missing or incomplete foreground prediction [483]. Figure 5.4 shows some examples of outcomes obtained with the cross-entropy loss function training.

Hence, we adopt a coarse-to-fine segmentation scheme such that, following the training with the cross-entropy loss function, we apply transfer learning and compared the performances of the two loss functions: 1) based on the dice coefficient, and 2) the combined absolute volumetric difference with the dice coefficient-based function.
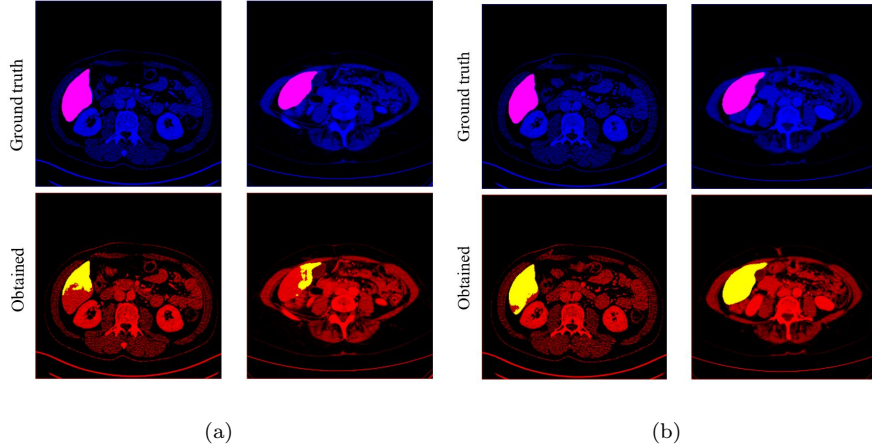
Figure 5.4: Liver segmentation (best viewed in color): a) miss-classification with the cross-entropy loss function and b) proposed loss function.

The dice coefficient function can be defined as

$$L_D(y,p) = \frac{2\sum_{i=1}^n y_i * p_i}{\sum_{i=1}^n y_i^2 + \sum_{i=1}^n p_i^2} \tag{5.5}$$

Likewise, absolute volumetric difference function can be defined as

$$L_V(y,p) = \frac{|\sum_{i=1}^n y_i - \sum_{i=1}^n p_i|}{\sum_{i=1}^n y_i} \tag{5.6}$$

The absolute volumetric difference ranges between 0 and 1, where 1 means no common region between reference and predicted output. Contrary, dice function ranges between 0 and 1, where 1 is for the perfect segmented output.

The combined loss function which we aim to minimize thus becomes

$$L_{Total} = [1 - L_D(y,p)] + L_V(y,p) \tag{5.7}$$

Figure 5.5 depicts the final 10 simulations outcome for accuracies of the dice- and proposed loss function. It is evident from the output that the training accuracies are comparable. However, the validation accuracies for the proposed objective function provides more stable and improved performance. Moreover, after performing transfer learning with the proposed loss function, we evaluated the liver segmentation as shown in Figure 5.4.

The initialization of weights in each convolution layer is important in order to avoid the excessive activation. Therefore, for the initial weights with the unit variance, we use a Gaussian distribution with a standard deviation of ($\sigma = \sqrt{2/n}$), where n denotes the number of inputs of one neuron [484]. For example, for a 3×3 convolution and 64 feature maps, we have $n = 9 * 64 = 576$. ADAM is used to optimize our loss function at a learning rate of 0.0001 [485].
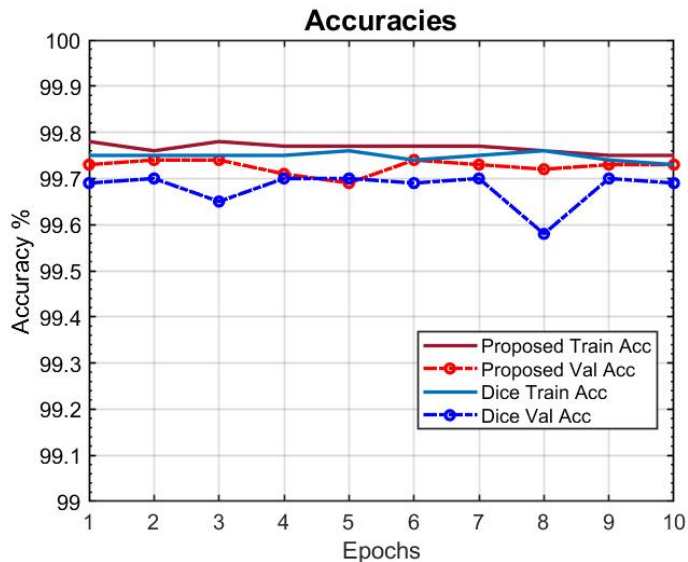
Figure 5.5: Comparison of accuracies for the dice- and proposed loss function.

5.2.3 Data Augmentation

As the anatomy of medical images varies highly from subject to subject in the shape and intensity of pixels, we trained our model with random elastic deformations to produce more generalized data for training. Hence, we perform data augmentation for each epoch in a dataset with rescaling equal to (1/255), rotation 40%, widening, height shift and shear in random range of $[-0.1, +0.1]$, zooming in the random range of $[-0.2, +0.2]$, and horizontal flip to images for each epoch. Introducing the batch normalization, layers before and after the bottleneck section provides further tacit data augmentation and showed great improvement in training and testing.

5.2.4 Dataset and Preprocessing

We use Dircadb [7], LiTS [9], Sliver07 [8] and Chaos [160] competition datasets to evaluate our proposed DNN for the liver segmentation. However, whereas only the Dircadb and LiTS datasets contain lesion CT scans with corresponding masks, we evaluate lesion segmentation only on those two publicly available datasets. The experimental dataset for liver contains a total of 180 subjects with approximately 55000 CT-scan images and 143 cases of lesion segmentation with approximately 22000 CT scans. Corresponding CT-scan masks are outlined by radiological experts. The input image size is $(512 \times 512 \times 3)$ with the variable voxel size from subject to subject. For the best visibility and uniform pixel intensity, we first apply histogram equalization and then truncate the image pixels intensity between $[180 - 255]$ to best differentiate the liver and tumor boundaries [471]. No further preprocessing (e.g. registration or noise removal) is performed. Figure 5.6 depicts the pipeline flow of our proposed methodology. The CT scans with corresponding liver masks are used to train the proposed neural network. Afterwards, the segmented liver-anatomical regions are considered

as input for lesion segmentation.
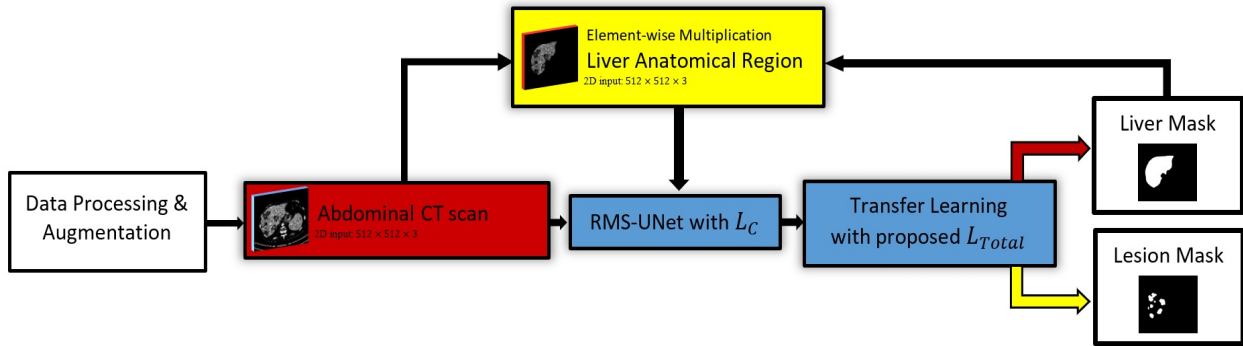
Figure 5.6: Proposed pipeline for liver and lesion segmentation.

5.2.5   Evaluation Metrics

Recent segmentation challenges (Dircadb , LiTS, Sliver07, Chaos) and previous studies [471, 472, 426, 49] have considered multiple metrics for the evaluation of segmentation performance. Hence, we extensively evaluate our liver and tumor segmentation outcomes based on all of the commonly and widely used metrics. Performance metrics are computed on $MATLAB^{©}$ in accordance with the Chaos grand challenge evaluation [180]. The details of these metrics are as follows.

Dice Similarity Coefficient (DSC)

It evaluates the overall performance of an algorithm by including the region-of-interest (ROI) pixels of segmentation [60, 486]. The DSC ranges from 0 to 1, where 0 means no overlap or common region between the segmented and reference images, and 1 means perfect segmentation. It can be defined as,

$$DSC = \left[\frac{2 \left| I_p \cap I_y \right|}{\left| I_p \right| + \left| I_y \right|} \right] \times 100 \tag{5.8}$$

where $|I_p|$ & $|I_y|$ are the cardinalities of pixels in the predicted and ground-truth segmented images, respectively.

Volumetric Overlap Error (VOE)

It measures the ratio of the common region between the segmented and reference images, as divided by their union [60]. It can be defined in terms of the complement of the Jaccard score, as

$$VOE = \left[1 - \frac{\left| I_p \cap I_y \right|}{\left| I_p \cup I_y \right|} \right] \times 100 \tag{5.9}$$

VOE is 0 for the perfect segmentation.

Relative Volume Difference (RVD)

It is the ratio of the total volume difference to the total ground-truth volume of an image. Its asymmetric metric is given as [60],

$$RVD = \left[\frac{|I_p| - |I_y|}{|I_p|} \right] \times 100 \tag{5.10}$$

RVD is 0 for perfect segmentation, whereas positive and negative for over- and under-segmentation, respectively.

Average Symmetric Surface Distance (ASSD)

The symmetric surface distance metric is used to measure the distance between the foregrounds of predicted and ground-truth images. The closest Euclidean distance between the predicted I_p image and the ground truth I_y image is measured alternatively [49]. Afterwards, the mean of all the obtained distances is computed, hence the name average symmetric surface distance (ASSD), with 0 mm for perfect segmentation. The shortest distance of an arbitrary voxel x to the set of surface voxels of I_y is defined as

$$d(x, I_y) = \min_{y \epsilon I_y} \|x - y\| \tag{5.11}$$

where $\|\cdot\|$ denotes the Euclidean distance between voxels x and y in real-world coordinates. Thus, the average of all sorted distances is computed as

$$ASSD = \frac{1}{|I_y| + |I_p|} \times \left(\sum_{x \epsilon I_p} d(x, I_y) + \sum_{y \epsilon I_y} d(y, I_p) \right) \tag{5.12}$$

Likewise, the Root Mean Square Symmetric Surface Distance ($RMSSD$) is calculated by computing the square root of the symmetric surface distance, with 0 mm for perfect segmentation.

Maximum Symmetric Surface Distance (MSSD)

The maximum symmetric surface distance is also known as Hausdorff distance, and is computed by alternatively calculating the maximum distance between the surface voxels of the predicted I_p and reference I_y images, with 0 mm for perfect segmentation [49]. Mathematically it can be written as,

$$MSSD = \max \left\{ \max_{x \epsilon I_p} d(x, I_y), \max_{y \epsilon I_y} d(y, I_p) \right\} \tag{5.13}$$

This metric is sensitive to outliers; thus, it is significant for applications such as surgical planning, where the true maximum error is more influential than the average error.

5.3   Experiments

We perform experiments on Python 3.7 with Keras and Tensorflow for DNN analysis. The machine is with the specification of an Intel(R) Xenon (R), 2.6GHz CPU with 30GB RAM and NVIDIA K40 GPU. The

training, validation and testing scheme is adopted to evaluate our DNN. 70% of the total data is utilized for training, while 15% of the data is utilized for validation and the 15% for testing. The stopping criterion is defined based on the training loss, such that iterations are to be stopped if a minimum change of 0.0001 is not experienced in loss for 5 consecutive epochs while restoring the best weights. Figure 5.7 depicts the proposed DNNs training and validation outcomes of the iterative process of liver and tumor segmentation. It is evident from the output responses that the proposed DNNs multi-scale context layers within each residual block explore more generalized features, which considerably improves performance and reduces losses to the second decimal place. Consequently, validation accuracies come close to training accuracies within a few iterations. Moreover, the obtained performance curves follow traditional trends that validate the optimal performance by the proposed algorithm. Likewise, it takes less than 2 *sec* to predict the liver or tumor mask corresponding to the CT-scan image.
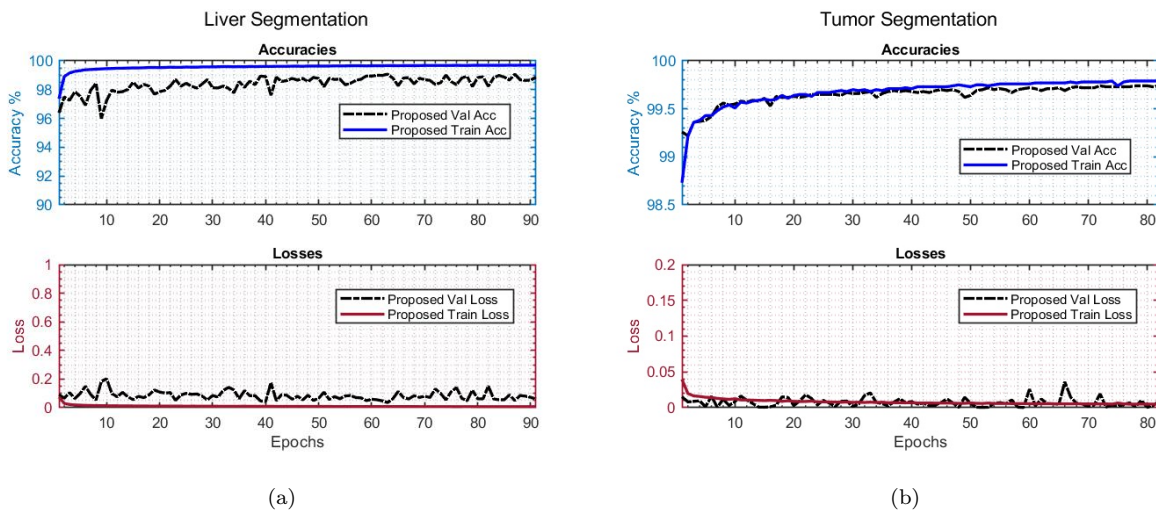
Figure 5.7: Proposed DNN performance for: a) liver and b) tumor segmentation.

We perform extensive comparative analysis with previously developed methodologies on publicly available liver and tumor segmentation datasets. Comprehensive results on liver segmentation are obtained with the Dircadb, LiTS, Sliver07 and Chaos test sets, whereas the lesion segmentation performance is evaluated with Dircadb and LiTS test sets only. Tables 5.1 and 5.2 illustrate the mean performance metrics with standard deviations. The best outcomes for each dataset are boldfaced for comparisons. Furthermore, individual dataset details and substantial analyses are as follows.

Dircadb Dataset

The Research Institute against Digestive Cancer (IRCAD) dataset, provided in 2016, is composed of 20 subjects CT scans of liver and hepatic tumors with an image resolution of (512×512) [7]. It is evident from Table 5.1 that several studies have utilized Dircadb database for the liver segmentation evaluation;

however, none of the previous studies solely performs the best. Zheng et al. [463] outperform on DSC 96.6%, MSSD 18.9% and RMSSD 2.1%, while Christ et al. [167] surpass on ASSD metric. This is due to the fact that they apply multiple semi-automatic algorithms, such as the level set and Chen-Vese model in cascaded form for coarse-to-ne segmentation. Semi-automatic methodologies lack in generalize performance due to heterogeneous background and ambiguous appearance. Similarly, Ahmad et al. [169] perform the best for VOE as 6.09% and Li et al. [471] for RVD as −0.01%. The former utilizes DNN with the active contour model for post-processing output images, whereas the latter employs cascaded 2-D and 3-D models of UNet [28] for refine outputs. We implement the 2-D model given the fact that CT slice thickness varies with variable voxel sizes. Sampling all slices with a similar voxel size may disregard minor lesions within the ROI [9]. In addition to this certainty, obtained outcomes for the liver segmentation surpass all previously claimed metrics except for RVD, where this outcome is comparable. Similarly, for the tumor segmentation results in Table 5.2, Sun et al. [164] performs the best for DSC and VOE, while Wu et al. [166] for ASSD and RMSSD. The former utilizes DNN with multi-channel input, whereas the latter applies semi-automatic approach to perform lesion segmentation. Our study shows margin in DSC, VOE and RVD with lagging in two distance measure metrics. This might have been due to the fact that Wu et al. [166] apply multiple semi-automatic approaches with the constrained optimization, whereas our results are consistent, as shown in Figure 5.8. The obtained liver masks are closely related to their references with minor discrepancies in the tumor masks.
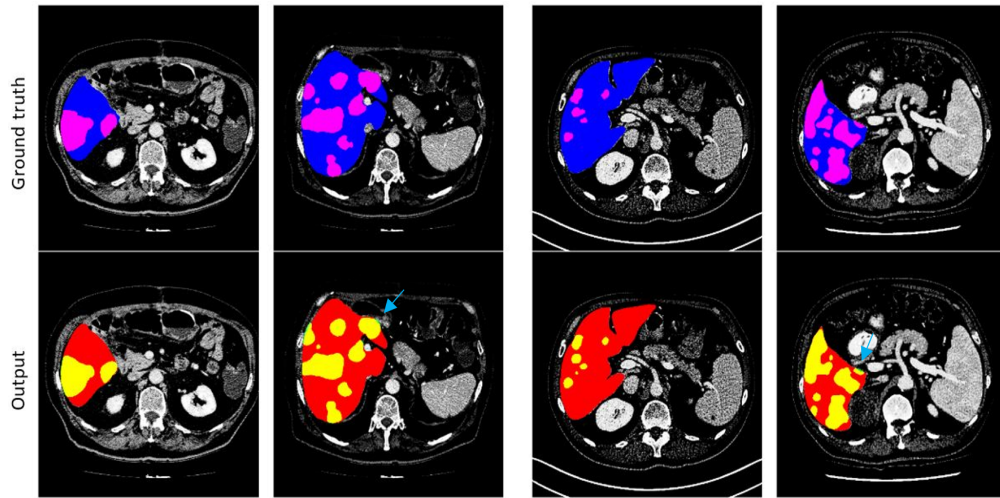
Figure 5.8: Liver and lesion segmentation outcomes for Dircadb test set. The arrows indicate the green under-segmented regions (best viewed in color).

LiTS Dataset

The Liver Tumor Segmentation Challenge 2017 was conducted in Australia by CodeLab. This dataset contains 130 subjects collected from six medical centers with varied numbers of CT scans [7]. In the case of the liver segmentation in Table 5.1, Qin et al. [163] and Chlebus et al. [460] perform collectively the best

among all of the discussed studies with the LiTS database. It is worth noting that the claimed results by the former study are based on only 100 selected CT scans, while the latter utilizes the shape based post-processing to improve 2-D DNN output. Comparatively, our metrics are based on a complete LiTS dataset analysis; thus, they could consider vague and generalized situations for segmentation. Apart from this, our algorithm outperforms the competing methods in terms of multiple commonly used metrics such as DSC, VOE, and others. Similarly, for the lesions in Table 5.2, our obtained outcomes are remarkably better than those of competing methods, considering the fact that we model the loss function based on two metrics. Likewise, taking into account distance-based metric can further improve performance for remaining metrics. Some outcomes from LiTS test set are depicted in Figure 5.9.

Figure 5.9: Liver and lesion segmentation outcomes for LiTS test set. The arrow indicates the green under-segmented region (best viewed in color).

Sliver07 Dataset

A Sliver07 competition was initiated with the title, Segmentation in the clinic: A grand challenge, as part of a workshop in 2007. Later, this challenge was conducted yearly under the auspices of the Medical Image Computing and Computer Assisted Intervention Society (MICCAI) [8]. It consists of 20 subjects with images of (638×638) resolution. Maklad et al. [492] show the best performance for most of the metrics in Table 5.1, based on only 10 subjects CT-scan images. Moreover, they apply a semi-automatic approach that lack in generalize performance. Our network utilizes multi-scale context layers within each residual block without largely increasing the computation cost. This helps us explore more generalized features with exceptional performance. Despite this fact, it outperforms among all other DNNs. Figure 5.10 shows some examples of Sliver07 dataset liver segmentation.

Table 5.1: Metrics outcome on liver segmentation datasets

Model	Year	DSC (%)	VOE (%)	RVD (%)	ASSD (mm)	MSSD (mm)	RMSSD (mm)
			Dircadb Dataset				
Erdt and Kirschner [153]*	2010	94.55	10.34±3.11	1.55±6.49	1.74±0.59	26.83±8.87	3.51±1.16
Chung and Delingette[487]*	2013	93.05	12.99±5.04	-5.66±5.59	2.24±1.08	25.74±8.85	-
Chartrand et al.[149]*	2014	96.48	6.8±1.2	1.7±1.6	1.6±0.3	24±5.9	2.9±0.7
Li et al.[488]*	2015	95.2	9.15±1.44	-0.07±3.64	1.55±0.39	28.22±8.31	3.15±0.98
Huang et al.[466]*	2016	95.92	7.84±2.95	3.42±2.11	1.97±1.02	37.05±9.82	4.71±1.96
Christ et al.[167]*	2016	94.3	10.7	-1.4	**1.5**	24	-
Zheng et al.[463]*	2018	96.64	6.5±2.8	2.1±0.8	1.9	18.9±8.3	2.1±0.8
Yuan et al.[489]*	2018	96.21±0.56	7.3±1.05	-0.01±2.50	-	24.23±6.97	2.37±0.73
Zhou et al.[165]*	2019	96.53	6.7±2.3	0.6±3.2	1.5±0.6	24.6±7.9	2.4±0.8
Han[155]	2017	93.8±0.02	11.65±4.06	-0.03±0.06	3.91±3.95	-	8.11±9.68
Lu et al.[168]	2017	95.09	9.36±3.34	0.97±3.26	1.89±1.08	33.14±16.36	4.15±3.16
Li et al.[471]	2018	94.7±0.01	10.02±3.44	**-0.01±0.05**	4.06±3.85	-	9.63±10.41
Ahmad et al.[169]	2019	91.83	6.09	5.59	-	-	-
Chen et al.[170]	2019	94.6±0.03	9.52±4.65	-0.02±0.07	8.43±9.37	-	14.21±5.71
Budak et al. [171]	2020	95.22	9.05	7.03	1.43	-	19.37
Proposed	**-**	**97.31±0.023**	**5.13±3.94**	0.04±0.05	2.00±1.92	**18.47±21.95**	**1.32±0.49**
			LiTS Dataset				
Qin et al.[163]	2018	97.31±0.36	5.24±0.69	1.97±1.7	1.77±0.49	**13.03±5.71**	-
Wang et al.[426]	2018	96.7	6	1	1.32	29.9	3.078
Li et al.[471]	2018	96.5	7.4	-1.8	1.45	27.118	3.15
Chlebus et al.[460]	2018	96.5	7.7	**-0.4**	**1.15**	24.49	2.42
Chen et al.[172]	2019	96.7	6.3	2	1.32	29.9	-
Proposed	**-**	**97.38±0.024**	**4.98±4.28**	-0.71±0.53	1.84±1.74	17.47±19.73	**1.27±0.46**
			Sliver07 Dataset				
Wimmer et al.[490]*	2008	96.65	6.47	1.04	1.57	18.32	2
Linguraru et al.[491]*	2012	96.71	6.37	2.26	1	20.75	1.92
Maklad et al.[492]*	2013	**97.78**	4.33±0.73	0.28±0.87	**0.63±0.16**	**14.01±2.88**	**1.19±0.28**
Gauriau et al.[493]*	2013	96.24	7.24	2.58	1.32	23.13	2.58
Salman Al-Shaikhli et al.[494]*	2015	90	7.2	1.25	1.16	22.88	2.14
Li et al.[488]*	2015	96.77	6.24±1.52	1.18±2.76	1.03±0.31	21.1±0.95	18.82±8.82
Wang et al.[152]*	2016	96.06	7.57	-1.83	1.57±0.34	21.11±4.6	2.33±0.46
Zheng et al.[463]*	2018	96.04	7.6	-0.1	0.8±0.5	20.8	1.5±1.1
Zhou et al.[165]*	2019	97.27	5.3±2.1	1.7±1.5	0.8±0.5	19.4±5.3	1.4±0.7
Hu et al.[173]	2016	97.25±0.65	5.35±1.23	-0.17±1.34	0.84±0.25	19.58±0.56	1.78±0.56
Dou et al.[175]	2016	97.24	5.37	1.32	-	29.63	1.48
Lu et al.[168]	2017	96.96	5.9	2.7	0.91	18.94	1.88
Dou et al.[174]	2017	97.24	5.37±0.73	1.32±1.35	0.67±0.12	29.63±16.31	1.48±0.21
Ahmad et al.[169]	2019	94.8	4.31	1.28	-	-	-
Proposed	**-**	97.39±0.0085	**3.89±1.61**	**0.09±1.7**	1.83±1.04	17.60±7.64	1.51±0.72
			Chaos Dataset				
Shvets et al.[176]	2018	90.01	18.16	-	-	-	-
Pham et al.[177]	2019	73.45±5.93	41.95	-	-	-	-
Proposed	**-**	**95.49±0.0093**	**10.87±1.77**	**1.2±3.0**	**3.88±4.58**	**26.10±28.83**	**1.78±0.86**

Note: * denotes the semi-automatic methodology; - denotes that the results are not reported for the corresponding metric.

Chaos Dataset

Combined Healthy Abdominal Organ Segmentation (Chaos) was conducted by randomly collecting data from hospitals. The dataset is comprised of 10 subjects with only corresponding masks of liver for CT and MRI scans [160]. For consistency, we utilize only CT scans in our experimentation. Table 5.1 illustrates some recent studies with a few reported metrics. Shvets et al. [176] show best results for DSC and VOE. Comparatively, our methodology outperforms the competing methods on this dataset as well due to the fact that we consider dual metrics for the objective loss function. Some segmented outcomes are depicted in Figure 5.11.
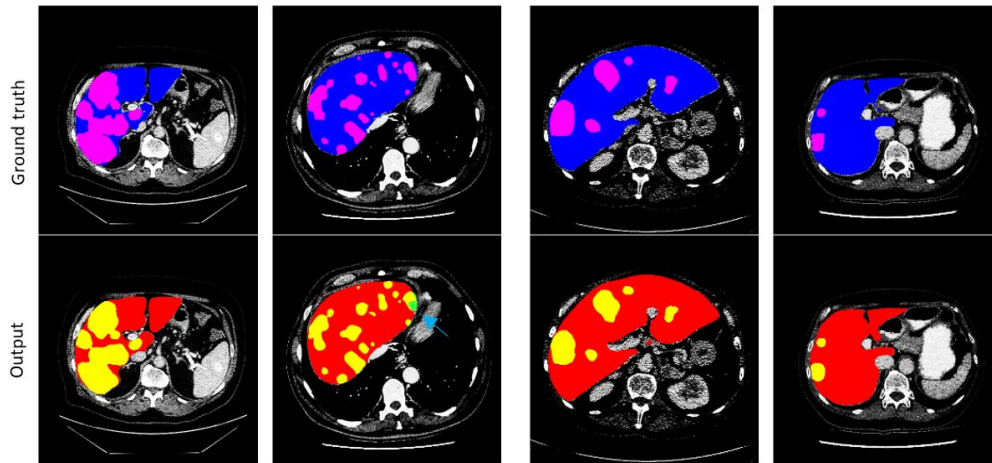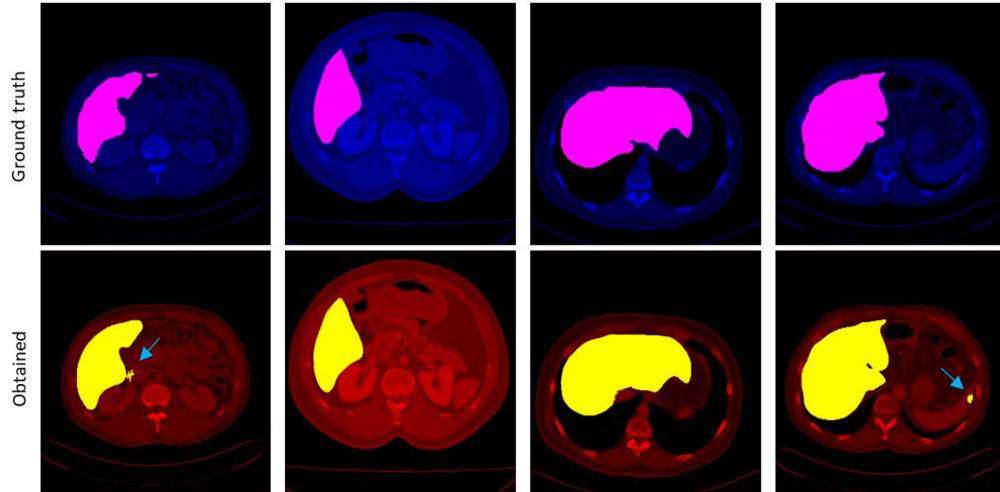
Figure 5.10: Liver segmentation outcomes for Sliver07 test set. The blue arrows indicate the over-segmented regions (best viewed in color).
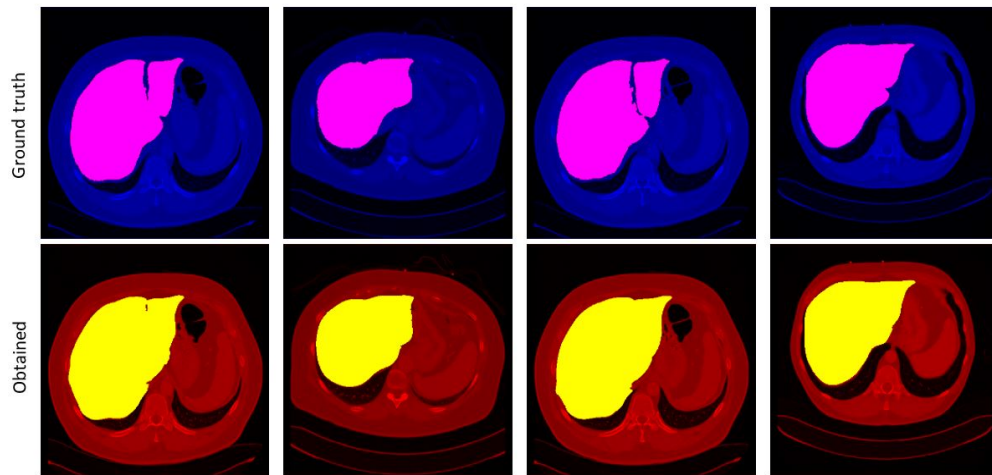
Figure 5.11: Liver segmentation outcomes for Chaos test set (best viewed in color).

5.4 Discussion

Clinical diagnosis requires identification and anatomical segmentation of lesions. Fully automatic algorithms can assist such tasks, provided that the training is generalized enough to tackle vague unknown situations. We propose a 35-layer-deep model with end-to-end training to probe more unique features with less computation cost. It can be utilized for liver and lesion segmentation with minor input changes such that only the liver-anatomical region can be considered for lesion segmentation. Except for intensity adjustment, no further pre-processing to input images or post-processing to output images is applied. We have explored the overall capability of the proposed DNN on multiple renowned and publicly available datasets, namely Dircadb, LiTS, Sliver07, and Chaos. Two of these datasets, Sliver07 and Chaos, relatively contain low-quality images in the

Table 5.2: Metrics outcome on liver lesion segmentation datasets

Model	Year	DSC (%)	VOE (%)	RVD (%)	ASSD (mm)	RMSSD (mm)
			Dircadb Dataset			
Foruzan and Chen [495]*	2016	82±0.07	30.61±10.44	15.97±12.04	4.18±9.60	5.09±10.71
Wu et al. [166]*	2017	83±0.06	29.04±8.16	-2.2±15.88	**0.72±0.33**	**1.10±0.49**
Han [155]	2017	60±0.12	56.47±13.62	-0.41±0.21	6.36±3.77	11.69±7.60
Sun et al. [164]	2017	91.54	15.6±4.33	5.8±3.5	2.0±0.9	2.9±1.5
Li et al.[471]	2018	65±0.02	49.72±5.2	-0.33±0.10	5.293±6.15	11.11±29.14
Chen et al.[170]	2019	66	41.54±4.32	**0.16±5.03**	2.04±4.32	2.12±5.52
Budak et al. [171]	2020	64.3±34.6	-	-	-	-
Proposed	-	**91.92±0.05**	**14.95±9.40**	-0.7±1.3	3.06±3.13	1.60±0.72
			LiTS Dataset			
Han [155]	2017	82	30.5	-	-	-
Vorontsov et al. [178]	2018	77.3	35.7	12.12	**1.075**	1.596
Li et al. [471]	2018	82.4	36.6	4.272	1.102	**1.595**
Chlebus et al. [460]	2018	79.6	38.3	46.4	1.143	1.728
Chen et al. [172]	2019	68.4	43.8	11	25	-
Proposed	-	**86.70±0.179**	**20.45±19.47**	**0.84±1.62**	5.33±17.87	1.68±1.58

Note: - denotes that the results are not reported for the corresponding metric;
* denotes the semi-automatic methodology.

non-processable form. Therefore, the proposed neural network performed an average for these datasets as illustrated in Table 5.1. Collectively, our method has achieved the best results in terms of several evaluation metrics with almost all of the databases.

5.4.1 Ablation Analysis

We perform an ablation analysis to verify the performance of our proposed neural network. The analysis of the proposed method and its variants include RMS-UNet with no dilation, RMS-UNet with dice loss, absolute volumetric difference loss, cross-entropy loss, and our proposed loss function. Figure 5.12 shows liver and lesion segmentation results with the mean DSC coefficient performance metric for the variants of the proposed network. It is evident from Figure 5.12 that dilated convolution explicitly improves results compared to non-dilated RMS-UNet by exploring novel features for improved performance. RMS-UNet without dilation performs under-segmentation or even miss-classification for very small tumors. Thus, results in lowest performance metric for liver and lesion segmentation. Similarly, the proposed novel loss function outperforms the competing loss functions with the mean DSC coefficient of 96.89 for liver segmentation and 89.31 for tumor segmentation.

Moreover, we conduct statistical tests to verify the significance of the proposed neural network with a novel loss function. The students t-test has been conducted at the significance level of $\alpha = 0.05$. Table 5.3 shows the significance level of the proposed neural network and loss function variants to the fourth decimal place. It is evident that the proposed neural network with a novel loss function outperforms among competing variants.

Table 5.3: Statistical significance outcomes of the proposed neural network and its variants.

Architecture	Loss Function	p-value
Liver Segmentation		
Non-dilated RMS-UNet vs. Dilated RMS-Unet	Cross-entropy	0
Dilated RMS-UNet	Cross-entropy vs. Proposed loss, Dice coefficient, Absolute volumetric difference	0.0441, 0, 0.0018
Dilated RMS-UNet	Dice coefficient vs. Proposed loss, Absolute volumetric difference	0, 0
Dilated RMS-UNet	Absolute volumetric difference vs. Proposed Loss	0.0016
Lesion Segmentation		
Non-dilated RMS-UNet vs. Dilated RMS-Unet	Cross-entropy	0
Dilated RMS-UNet	Cross-entropy vs. Proposed loss, Dice coefficient, Absolute volumetric difference	0.8365, 0.0001, 0.0012
Dilated RMS-UNet	Dice coefficient vs. Proposed loss, Absolute volumetric difference	0.0002, 0.3499
Dilated RMS-UNet	Absolute volumetric difference vs. Proposed Loss	0.00003

5.4.2 Comparative Analysis

We have discussed a few of most of the relevant architectures [471, 472, 426, 474, 475, 155, 483, 496, 481] for extensive comparative analysis, mainly from the following three perspectives.

We have analyzed studies [164, 471, 426, 475, 476, 173, 496] to probe more unique features for DNN training. In this context, Sun et al. [164] and Li et al. [471] utilize multi-channel features by considering several consecutive scans as channels for input. Due to the variation in CT-scan voxel size, resampling several scans on the same scale can pass over small regions. Therefore, to obtain the benefit of multi-channels input [170], we utilize only colored, three-channel (RGB) input images. In a few earlier studies [426, 475, 476], various form of dilation are utilized to explore novel features, while Liu et al. [496] utilize multi-kernel methodology to fulfill the same task. However, using multi-kernels can exponentially increase computation cost. Therefore, we utilize only dual-scale dilation layers to examine unique features without significant increase in computation cost.

Three previous studies [471, 472, 483] utilize custom loss-functions in their proposed methodologies due to the possible local minima trap (discussed in sub-section 5.2.2). Nevertheless, we have compared the performance of the universally used dice loss-function with that of the proposed model, and found that the additional volume-measuring metric can increase accuracy, and improves segmentation performance.

Stacking several hidden layers up to some limit can incur the vanishing gradient problems related to the training error. Therefore, multiple studies [471, 472, 155, 481, 10, 349] have utilized residual and dense blocks

to keep some gradients with identity mapping. We have customized the UNet [441] DNN with several dilated layers to increase the receptive view. Accordingly, there are residual blocks to address the vanishing gradient problem.

5.5   Conclusion

In this study, we have modified the UNet deep neural network (DNN) to minimize training losses and improve the segmentation performance. We probe inter-slice features by considering three-channel RGB abdominal CT-scan input images. To increase the valuable and unique information for each layer, multi-scale context layers with various dilation rates are considered instead of utilizing multiple kernel sizes. Consequently, residual blocks are added to compensate for training losses due to the increase in the number of convolution layers. A normalization scheme instead of dropout is preferred to avoid over-fitting and enable more generalized training without loss of any information. Moreover, to avoid fore- and background imbalance, a novel objective function is adopted, based on dice and absolute volumetric difference, to reduce output losses. To explore the proposed DNN performance, four commonly used publicly available datasets are used, namely Dircadb, LiTS, Sliver07, and Chaos. Mean dice similarity coefficient (DSC) scores of 97.31, 97.38, 97.39, and 95.49% and average symmetric surface distance (ASSDs) of 2.00, 1.84, 1.83 and $3.88mm$ are achieved for Dircadb, LiTS, Sliver07 and Chaos of 15% liver test sets, respectively. Likewise, DSC scores of 91.92, 86.70% and ASSDs of 3.06, $5.33mm$ for the 15% lesion test sets are achieved. It should be noted here that most of the previous studies utilize limited test sets (as discussed in section 5.2.2) to validate their methodologies, whereas we consider extensive validation and test sets. Also, the highest mean scores for lesions, and comparative scores for liver segmentation, have been attained. The comprehensive comparative analysis illustrated significant improvement in segmentation performance and minimization of training losses.

Reference	RMS-UNet without dilation/ Cross-entropy loss	RMS-UNet/ Cross-entropy loss	RMS-UNet/ Dice loss	RMS-UNet/ Absolute Volumetric difference loss	RMS-UNet/ Proposed loss
Liver DSC/ Tumor DSC	79.08/ 80.78	81.72/ 82.07	85.02/ 85.74	93.35/ 86.06	96.89/ 89.31
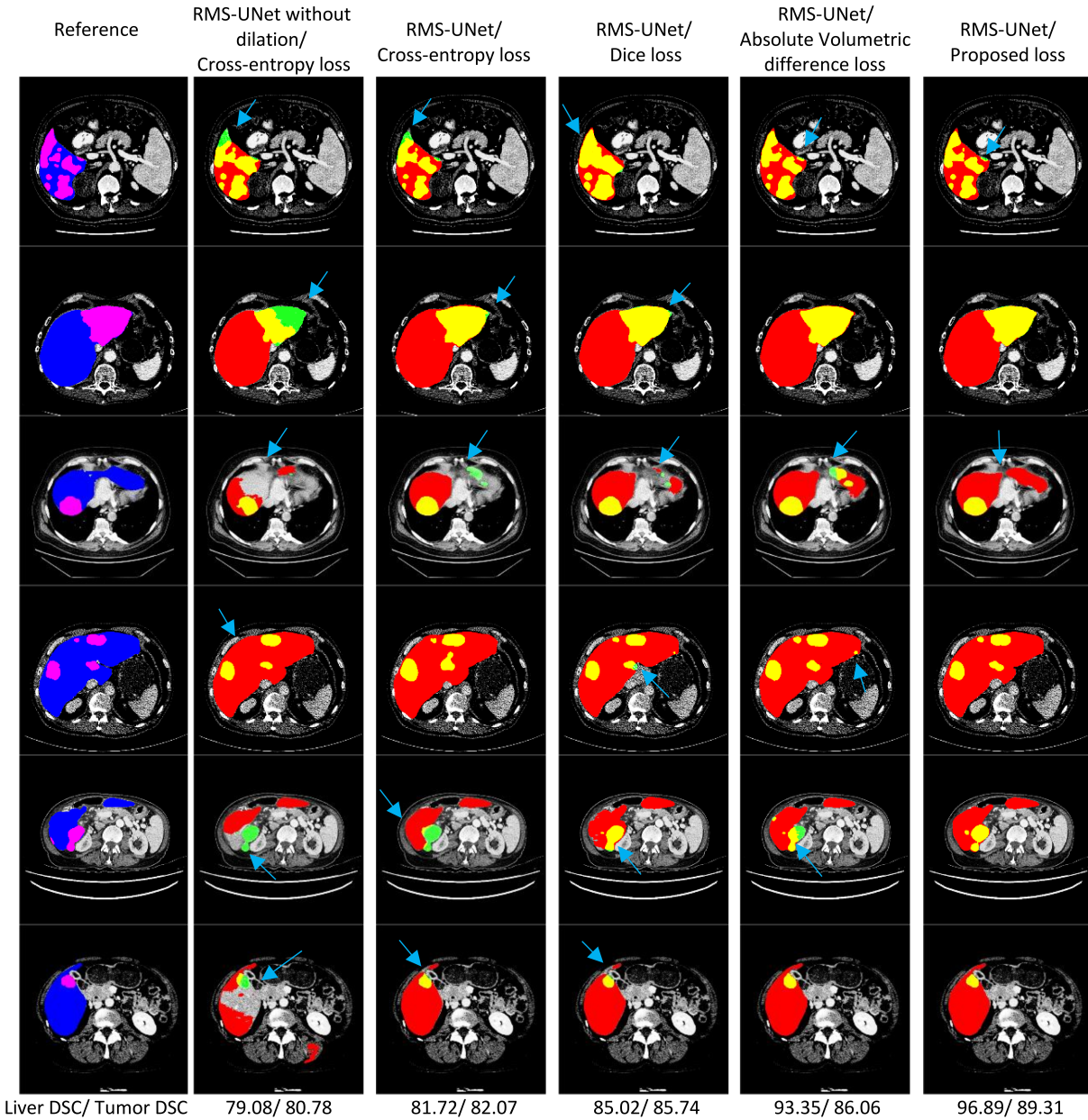
Figure 5.12: An ablation analysis of RMS-UNet and its variants in terms of mean DSC coefficient for liver and lesion segmentation. The arrows indicate the under-segmented regions for liver and lesion segmentation (best viewed in color).

6 A Multi-modal Deep Neural Network for Multi-class Liver Cancer Diagnosis

After extracting the region of interest (chapter 5), a model is trained to classify malignant liver between cancer variants. However, the liver cancer pathology data is correlated with imagery scans. In accordance with the medical perspective, I integrate the pathology data with CT scans to perform diagnosis. Therefore, this chapter proposes a multi-modal neural network to diagnose liver cancer variants. This chapter covers Objective 5 of this dissertation.

Abstract

Liver disease is a potentially asymptomatic clinical entity that may progress to patient death. This study proposes a multi-modal deep neural network for multi-class malignant liver diagnosis. In parallel with the portal venous computed tomography (CT) scans, pathology data is utilized to prognosticate primary liver cancer variants and metastasis. The processed CT scans are fed to the deep dilated convolution neural network to explore salient features. The residual connections are further added to address vanishing gradient problems. Correspondingly, five pathological features are learned using a wide and deep network that gives a benefit of memorization with generalization. The down-scaled hierarchical features from CT scan and pathology data are concatenated to pass through fully connected layers for classification between liver cancer variants. In addition, the transfer learning of pre-trained deep dilated convolution layers assists in handling insufficient and imbalanced dataset issues. The fine-tuned network can predict three-class liver cancer variants with an average accuracy of 96.06% and an Area Under Curve (AUC) of 0.832. To the best of our knowledge, this is the first study to classify liver cancer variants by integrating pathology and image data, hence following the medical perspective of malignant liver diagnosis. The comparative analysis on the benchmark dataset shows that the proposed multi-modal neural network outperformed most of the liver diagnostic studies and is comparable to others.

6.1 Introduction

Liver diseases are potentially fatal due to their poor clinical prognosis and to potential diagnostic delays in asymptomatic patients. Liver diseases can be categorized as focal or diffuse. Focal liver diseases may include benign entities such as liver cysts or malignant lesions related to primary or secondary tumors. In contrast, diffuse liver diseases include viral hepatitis, fibrosis, cirrhosis, hepatic steatosis, etc.

Liver malignancy is divided into primary- and secondary cancer and has a very high mortality to the incidence rate of 0.95, resulting in low rates of patient 6-year survival. Primary cancer begins with the accumulation of cancer cells at single or multiple liver sites. It includes Hepatoblastoma, Cholangiocarcinoma (CC), Angiosarcoma, and Hepatocellular carcinoma (HCC), with HCC accounting for 75% of primary cases. Secondary cancer is commonly known as metastasis and occurs when cancer cells from other primary organs of the body deposit inside the liver [15]. Liver tumors may be solitary or multiple. Multiple liver lesions may result from multifocal HCC (MO) or hepatic metastases (IM), originate from primary colorectum cancer. A common cause for multiple liver metastases is colorectal cancer (mCRC) [497].

Due to poor differentiation between cancer variants, as shown in Figure 6.1, multiple possible tests are performed for diagnosis, including biopsy, laparoscopy, blood tests, imaging scans, etc. These tests are laborious, cumbersome, and time-consuming, while the results are often not completely reliable. Additionally, specimens obtained for histopathology require biopsy techniques that are associated with potential patient morbidity or mortality. However, non-invasive methods such as medical imaging may be used to develop reliable diagnosis algorithms to identify different clinical outcomes. Thus, computer-aided algorithms can assist experts, such as radiologists, assess and categorize liver lesions [20].

Medical imaging of liver disease may utilize three different imaging modalities, ultrasound (US), magnetic resonance imaging (MRI), and computed tomography (CT). Nevertheless, CT is more common for abdomen image acquisition due to its operational independency, user-friendly interface, and better resolution than US [21, 22, 23, 24, 25, 26, 27]. However, CT images consist of a series of frames containing noise, heterogeneous contrast, and blurriness that can adversely affect diagnosis performance. Therefore, conventional or deep learning models are designed to process acquired images with contrast adjustment techniques [52, 110, 14, 11, 78, 77]. Segmentation is then performed to determine the region of interest (ROI) [55, 13]. Afterward, unique meaningful features are extracted to assist designed models in classifying between multiple outcomes. Several machine learning-based algorithms have been proposed to perform computer-aided diagnosis (CAD) [498, 20, 41, 57, 58, 59, 60, 61]. However, deep neural networks have proved to perform the best with better generalization and less feature engineering [231, 236, 245, 36]. A typical workflow of CAD is shown in Figure 6.2.

This study proposes a multi-modal deep neural network for multi-class liver cancer diagnosis. A deep neural network is designed to classify primary- and secondary liver cancer, including HCC Multiple, HCC
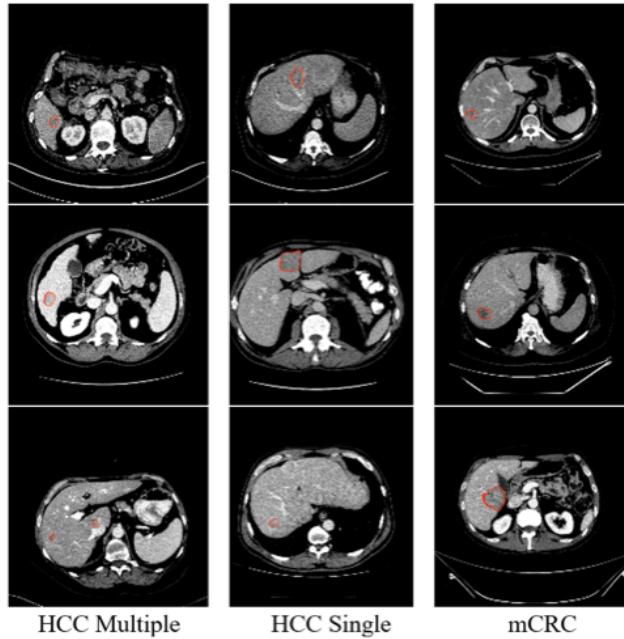
Figure 6.1: Abdominal portal venous CT scans of HCC variants and metastasis. The red traced lines by a radiologist indicate the perimeter of the lesions.
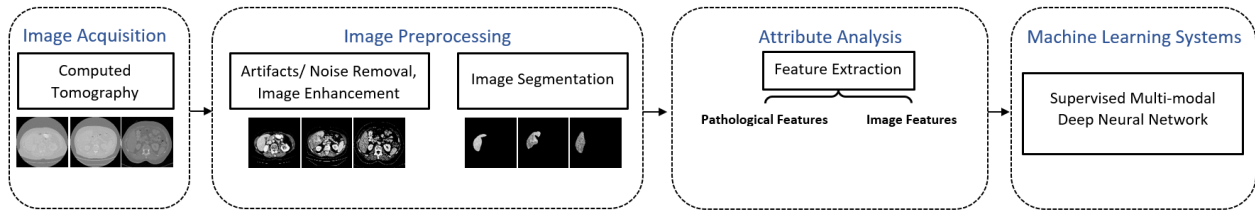
Figure 6.2: The workflow of a computer-aided system for liver cancer diagnosis.

Single, and mCRC. After contrast adjustment and selection of ROI with the help of our previously developed networks, we process CT scan images to extract image-based high-order and pathology-based low-order features. The high-order features are achieved by considering several dilated convolution layers that increase the receptive view of input images without affecting computation cost. Consequently, residual connections are added to address vanishing gradient problems. Similarly, five pathological low-order features are explored using a wide and deep network. The wide and deep network helps memorize low-order features with generalized learning. Finally, the processed features (high-order and low-order) are concatenated to perform a multi-class liver cancer diagnosis.

Typically, pathology reports are correlated with imaging studies to predict malignancy. Keeping this medical perspective in view, we propose a deep neural network that integrates imaging and pathology data for a reliable prognosis between liver cancer variants such as HCC Multiple, HCC Single, and mCRC. The results show a competitive performance of the proposed network in liver cancer diagnosis. In summary, this study contributes as follows.

1. The multi-modal deep neural network framework with transfer learning, which considers image and pathology data for diagnosis of liver cancer variants, namely, HCC Multiple, HCC Single, and mCRC.

2. The proposed deep neural network is pre-trained on binary dataset for the transfer learning.

3. A comparative analysis is performed with state-of-the-art methods on the benchmark dataset.

6.2   Related Work

We have searched CAD studies for a focal liver disorder, particularly CT scan based liver malignancy. Several studies have addressed diffuse- and focal liver disease diagnosis problems [20]. However, very few have discussed liver malignancy diagnosis, especially hepatocellular carcinoma variants (i.e., HCC Multiple and HCC Single) and metastasis. Since liver cancer variants have similar homogeneous CT density and may be small lesions, it is difficult for traditional machine learning algorithms to classify them with good performance. Nevertheless, neural networks can be wisely designed to overcome such limitations with better generalized learning.

Balagourouchetty et al. [36] modify a GoogleNet neural network to classify liver diseases into several classes: normal, HCC, hemangioma, cyst, abscess, and liver metastasis. CT images are enhanced using decorrelation stretching and segmented by the semi-automatic active contour method with transfer learning to cater data limitations. Consequently, very few images are considered compared to the multi-class problem. Li et al. [35] propose a 9-layer convolution neural network (CNN) to diagnose diffuse, nodular, and massive tumors while including only 165 venous phase CT scans. Similarly, Renukadevi and Karunakaran [34] utilize handcrafted features with principal component analysis (PCA) and grasshopper optimization to select valuable information. A deep belief network is then applied to perform classification on 583 CT scans between various focal- and diffuse liver diseases.

Likewise, Vivanti et al. [269] utilize a CNN to segment and quantify 246 tumors. The identified tumors are used to extract texture, geometric and statistical features for machine learning-based classification. Wang et al. [236] apply a residual convolutional neural network to diagnose four-class focal liver lesions with 388 CT scans. Midya et al. [273] implement an Inception-v3 network to perform binary primary liver cancer classification. The cropped CT scan images of 223 patients are employed to prognosticate CC and HCC. Yasaka et al. [242] consider a custom CNN to classify HCC, hemangioma, cyst, and lesions. Relatively, a large number of 1068 CT scans are engaged with augmentation for generalized training. Kutlu and Avcı [32] combine a CNN and long short-term memory (LSTM) for binary classification. The CNN-based hierarchical features are extracted from 122 CT scans and are fed to the LSTM for benign versus malignant identification. Likewise, Chen et al. [276] propose a dual attention dilated residual network with spatial and spectral attention modules for binary classification. Romero et al. [277] apply an Inception-v3 network with 230 CT scans. Moreover, Yao et al. [240] utilize the largest dataset with more than 76 thousand images to train a dense neural network and predict 13-class liver function test indicators.

After reviewing the literature, several loopholes were brought to our attention, including small datasets, manual preprocessing methods (such as contrast adjustment and segmentation), and attribute analysis. Biomedical studies lack the availability of large datasets, making the proposed model performance doubtful. However, very few studies address such limitations by utilizing synthetic data generation or transfer learning techniques [230, 245, 499]. Nevertheless, we consider a pre-trained network besides a large training dataset and extensive augmentation. Moreover, automatic preprocessing steps, including contrast adjustment and segmentation [13, 12], are performed to develop labeled datasets. Furthermore, hierarchical features are explored instead of handcrafted features for multi-class diagnosis.

6.3 Materials

This section briefly discusses data processing details, including data imputation, enhancement, segmentation, and augmentation.

6.3.1 Dataset

We explored the Royal University Hospital (RUH), Saskatoon, Saskatchewan, Canada repository for the required dataset. The database is searched from 2008 onwards for HCC Multiple, HCC Single, mCRC, and normal cases. The shortlisted dataset consists of 68 pathological confirmed subjects, including 7 HCC Multiple, 11 HCC Single, 12 mCRC, and 38 normal individuals between the ages of 16 and 93. The portal venous CT scans with corresponding pathology reports are collected close to the surgery date so that both modalities information remains comparable. Moreover, we developed a binary (malignant versus normal) dataset with 180 subjects using publicly available Dircadb [7], LiTS [9], Sliver07 [8], and Chaos [160] databanks to obtain a pre-trained network for the multi-class problem. Therefore, this study utilizes 248 subjects collectively to develop the diagnosis algorithm, while 30% of them is used for testing and validation. The ethics board of the University of Saskatchewan has approved this study under file number Bio-1493.

6.3.2 Data Processing

The CT scans consist of sequential images with equal intervals. However, each image can have variable contrast due to changing depth of field, exposure, focus, subject movement, etc. Some, or all, of these factors can combine to produce images with variable intensities. Therefore, we consider our previously developed multi-level GAN network to enhance portal venous CT scan images with uniform intensities. Our multi-level GAN consists of three auto-encoder blocks to process images at three variant scales, while each preceding block is linked with the succeeding block [12]. Figure 6.3 shows various portal venous CT scans enhanced with our multi-level GAN. Afterward, we utilize our RMS-Unet for segmentation [13]. As discussed in section 6.2, manual or semi-automatic segmentation is commonly performed for CAD. However, we designed a Unet-

based, fully automatic neural network to extract liver regions from lower abdominal images, as shown in Figure 6.3. Thereafter, the segmented images are cropped and resized for background removal.

Moreover, we apply data augmentation with zoom operation in the range of $[-0.2, +0.2]$, rotation 5%, height-, widening- shift, and shear in the random range of $[-0.1, +0.1]$, rescaling $\left(\frac{1}{255}\right)$, and horizontal-, vertical- flips to every image of epoch. Alternatively, pathology data is processed for five features, namely age, gender, length, width, and depth of the largest liver tumor. Whereas the Gaussian distribution $N\left(\mu, \sigma^2\right)$ is utilized for data imputation. Some augmented images are depicted in Figure 6.3.

Raw Input   Enhancement   Segmentation   Cropping   Data Augmentation
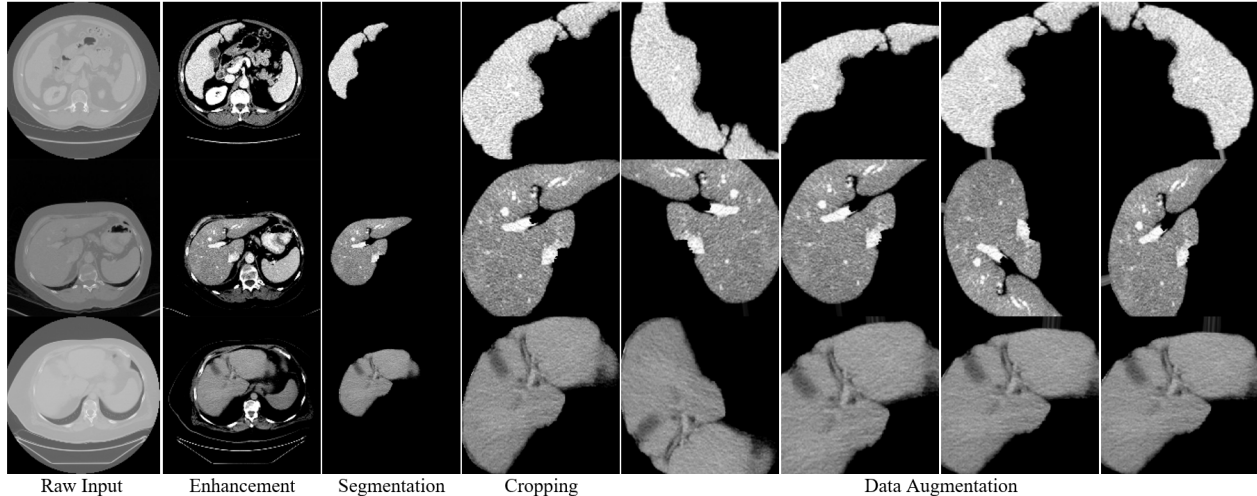
Figure 6.3: The image processing layout for portal venous CT scans [11, 12, 13].

6.4   Methodology

6.4.1   Our Network

We modify to propose our multi-modal deep neural network base on AlexNet [228]. It requires image-based high-order features and pathology-based low-order features to perform categorical classification. The schematic of the proposed multi-modal network is depicted in Figure 6.4. After image processing (discussed in section 6.3.2), the single-channel images are fed to the five blocks network with three valid convolution layers in each block having different dilation rates $\{1, 2, 4\}$ and the same feature map. In contrast, each layer utilizes a 3×3 kernel size with rectified linear (ReLU) activation function. Convolution layers with different dilation rates assist the network in identifying unique features without increasing kernel size. However, adding more layers can cause vanishing gradient problems. Therefore, the residual connection is included in four blocks to avoid such degradation issues. Mathematically, suppose we have y_{res}^{B+1} feature map output after applying input x^B to B-th block with the residual mapping written as,

$$y_{res}^{B+1}(x, W) = F(x^B, w_i^B) + r(w_i^B x^B) + b^B \tag{6.1}$$

where B is the block number, r indicates rectified linear activation function, $F\left(x^B, w_i^B\right)$ represents residual mapping, w_i^B is the i-th weighted dilated convolution layer, and b^B is the bias at B-th block. Similarly, $F\left(x^B, w_i^B\right)$ is defined as

$$F(x^B, w_i^B) = r\left(w_{i+2}^B r\left(w_{i+1}^B[r(w_i^B x^B)]\right)\right) \tag{6.2}$$

After that, block output is regularized with batch normalization and downscaled using a 2×2 max-pooling layer. Consequently, we get 8^2 layer size concatenated with low-order features as shown in Figure 6.4.
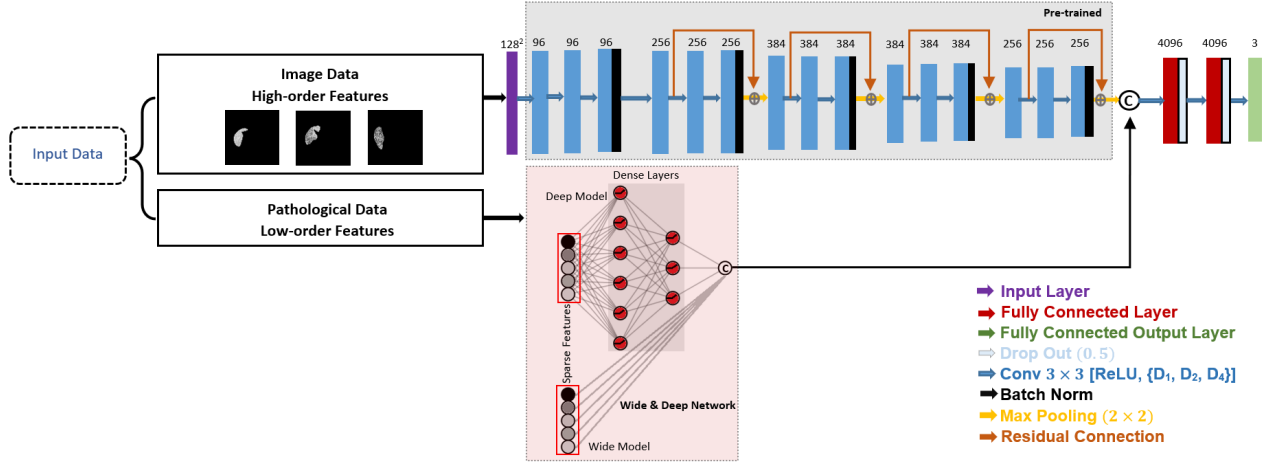
Figure 6.4: The schematic of the proposed multi-modal deep neural network.

The pathology data is processed with a wide and deep neural network [500]. The wide and deep network simultaneously works on a principal of memorization and generalization. The wide network learns the correlation between the most frequent co-occurrence of features, which is noted as memorization. It can be presented as a generalized linear model of the form,

$$y_{wide}(V, W) = r(W_{wide}^T V) + b \tag{6.3}$$

where V is a vector of $n = 5$ features, and W is the weights vector.

In contrast, a deep network tries to explore and learn novel combinations that have rarely occurred in the past. The deep network component uses two hidden layers with 64 and 32 feature maps, batch normalization, and a drop-out layer. Thus, each hidden layer can be computed as,

$$y_{deep}^{l+1} = r(w_{deep}^l x^l) + b^l \tag{6.4}$$

where l is the layer number.

Lastly, the two fully connected layers process the high- and low-order features together, given as,

$$\hat{y} = \sigma\{rw_{i+1}[rw_i(y_{res} + y_{wide} + y_{deep})]\} \tag{6.5}$$

where $\hat{y}$ is the multi-modal deep neural network output and σ represents the Softmax activation function.

6.5 Experimental Setup

6.5.1 Transfer Learning and Fine-tuning

Initially, the proposed deep neural network is trained for binary classification without considering pathological input (i.e., image input only). The binary (normal versus malignant) dataset contains 23 thousand original and augment images (discussed in section 6.3.2) for every epoch. Afterward, the weight of the pre-trained layers are transferred (i.e., for transfer learning) and fine-tuned for multi-class diagnosis with multi-modal input. Suppose we have n images in a training dataset X, then fine-tuning tries to find optimal weights w in an iterative process by minimizing the proposed neural network error rate (i.e., empirical loss) given as [248],

$$L(X, w) = \frac{1}{n} \sum_{i=1}^{n} \ell(\hat{y}_i(x_i, w), y_i) \tag{6.6}$$

where x_i is the i-th image of X, $\hat{y}_i(x_i, w)$ is the output prediction function of the multi-modal neural network, y_i is the reference or ground-truth class of the i-th image, and $\ell(\hat{y}_i, y_i)$ is a function of penalty for predicting $\hat{y}_i$ instead of y_i which is set to the categorical cross-entropy loss function [240].

We utilize a mini-batch Adam optimizer to find optimal weights w. Let $X_m \subset X$ be a subset of m images; then we may call X_m be a mini-batch of X with batch size m. A randomized set of mini-batches are generated that covers all the samples of X. Thus, each iteration can be termed as the completion of one mini-batch. However an epoch is one training pass completed after using all elements of X or all mini-batches. Conventionally, each epoch contains $\frac{n}{m}$ mini-batches. However, if m is not a factor of n then any mini-batch can have less than m samples [248].

The network is run for a number of epochs with mini-batches until the error minimizes and does not improves for 10 epochs. In contrast, weights w are calculated after each iteration. The revised weights w_{t+1} are computed by the gradient of the loss function L over mini-batch X_m with current weights w_t defined as

$$w_{t+1} = w_t + l \left[\wp \Delta w_t - \frac{\partial L(w_t, X_m)}{\partial w_t} - dw_t \right] \tag{6.7}$$

where l is the learning rate for controlling the change of the weights, Δw_t is the change in weights from the previous iteration, $\wp$ is the momentum coefficient that suppresses the weight change fluctuations between successive iterations by adding some part of previous or current weights to the revised weights; thus, speeding up the learning with uniform weights update, and d is the weight decay to find minimal optimal weights.

The weights are updated only if they perform better than the previous epochs. Figure 6.5 depicts the learning curves of proposed neural network training from scratch and with transfer learning. It is evident from the trends that training and validation losses gradually decrease while following the traditional fashion. The responses also confirm that the fine-tuned network considerably improves its performance with transfer learning. Moreover, the similarity between training and validation graphs indicates a well-trained model.

(a)   (b)
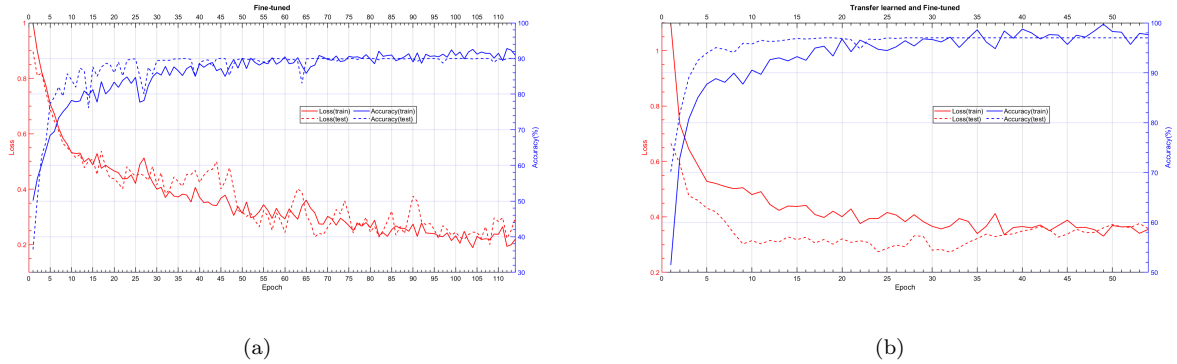
Figure 6.5: The proposed multi-modal deep neural network learning curves with; a) training from scratch and b) transfer learning with fine-tuning.

6.5.2 Fine-tuning Parameter Selection

Table 6.1 shows the training summary of the proposed network with and without transfer learning. As discussed in the section 6.5.1, our proposed multi-modal network takes less than 60 epochs to fine-tune with the pre-trained layers and a set of parameters that are explained below.

We use $Intel^{©}$ $Xenon^{©}$ 2.6 GHz CPU with K40 NVIDIA GPU to train and evaluate our proposed network with the batch size of $m = 5$. The learning rate is kept uniform at $l = 1 \times 10^{-5}$. Higher rates usually lead to over-fitting. In contrast, lower rates can affect computation costs by causing little change across epochs. Therefore, this value is determined by empirically analyzing validation errors during fine-tuning in our preliminary works related to image processing [11, 12, 13].

The momentum coefficient $\wp$ controls the fluctuation of weights by adding some parts of the previous iteration to the current iteration. Its value is kept low in early iterations due to drastic changes and non-optimal learning of the network. However, with transfer learning, the fluctuations are minimized by encouraging the changes similar to previous iterations. Thus, leading towards uniform and earlier convergence of optimal weights. We consider $\wp = 0.9$ for all epochs of the pre-trained network, which is trained for binary classification but yet to be trained and fine-tuned for optimal weights of multi-class classification. Similarly, the weight decay is set to $d = 1$, which helps in regularizing the gradient descent [248].

Table 6.1: The proposed neural network training summary with and without transfer learning.

Architecture	#Parameters	Epochs	Time(s)
Fine-tuned		114	14,820
Transfer learned & fine-tuned	113, 360, 995	54	7,020

6.6 Results and Analysis

An ablation analysis is performed to investigate the proposed multi-modal deep neural network and its variants, as illustrated in Table 6.2. The suggested changes are introduced step-wise, including multi-modal input (i.e., image and pathological data), dilated convolution, residual connections, and transfer learning. It is evident from Table 6.2 that pathology data makes a considerable improvement. Likewise, dilation and residual connections are added to the gradual increase in accuracy. However, including pre-trained layer weight (discussed in section 6.5.1) further improves the performance of the proposed network with less computation time such that it takes only 54 epochs to converge. Thus, we can infer that transfer learning considerably improves model performance.

Table 6.2: An ablation analysis of the proposed multi-modal deep neural network.

Architecture	Accuracy(%)
Proposed neural network (Image input only)	78.37
+Pathological data input	85.21
+Dilated convolution	86.01
+Residual connection	89.93
+Transfer learning & fine-tuned	96.09

Similarly, we compute various evaluation metrics for the multi-modal deep neural network. These metrics include accuracy, precision, sensitivity, specificity, and F1-score, as shown in Table 6.3. Average accuracy of 96.09% is achieved for the proposed model. Contrarily, HCC Single has better precision and sensitivity among three classes. It implies that the trained model has relatively better reproducibility and a true positive rate for the HCC Single class. In contrast, the HCC Multiple has higher specificity. Favorably, all three classes have comparable performance outcomes that indicate a trained model is less bias towards any particular class. However, besides accuracy-based metrics, to further evaluate model performance, we plot receiver operating curves (ROC) for the multi-class problem using the one versus rest (OvR) strategy as depicted in Figure 6.6. The area under ROC curves (AUC) shows the trained model performs the best for the HCC Single and is comparable for the mCRC class, while for HCC Multiple it is satisfactory. It also verifies that the proposed model is well-trained to classify liver cancer variants.

Table 6.3: The performance metrics evaluation of the proposed multi-modal deep neural network.

	HCC Multiple	HCC Single	mCRC
Precision	0.9523	0.9647	0.9623
Sensitivity	0.9424	0.9678	0.9654
Specificity	0.9856	0.9783	0.9761
F1-Score	0.9473	0.9662	0.9638

Furthermore, Table 6.4 illustrates a comparative analysis of several deep learning studies on our benchmark dataset. To reproduce results with fair comparative analysis, we only review some well-known diagnosis studies that consider neural network models with hierarchical features. Nevertheless, it is evident from the
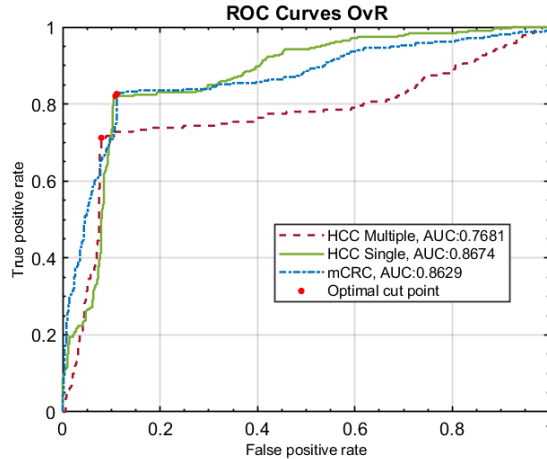
Figure 6.6: The ROC curves for one versus rest class.

analyses that our proposed network outperforms all previously developed methods. However, Das et al. [33] give competitive results for our three-class diagnosis.

Table 6.4: Prediction performance of various models for liver cancer diagnosis.

Study	Year	Accuracy(%)
AlexNet [228]	2012	76.56
VGG [339]	2014	66.21
Wu et al. [258]	2014	81.19
ResNet-152 [10]	2015	82.02
Meng et al. [22]	2017	87.34
Vivanti et al. [269]	2017	81.29
Wang et al. [236]	2018	89.49
Yasaka et al. [242]	2018	86.90
Yasaka et al. [267]	2018	79.55
Reddy et al. [29]	2018	84.09
Yamakawa et al. [30]	2019	89.82
Wang et al. [31]	2019	80.53
Kutlu and Avcı [32]	2019	90.10
Das et al. [33]	2019	95.56
Renukadevi and Karunakaran [34]	2020	94.71
Li et al. [35]	2020	85.28
Balagourouchetty et al. [36]	2020	93.85
Ours	**-**	**96.09**

Our experimental findings validate our proposed model performance compared to various baseline models. Several more studies [70, 73, 204, 32, 219, 31, 208, 262] have proposed different algorithms to predict focal and diffuse liver disorders. However, these studies commonly face numerous constraints, including insufficient and imbalance dataset, multi-modal acquisition problems, etc. All these factors can immensely affect algorithm performance for a reliable diagnosis. Nevertheless, we examine 248 subjects with transfer learning and fine-tuning for our analysis. Transfer learning assists to cater imbalance and quantity related issues by learning generic features on another dataset. Whereas fine-tuning further improves by learning features specific to our multi-class dataset. Likewise, neural networks have been commonly applied for image-based diagnosis [143],

while definite medical analyses are still performed with reference to clinical reports. In contrast, we propose a multi-modal network that integrate pathology data with image data, thus making the diagnosis according to a medical perspective. Moreover, pathology data can also assist in making some novel prognoses such as mortality rate and cancer reoccurrence [451].

6.7   Conclusion

In this study, we present a multi-modal deep neural network that integrate images with pathology data for multi-class liver cancer diagnosis. The preprocessed portal venous CT scans are fed to the proposed network that frequently uses dilation, residual connection, and batch normalization to explore novel features for reliable diagnosis meanwhile addressing vanishing gradient problems. The pre-trained layers further help counter imbalance and insufficient dataset issues with transfer learning. Alternatively, the wide and deep network is utilized to process pathology data with a concept of memorization and generalization. The algorithm is trained and tested on 248 subjects (more than double the previous studies). The fine-tuned network can classify malignant liver among HCC Multiple, HCC Single, and mCRC with an AUC of 0.77, 0.87, and 0.86 and an average accuracy of 96.06%. Our experiments validate the proposed deep neural network surpasses the existing models with subtle differences over the same benchmark dataset. Future work involves prognoses of reoccurrence risk and mortality rate with deep neural networks.

7 Summary and Future Work

7.1 Summary

Chronic liver diseases are considered potentially fatal due to asymptotic behavior and being progressive in nature. This dissertation research aims to lower the mortality rate by diagnosing liver cancer variants in a timely manner. Novel methodologies have been proposed to preprocess images that contain artifacts and noise due to contrast enhancement injection, unintentional patient movements, and poor optical quality. Similarly, the obtained CT scans may contain vague boundaries of body organs due to similar intensity levels and could lead to imprecise segmentation. Likewise, computer-aided diagnosis algorithms have been developed using only image data without considering clinical information correlated to image data. Here I proposed five objectives in Chapter 1 to address these issues; Chapters 2 to 6 have achieved these objectives.

Chapter 2 presented a comprehensive review of machine learning models for liver disease diagnosis based on US, CT, and MRI. Advances in preliminary steps, including image acquisition, preprocessing, attribute analysis, and machine learning models, have been discussed in detail with a comparative analysis of best-performed studies.

Chapter 3 proposed a multi-scale GAN for the deblurring of non-homogeneous blur. Residual image learning was also performed in the cascaded form to recover some fine details lost in deblurring the images.

Chapter 4 proposed a multi-level GAN to enhance low-contrast CT scans, while the conventional techniques resulted in non-uniform contrast images that could adversely affect diagnostic performance.

Chapter 5 proposed a multi-scale Unet with a novel loss function to perform liver and lesion segmentation. The proposed technique is capable of segmenting small lesions that are misclassified with conventional models.

Chapter 6 proposed a multi-modal neural network to diagnose liver cancer variants. The pathological data was integrated with imaging data to perform diagnosis.

With the development of these models, liver cancer diagnosis may be more accurately identified and diagnosed. However, it is to be noted that my methods are not optimal. For instance, Chapter 5 utilized a supervised neural network for segmentation, which can lead to an over-trained model with limitations of training data. Therefore, an unsupervised neural network could be developed to address such issues. Similarly, due to data constraints, Chapter 6 utilized limited pathology features for prediction. In contrast, multiple clinical factors could play a significant role in diagnosis.

7.2 Contributions

This dissertation provides innovative methodological and analytical approaches for diagnosing liver cancer variants, which are explained as follows.

1. Image deblurring is commonly performed using conventional iterative and deep learning-based techniques that limit computational time, output spatial resolution, and structural similarity due to variation in homogeneous and heterogeneous blur. In Chapter 3, I incorporated gradient input with a generative adversarial network and performed residual image learning to improve the restoration of structural details in variant and non-variant blur. Residual image learning was introduced to recover high-resolution details in the generated images.

2. Contrast adjustment was applied to enhance low-contrast images that led to a series of variable contrast images, while neural networks with supervised learning resulted in blurred images. In Chapter 4, I developed a generative adversarial network to enhance low-contrast CT scanned images. The network processed images at various scales in a coarse-to-fine manner.

3. Segmentation is required for the extraction of malignant regions. However, CT scans with multiple lesions may result in poor segmentation performance due to diverse anatomy and homogeneous contrast. In Chapter 5, I designed a novel loss function with improved Unet to perform automatic segmentation of malignant liver and lesions. The network, trained on several datasets, segmented minute lesions with improved precision.

4. Previously, imagery data were frequently used to perform computer-aided diagnoses. In Chapter 6, I proposed a multi-modal neural network that integrates correlated pathological and imagery data to diagnose liver cancer variants.

7.3 Future Work

Based on the algorithms proposed in this dissertation, I find multiple future directions that could improve diagnosis:

1. The human liver is mainly divided into two lobes. These lobes are further sub-divided into sections as left lobe with segments II and III. A medial sector with segment IV and right lobe with segments V, VI, VII and $VIII$. The identification of these segments in scans of the liver is a bit of a complicated task and requires expert knowledge. In frontal view, segments VI and VII are not visible therefore transverse images at various positions are acquired for all segments. Therefore, malignant CT scans can be classified into segments to assist in surgery.

2. Tumor staging (resectable, borderline resectable or unresectable) is performed to decide the feasibility of surgery. Estimation of tumor size could become a valuable estimator in deciding accurate staging.

3. Traditionally, TNM (Tumor, lymph Node, Metastasis) staging is used to predict patient survival in the majority of cancers. However, not all information for TNM staging can be accurately available before the final pathology, often from surgery. In contrast, survival prediction and recurrence rate could be vital in determining if patients should undergo surgery. Accordingly, it would be useful to integrate image and clinical data with TNM staging to predict survival outcomes and recurrence rate of liver cancer.

4. Several diagnosis studies validate their algorithms on limited non-public datasets, making it difficult for comparative analysis between different discussed methodologies. Keeping in view these constraints, the developed CAD systems still lack the fully commercial use, hence unsupervised models could be used to cater data constraints and develop reliable CAD systems for commercial use.

References

[1] M. Bansal, V. Sharma, S. Singh, Comparison of texture models for efficient ultrasound image retrieval, Medical Imaging 2013: Computer-Aided Diagnosis 8670 (2013) 86702C. `doi:10.1117/12.2007837`. viii, 14, 15

[2] V. Sharma, K. C. Juglan, Automated classification of fatty and normal liver ultrasound images based on mutual information feature selection, Irbm 39 (5) (2018) 313–323. `doi:10.1016/j.irbm.2018.09.006`. viii, 15, 23, 50

[3] M. Subramanya, J. Virmani, Kriti, Chapter 3 - A DEFS based system for differential diagnosis between severe fatty liver and cirrhotic liver using ultrasound images, in: N. Dey, S. Borra, A. S. Ashour, F. Shi (Eds.), Machine Learning in Bio-Signal Analysis and Diagnostic Imaging, Academic Press, 2019, pp. 53 – 72. `doi:10.1016/B978-0-12-816086-2.00003-5`. viii, 13, 15, 18, 19, 23

[4] S. Nah, T. H. Kim, K. M. Lee, Deep multi-scale convolutional neural network for dynamic scene deblurring, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2017-Janua, 2017, pp. 257–265. `doi:10.1109/CVPR.2017.35`. viii, ix, 10, 28, 30, 37, 39, 40, 41, 42, 43, 44, 51, 53, 134, 135

[5] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, S. Harmeling, Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database, in: European Conference on Computer Vision, 2012, pp. 27–40. `doi:10.1007/978-3-642-33786-4_3`. ix, x, 28, 37, 39, 41, 42, 134, 135

[6] W.-S. Lai, J.-B. Huang, Z. Hu, N. Ahuja, M.-H. Yang, A comparative study for single image blind deblurring, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1701–1709. `doi:10.1109/CVPR.2016.188`. ix, x, 28, 38, 39, 41, 43, 134, 136

[7] J. Moreau, A. Hostettler, L. Soler, Dircadb dataset: Research institute against digestive cancer, `https://www.ircad.fr/research/3dircadb/`, Accessed on: Jun. 2020 (2016). ix, 12, 48, 56, 58, 59, 71, 74, 75, 87

[8] V. G. Bram, H. Tobias, M. Styner, Segmentation in the clinic: A grand challenge (SLIVER07) database, `https://sliver07.grand-challenge.org/`, Accessed on: Jun. 2020 (2007). ix, 12, 48, 56, 59, 60, 71, 76, 87

[9] Codelab, Liver tumor segmentation challenge (LiTS) database, `http://www.lits-challenge.com/`, Accessed on: Jun. 2020 (2017). ix, 12, 48, 56, 59, 61, 71, 87

[10] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016, pp. 770–778. `doi:10.1109/CVPR.2016.90`. ix, 17, 34, 67, 68, 80, 93

[11] R. A. Khan, Y. Luo, F.-X. Wu, Multi-scale GAN with residual image learning for removing heterogeneous blur, IET Image Processing 16 (9) (2022) 2412–2431. `doi:10.1049/ipr2.12497`. x, 2, 47, 84, 88, 91

[12] R. A. Khan, Y. Luo, F.-X. Wu, Multi-level GAN based enhanced CT scans for liver cancer diagnosis, Biomedical Signal Processing and Control (2022). x, 2, 87, 88, 91

[13] R. A. Khan, Y. Luo, F.-X. Wu, RMS-UNet: Residual multi-scale UNet for liver and lesion segmentation, Artificial Intelligence in Medicine 124 (2022) 102–231. `doi:10.1016/j.artmed.2021.102231`. x, 2, 30, 60, 84, 87, 88, 91

[14] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, J. Matas, Deblurgan: Blind motion deblurring using conditional adversarial networks, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 8183–8192. `doi:10.1109/CVPR.2018.00854`. x, 10, 22, 28, 31, 32, 33, 34, 35, 36, 40, 41, 42, 43, 44, 48, 51, 53, 84, 134, 135

[15] World Cancer Research, American institute for cancer research, `https://www.wcrf.org/dietandcancer/liver-cancer`, Accessed on: Apr. 2020 (2018). 1, 5, 21, 64, 84

[16] S. Asokan, A. Suruliandi, Performance Evaluation of Tumor Detection Techniques, International Journal on Computational Science & Applications 5 (3) (2015) 25–36. `doi:10.5121/ijcsa.2015.5303`. 1, 64

[17] D. T. Fakultät, A. Militzer, Boosting Methods for Automatic Segmentation of Focal Liver Lesions Boosting-Verfahren zur automatischen Segmentierung fokaler Leberläsionen, Ph.D. thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (2014). 1

[18] M. G. Linguraru, W. J. Richbourg, J. M. Watt, V. Pamulapati, R. M. Summers, Liver and tumor segmentation and analysis from CT of diseased patients via a generic affine invariant shape parameterization and graph cuts, in: International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging, Springer, 2011, pp. 198–206. `doi:10.1007/978-3-642-28557-8_25`. 1, 64

[19] R. Tomas, Automatic localization and classification of liver lesions, Ph.D. thesis, University of West Bohemia, Accessed on: Jun. 2020 (2016).
URL `https://dspace5.zcu.cz/bitstream/11025/28548/1/doctoral_thesis_Ryba.pdf` 1, 64

[20] R. A. Khan, Y. Luo, F.-X. Wu, Machine learning based liver disease diagnosis: A systematic review, Neurocomputing 468 (2022) 492–509. `doi:10.1016/j.neucom.2021.08.138`. 1, 2, 30, 60, 64, 84, 86

[21] S. Kondo, K. Takagi, M. Nishida, T. Iwai, Y. Kudo, K. Ogawa, T. Kamiyama, H. Shibuya, K. Kahata, C. Shimizu, Computer-aided diagnosis of focal liver lesions using contrast-enhanced ultrasonography with perflubutane microbubbles, IEEE Transactions on Medical Imaging 36 (7) (2017) 1427–1437. `doi:10.1109/TMI.2017.2659734`. 1, 5, 14, 23, 84

[22] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, B. Hu, Liver fibrosis classification based on transfer learning and FCNet for ultrasound images, IEEE Access 5 (1) (2017) 5804–5810. `doi:10.1109/ACCESS.2017.2689058`. 1, 5, 21, 23, 84, 93

[23] L. Yu, C. Wang, S. Cheng, L. Guo, Establishment of computer-aided diagnosis system for liver tumor CT based on SVM, in: 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), IEEE, 2018, pp. 710–715. `doi:10.1109/DSC.2018.00113`. 1, 5, 16, 24, 47, 84

[24] E. Trivizakis, G. C. Manikis, K. Nikiforaki, K. Drevelegas, M. Constantinides, A. Drevelegas, K. Marias, Extending 2-D convolutional neural networks to 3-D for advancing deep learning cancer classification with application to MRI liver tumor differentiation, IEEE Journal of Biomedical and Health Informatics 23 (3) (2019) 923–930. `doi:10.1109/JBHI.2018.2886276`. 1, 5, 11, 24, 84

[25] M. Gletsos, S. G. Mougiakakou, G. K. Matsopoulos, K. S. Nikita, A. S. Nikita, D. Kelekis, A Computer-aided diagnostic system to characterize CT focal liver lesions: design and optimization of a neural network classifier, IEEE Transactions on Information Technology in Biomedicine 7 (3) (2003) 153–162. `doi:10.1109/TITB.2003.813793`. 1, 5, 21, 84

[26] K. M. Horton, D. A. Bluemke, R. H. Hruban, P. Soyer, E. K. Fishman, CT and MR imaging of benign hepatic and biliary tumors, RadioGraphics 19 (2) (1999) 431–451. `doi:10.1148/radiographics.19.2.g99mr04431`. 1, 5, 7, 8, 84

[27] A. Oyama, Y. Hiraoka, I. Obayashi, Y. Saikawa, S. Furui, K. Shiraishi, S. Kumagai, T. Hayashi, J. Kotoku, Hepatic tumor classification using texture and topology analysis of non-contrast-enhanced three-dimensional T1-weighted MR images with a radiomics approach, Scientific Reports 9 (1) (2019) 2–11. `doi:10.1038/s41598-019-45283-z`. 1, 5, 13, 15, 22, 84

[28] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA: A Cancer Journal for Clinicians 61 (2) (2011) 69–90. `doi:10.3322/caac.20107`. 2, 5, 6, 20

[29] D. S. Reddy, R. Bharath, P. Rajalakshmi, Classification of nonalcoholic fatty liver texture using convolution neural networks, 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services, Healthcom 2018 (2018). `doi:10.1109/HealthCom.2018.8531193`. 2, 13, 15, 23, 93

[30] M. Yamakawa, T. Shiina, N. Nishida, M. Kudo, Computer aided diagnosis system developed for ultrasound diagnosis of liver lesions using deep learning, IEEE International Ultrasonics Symposium, IUS 2019-Octob (2019) 2330–2333. `doi:10.1109/ULTSYM.2019.8925698`. 2, 23, 93

[31] C. J. Wang, C. A. Hamm, B. S. Letzen, J. S. Duncan, A probabilistic approach for interpretable deep learning in liver cancer diagnosis, in: Medical Imaging 2019: Computer-Aided Diagnosis, SPIE, 2019, p. 29. `doi:10.1117/12.2512473`. 2, 18, 24, 47, 93

[32] H. Kutlu, E. Avcı, A novel method for classifying liver and brain tumors using convolutional neural networks, discrete wavelet transform and long short-term memory networks, Sensors 19 (9) (2019) 1992. `doi:10.3390/s19091992`. 2, 14, 18, 20, 24, 86, 93

[33] A. Das, P. Das, S. Panda, S. Sabut, Detection of liver cancer using modified fuzzy clustering and decision tree classifier in CT images, Pattern Recognition and Image Analysis 29 (2) (2019) 201–211. `doi:10.1134/S1054661819020056`. 2, 25, 93

[34] T. Renukadevi, S. Karunakaran, Optimizing deep belief network parameters using grasshopper algorithm for liver disease classification, International Journal of Imaging Systems and Technology 30 (1) (2020) 168–184. `doi:10.1002/ima.22375`. 2, 13, 15, 16, 25, 55, 86, 93

[35] J. Li, Y. Wu, N. Shen, J. Zhang, E. Chen, J. Sun, Z. Deng, Y. Zhang, A fully automatic computer-aided diagnosis system for hepatocellular carcinoma using convolutional neural networks, Biocybernetics and Biomedical Engineering 40 (1) (2020) 238–248. `doi:10.1016/J.BBE.2019.05.008`. 2, 25, 55, 86, 93

[36] L. Balagourouchetty, J. K. Pragatheeswaran, B. Pottakkat, G. Ramkumar, GoogLeNet-based ensemble FCNet classifier for focal liver lesion diagnosis, IEEE Journal of Biomedical and Health Informatics 24 (6) (2020) 1686–1694. `doi:10.1109/JBHI.2019.2942774`. 2, 20, 25, 55, 60, 84, 86, 93

[37] M. C. Hulvat, Cancer incidence and trends, Surgical Clinics of North America 100 (3) (2020) 469–481. `doi:10.1016/j.suc.2020.01.002`. 5

[38] A. C. Botero, S. M. Strasberg, Division of the left hemiliver in man - segments, sectors, or sections, Liver Transplantation and Surgery 4 (3) (1998) 226–231. `doi:10.1002/lt.500040307`. 5, 22

[39] J. C. Mejia, J. Pasko, Primary liver cancers, Surgical Clinics of North America 100 (3) (2020) 535–549. `doi:10.1016/j.suc.2020.02.013`. 5

[40] N. Rashidian, A. Alseidi, R. C. Kirks, Cancers metastatic to the liver, Surgical Clinics of North America 100 (3) (2020) 551–563. `doi:10.1016/j.suc.2020.02.005`. 5

[41] A. K. Siriwardena, J. M. Mason, S. Mullamitha, H. C. Hancock, S. Jegatheeswaran, Management of colorectal cancer presenting with synchronous liver metastases, Nature Reviews Clinical Oncology 11 (8) (2014) 446–459. `doi:10.1038/nrclinonc.2014.90`. 5, 6, 84

[42] S. S. Kumar, D. Devapal, Survey on recent CAD system for liver disease diagnosis, in: 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, 2014, pp. 763–766. `doi:10.1109/ICCICCT.2014.6993061`. 5, 6

[43] C. Li, R. Li, W. Zhang, Progress in non-invasive detection of liver fibrosis, Cancer Biology and Medicine 15 (2) (2018) 124–136. `doi:10.20892/j.issn.2095-3941.2018.0018`. 5

[44] F. Gao, Y. Zhu, J. Zhang, Artificial intelligence in computer-aided diagnosis of abdomen diseases, Science China Life Sciences 62 (10) (2019) 1396–1399. `doi:10.1007/s11427-019-1556-7`. 5, 47

[45] X. Zhang, H. Fujita, T. Qin, J. Zhao, M. Kanematsu, T. Hara, X. Zhou, R. Yokoyama, H. Kondo, H. Hoshi, CAD on liver using CT and MRI, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 4987 LNCS, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 367–376. `doi:10.1007/978-3-540-79490-5_43`. 5

[46] Y. Kadah, A. Farag, J. Zurada, A. Badawi, A.-B. Youssef, Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images, IEEE Transactions on Medical Imaging 15 (4) (1996) 466–478. `doi:10.1109/42.511750`. 5

[47] E. Kyriacou, S. Pavlopoulos, G. Konnis, D. Koutsouris, P. Zoumpoulis, L. Theotokas, I. Theotokas, Computer assisted characterization of diffused liver disease using image texture analysis techniques on B-scan images, in: 1997 IEEE Nuclear Science Symposium Conference Record, Vol. 2, IEEE, 1997, pp. 1479–1483. `doi:10.1109/NSSMIC.1997.670599`. 5, 14

[48] A. Myronenko, X. Song, Intensity-based image registration by minimizing residual complexity, IEEE Transactions on Medical Imaging 29 (11) (2010) 1882–1891. `doi:10.1109/TMI.2010.2053043`. 5

[49] T. Heimann, B. Van Ginneken, M. A. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, F. Bello, G. Binnig, H. Bischof, A. Bornik, P. M. Cashman, Y. Chi, A. Córdova, B. M. Dawant, M. Fidrich, J. D. Furst, D. Furukawa, L. Grenacher, J. Hornegger, D. Kainmüller, R. I. Kitney, H. Kobatake, H. Lamecker, T. Lange, J. Lee, B. Lennon, R. Li, S. Li, H. P. Meinzer, G. Németh, D. S. Raicu, A. M. Rau, E. M. Van Rikxoort, M. Rousson, L. Ruskó, K. A. Saddi, G. Schmidt, D. Seghers, A. Shimizu, P. Slagmolen, E. Sorantin, G. Soza, R. Susomboon, J. M. Waite, A. Wimmer, I. Wolf, Comparison and evaluation of methods for liver segmentation from CT datasets, IEEE Transactions on Medical Imaging 28 (8) (2009) 1251–1265. `doi:10.1109/TMI.2009.2013851`. 5, 11, 12, 72, 73

[50] J. Zhang, Y. Xia, Y. Xie, M. Fulham, D. D. Feng, Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features, IEEE Journal of Biomedical and Health Informatics 22 (5) (2018) 1521–1530. `doi:10.1109/JBHI.2017.2775662`. 5, 6

[51] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), Narosa Publishing House, 1998, pp. 839–846. `doi:10.1109/ICCV.1998.710815`. 6, 9

[52] V. Bhateja, M. Misra, S. Urooj, Non-linear polynomial filters for edge enhancement of mammogram lesions, Computer Methods and Programs in Biomedicine 129 (2016) 125–134. `doi:10.1016/j.cmpb.2016.01.007`. 6, 61, 84

[53] A. Gotra, L. Sivakumaran, G. Chartrand, K.-N. N. Vu, F. Vandenbroucke-Menu, C. Kauffmann, S. Kadoury, B. Gallix, J. A. de Guise, A. Tang, Liver segmentation: indications, techniques and future directions, Insights into Imaging 8 (4) (2017) 377–392. `doi:10.1007/s13244-017-0558-1`. 6, 11, 64

[54] A. S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on MRI, Zeitschrift für Medizinische Physik 29 (2) (2019) 102–127. `doi:10.1016/j.zemedi.2018.11.002`. 6, 11, 17

[55] L. Liu, J. Cheng, Q. Quan, F.-X. Wu, Y.-P. Wang, J. Wang, A survey on u-shaped networks in medical image segmentations, Neurocomputing 409 (2020) 244–258. `doi:10.1016/j.neucom.2020.05.070`. 6, 12, 68, 84

[56] L. Ali, K. Khelil, S. K. Wajid, Z. U. Hussain, M. A. Shah, A. Howard, A. Adeel, A. A. Shah, U. Sudhakar, N. Howard, A. Hussain, Machine learning based computer-aided diagnosis of liver tumours, in: 2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), IEEE, 2017, pp. 139–145. `doi:10.1109/ICCI-CC.2017.8109742`. 6, 16, 21

[57] S. A. Azer, Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: A systematic review, World Journal of Gastrointestinal Oncology 11 (12) (2019) 1218–1230. doi:10.4251/wjgo.v11.i12.1218. 6, 8, 16, 84

[58] J. Xu, M. Jing, S. Wang, C. Yang, X. Chen, A review of medical image detection for cancers in digestive system based on artificial intelligence, Expert Review of Medical Devices 16 (10) (2019) 877–889. doi:10.1080/17434440.2019.1669447. 6, 16, 84

[59] O. Faust, U. R. Acharya, K. M. Meiburger, F. Molinari, J. E. Koh, C. H. Yeong, P. Kongmebhol, K. H. Ng, Comparative assessment of texture features for the identification of cancer in ultrasound images: a review, Biocybernetics and Biomedical Engineering 38 (2) (2018) 275–296. doi:10.1016/j.bbe.2018.01.001. 6, 14, 21, 84

[60] M. Moghbel, S. Mashohor, R. Mahmud, M. I. B. Saripan, Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography, Artificial Intelligence Review 50 (4) (2018) 497–537. doi:10.1007/s10462-017-9550-x. 6, 20, 64, 72, 73, 84

[61] P. V. Nayantara, S. Kamath, K. Manjunath, K. Rajagopal, Computer-Aided Diagnosis of liver lesions using CT images: A systematic review, Computers in Biology and Medicine (2020) 104035doi:10.1016/j.compbiomed.2020.104035. 6, 84

[62] K. M. Elsayes, A. Z. Kielar, V. Chernyak, A. Morshid, A. Furlan, W. R. Masch, R. M. Marks, A. Kamaya, R. K. Do, Y. Kono, et al., Li-rads: a conceptual and historical review from its beginning to its recent integration into aasld clinical practice guidance, Journal of hepatocellular carcinoma 6 (2019) 49. doi:10.2147/JHC.S186239. 7

[63] J.-H. Lee, H. Y. Kim, Y. J. Kim, J.-H. Yoon, J. W. Chung, H.-S. Lee, Barcelona clinic liver cancer staging system and survival of untreated hepatocellular carcinoma in a hepatitis b virus endemic area, Journal of gastroenterology and hepatology 30 (4) (2015) 696–705. doi:10.1111/jgh.12788. 7

[64] R. C. N. Melo, M. W. D. Raas, C. Palazzi, V. H. Neves, K. K. Malta, T. P. Silva, Whole slide imaging and its applications to histopathological studies of liver disorders, Frontiers in Medicine 6 (2020) 310. doi:10.3389/fmed.2019.00310. 7

[65] X. Zhang, X. Gao, B. J. Liu, K. Ma, W. Yan, L. Liling, H. Yuhong, H. Fujita, Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging?, Computerized Medical Imaging and Graphics 46 (June) (2015) 227–236. doi:10.1016/j.compmedimag.2015.09.003. 7, 11, 13, 14

[66] S. G. Stanciu, S. Xu, Q. Peng, J. Yan, G. A. Stanciu, R. E. Welsch, P. T. So, G. Csucs, H. Yu, Experimenting liver fibrosis diagnostic by two photon excitation microscopy and bag-of-features image classification, Scientific reports 4 (1) (2014) 1–12. doi:10.1038/srep04636. 7

[67] J. K. Dowman, J. Tomlinson, P. Newsome, Systematic review: Diagnosis and staging of non-alcoholic fatty liver disease (nafld)/non-alcoholic steatohepatitis (nash)', Alimentary Pharmacology and Therapeutics 33 (2010) 525–540. doi:10.1111/j.1365-2036.2010.04556.x. 7

[68] T. J. Holmes, D. Biggs, A. Abu-Tarif, Blind deconvolution, in: J. B. Pawley (Ed.), Handbook Of Biological Confocal Microscopy, Springer US, Boston, MA, 2006, pp. 468–487. doi:10.1007/978-0-387-45524-2_24. 8, 47

[69] M. Y. Jabarulla, H. N. Lee, Computer aided diagnostic system for ultrasound liver images: A systematic review, Optik 140 (2017) 1114–1126. doi:10.1016/j.ijleo.2017.05.013. 8, 9, 16, 18

[70] B. Xia, H. Jiang, H. Liu, D. Yi, A novel hepatocellular carcinoma image classification method based on voting ranking random forests, Computational and Mathematical Methods in Medicine 2016 (2016). doi:10.1155/2016/2628463. 8, 14, 18, 19, 20, 21, 24, 93

[71] T. M. Hassan, M. Elmogy, E. Sallam, A classification framework for diagnosis of focal liver diseases, in: 2015 Tenth International Conference on Computer Engineering & Systems (ICCES), IEEE, 2015, pp. 395–401. `doi:10.1109/ICCES.2015.7393083`. 8, 16, 23

[72] M. Kachelrieß, O. Watzke, W. A. Kalender, Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam ct, Medical physics 28 (4) (2001) 475–490. `doi:10.1118/1.1358303`. 8, 9

[73] K. Raghesh Krishnan, S. Radhakrishnan, Hybrid approach to classification of focal and diffused liver disorders using ultrasound images with wavelets and texture features, IET Image Processing 11 (7) (2017) 530–538. `doi:10.1049/iet-ipr.2016.1072`. 8, 15, 18, 23, 93

[74] T. Lossau, H. Nickisch, T. Wissel, R. Bippus, H. Schmitt, M. Morlock, M. Grass, Motion artifact recognition and quantification in coronary CT angiography using convolutional neural networks, Medical Image Analysis 52 (2019) 68–79. `doi:10.1016/j.media.2018.11.003`. 8, 9, 10

[75] I. Fantini, L. Rittner, C. Yasuda, R. Lotufo, Automatic detection of motion artifacts on mri using deep cnn, in: 2018 International Workshop on Pattern Recognition in Neuroimaging (PRNI), IEEE, 2018, pp. 1–4. `doi:10.1109/PRNI.2018.8423948`. 8, 9

[76] H. Wu, S. Zheng, J. Zhang, K. Huang, Gp-gan: Towards realistic high-resolution image blending, in: Proceedings of the 27th ACM international conference on multimedia, 2019, pp. 2487–2495. `doi:10.1145/3343031.3350944`. 8, 10

[77] K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu, J. Jiang, Edge-enhanced GAN for remote sensing image superresolution, IEEE Transactions on Geoscience and Remote Sensing 57 (8) (2019) 5799–5812. `doi:10.1109/TGRS.2019.2902431`. 8, 10, 31, 35, 50, 84

[78] C. You, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li, S. Ju, Z. Zhao, Z. Zhang, W. Cong, et al., CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE), IEEE transactions on medical imaging 39 (1) (2019) 188–203. `doi:10.1109/TMI.2019.2922960`. 8, 10, 84

[79] A. H. Mahnken, R. Raupach, J. E. Wildberger, B. Jung, N. Heussen, T. G. Flohr, R. W. Gnther, S. Schaller, A new algorithm for metal artifact reduction in computed tomography, Investigative Radiology 38 (12) (2003) 769–775. `doi:10.1097/01.rli.0000086495.96457.54`. 8

[80] P. B. Heffernan, R. A. Robb, Image reconstruction from incomplete projection data: Iterative reconstruction-reprojection techniques, IEEE Transactions on Biomedical Engineering BME-30 (12) (1983) 838–841. `doi:10.1109/TBME.1983.325089`. 8

[81] W. A. Kalender, R. Hebel, J. Ebersberger, Reduction of CT artifacts caused by metallic implants., Radiology 164 (2) (1987) 576–577. `doi:10.1148/radiology.164.2.3602406`. 8

[82] G. W. Goerres, C. Burger, E. Kamel, B. Seifert, A. H. Kaim, A. Buck, T. C. Buehler, G. K. von Schulthess, Respiration-induced attenuation artifact at PET/CT: Technical considerations, Radiology 226 (3) (2003) 906–910. `doi:10.1148/radiol.2263011732`. 9

[83] C. Crawford, K. King, C. Ritchie, J. Godwin, Respiratory compensation in projection imaging using a magnification and displacement model, IEEE Transactions on Medical Imaging 15 (3) (1996) 327–332. `doi:10.1109/42.500141`. 9

[84] X. Li, Fast and efficient block motion estimation in the wavelet space, in: Proceedings ITCC 2003. International Conference on Information Technology: Coding and Computing, IEEE Comput. Soc, 2003, pp. 488–494. `doi:10.1109/ITCC.2003.1197578`. 9, 11

[85] K. Taguchi, H. Aradate, Algorithm for image reconstruction in multi-slice helical CT, Medical Physics 25 (4) (1998) 550–561. `doi:10.1118/1.598230`. 9

[86] T. Elss, H. Nickisch, T. Wissel, H. Schmitt, M. Vembar, M. Morlock, M. Grass, Deep-learning-based CT motion artifact recognition in coronary arteries, in: Medical Imaging 2018: Image Processing, Vol. 10574, International Society for Optics and Photonics, 2018, p. 1057416. `doi:10.1117/12.2292882`. 9

[87] J. E. Iglesias, G. Lerma-Usabiaga, L. C. Garcia-Peraza-Herrera, S. Martinez, P. M. Paz-Alonso, Retrospective head motion estimation in structural brain MRI with 3D CNNs, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2017, pp. 314–322. `doi:10.1007/978-3-319-66185-8_36`. 9

[88] T. Küstner, A. Liebgott, L. Mauch, P. Martirosian, F. Bamberg, K. Nikolaou, B. Yang, F. Schick, S. Gatidis, Automated reference-free detection of motion artifacts in magnetic resonance images, Magnetic Resonance Materials in Physics, Biology and Medicine 31 (2) (2018) 243–256. `doi: 10.1007/s10334-017-0650-z`. 9

[89] K. M. van Hespen, J. J. Zwanenburg, J. W. Dankbaar, M. I. Geerlings, J. Hendrikse, H. J. Kuijf, An anomaly detection approach to identify chronic brain infarcts on MRI, Scientific Reports 11 (1) (2021) 1–10. `doi:10.1038/s41598-021-87013-4`. 9

[90] H. Ma, E. Gros, A. Szabo, S. G. Baginski, Z. R. Laste, N. M. Kulkarni, D. Okerlund, T. G. Schmidt, Evaluation of motion artifact metrics for coronary CT angiography, Medical physics 45 (2) (2018) 687–702. `doi:10.1002/mp.12720`. 9

[91] R. C. Gonzalez, R. E. Woods, Digital image processing (3rd Edition), Prentice-Hall, Inc., USA, 2006. 9, 48

[92] J. Benesty, J. Chen, Y. Huang, Study of the widely linear Wiener filter for noise reduction, in: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2010, pp. 205–208. `doi:10.1109/ICASSP.2010.5496033`. 9

[93] R. Yang, L. Yin, M. Gabbouj, J. Astola, Y. Neuvo, Optimal weighted median filtering under structural constraints, IEEE Transactions on Signal Processing 43 (3) (1995) 591–604. `doi:10.1109/78.370615`. 9

[94] A. Mendrik, E.-J. Vonken, A. Rutten, M. Viergever, B. van Ginneken, Noise reduction in computed tomography scans using 3-D anisotropic hybrid diffusion with continuous switch, IEEE Transactions on Medical Imaging 28 (10) (2009) 1585–1594. `doi:10.1109/TMI.2009.2022368`. 9

[95] J. V. Manjón, P. Coupé, L. Martí-Bonmatí, D. L. Collins, M. Robles, Adaptive non-local means denoising of MR images with spatially varying noise levels, Journal of Magnetic Resonance Imaging 31 (1) (2010) 192–203. `doi:10.1002/jmri.22003`. 9

[96] M. Mafi, H. Martin, M. Cabrerizo, J. Andrian, A. Barreto, M. Adjouadi, A comprehensive survey on impulse and Gaussian denoising filters for digital images, Signal Processing 157 (2019) 236–260. `doi:10.1016/j.sigpro.2018.12.006`. 9, 29, 47, 54

[97] W.-S. Lai, J.-J. Ding, Y.-Y. Lin, Y.-Y. Chuang, Blur kernel estimation using normalized color-line priors, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2015, pp. 64–72. `doi:10.1109/CVPR.2015.7298601`. 9, 27

[98] J. J. Lizhe Tan, Digital signal processing, Academic Press, 2018. 9, 27

[99] J. Proakis, D. Manolakis, Digital signal processing: Principles, algorithms, and applications, Prentice Hall, 2007. 9, 27

[100] A. Oppenheim, R. Schafer, Discrete-time signal processing, Prentice Hall, 2009. 9, 27

[101] A. Buades, B. Coll, J.-M. Morel, A review of image denoising algorithms, with a new one, Multiscale Modeling & Simulation 4 (2) (2005) 490–530. `doi:10.1137/040616024`. 9

[102] J. Jiao, W.-C. Tu, S. He, R. W. Lau, Formresnet: Formatted residual learning for image restoration, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 38–46. `doi:10.1109/CVPRW.2017.140`. 9, 28, 31, 33, 35, 44, 48

[103] J. Jiao, W.-C. Tu, D. Liu, S. He, R. W. H. Lau, T. S. Huang, FormNet: Formatted learning for image restoration, IEEE Transactions on Image Processing 29 (2020) 6302–6314. `doi:10.1109/TIP.2020.2990603`. 9, 27, 28, 31, 33, 35, 44, 48

[104] Q. Yuan, Q. Zhang, J. Li, H. Shen, L. Zhang, Hyperspectral image denoising employing a spatialspectral deep residual convolutional neural network, IEEE Transactions on Geoscience and Remote Sensing 57 (2) (2019) 1205–1218. `doi:10.1109/TGRS.2018.2865197`. 9

[105] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Advances in neural information processing systems 27 (2014). `doi:10.1145/3422622`. 9, 17, 27, 28, 31, 32, 48, 50

[106] L. Xu, S. Zheng, J. Jia, Unnatural L0 sparse representation for natural image deblurring, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2013, pp. 1107–1114. `doi:10.1109/CVPR.2013.147`. 10, 27, 29, 40, 41, 43, 44, 134, 136

[107] A. Levin, R. Fergus, F. Durand, W. T. Freeman, Image and depth from a conventional camera with a coded aperture, ACM transactions on graphics (TOG) 26 (3) (2007) 70–es. `doi:10.1145/1275808.1276464`. 10, 27

[108] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, W. T. Freeman, Removing camera shake from a single photograph, in: ACM SIGGRAPH 2006 Papers on - SIGGRAPH '06, ACM Press, New York, New York, USA, 2006, p. 787. `doi:10.1145/1179352.1141956`. 10, 27, 29, 41, 42, 135, 136

[109] J. Portilla, V. Strela, M. J. Wainwright, E. P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, IEEE Transactions on Image Processing 12 (11) (2003) 1338–1351. `doi:10.1109/TIP.2003.818640`. 10

[110] M. B. Cannell, A. McMorland, C. Soeller, Image enhancement by deconvolution, in: J. B. Pawley (Ed.), Handbook Of Biological Confocal Microscopy, Springer US, Boston, MA, 2006, pp. 488–500. `doi:10.1007/978-0-387-45524-2_25`. 10, 84

[111] R. Bassett, J. Deride, Maximum a posteriori estimators as a limit of Bayes estimators, Mathematical Programming 174 (1-2) (2019) 129–144. `doi:10.1007/s10107-018-1241-0`. 10

[112] W. Dong, G. Shi, X. Li, Nonlocal image restoration with bilateral variance estimation: A low-rank approach, IEEE Transactions on Image Processing 22 (2) (2013) 700–711. `doi:10.1109/TIP.2012.2221729`. 10

[113] H. C. Andrews, B. R. Hunt, Digital image restoration, Prentice Hall, 1977. 10, 132

[114] M. Potmesil, I. Chakravarty, Modeling motion blur in computer-generated images, Proceedings of the 10th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1983 (1983) 389–399`doi:10.1145/800059.801169`. 10, 27, 132

[115] L. Sun, S. Cho, J. Wang, H. James, Edge-based blur kernel estimation using patch priors, in: IEEE International Conference on Computational Photography (ICCP), IEEE, 2013, pp. 1–8. `doi:10.1109/ICCPhot.2013.6528301`. 10, 27, 30, 41, 43, 136

[116] L. Yuan, J. Sun, L. Quan, H.-Y. Shum, Image deblurring with blurred/noisy image pairs, ACM Transactions on Graphics 26 (99) (2007) 1. `doi:10.1145/1276377.1276379`. 10, 27, 28, 29, 33, 35, 44

[117] D.-B. Lee, S.-C. Jeong, Y.-G. Lee, B. C. Song, Video deblurring algorithm using accurate blur kernel estimation and residual deconvolution based on a blurred-unblurred frame pair, IEEE Transactions on Image Processing 22 (3) (2013) 926–940. `doi:10.1109/TIP.2012.2222898`. 10, 28, 30, 33, 35, 44, 48, 61

[118] Y. Lu, J. Sun, L. Quan, H.-Y. Shum, Blurred/non-blurred image alignment using sparseness prior, in: 2007 IEEE 11th International Conference on Computer Vision, IEEE, 2007, pp. 1–8. `doi:10.1109/ICCV.2007.4408915`. 10, 29, 30

[119] A. Goldstein, R. Fattal, Blur-kernel estimation from spectral irregularities, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7576 LNCS (PART 5) (2012) 622–635. `doi:10.1007/978-3-642-33715-4_45`. 10, 27

[120] T. S. Cho, S. Paris, B. K. P. Horn, W. T. Freeman, Blur kernel estimation using the radon transform, in: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1076, IEEE, 2011, pp. 241–248. `doi:10.1109/CVPR.2011.5995479`. 10, 27

[121] X. Xu, J. Pan, Y. J. Zhang, M. H. Yang, Motion blur kernel estimation via deep learning, IEEE Transactions on Image Processing 27 (1) (2018) 194–205. `doi:10.1109/TIP.2017.2753658`. 10, 48

[122] K. Zhang, W. Zuo, L. Zhang, Deep plug-and-play super-resolution for arbitrary blur kernels, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 1671–1681. `doi:10.1109/CVPR.2019.00177`. 10

[123] K. Meding, A. Loktyushin, M. Hirsch, Automatic detection of motion artifacts in MR images using CNNS, in: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2017, pp. 811–815. `doi:10.1109/ICASSP.2017.7952268`. 10

[124] S. Ramakrishnan, S. Pachori, A. Gangopadhyay, S. Raman, Deep generative filter for motion deblurring, in: Proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017, 2017, pp. 2993–3000. `doi:10.1109/ICCVW.2017.353`. 10, 31

[125] D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, A. Van Den Hengel, Q. Shi, From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3806–3815. `doi:10.1109/CVPR.2017.405`. 10, 31, 41, 44, 134

[126] X. Tao, H. Gao, X. Shen, J. Wang, J. Jia, Scale-recurrent network for deep image deblurring, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8174–8182. `doi:10.1109/CVPR.2018.00853`. 10, 28, 32, 40, 41, 42, 43, 44, 51, 53, 134, 135

[127] M. Ye, D. Lyu, G. Chen, Scale-iterative upscaling network for image deblurring, IEEE Access 8 (2020) 18316–18325. `doi:10.1109/ACCESS.2020.2967823`. 10, 32, 40, 41, 42, 43, 44, 134

[128] J. L. Barron, D. J. Fleet, S. S. Beauchemin, T. A. Burkitt, Performance of optical flow techniques, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1992-June, 1992, pp. 236–242. `doi:10.1109/CVPR.1992.223269`. 10

[129] J. Sun, W. Cao, Z. Xu, J. Ponce, Learning a convolutional neural network for non-uniform motion blur removal, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, pp. 769–777. `doi:10.1109/CVPR.2015.7298677`. 10, 30, 41, 44, 134

[130] J. B. Maintz, M. A. Viergever, A survey of medical image registration, Medical Image Analysis 2 (1) (1998) 1–36. `doi:10.1016/S1361-8415(01)80026-8`. 11

[131] M. A. Viergever, J. B. Maintz, S. Klein, K. Murphy, M. Staring, J. P. Pluim, A survey of medical image registration under review, Medical Image Analysis 33 (2016) 140–144. `doi:10.1016/j.media.2016.06.030`. 11

[132] A. Sotiras, C. Davatzikos, N. Paragios, Deformable medical image registration: A survey, IEEE Transactions on Medical Imaging 32 (7) (2013) 1153–1190. `doi:10.1109/TMI.2013.2265603`. 11

[133] F. P. Oliveira, J. M. R. Tavares, Medical image registration: A review, Computer Methods in Biomechanics and Biomedical Engineering 17 (2) (2014) 73–93. `doi:10.1080/10255842.2012.670855`. 11

[134] P. K. Saha, R. Strand, G. Borgefors, Digital topology and geometry in medical imaging: A survey, IEEE Transactions on Medical Imaging 34 (9) (2015) 1940–1964. `doi:10.1109/TMI.2015.2417112.` 11

[135] E. Ferrante, N. Paragios, Slice-to-volume medical image registration: A survey, Medical Image Analysis 39 (2017) 101–123. `doi:10.1016/j.media.2017.04.010.` 11

[136] L. G. Brown, A survey of image registration techniques, ACM computing surveys (CSUR) 24 (4) (1992) 325–376. `doi:10.1145/146370.146374.` 11

[137] B. Reddy, B. Chatterji, An FFT-based technique for translation, rotation, and scale-invariant image registration, IEEE Transactions on Image Processing 5 (8) (1996) 1266–1271. `doi:10.1109/83.506761.` 11

[138] H. F. Silverman, A class of algorithms for fast digital image registration, IEEE Transactions on Computers C-21 (2) (1972) 179–186. `doi:10.1109/TC.1972.5008923.` 11

[139] E. De Castro, C. Morandi, Registration of translated and rotated images using finite Fourier transforms, IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9 (5) (1987) 700–703. `doi:10.1109/TPAMI.1987.4767966.` 11

[140] G. Wu, M. Kim, Q. Wang, B. C. Munsell, D. Shen, Scalable high-performance image registration framework by unsupervised deep feature representations learning, IEEE Transactions on Biomedical Engineering 63 (7) (2016) 1505–1516. `doi:10.1109/TBME.2015.2496253.` 11

[141] X. Yang, R. Kwitt, M. Styner, M. Niethammer, Quicksilver: Fast predictive image registration A deep learning approach, NeuroImage 158 (2017) 378–396. `doi:10.1016/j.neuroimage.2017.07.008.` 11

[142] A. P. Keszei, B. Berkels, T. M. Deserno, Survey of non-rigid registration tools in medicine, Journal of Digital Imaging 30 (1) (2017) 102–116. `doi:10.1007/s10278-016-9915-8.` 11

[143] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis 42 (2012) (2017) 60–88. `doi:10.1016/j.media.2017.07.005.` 11, 17, 93

[144] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, I. Išgum, A deep learning framework for unsupervised affine and deformable image registration, Medical Image Analysis 52 (2019) 128–143. `doi:10.1016/j.media.2018.11.010.` 11

[145] S. S. Mohseni Salehi, S. Khan, D. Erdogmus, A. Gholipour, Real-time deep pose estimation with geodesic loss for image-to-template rigid registration, IEEE Transactions on Medical Imaging 38 (2) (2019) 470–481. `doi:10.1109/TMI.2018.2866442.` 11

[146] G. Balakrishnan, A. Zhao, M. R. Sabuncu, A. V. Dalca, J. Guttag, An unsupervised learning model for deformable medical image registration, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018, pp. 9252–9260. `doi:10.1109/CVPR.2018.00964.` 11

[147] K. Ungru, X. Jiang, Dynamic programming based segmentation in biomedical imaging, Computational and Structural Biotechnology Journal 15 (2017) 255 – 264. `doi:10.1016/j.csbj.2017.02.001.` 11, 64

[148] Y. Tong, L. Huang, L. Cao, J. K. Udupa, D. Odhner, D. A. Torigian, A general approach to liver lesion segmentation in CT images, in: R. J. Webster, Z. R. Yaniv (Eds.), Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling, Vol. 9786, 2016, p. 978623. `doi:10.1117/12.2217778.` 11, 64

[149] G. Chartrand, T. Cresson, R. Chav, A. Gotra, A. Tang, J. DeGuise, Semi-automated liver CT segmentation using Laplacian meshes, in: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), 2014, pp. 641–644. `doi:10.1109/ISBI.2014.6867952.` 11, 77

[150] G. Li, X. Chen, F. Shi, W. Zhu, J. Tian, D. Xiang, Automatic liver segmentation based on shape constraints and deformable graph cut in CT images, IEEE Transactions on Image Processing 24 (12) (2015) 5315–5329. `doi:10.1109/TIP.2015.2481326`. 11

[151] J. Liu, Z. Wang, R. Zhang, Liver cancer ct image segmentation methods based on watershed algorithm, in: 2009 International Conference on Computational Intelligence and Software Engineering, 2009, pp. 1–4. `doi:10.1109/CISE.2009.5364437`. 11

[152] J. Wang, Y. Cheng, C. Guo, Y. Wang, S. Tamura, Shapeintensity prior level set combining probabilistic atlas and probability map constrains for automatic liver segmentation from abdominal CT images, International Journal of Computer Assisted Radiology and Surgery 11 (5) (2016) 817–826. `doi:10.1007/s11548-015-1332-9`. 11, 77

[153] M. Erdt, M. Kirschner, Fast automatic liver segmentation combining learned shape priors with observed shape deviation, in: 2010 IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS), IEEE, 2010, pp. 249–254. `doi:10.1109/CBMS.2010.6042650`. 11, 64, 65, 77

[154] O. I. Alirr, A. A. A. Rahni, E. Golkar, An automated liver tumour segmentation from abdominal CT scans for hepatic surgical planning, International Journal of Computer Assisted Radiology and Surgery 13 (8) (2018) 1169–1176. `doi:10.1007/s11548-018-1801-z`. 11

[155] X. Han, Automatic liver lesion segmentation using a deep convolutional neural network method (2017). `arXiv:1704.07239`. 11, 13, 66, 68, 77, 79, 80

[156] G. Chlebus, H. Meine, J. H. Moltz, A. Schenk, Neural network-based automatic liver tumor segmentation with random forest-based candidate filtering (2017). `arXiv:1706.00842`. 11

[157] P. F. Christ, F. Ettlinger, F. Grün, M. E. A. Elshaera, J. Lipkova, S. Schlecht, F. Ahmaddy, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, F. Hofmann, M. D. Anastasi, S.-A. Ahmadi, G. Kaissis, J. Holch, W. Sommer, R. Braren, V. Heinemann, B. Menze, Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks (2017). `arXiv:1702.05970`. 11

[158] Y. Yu, J. Wang, C. W. Ng, Y. Ma, S. Mo, E. L. S. Fong, J. Xing, Z. Song, Y. Xie, K. Si, A. Wee, R. E. Welsch, P. T. C. So, H. Yu, Deep learning enables automated scoring of liver fibrosis stages, Scientific Reports 8 (1) (2018) 16016. `doi:10.1038/s41598-018-34300-2`. 11, 21

[159] Q. Zeng, D. Karimi, E. H. T. Pang, S. Mohammed, C. Schneider, M. Honarvar, S. E. Salcudean, Liver segmentation in magnetic resonance imaging via mean shape fitting with fully convolutional neural networks, in: D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, A. Khan (Eds.), Proceeding of the International Conference on Medical Image Computing and Computer Assisted Interventions, Vol. 11765, Springer International Publishing, Cham, 2019, pp. 246–254. `doi:10.1007/978-3-030-32245-8_28`. 11

[160] A. E. Kavur, A. M. Selver, O. D. Gezer, M. Bar, N. S. Gezer, CHAOS - Combined (CT-MR) healthy abdominal organ segmentation challenge data, `https://chaos.grand-challenge.org/`, Accessed on: Jun. 2020 (2019). 12, 71, 77, 87

[161] L. Li, J. Pan, W. S. Lai, C. Gao, N. Sang, M. H. Yang, Learning a discriminative prior for blind image deblurring, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 6616–6625. `doi:10.1109/CVPR.2018.00692`. 12, 13, 31, 39, 48, 49, 58, 59, 60, 61

[162] G. Chlebus, A. Schenk, J. H. Moltz, B. van Ginneken, H. K. Hahn, H. Meine, Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing, Scientific Reports 8 (1) (2018) 15497. `doi:10.1038/s41598-018-33860-7`. 12, 13

[163] W. Qin, J. Wu, F. Han, Y. Yuan, W. Zhao, B. Ibragimov, J. Gu, L. Xing, Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation, Physics in Medicine & Biology 63 (9) (2018) 095017. `doi:10.1088/1361-6560/aabd19`. 12, 13, 75, 77

[164] C. Sun, S. Guo, H. Zhang, J. Li, M. Chen, S. Ma, L. Jin, X. Liu, X. Li, X. Qian, Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs, Artificial Intelligence in Medicine 83 (March 2018) (2017) 58–66. `doi:10.1016/j.artmed.2017.03.008`. 12, 13, 47, 64, 66, 75, 79, 80

[165] L. Zhou, L. Wang, W. Li, S. Liang, Q. Zhang, Automatic segmentation of liver from CT scans with CCPTSPM algorithm, International Journal of Pattern Recognition and Artificial Intelligence 33 (14) (2019) 1957005. `doi:10.1142/S0218001419570052`. 12, 13, 67, 77

[166] W. Wu, S. Wu, Z. Zhou, R. Zhang, Y. Zhang, 3D Liver tumor segmentation in CT images using improved Fuzzy C-means and graph cuts, BioMed Research International 2017 (2017) 1–11. `doi:10.1155/2017/5207685`. 12, 13, 64, 65, 75, 79

[167] P. F. Christ, M. E. A. Elshaer, F. Ettlinger, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, M. Armbruster, F. Hofmann, M. D'Anastasi, W. H. Sommer, S.-A. A. Ahmadi, B. H. Menze, F. Grün, M. E. A. Elshaera, J. Lipkova, S. Schlecht, F. Ahmaddy, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, F. Hofmann, M. D. Anastasi, S.-A. A. Ahmadi, G. Kaissis, J. Holch, W. H. Sommer, R. Braren, V. Heinemann, B. H. Menze, Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, pp. 415–423. `doi:10.1007/978-3-319-46723-8_48`. 13, 66, 67, 75, 77

[168] F. Lu, F. Wu, P. Hu, Z. Peng, D. Kong, Automatic 3D liver location and segmentation via convolutional neural network and graph cut, International Journal of Computer Assisted Radiology and Surgery 12 (2) (2017) 171–182. `doi:10.1007/s11548-016-1467-3`. 13, 77

[169] M. Ahmad, D. Ai, G. Xie, S. F. Qadri, H. Song, Y. Huang, Y. Wang, J. Yang, Deep belief network modeling for automatic liver segmentation, IEEE Access 7 (2019) 20585–20595. `doi:10.1109/ACCESS.2019.2896961`. 13, 64, 65, 75, 77

[170] Y. Chen, K. Wang, X. Liao, Y. Qian, Q. Wang, Z. Yuan, P.-A. Heng, Channel-Unet: A spatial channel-wise convolutional neural network for liver and tumors segmentation, Frontiers in Genetics 10 (November) (2019) 1–13. `doi:10.3389/fgene.2019.01110`. 13, 77, 79, 80

[171] Ü. Budak, Y. Guo, E. Tanyildizi, A. Şengür, Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation, Medical hypotheses 134 (2020) 109431. `doi:10.1016/j.mehy.2019.109431`. 13, 77, 79

[172] L. Chen, H. Song, C. Wang, Y. Cui, J. Yang, X. Hu, L. Zhang, Liver tumor segmentation in CT volumes using an adversarial densely connected network, BMC Bioinformatics 20 (S16) (2019) 587. `doi:10.1186/s12859-019-3069-x`. 13, 77, 79

[173] P. Hu, F. Wu, J. Peng, P. Liang, D. Kong, Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution, Physics in Medicine and Biology 61 (24) (2016) 8676–8698. `doi:10.1088/1361-6560/61/24/8676`. 13, 77, 80

[174] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, P.-A. Heng, 3D deeply supervised network for automated segmentation of volumetric medical images, Medical Image Analysis 41 (2017) 40–54. `doi:10.1016/j.media.2017.05.001`. 13, 77

[175] Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, P.-A. Heng, 3D deeply supervised network for automatic liver segmentation from CT volumes, in: International conference on medical image computing and computer-assisted intervention, Springer, 2016, pp. 149–157. `doi:10.1007/978-3-319-46723-8_18`. 13, 77

[176] A. A. Shvets, A. Rakhlin, A. A. Kalinin, V. I. Iglovikov, Automatic instrument segmentation in robot-assisted surgery using deep learning, in: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, 2018, pp. 624–628. `doi:10.1109/ICMLA.2018.00100`. 13, 77

[177] D. D. Pham, G. Dovletov, S. Warwas, S. Landgraeber, M. Jager, J. Pauli, Deep learning with anatomical priors: Imitating enhanced autoencoders in latent space for improved pelvic bone segmentation in MRI, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, 2019, pp. 1166–1169. `doi:10.1109/ISBI.2019.8759221`. 13, 65, 77

[178] E. Vorontsov, A. Tang, C. Pal, S. Kadoury, Liver lesion segmentation informed by joint liver segmentation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 1332–1335. `doi:10.1109/ISBI.2018.8363817`. 13, 79

[179] X. Li, L. Yu, H. Chen, C.-W. Fu, L. Xing, P.-A. Heng, Transformation-consistent self-ensembling model for semisupervised medical image segmentation, IEEE Transactions on Neural Networks and Learning Systems 32 (2) (2021) 523–534. `doi:10.1109/TNNLS.2020.2995319`. 12

[180] A. E. Kavur, A. M. Selver, O. D. Gezer, M. Bar, N. S. Gezer, CHAOS - Combined (CT-MR) healthy abdominal organ segmentation challenge data evaluation, `https://chaos.grand-challenge.org/Evaluation/`, Accessed on: Jun. 2020 (2019). 12, 72

[181] A. Nayak, E. Baidya Kayal, M. Arya, J. Culli, S. Krishan, S. Agarwal, A. Mehndiratta, Computer-aided diagnosis of cirrhosis and hepatocellular carcinoma using multi-phase abdomen CT, International Journal of Computer Assisted Radiology and Surgery 14 (8) (2019) 1341–1352. `doi:10.1007/s11548-019-01991-5`. 13, 24, 47

[182] . Kayaalti, B. H. Aksebzeci, I. . Karahan, K. Deniz, M. ztrk, B. Yilmaz, S. Kara, M. H. Asyali, Staging of the liver fibrosis from ct images using texture features, in: 2012 7th International Symposium on Health Informatics and Bioinformatics, 2012, pp. 47–52. `doi:10.1109/HIBIT.2012.6209041`. 13, 14, 15, 20, 21, 25

[183] W.-L. Lee, Y.-C. Chen, K.-S. Hsieh, Ultrasonic liver tissues classification by fractal feature vector based on m-band wavelet transform, in: ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems (Cat. No.01CH37196), Vol. 2, 2001, pp. 1–4 vol. 2. `doi:10.1109/ISCAS.2001.920991`. 13, 21

[184] J. Virmani, V. Kumar, N. Kalra, N. Khandelwal, Prediction of cirrhosis based on singular value decomposition of gray level co-occurence marix and a neural network classifier, Proceedings - 4th International Conference on Developments in eSystems Engineering, DeSE 2011 (2011) 146–151`doi:10.1109/DeSE.2011.56`. 13, 14

[185] A. N. Karahaliou, I. S. Boniatis, S. G. Skiadopoulos, F. N. Sakellaropoulos, N. S. Arikidis, L. E. Likaki, G. S. Panayiotakis, L. I. Costaridou, Breast cancer diagnosis: Analyzing texture of tissue surrounding microcalcifications, IEEE Transactions on Information Technology in Biomedicine 12 (6) (2008) 731–738. `doi:10.1109/TITB.2008.920634`. 13, 15, 21

[186] J. Virmani, V. Kumar, N. Kalra, N. Khadelwal, A rapid approach for prediction of liver cirrhosis based on first order statistics, 2011 International Conference on Multimedia, Signal Processing and Communication Technologies, IMPACT 2011 (2011) 212–215`doi:10.1109/MSPCT.2011.6150477`. 13, 21

[187] R. M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Transactions on Systems, Man, and Cybernetics SMC-3 (6) (1973) 610–621. `doi:10.1109/TSMC.1973.4309314`. 13, 15, 21

[188] S. S.-D. Xu, C.-C. Chang, C.-T. Su, P. Q. Phu, Classification of liver diseases based on ultrasound image texture features, Applied Sciences 9 (2) (2019) 342. `doi:10.3390/app9020342`. 14, 15, 23

[189] C.-c. Lee, S.-h. Chen, Gabor wavelets and SVM classifier for liver diseases classification from CT images, in: 2006 IEEE International Conference on Systems, Man and Cybernetics, IEEE, 2006, pp. 548–552. `doi:10.1109/ICSMC.2006.384441`. 14, 15

[190] W. Chung-Ming, C. Yung-Chang, H. Kai-Sheng, Texture features for classification of ultrasonic liver images, IEEE Transactions on Medical Imaging 11 (2) (1992) 141–152. `doi:10.1109/42.141636`. 14, 22

[191] M. C. Yang, W. K. Moon, Y. C. F. Wang, M. S. Bae, C. S. Huang, J. H. Chen, R. F. Chang, Robust texture analysis using multi-resolution gray-scale invariant features for breast sonographic tumor diagnosis, IEEE Transactions on Medical Imaging 32 (12) (2013) 2262–2273. `doi:10.1109/TMI.2013.2279938`. 14, 21

[192] D. Mitrea, P. Mitrea, S. Nedevschi, R. Badea, M. Lupsor, M. Socaciu, A. Golea, C. Hagiu, L. Ciobanu, Abdominal tumor characterization and recognition using superior-order cooccurrence matrices, based on ultrasound images, Computational and Mathematical Methods in Medicine 2012 (2012). `doi:10.1155/2012/348135`. 14, 23

[193] P. A. Megha, G. Ram Mohana Reddy, Recent advances and future potential of computer aided diagnosis of liver cancer on computed tomography images, Communications in Computer and Information Science 157 CCIS (2011) 246–251. `doi:10.1007/978-3-642-22786-8_31`. 14, 16

[194] K. I. Laws, Rapid Texture Identification, in: T. F. Wiener (Ed.), Image processing for missile guidance, Vol. 0238, International Society for Optics and Photonics, SPIE, 1980, pp. 376 – 381. `doi:10.1117/12.959169`. 14, 15

[195] Y. N. Hwang, J. H. Lee, G. Y. Kim, Y. Y. Jiang, S. M. Kim, Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network, Bio-Medical Materials and Engineering 26 (2015) S1599–S1611. `doi:10.3233/BME-151459`. 14, 15, 16, 19, 23

[196] X. Li, B. Tian, N. Xu, Perceptual hash image classification algorithm based on sift feature, in: 2nd International Conference on Mechatronics and Information Technology (ICMIT 2017), 2017, pp. 411–415. `doi:10.25236/icmit.2017.76`. 14, 15

[197] A. Mojsilovic, M. Popovic, S. Markovic, M. Krstic, Characterization of visually similar diffuse diseases from B-scan liver images using nonseparable wavelet transform, IEEE Transactions on Medical Imaging 17 (4) (1998) 541–549. `doi:10.1109/42.730399`. 14

[198] Z. Xizhi, The application of wavelet transform in digital image processing, in: 2008 International Conference on MultiMedia and Information Technology, 2008, pp. 326–329. `doi:10.1109/MMIT.2008.134`. 14

[199] C.-C. Lee, S.-H. Chen, H.-M. Tsai, P.-C. Chung, Y.-C. Chiang, Discrimination of liver diseases from CT images based on Gabor filters, in: 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06), Vol. 2006, IEEE, 2006, pp. 203–206. `doi:10.1109/CBMS.2006.77`. 14, 24

[200] J. Virmani, V. Kumar, N. Kalra, N. Khandelwal, SVM-based characterization of liver ultrasound images using wavelet packet texture descriptors, Journal of Digital Imaging 26 (3) (2013) 530–543. `doi:10.1007/s10278-012-9537-8`. 14, 16, 23

[201] L.-H. Guo, D. Wang, Y.-Y. Qian, X. Zheng, C.-K. Zhao, X.-L. Li, X.-W. Bo, W.-W. Yue, Q. Zhang, J. Shi, et al., A two-stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images, Clinical hemorheology and microcirculation 69 (3) (2018) 343–354. `doi:10.3233/CH-170275`. 15

[202] F. Meng, J. Shi, B. Gong, Q. Zhang, L. Guo, D. Wang, H. Xu, B-Mode ultrasound based diagnosis of liver cancer with CEUS images as privileged information, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 3124–3127. `doi:10.1109/EMBC.2018.8513062`. 15

[203] J. S. Weszka, C. R. Dyer, A. Rosenfeld, A comparative study of texture measures for terrain classification, IEEE Transactions on Systems, Man and Cybernetics SMC-6 (4) (1976) 269–285. `doi:10.1109/TSMC.1976.5408777`. 15

[204] P. Bharti, D. Mittal, R. Ananthasivan, Preliminary study of chronic liver classification on ultrasound images using an ensemble model, Ultrasonic Imaging 40 (6) (2018) 357–379. `doi:10.1177/0161734618787447`. 15, 16, 18, 19, 23, 93

[205] K. Kalyan, B. Jakhia, R. D. Lele, M. Joshi, A. Chowdhary, Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images, Advances in Bioinformatics 2014 (2014). `doi:10.1155/2014/708279`. 15, 23

[206] D. Mittal, V. Kumar, S. C. Saxena, N. Khandelwal, N. Kalra, Neural network based focal liver lesion diagnosis using ultrasound images, Computerized Medical Imaging and Graphics 35 (4) (2011) 315–323. `doi:10.1016/j.compmedimag.2011.01.007`. 15, 20, 23

[207] A. Takemura, A. Shimizu, K. Hamamoto, Discrimination of Breast Tumors in Ultrasonic Images Using an Ensemble Classifier Based on the AdaBoost Algorithm With Feature Selection, IEEE Transactions on Medical Imaging 29 (3) (2010) 598–609. `doi:10.1109/TMI.2009.2022630`. 15, 22

[208] J. Virmani, V. Kumar, N. Kalra, N. Khandelwal, Prediction of cirrhosis from liver ultrasound B-mode images based on Laws' masks analysis, in: 2011 International Conference on Image Information Processing, IEEE, 2011, pp. 1–5. `doi:10.1109/ICIIP.2011.6108894`. 15, 18, 21, 93

[209] U. R. Acharya, U. Raghavendra, J. E. W. Koh, K. M. Meiburger, E. J. Ciaccio, Y. Hagiwara, F. Molinari, W. L. Leong, A. Vijayananthan, N. A. Yaakup, M. K. B. M. Fabell, C. H. Yeong, Automated detection and classification of liver fibrosis stages using contourlet transform and nonlinear features, Computer Methods and Programs in Biomedicine 166 (2018) 91–98. `doi:10.1016/j.cmpb.2018.10.006`. 14, 15, 19, 23

[210] S. S. Kumar, R. S. Moni, J. Rajeesh, An automatic computer-aided diagnosis system for liver tumours on computed tomography images, Computers and Electrical Engineering 39 (5) (2013) 1516–1526. `doi:10.1016/j.compeleceng.2013.02.008`. 15, 24

[211] X. Liu, J. Song, S. Wang, J. Zhao, Y. Chen, Learning to diagnose cirrhosis with liver capsule guided ultrasound image classification, Sensors 17 (12) (2017) 149. `doi:10.3390/s17010149`. 14, 15, 19, 21, 23

[212] D. Liang, L. Lin, H. Hu, Q. Zhang, Q. Chen, X. Han, Y.-W. Chen, et al., Residual convolutional neural networks with global and local pathways for classification of focal liver lesions, in: Pacific rim international conference on artificial intelligence, Springer, 2018, pp. 617–628. `doi:10.1007/978-3-319-97304-3_47`. 14, 19, 24

[213] T. M. Hassan, M. Elmogy, E. S. Sallam, Diagnosis of focal liver diseases based on deep learning technique for ultrasound images, Arabian Journal for Science and Engineering 42 (8) (2017) 3127–3140. `doi:10.1007/s13369-016-2387-9`. 14, 16

[214] Q. Y. Shan, L. D. Chen, L. Y. Zhou, Z. Wang, G. J. Liu, Y. Huang, W. Li, J. Y. Liu, X. Y. Xie, M. D. Lu, J. Liu, W. Wang, Focal lesions in fatty liver: If quantitative analysis facilitates the differentiation of atypical benign from malignant lesions, Scientific Reports 6 (26) (2016) 1–10. `doi:10.1038/srep18640`. 14, 23

[215] C.-C. Wu, W.-L. Lee, Y.-C. Chen, C.-H. Lai, K.-S. Hsieh, Ultrasonic liver tissue characterization by feature fusion, Expert Systems with Applications 39 (10) (2012) 9389–9397. `doi:10.1016/j.eswa.2012.02.128`. 14, 16, 23

[216] M. Pagano, K. Gauvreau, Principles of biostatistics, CRC Press, 2018. 15

[217] W. C. Navidi, Statistics for engineers and scientists, McGraw-Hill Higher Education New York, NY, USA, 2008. 15

[218] M. Mitchell, An introduction to genetic algorithms, MIT Press, 1998. 15

[219] A. Das, U. R. Acharya, S. S. Panda, S. Sabut, Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques, Cognitive Systems Research 54 (2019) 165–175. `doi:10.1016/j.cogsys.2018.12.009`. 15, 18, 20, 21, 24, 93

[220] S. Poonguzhali, B. Deepalakshmi, G. Ravindran, Optimal feature selection and automatic classification of abnormal masses in ultrasound liver images, in: 2007 International Conference on Signal Processing, Communications and Networking, IEEE, 2007, pp. 503–506. `doi:10.1109/ICSCN.2007.350789`. 16

[221] M. Prasad, A. Sowmya, I. Koch, Efficient feature selection based on Independent Component Analysis, Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference, ISSNIP '04 (2004) 427–432`doi:10.1109/issnip.2004.1417499`. 16

[222] M. Wu, L. Li, J. Wang, Y. Zhang, Q. Guo, X. Li, X. Zhang, Contrast-enhanced US for characterization of focal liver lesions: A comprehensive meta-analysis, European Radiology 28 (5) (2018) 2077–2088. `doi:10.1007/s00330-017-5152-x`. 16

[223] J. Santos, J. S. Silva, A. A. Santos, P. Belo-Soares, Detection of pathologic liver using ultrasound images, Biomedical Signal Processing and Control 14 (2014) 248–255. `doi:10.1016/j.bspc.2014.08.009`. 16, 23

[224] R. A. Khan, N. Naseer, M. J. Khan, Drowsiness detection during a driving task using fNIRS, in: Neuroergonomics, Elsevier, 2019, pp. 79–85. `doi:10.1016/B978-0-12-811926-6.00013-0`. 16

[225] R. A. Khan, N. Naseer, N. K. Qureshi, F. M. Noori, H. Nazeer, M. U. Khan, fNIRS-based neurorobotic interface for gait rehabilitation, Journal of NeuroEngineering and Rehabilitation 15 (1) (2018) 7. `doi:10.1186/s12984-018-0346-2`. 16

[226] R. A. Khan, N. Naseer, S. Saleem, N. K. Qureshi, F. M. Noori, M. J. Khan, Cortical tasks-based optimal filter selection: An fNIRS study, Journal of Healthcare Engineering 2020 (2020) 1–15. `doi:10.1155/2020/9152369`. 16

[227] R. Khan, N. Naseer, H. Nazeer, M. N. Khan, Control of a prosthetic leg based on walking intentions for gait rehabilitation: an fnirs study, Frontiers in Human Neuroscience 12 (2018). `doi:10.3389/conf.fnhum.2018.227.00144`. 16

[228] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, Communications of the ACM 60 (6) (2017) 84–90. `doi:10.1145/3065386`. 17, 48, 56, 88, 93

[229] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1–9. `doi:10.1109/CVPR.2015.7298594`. 17

[230] A. Ben-Cohen, E. Klang, S. P. Raskin, S. Soffer, S. Ben-Haim, E. Konen, M. M. Amitai, H. Greenspan, Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection, Engineering Applications of Artificial Intelligence 78 (2018) (2019) 186–194. `doi:10.1016/j.engappai.2018.11.013`. 17, 87

[231] A. Ben-Cohen, E. Klang, A. Kerpel, E. Konen, M. M. Amitai, H. Greenspan, Fully convolutional network and sparsity-based dictionary learning for liver lesion detection in CT examinations, Neurocomputing 275 (2018) 1585–1594. `doi:10.1016/j.neucom.2017.10.001`. 17, 24, 84

[232] D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis, Annual Review of Biomedical Engineering 19 (1) (2017) 221–248. `doi:10.1146/annurev-bioeng-071516-044442`. 17

[233] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (8) (2013) 1798–1828. `doi:10.1109/TPAMI.2013.50`. 17

[234] H. C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Transactions on Medical Imaging 35 (5) (2016) 1285–1298. `doi:10.1109/TMI.2016.2528162`. 17, 21

[235] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, G.-Z. Yang, Deep learning for health informatics, IEEE Journal of Biomedical and Health Informatics 21 (1) (2017) 4–21. `doi:10.1109/JBHI.2016.2636665`. 17, 24

[236] W. Wang, Y. Iwamoto, X. Han, Y.-W. Chen, Q. Chen, D. Liang, L. Lin, H. Hu, Q. Zhang, Classification of focal liver lesions using deep learning with fine-tuning, in: Proceedings of the 2018 International Conference on Digital Medicine and Image Processing - DMIP '18, ACM Press, New York, New York, USA, 2018, pp. 56–60. `doi:10.1145/3299852.3299860`. 20, 24, 84, 86, 93

[237] G. Sethi, B. S. Saini, Computer aided diagnosis system for abdomen diseases in computed tomography images, Biocybernetics and Biomedical Engineering 36 (1) (2016) 42–55. `doi:10.1016/j.bbe.2015.10.008`. 20

[238] C.-C. C. Chang, H.-H. H. Chen, Y.-C. C. Chang, M.-Y. Y. Yang, C.-M. M. Lo, W.-C. C. Ko, Y.-F. F. Lee, K.-L. L. Liu, R.-F. F. Chang, Computer-aided diagnosis of liver tumors on computed tomography images, Computer Methods and Programs in Biomedicine 145 (2017) 45–51. `doi:10.1016/j.cmpb.2017.04.008`. 20, 21, 24

[239] K. Aggarwal, M. S. Bhamrah, H. S. Ryait, Detection of cirrhosis through ultrasound imaging, Advances in Intelligent Systems and Computing 1073 (2020) 245–258. `doi:10.1007/978-3-030-33582-3_24`. 20, 23

[240] Z. Yao, J. Li, Z. Guan, Y. Ye, Y. Chen, Liver disease screening based on densely connected deep neural networks, Neural Networks 123 (2020) 299–304. `doi:10.1016/j.neunet.2019.11.005`. 20, 21, 25, 47, 55, 86, 90

[241] A. Ben-Cohen, E. Klang, I. Diamant, N. Rozendorn, S. P. Raskin, E. Konen, M. M. Amitai, H. Greenspan, CT image-based decision support system for categorization of liver metastases into primary cancer sites, Academic Radiology 24 (12) (2017) 1501–1509. `doi:10.1016/j.acra.2017.06.008`. 20, 24

[242] K. Yasaka, H. Akai, O. Abe, S. Kiryu, Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: A preliminary study, Radiology 286 (3) (2018) 887–896. `doi:10.1148/radiol.2017170706`. 20, 24, 86, 93

[243] A. I. Awad, M. Hassaballah, Image feature detectors and descriptors, Studies in Computational Intelligence 630 (2016). `doi:10.1007/978-3-319-28854-3`. 21

[244] W. Chen, M. L. Giger, U. Bick, G. M. Newstead, Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI, Medical Physics 33 (8) (2006) 2878–2887. `doi:10.1118/1.2210568`. 21

[245] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, Neurocomputing 321 (2018) 321–331. `doi:10.1016/j.neucom.2018.09.013`. 21, 22, 24, 60, 67, 84, 87

[246] B. Krawczyk, M. Galar, L. Jeleń, F. Herrera, Evolutionary undersampling boosting for imbalanced classification of breast cancer malignancy, Applied Soft Computing Journal 38 (2016) 714–726. `doi:10.1016/j.asoc.2015.08.060`. 21

[247] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: Full training or fine tuning?, IEEE Transactions on Medical Imaging 35 (5) (2016) 1299–1312. `doi:10.1109/TMI.2016.2535302`. 21

[248] A. Kumar, J. Kim, D. Lyndon, M. Fulham, D. Feng, An ensemble of fine-tuned convolutional neural networks for medical image classification, IEEE Journal of Biomedical and Health Informatics 21 (1) (2017) 31–40. `doi:10.1109/JBHI.2016.2635663`. 21, 90, 91

[249] X. Han, MR-based synthetic CT generation using a deep convolutional neural network method, Medical Physics 44 (4) (2017) 1408–1419. `doi:10.1002/mp.12155`. 22

[250] L. Xiang, Y. Qiao, D. Nie, L. An, W. Lin, Q. Wang, D. Shen, Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI, Neurocomputing 267 (2017) 406–416. `doi:10.1016/j.neucom.2017.06.048`. 22

[251] Y. Zhang, S. Y. Ma, X. Zhang, L. Li, W. H. Ip, K. L. Yung, EDGAN: motion deblurring algorithm based on enhanced generative adversarial networks, The Journal of Supercomputing 76 (11) (2020) 8922–8937. `doi:10.1007/s11227-020-03189-y`. 22, 32, 39, 40, 41

[252] I. Diamant, A. Hoogi, C. F. Beaulieu, M. Safdari, E. Klang, M. Amitai, H. Greenspan, D. L. Rubin, Improved patch-based automated liver lesion classification by separate analysis of the interior and boundary regions, IEEE Journal of Biomedical and Health Informatics 20 (6) (2016) 1585–1594. `doi:10.1109/JBHI.2015.2478255`. 22

[253] M. H. Horng, An ultrasonic image evaluation system for assessing the severity of chronic liver disease, Computerized Medical Imaging and Graphics 31 (7) (2007) 485–491. `doi:10.1016/j.compmedimag.2007.05.001`. 23

[254] C. T. Streba, M. Ionescu, D. I. Gheonea, L. Sandulescu, T. Ciurea, A. Saftoiu, C. C. Vere, I. Rogoveanu, Contrast-enhanced ultrasonography parameters in neural network diagnosis of liver tumors, World Journal of Gastroenterology 18 (32) (2012) 4427–4434. `doi:10.3748/wjg.v18.i32.4427`. 23

[255] R. Ribeiro, R. T. Marinho, J. S. Suri, J. M. Sanches, Classification and staging of chronic liver disease based on ultrasound, laboratorial, and clinical data, in: Ultrasound Imaging, Springer US, Boston, MA, 2012, pp. 255–282. `doi:10.1007/978-1-4614-1180-2_11`. 23

[256] W. L. Lee, An ensemble-based data fusion approach for characterizing ultrasonic liver tissue, Applied Soft Computing Journal 13 (8) (2013) 3683–3692. `doi:10.1016/j.asoc.2013.03.009`. 23, 48, 49

[257] J. Virmani, V. Kumar, N. Kalra, N. Khandelwal, Characterization of primary and secondary malignant liver lesions from B-mode ultrasound, Journal of Digital Imaging 26 (6) (2013) 1058–1070. `doi:10.1007/s10278-013-9578-7`. 23

[258] K. Wu, X. Chen, M. Ding, Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound, Optik 125 (15) (2014) 4057–4063. `doi:10.1016/j.ijleo.2014.01.114`. 23, 93

[259] I. Gatos, S. Tsantis, S. Spiliopoulos, A. Skouroliakou, I. Theotokas, P. Zoumpoulis, J. D. Hazle, G. C. Kagadis, A new automated quantification algorithm for the detection and evaluation of focal liver lesions with contrast-enhanced ultrasound, Medical Physics 42 (7) (2015) 3948–3959. `doi:10.1118/1.4921753`. 23

[260] L. Guo, D. Wang, H. Xu, Y. Qian, C. Wang, X. Zheng, Q. Zhang, J. Shi, CEUS-based classification of liver tumors with deep canonical correlation analysis and multi-kernel learning, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2017, pp. 1748–1751. `doi:10.1109/EMBC.2017.8037181`. 23, 49, 56, 57, 58, 59, 60, 61

[261] F. Pan, Q. Huang, X. Li, Classification of liver tumors with CEUS based on 3D-CNN, 2019 4th IEEE International Conference on Advanced Robotics and Mechatronics, ICARM 2019 (2019) 845–849`doi:10.1109/ICARM.2019.8834190`. 23

[262] R. T. Ribeiro, R. T. Marinho, J. M. Sanches, Classification and staging of chronic liver disease from multimodal data, IEEE Transactions on Biomedical Engineering 60 (5) (2013) 1336–1344. `doi:10.1109/TBME.2012.2235438`. 23, 93

[263] M. Singh, S. Singh, S. Gupta, An information fusion based method for liver classification using texture analysis of ultrasound images, Information Fusion 19 (2014) 91–96. `doi:10.1016/j.inffus.2013.05.007`. 23

[264] U. R. Acharya, H. Fujita, S. Bhat, U. Raghavendra, A. Gudigar, F. Molinari, A. Vijayananthan, K. Hoong Ng, Decision support system for fatty liver disease using GIST descriptors extracted from ultrasound images, Information Fusion 29 (2016) 32–39. `doi:10.1016/j.inffus.2015.09.006`. 23

[265] U. R. Acharya, J. E. W. Koh, Y. Hagiwara, J. H. Tan, A. Gertych, A. Vijayananthan, N. A. Yaakup, B. J. J. Abdullah, M. K. Bin Mohd Fabell, C. H. Yeong, Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features, Computers in Biology and Medicine 94 (2018) 11–18. `doi:10.1016/j.compbiomed.2017.12.024`. 23

[266] L. Meng, C. Wen, G. Li, Support vector machine based liver cancer early detection using magnetic resonance images, in: 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), 2014, pp. 861–864. `doi:10.1109/ICARCV.2014.7064417`. 24

[267] K. Yasaka, H. Akai, A. Kunimatsu, O. Abe, S. Kiryu, Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid–enhanced hepatobiliary phase mr images, Radiology 287 (1) (2018) 146–155. `doi:10.1148/radiol.2017171928`. 24, 93

[268] A. Adcock, D. Rubin, G. Carlsson, Classification of hepatic lesions using the matching metric, Computer Vision and Image Understanding 121 (2014) 36–42. `doi:10.1016/j.cviu.2013.10.014`. 24

[269] R. Vivanti, A. Szeskin, N. Lev-Cohain, J. Sosna, L. Joskowicz, Automatic detection of new tumors and tumor burden evaluation in longitudinal liver CT scan studies, International Journal of Computer Assisted Radiology and Surgery 12 (11) (2017) 1945–1957. `doi:10.1007/s11548-017-1660-z`. 24, 86, 93

[270] A. Krishna, D. Edwin, S. Hariharan, Classification of liver tumor using sfta based nave bayes classifier and support vector machine, in: 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2017, pp. 1066–1070. `doi:10.1109/ICICICT1.2017.8342716`. 24

[271] P. Sreeja, S. Hariharan, Image analysis for the detection and diagnosis of hepatocellular carcinoma from abdominal CT images, Lecture Notes in Networks and Systems 19 (2018) 107–117. `doi:10.1007/978-981-10-5523-2_11`. 24

[272] W.-J. Kuo, Computer-aided diagnosis for feature selection and classification of liver tumors in computed tomography images, in: 2018 IEEE International Conference on Applied System Invention (ICASI), IEEE, 2018, pp. 1207–1210. `doi:10.1109/ICASI.2018.8394505`. 24, 47

[273] A. Midya, J. Chakraborty, L. M. Pak, J. Zheng, W. R. Jarnagin, R. K. Do, A. L. Simpson, Deep convolutional neural network for the classification of hepatocellular carcinoma and intrahepatic cholangiocarcinoma, in: Medical Imaging 2018: Computer-Aided Diagnosis, Vol. 10575, International Society for Optics and Photonics, 2018, p. 1057528. `doi:10.1117/12.2293683`. 24, 86

[274] A. M. Anter, A. E. Hassenian, Normalized multiple features fusion based on PCA and multiple classifiers voting in CT liver tumor recognition, in: Advances in Soft Computing and Machine Learning in Image Processing, Springer, 2018, pp. 113–129. `doi:10.1007/978-3-319-63754-9_6`. 24

[275] A. Doğantekin, F. Özyurt, E. Avcı, M. Koc, A novel approach for liver image classification: Ph-c-elm, Measurement 137 (2019) 332–338. `doi:10.1016/j.measurement.2019.01.060`. 24

[276] X. Chen, L. Lin, D. Liang, H. Hu, Q. Zhang, Y. Iwamoto, X.-H. Han, Y.-W. Chen, R. Tong, J. Wu, A dual-attention dilated residual network for liver lesion classification and localization on ct images, in: 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 235–239. `doi:10.1109/ICIP.2019.8803009`. 24, 86

[277] F. P. Romero, A. Diler, G. Bisson-Gregoire, S. Turcotte, R. Lapointe, F. Vandenbroucke-Menu, A. Tang, S. Kadoury, End-to-end discriminative deep network for liver lesion classification, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 1243–1246. `doi:10.1109/ISBI.2019.8759257`. 25, 86

[278] J. Muthuswamy, Extraction and classification of liver abnormality based on neutrosophic and SVM classifier, in: Progress in Advanced Computing and Intelligent Engineering, Springer, 2019, pp. 269–279. `doi:10.1007/978-981-13-1708-8_25`. 25

[279] K. Mala, V. Sadasivam, Wavelet based texture analysis of liver tumor from computed tomography images for characterization using linear vector quantization neural network, Proceedings - 2006 14th International Conference on Advanced Computing and Communications, ADCOM 2006 (2006) 267–270`doi:10.1109/ADCOM.2006.4289897`. 25

[280] Y.-W. Chen, J. Luo, C. Dong, X. Han, T. Tateyama, A. Furukawa, S. Kanasaki, Computer-aided diagnosis and quantification of cirrhotic livers based on morphological analysis and machine learning, Computational and Mathematical Methods in Medicine 2013 (2013) 1–8. `doi:10.1155/2013/264809`. 25

[281] O. Demirkaya, Image Processing with MATLAB, CRC Press, 2008. `doi:10.1201/9781420008937`. 27, 29, 47

[282] H. G. Bock, T. Carraro, W. Jäger, S. Körkel, R. Rannacher, J. P. Schlöder, Model based parameter estimation: theory and applications, Vol. 4, Springer Science & Business Media, 2013. `doi:10.1007/978-3-642-30367-8`. 27, 29

[283] N. W. Churchill, R. Spring, B. Afshin-Pour, F. Dong, S. C. Strother, An automated, adaptive framework for optimizing preprocessing pipelines in task-based functional MRI, PLoS ONE 10 (7) (2015) 1–25. `doi:10.1371/journal.pone.0131520`. 27, 29

[284] S. Sun, H. Zhao, B. Li, M. Hao, J. Lv, Kernel estimation for robust motion deblurring of noisy and blurry images, Journal of Electronic Imaging 25 (3) (2016) 033019. `doi:10.1117/1.JEI.25.3.033019`. 27, 29, 47

[285] F. Strappini, E. Gilboa, S. Pitzalis, K. Kay, M. McAvoy, A. Nehorai, A. Z. Snyder, Adaptive smoothing based on Gaussian processes regression increases the sensitivity and specificity of fMRI data, Human Brain Mapping 38 (3) (2017) 1438–1459. `doi:10.1002/hbm.23464`. 27, 29

[286] D. P. Kingma, M. Welling, Auto-encoding variational bayes, arXiv (2013). `arXiv:1312.6114v10`. 27, 48

[287] M. Rosca, B. Lakshminarayanan, D. Warde-Farley, S. Mohamed, Variational approaches for auto-encoding generative adversarial networks, arXiv (2017). `arXiv:1706.04987v2`. 27, 48

[288] S. A. Bigdeli, M. Zwicker, Image restoration using autoencoding priors, arXiv (2017). `arXiv:1703.09964v1`. 27

[289] C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda, S. Muramatsu, Y. Furukawa, G. Mauri, H. Nakayama, Gan-based synthetic brain mr image generation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 734–738. `doi:10.1109/ISBI.2018.8363678`. 27

[290] X. Xu, D. Sun, J. Pan, Y. Zhang, H. Pfister, M.-H. Yang, Learning to super-resolve blurry face and text images, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 251–260. `doi:10.1109/ICCV.2017.36`. 27, 48

[291] J. Gu, Y. Shen, B. Zhou, Image processing using multi-code gan prior, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 3012–3021. `doi:10.1109/CVPR42600.2020.00308`. 27

[292] J. Wu, X. Di, Integrating neural networks into the blind deblurring framework to compete with the end-to-end learning-based methods, IEEE Transactions on Image Processing 29 (2020) 6841–6851. `doi:10.1109/TIP.2020.2994413`. 27, 32

[293] A. Zhen, R. L. Stevenson, et al., Gan based image deblurring using dark channel prior, Electronic Imaging 2019 (13) (2019) 136–1. 27

[294] M. Chen, Y. Chang, S. Cao, L. Yan, Learning blind denoising network for noisy image deblurring, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 2533–2537. `doi:10.1109/ICASSP40776.2020.9053539`. 27, 50

[295] A. Li, J. Li, Q. Lin, C. Ma, B. Yan, Deep image quality assessment driven single image deblurring, in: IEEE International Conference on Multimedia and Expo (ICME), 2020, pp. 1–6. `doi:10.1109/ICME46284.2020.9102899`. 27

[296] K.-H. Liu, C.-H. Yeh, J.-W. Chung, C.-Y. Chang, A motion deblur method based on multi-scale high frequency residual image learning, IEEE Access 8 (2020) 66025–66036. `doi:10.1109/ACCESS.2020.2985220`. 27, 28, 31, 33, 35, 44, 48

[297] B. Wei, Y. Yuan, Q. Wang, MSPNET: Multi-supervised parallel network for crowd counting, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 2418–2422. `doi:10.1109/ICASSP40776.2020.9054479`. 27, 31

[298] F. Yang, L. Xiao, J. Yang, Video deblurring via 3D CNN and Fourier accumulation learning, in: IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings (ICASSP), 2020, pp. 2443–2447. `doi:10.1109/ICASSP40776.2020.9054514`. 27, 50

[299] D. Ren, K. Zhang, Q. Wang, Q. Hu, W. Zuo, Neural blind deconvolution using deep priors, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 3341–3350. `doi:10.1109/CVPR42600.2020.00340`. 27

[300] X. Zhang, Y. Lv, Y. Li, Y. Liu, P. Luo, A modified image processing method for deblurring based on GAN networks, in: 5th International Conference on Big Data and Information Analytics (BigDIA), 2019, pp. 29–34. `doi:10.1109/BigDIA.2019.8802800`. 28, 32, 33, 36

[301] L. Yuan, J. Sun, L. Quan, H.-Y. Shum, Progressive inter-scale and intra-scale non-blind image deconvolution, ACM Transactions on Graphics 27 (3) (2008) 1. `doi:10.1145/1360612.1360673`. 28, 33, 35

[302] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image Quality Assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing 13 (4) (2004) 600–612. `doi:10.1109/TIP.2003.819861`. 28, 38, 48, 55

[303] N. Venkatanath, D. Praneeth, B. H. Maruthi Chandrasekhar, S. S. Channappayya, S. S. Medasani, Blind image quality evaluation using perception based features, in: 2015 21st National Conference on Communications, NCC 2015, IEEE, 2015, pp. 1–6. `doi:10.1109/NCC.2015.7084843`. 28, 39

[304] L. Liu, B. Liu, H. Huang, A. C. Bovik, No-reference image quality assessment based on spatial and spectral entropies, Signal Processing: Image Communication 29 (8) (2014) 856–863. `doi:10.1016/j.image.2014.06.006`. 28, 39

[305] M. M. Ghazi, H. Erdogan, Image noise level estimation based on higher-order statistics, Multimedia Tools and Applications 76 (2) (2017) 2379–2397. `doi:10.1007/s11042-015-3169-1`. 29, 47

[306] S. Joshi, R. K. Kulkarni, Medical Image Enhancement Using Hybrid Techniques for Accurate Anomaly Detection And Malignancy Predication, in: X.-S. Yang, S. Sherratt, N. Dey, A. Joshi (Eds.), Advances in Intelligent Systems and Computing, Vol. 797, Springer Singapore, Singapore, 2019, pp. 951–961. `doi:10.1007/978-981-13-1165-9_87`. 29, 47, 55, 61

[307] H. Liu, S. Liu, D. Guo, Y. Zheng, P. Tang, G. Dan, Original intensity preserved inhomogeneity correction and segmentation for liver magnetic resonance imaging, Biomedical Signal Processing and Control 47 (2019) 231–239. `doi:10.1016/j.bspc.2018.08.005`. 29, 47, 49, 61

[308] S. Sahu, M. K. Lenka, P. K. Sa, Blind deblurring using deep learning: A survey, arXiv (2019). `arXiv:1907.10128v1`. 29

[309] Z. Hu, J.-B. Huang, M.-H. Yang, Single image deblurring with adaptive dictionary learning, in: IEEE International Conference on Image Processing, 2010, pp. 1169–1172. `doi:10.1109/ICIP.2010.5651892`. 29, 48

[310] D. Perrone, P. Favaro, Total variation blind deconvolution: The devil is in the details, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2909–2916. `doi:10.1109/CVPR.2014.372`. 29, 41, 43, 136

[311] S. Cho, S. Lee, Fast motion deblurring, in: ACM SIGGRAPH Asia 2009 Papers, 2009, pp. 1–8. `doi:10.1145/1618452.1618491`. 29, 30, 41, 44

[312] S. Harmeling, H. Michael, B. Schölkopf, Space-variant single-image blind deconvolution for removing camera shake, Advances in Neural Information Processing Systems 23 (2010) 829–837. 29, 41

[313] L. Xu, J. Jia, Two-phase kernel estimation for robust motion deblurring, in: European Conference on Computer Vision, 2010, pp. 157–170. `doi:10.1007/978-3-642-15549-9_12`. 29, 41, 42, 43, 135, 136

[314] A. Levin, Y. Weiss, F. Durand, W. T. Freeman, Efficient marginal likelihood optimization in blind deconvolution, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 2657–2664. `doi:10.1109/CVPR.2011.5995308`. 29, 41, 43, 136

[315] D. Krishnan, T. Tay, R. Fergus, Blind deconvolution using a normalized sparsity measure, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 233–240. `doi:10.1109/CVPR.2011.5995521`. 29, 41, 42, 43, 135, 136

[316] H. Zhang, D. Wipf, Y. Zhang, Multi-image blind deblurring using a coupled adaptive sparse prior, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1051–1058. `doi:10.1109/CVPR.2013.140`. 29, 41, 43, 136

[317] J. Pan, Z. Hu, Z. Su, M.-H. Yang, Deblurring text images via l0-regularized intensity and gradient prior, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2901–2908. `doi:10.1109/CVPR.2014.371`. 29, 41, 43, 136

[318] O. Whyte, J. Sivic, A. Zisserman, Deblurring Shaken and Partially Saturated Images, International Journal of Computer Vision 110 (2) (2014) 185–201. `doi:10.1007/s11263-014-0727-3`. 29, 42, 135

[319] Q. Shan, J. Jia, A. Agarwala, High-quality motion deblurring from a single image, Acm Transactions on Graphics (tog) 27 (3) (2008) 1–10. `doi:10.1145/1360612.1360672`. 29, 41, 42, 135

[320] L. Zhong, S. Cho, D. Metaxas, S. Paris, J. Wang, Handling noise in single image deblurring using directional filters, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 612–619. `doi:10.1109/CVPR.2013.85`. 30, 41, 43, 136

[321] T. Hyun Kim, K. Mu Lee, Segmentation-free dynamic scene deblurring, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2766–2773. `doi:10.1109/CVPR.2014.348`. 30, 39, 40, 41

[322] H. Liu, J. Gu, M. Q.-H. Meng, W.-S. Lu, Fast weighted total variation regularization algorithm for blur identification and image restoration, IEEE Access 4 (2016) 6792–6801. `doi:10.1109/ACCESS.2016.2516949`. 30

[323] P.-H. Huang, Y.-M. Lin, S.-H. Lai, Image deblurring with blur kernel estimation from a reference image patch, in: International Conference on Pattern Recognition, 2008, pp. 1–4. `doi:10.1109/ICPR.2008.4761416`. 30

[324] W. H. Richardson, Bayesian-based iterative method of image restoration, JoSA 62 (1) (1972) 55–59. `doi:10.1364/JOSA.62.000055`. 30

[325] O. Whyte, J. Sivic, A. Zisserman, J. Ponce, Non-uniform Deblurring for Shaken Images, International Journal of Computer Vision 98 (2) (2012) 168–186. `doi:10.1007/s11263-011-0502-7`. 30, 40, 41, 43, 44, 134, 136

[326] T. Michaeli, M. Irani, Blind deblurring using internal patch recurrence, in: European Conference on Computer Vision, 2014, pp. 783–798. `doi:10.1007/978-3-319-10578-9_51`. 30, 41, 43, 136

[327] J. Pan, D. Sun, H. Pfister, M.-H. Yang, Blind image deblurring using dark channel prior, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1628–1636. `doi:10.1109/CVPR.2016.180`. 30, 41, 44, 134

[328] L. Xu, J. S. Ren, C. Liu, J. Jia, Deep convolutional neural network for image deconvolution, in: Advances in Neural Information Processing Systems, Vol. 27, 2014, pp. 1790–1798. 30, 43

[329] C. J. Schuler, M. Hirsch, S. Harmeling, B. Schölkopf, Learning to deblur, IEEE Transactions on Pattern Analysis and Machine Intelligence 38 (7) (2015) 1439–1451. `doi:10.1109/TPAMI.2015.2481418`. 30

[330] J. Pan, W. Ren, Z. Hu, M.-H. Yang, Learning to deblur images with exemplars, IEEE Transactions on Pattern analysis and Machine Intelligence 41 (6) (2018) 1412–1425. `doi:10.1109/TPAMI.2018.2832125`. 31

[331] B. Lu, J.-C. Chen, R. Chellappa, UID-GAN: Unsupervised image deblurring via disentangled representations, IEEE Transactions on Biometrics, Behavior, and Identity Science 2 (1) (2019) 26–39. `doi:10.1109/TBIOM.2019.2959133`. 31, 44

[332] Z. Shen, W.-s. Lai, T. Xu, J. Kautz, M.-h. Yang, Deep semantic face deblurring, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8260–8269. `doi:10.1109/CVPR.2018.00862`. 31, 43

[333] R. Yasarla, F. Perazzi, V. M. Patel, Deblurring face images using uncertainty guided multi-stream semantic networks, IEEE Transactions on Image Processing 29 (2020) 6251–6263. `doi:10.1109/TIP.2020.2990354`. 31, 43

[334] S. Vasu, V. R. Maligireddy, A. Rajagopalan, Non-blind deblurring: Handling kernel uncertainty with CNNs, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 3272–3281. `doi:10.1109/CVPR.2018.00345`. 31

[335] T. M. Nimisha, A. Kumar Singh, A. N. Rajagopalan, Blur-invariant deep learning for blind-deblurring, in: IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4752–4760. `doi:10.1109/ICCV.2017.509`. 31, 32

[336] H. Wu, S. Zheng, J. Zhang, K. Huang, Gp-gan: Towards realistic high-resolution image blending, in: Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 2487–2495. `doi:10.1145/3343031.3350944`. 31

[337] Y. Zhang, Y. Xiang, L. Bai, Generative adversarial network for deblurring of remote sensing image, in: 26th International Conference on Geoinformatics, 2018, pp. 1–4. `doi:10.1109/GEOINFORMATICS.2018.8557110`. 31

[338] J. Liu, W. Sun, M. Li, Recurrent conditional generative adversarial network for image deblurring, IEEE Access 7 (2018) 6186–6193. `doi:10.1109/ACCESS.2018.2888885`. 31

[339] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv (2014). `arXiv:1409.1556`. 32, 36, 54, 93

[340] S. Zheng, Z. Zhu, J. Cheng, Y. Guo, Y. Zhao, Edge heuristic GAN for non-uniform blind deblurring, IEEE Signal Processing Letters 26 (10) (2019) 1546–1550. `doi:10.1109/LSP.2019.2939752`. 32

[341] C. Wu, H. Du, Q. Wu, S. Zhang, Image text deblurring method based on generative adversarial network, Electronics 9 (2) (2020) 220. `doi:10.3390/electronics9020220`. 32

[342] O. Kupyn, T. Martyniuk, J. Wu, Z. Wang, Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better, in: IEEE International Conference on Computer Vision (ICCV), 2019, pp. 8878–8887. `doi:10.1109/ICCV.2019.00897`. 32

[343] J. Cai, W. Zuo, L. Zhang, Dark and bright channel prior embedded network for dynamic scene deblurring, IEEE Transactions on Image Processing 29 (2020) 6885–6897. `doi:10.1109/TIP.2020.2995048`. 32, 40, 41, 44, 134

[344] J. Zhang, J. Pan, J. Ren, Y. Song, L. Bao, R. W. Lau, M.-H. Yang, Dynamic scene deblurring using spatially variant recurrent neural networks, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 2521–2529. `doi:10.1109/CVPR.2018.00267`. 32, 40, 41, 43, 44, 134

[345] H. Zhang, Y. Dai, H. Li, P. Koniusz, Deep stacked hierarchical multi-patch network for image deblurring, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 5978–5986. `doi:10.1109/CVPR.2019.00613`. 32, 41, 42, 43, 44, 49, 58, 59, 60, 61

[346] K. Zhang, W. Luo, Y. Zhong, L. Ma, W. Liu, H. Li, Adversarial spatio-temporal learning for video deblurring, IEEE Transactions on Image Processing 28 (1) (2019) 291–301. `doi:10.1109/TIP.2018.2867733`. 32, 51

[347] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv (2014). `arXiv:1411.1784v1`. 32

[348] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein generative adversarial networks, in: International Conference on Machine Learning, 2017, pp. 214–223. 33, 36, 54

[349] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017-Janua (2017) 2261–2269. `doi:10.1109/CVPR.2017.243`. 34, 66, 80

[350] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu, Residual Dense Network for Image Super-Resolution, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2018) 2472–2481`arXiv:1812.10477`, `doi:10.1109/CVPR.2018.00262`. 34, 40

[351] C. Li, M. Wand, Precomputed real-time texture synthesis with markovian generative adversarial networks, in: European Conference on Computer Vision, 2016, pp. 702–716. `doi:10.1007/978-3-319-46487-9_43`. 35

[352] V. A. Knyaz, V. V. Kniaz, F. Remondino, Image-to-voxel model translation with conditional adversarial networks, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0. `doi:10.1007/978-3-030-11009-3_37`. 35

[353] E. Ong, W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, F. Moschetti, A no-reference quality metric for measuring image blur, Proceedings - 7th International Symposium on Signal Processing and Its Applications, ISSPA 2003 1 (4) (2003) 469–472. `doi:10.1109/ISSPA.2003.1224741`. 39

[354] T. J. Chen, K. S. Chuang, J. H. Chang, Y. H. Shiao, C. C. Chuang, A blurring index for medical images, Journal of Digital Imaging 19 (2) (2006) 118–125. `doi:10.1007/s10278-005-8736-y`. 39

[355] M.-J. Chen, A. C. Bovik, No-reference image blur assessment using multiscale gradient, EURASIP Journal on image and video processing 2011 (1) (2011) 1–11. `doi:10.1109/QOMEX.2009.5246973`. 39

[356] L. Liang, J. Chen, S. Ma, D. Zhao, W. Gao, A no-reference perceptual blur metric using histogram of gradient profile sharpness, Proceedings - International Conference on Image Processing, ICIP 60833013 (60833013) (2009) 4369–4372. `doi:10.1109/ICIP.2009.5413545`. 39

[357] A. Liu, W. Lin, M. Narwaria, Image quality assessment based on gradient similarity, IEEE Transactions on Image Processing 21 (4) (2012) 1500–1512. `doi:10.1109/TIP.2011.2175935`. 39

[358] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, A. C. Kot, No-reference image blur assessment based on discrete orthogonal moments, IEEE Transactions on Cybernetics 46 (1) (2016) 39–50. `doi:10.1109/TCYB.2015.2392129`. 39

[359] M. E. Osadebey, M. Pedersen, D. L. Arnold, K. E. Wendel-Mitoraj, Blind blur assessment of MRI images using parallel multiscale difference of Gaussian filters, BioMedical Engineering Online 17 (1) (2018) 1–22. `doi:10.1186/s12938-018-0514-4`. 39

[360] Y. Zhan, R. Zhang, No-Reference Image Sharpness Assessment Based on Maximum Gradient and Variability of Gradients, IEEE Transactions on Multimedia 20 (7) (2018) 1796–1808. `doi:10.1109/TMM.2017.2780770`. 39

[361] S. Vasu, P. D. Vergara, Z. Zhong, A. Arora, Awesome deblurring, `https://github.com/subeeshvasu/Awesome-Deblurring/`, Accessed on: Jun. 2021 (2019). 39, 134

[362] L. Zhou, Z. Liu, Blind deblurring based on a single luminance channel and l1-norm, IEEE Access 9 (2021) 126717–126727. `doi:10.1109/ACCESS.2021.3112491`. 41

[363] H. Feng, J. Guo, H. Xu, S. S. Ge, SharpGAN: Dynamic scene deblurring method for smart ship based on receptive field block and generative adversarial networks, Sensors 21 (11) (2021) 3641. `doi:10.3390/s21113641`. 41

[364] Y. Wen, J. Chen, B. Sheng, Z. Chen, P. Li, P. Tan, T.-Y. Lee, Structure-aware motion deblurring using multi-adversarial optimized cyclegan, IEEE Transactions on Image Processing 30 (2021) 6142–6155. `doi:10.1109/TIP.2021.3092814`. 41

[365] Y. Wu, P. Qian, X. Zhang, Two-level wavelet-based convolutional neural network for image deblurring, IEEE Access 9 (2021) 45853–45863. `doi:10.1109/ACCESS.2021.3067055`. 41

[366] Z. Pan, Q. Lv, Z. Tan, A two-stage network for image deblurring, IEEE Access (2021). `doi:10.1109/ACCESS.2021.3082211`. 41

[367] A. Kaufman, R. Fattal, Deblurring using analysis-synthesis networks pair, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 5811–5820. `doi:10.1109/CVPR42600.2020.00585`. 41

[368] M. Hirsch, C. J. Schuler, S. Harmeling, B. Schölkopf, Fast removal of non-uniform camera shake, in: International Conference on Computer Vision (ICCV), 2011, pp. 463–470. `doi:10.1109/ICCV.2011.6126276`. 42, 135

[369] T. S. Cho, A. Levin, F. Durand, W. T. Freeman, Motion blur removal with orthogonal parabolic exposures, in: IEEE International Conference on Computational Photography (ICCP), 2010, pp. 1–8. `doi:10.1109/ICCPHOT.2010.5585100`. 42, 43, 135, 136

[370] D. N. Thanh, V. S. Prasath, S. Dvoenko, et al., An adaptive method for image restoration based on high-order total variation and inverse gradient, Signal, Image and Video Processing 14 (6) (2020) 1189–1197. `doi:10.1007/s11760-020-01657-9`. 43

[371] L. T. Thanh, D. N. H. Thanh, V. B. S. Prasath, Adaptive texts deconvolution method for real natural images, in: 2019 25th Asia-Pacific Conference on Communications (APCC), 2019, pp. 110–115. `doi:10.1109/APCC47188.2019.9026515`. 43

[372] V. B. S. Prasath, D. N. H. Thanh, Structure tensor adaptive total variation for image restoration, Turkish Journal of Electrical Engineering & Computer Sciences 27 (2) (2019) 1147–1156. 43

[373] B. Fu, Y. Dong, S. Fu, Y. Mao, D. N. Thanh, Learning domain transfer for unsupervised magnetic resonance imaging restoration and edge enhancement, International Journal of Imaging Systems and Technology 32 (1) (2022) 144–154. `doi:10.1002/ima.22658`. 43

[374] K. Zhang, W. Luo, Y. Zhong, L. Ma, B. Stenger, W. Liu, H. Li, Deblurring by realistic blurring, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2734–2743. `doi:10.1109/CVPR42600.2020.00281`. 44

[375] W. Niu, K. Zhang, W. Luo, Y. Zhong, H. Li, Deep robust image deblurring via blur distilling and information comparison in latent space, Neurocomputing 466 (2021) 69–79. `doi:10.1016/j.neucom.2021.09.019`. 44

[376] E. H. Land, J. J. McCann, Lightness and retinex theory, Journal of the Optical Society of America 61 (1) (1971) 1. `doi:10.1364/JOSA.61.000001`. 47, 49

[377] Y. Wang, W. Song, G. Fortino, L.-Z. Qi, W. Zhang, A. Liotta, An experimental-based review of image enhancement and image restoration methods for underwater imaging, IEEE Access 7 (2019) 140233–140251. `doi:10.1109/ACCESS.2019.2932130`. 47

[378] C. Li, C. Guo, L. Han, J. Jiang, M.-M. Cheng, J. Gu, C. C. Loy, Low-light image and video enhancement using deep learning: A survey, arXiv (2021). `arXiv:2104.10729v3`. 47

[379] S. H. Malik, T. A. Lone, S. Quadri, Contrast enhancement and smoothing of CT images for diagnosis, in: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), IEEE, 2015, pp. 2214–2219. 48

[380] N. Taguchi, S. Oda, T. Nakaura, D. Utsunomiya, Y. Funama, M. Imuta, H. Yuki, Y. Nagayama, M. Kidoh, K. Hirata, et al., Contrast enhancement in abdominal computed tomography: Influence of photon energy of different scanners, The British journal of radiology 91 (1081) (2018) 20170285. `doi:10.1259/bjr.20170285`. 48

[381] T. Qiu, C. Wen, K. Xie, F.-Q. Wen, G.-Q. Sheng, X.-G. Tang, Efficient medical image enhancement based on CNN-FBB model, IET Image Processing 13 (10) (2019) 1736–1744. `doi:10.1049/iet-ipr.2018.6380`. 48

[382] M. Charytanowicz, P. Kulczycki, S. Łukasik, P. A. Kowalski, Image enhancement with applications in biomedical processing, in: Conference on Information Technology, Systems Research and Computational Physics, Springer, 2018, pp. 97–106. `doi:10.1007/978-3-030-18058-4_8`. 48

[383] M. Inoue, T. Freel, A. Van Avermaete, W. M. Leevy, Color enhancement strategies for 3D printing of X-ray computed tomography bone data for advanced anatomy teaching models, Applied Sciences 10 (5) (2020) 1571. `doi:10.3390/app10051571`. 48

[384] P. Cascarano, E. L. Piccolomini, E. Morotti, A. Sebastiani, Plug-and-play gradient-based denoisers applied to ct image enhancement, Applied Mathematics and Computation 422 (2022) 126967. `doi:10.1016/j.amc.2022.126967`. 48

[385] K. de Haan, Y. Rivenson, Y. Wu, A. Ozcan, Deep-learning-based image reconstruction and enhancement in optical microscopy, Proceedings of the IEEE 108 (1) (2019) 30–50. `doi:10.1109/JPROC.2019.2949575`. 48

[386] F. Chen, J. Liu, D. Gou, X. Zhang, L. Chen, H. Liao, An accurate and universal approach for short-exposure-time microscopy image enhancement, Computerized Medical Imaging and Graphics 83 (2020) 101743. `doi:10.1016/j.compmedimag.2020.101743`. 48, 57

[387] V. Georgieva, P. Petrov, R. Mironov, A. Mihaylova, An approach for microscopy image restoration, in: Proceedings of the Eighth International Conference on Telecommunications and Remote Sensing (ICTRS), ACM Press, 2019, pp. 35–39. `doi:10.1145/3357767.3357772`. 48

[388] A. Mohammed, I. Farup, M. Pedersen, Ø. Hovde, S. Yildirim Yayilgan, Stochastic capsule endoscopy image enhancement, Journal of Imaging 4 (6) (2018) 75. `doi:10.3390/jimaging4060075`. 48

[389] K. Munadi, K. Muchtar, N. Maulina, B. Pradhan, Image enhancement for tuberculosis detection using deep learning, IEEE Access 8 (2020) 217897–217907. `doi:10.1109/ACCESS.2020.3041867`. 48

[390] T. Arici, S. Dikbas, Y. Altunbasak, A histogram modification framework and its application for image contrast enhancement, IEEE Transactions on Image Processing 18 (9) (2009) 1921–1935. `doi:10.1109/TIP.2009.2021548.` 48

[391] T. Celik, T. Tjahjadi, Contextual and variational contrast enhancement, IEEE Transactions on Image Processing 20 (12) (2011) 3431–3441. `doi:10.1109/TIP.2011.2157513.` 48

[392] T. Huynh-The, T. Le-Tien, Brightness preserving weighted dynamic range histogram equalization for image contrast enhancement, in: 2013 International Conference on Advanced Technologies for Communications (ATC), IEEE, 2013, pp. 386–391. `doi:10.1109/ATC.2013.6698142.` 48

[393] K. Nakai, Y. Hoshi, A. Taguchi, Color image contrast enhancement method based on differential intensity/saturation gray-levels histograms, in: 2013 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), IEEE, 2013, pp. 445–449. `doi:10.1109/ISPACS.2013.6704591.` 48

[394] S. Wang, J. Zheng, H.-M. Hu, B. Li, Naturalness preserved enhancement algorithm for non-uniform illumination images, IEEE Transactions on Image Processing 22 (9) (2013) 3538–3548. `doi:10.1109/TIP.2013.2261309.` 48, 56, 57, 58, 59, 60, 61

[395] B. Cai, X. Xu, K. Guo, K. Jia, B. Hu, D. Tao, A joint intrinsic-extrinsic prior model for retinex, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4020–4029. `doi:10.1109/ICCV.2017.431.` 48, 49, 58, 59, 60, 61

[396] Y. Zhou, K. Panetta, S. Agaian, Mammogram enhancement using alpha weighted quadratic filter, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2009, pp. 3681–3684. `doi:10.1109/IEMBS.2009.5334723.` 48

[397] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, S. M. Pizer, Contrast limited adaptive Histogram Equalization image processing to improve the detection of simulated spiculations in dense mammograms, Journal of Digital Imaging 11 (4) (1998) 193–200. `doi:10.1007/BF03178082.` 48

[398] S. Somal, Image enhancement using local and global Histogram Equalization technique and their comparison, in: First International Conference on Sustainable Technologies for Computational Intelligence, Springer, 2020, pp. 739–753. `doi:10.1007/978-981-15-0029-9_58.` 49

[399] D. J. Jobson, Z.-u. Rahman, G. A. Woodell, Properties and performance of a center/surround retinex, IEEE Transactions on Image Processing 6 (3) (1997) 451–462. `doi:10.1109/83.557356.` 49

[400] D. J. Jobson, Z.-u. Rahman, G. A. Woodell, A multiscale retinex for bridging the gap between color images and the human observation of scenes, IEEE Transactions on Image processing 6 (7) (1997) 965–976. `doi:10.1109/83.597272.` 49

[401] Z. Ying, G. Li, Y. Ren, R. Wang, W. Wang, A new image contrast enhancement algorithm using exposure fusion framework, in: International Conference on Computer Analysis of Images and Patterns, Springer, 2017, pp. 36–46. `doi:10.1007/978-3-319-64698-5_4.` 49, 58, 59, 60, 61

[402] X. Fu, Y. Liao, D. Zeng, Y. Huang, X.-P. Zhang, X. Ding, A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation, IEEE Transactions on Image Processing 24 (12) (2015) 4965–4977. `doi:10.1109/TIP.2015.2474701.` 49, 58, 59, 60, 61

[403] V. Pătraşcu, Color image enhancement using the support fuzzification, in: International Fuzzy Systems Association World Congress, Springer, 2003, pp. 412–419. `doi:10.1007/3-540-44967-1_49.` 49

[404] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, X. Ding, A weighted variational model for simultaneous reflectance and illumination estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2782–2790. `doi:10.1109/CVPR.2016.304.` 49

[405] H. Lin, C. Wei, N. Cao, H. Chen, G. Wang, J. Chen, G. Chen, S. Zhuo, A novel low-signal image enhancement method for multiphoton microscopy, Journal of Physics D: Applied Physics 52 (28) (2019) 285401. `doi:10.1088/1361-6463/ab1cad`. 49

[406] Z. Ying, G. Li, Y. Ren, R. Wang, W. Wang, A new low-light image enhancement algorithm using camera response model, in: Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW), 2017, pp. 3015–3022. `doi:10.1109/ICCVW.2017.356`. 49, 58, 59, 60, 61

[407] K. G. Lore, A. Akintayo, S. Sarkar, LLNet: A deep autoencoder approach to natural low-light image enhancement, Pattern Recognition 61 (2017) 650–662. `doi:10.1016/j.patcog.2016.06.008`. 49

[408] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, F. Durand, Deep bilateral learning for real-time image enhancement, ACM Transactions on Graphics (TOG) 36 (4) (2017) 1–12. `doi:10.1145/3072959.3073592`. 49

[409] S. Anwar, C. Li, F. Porikli, Deep underwater image enhancement, arXiv (2018). `arXiv:1807.03528v1`. 49

[410] C. Wei, W. Wang, W. Yang, J. Liu, Deep retinex decomposition for low-light enhancement, arXiv (2018). `arXiv:1808.04560`. 49, 56, 57, 58, 59, 60, 61

[411] Q. Fu, X. Di, Y. Zhang, Learning an adaptive model for extreme low-light raw image processing, arXiv (2020). `arXiv:2004.10447v1`. 49

[412] K. Xu, X. Yang, B. Yin, R. W. Lau, Learning to restore low-light images via decomposition-and-enhancement, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2278–2287. `doi:10.1109/CVPR42600.2020.00235`. 50

[413] F. Lv, Y. Li, F. Lu, Attention guided low-light image enhancement with a large scale low-light simulation dataset, International Journal of Computer Vision 129 (7) (2021) 2175–2193. `doi:10.1007/s11263-021-01466-8`. 50

[414] Z. Shen, H. Fu, J. Shen, L. Shao, Modeling and enhancing low-quality Retinal Fundus images, IEEE Transactions on Medical Imaging 40 (3) (2021) 996–1006. `doi:10.1109/TMI.2020.3043495`. 50

[415] Y. Deng, C. C. Loy, X. Tang, Aesthetic-driven image enhancement by adversarial learning, in: Proceedings of the 26th ACM international Conference on Multimedia, 2018, pp. 870–878. `doi:10.1145/3240508.3240531`. 50

[416] E. de Stoutz, A. Ignatov, N. Kobyshev, R. Timofte, L. Van Gool, Fast perceptual image enhancement, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0. 50

[417] M. J. Islam, Y. Xia, J. Sattar, Fast underwater image enhancement for improved visual perception, IEEE Robotics and Automation Letters 5 (2) (2020) 3227–3234. `doi:10.1109/LRA.2020.2974710`. 50, 54

[418] P. M. Uplavikar, Z. Wu, Z. Wang, All-in-One underwater image enhancement using domain-adversarial learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 1–8. 50, 55

[419] A. S. Shamsabadi, C. Oh, A. Cavallaro, Edgefool: an adversarial image enhancement filter, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020, pp. 1898–1902. `doi:10.1109/ICASSP40776.2020.9054368`. 50

[420] Z. Ni, W. Yang, S. Wang, L. Ma, S. Kwong, Towards unsupervised deep image enhancement with generative adversarial network, IEEE Transactions on Image Processing 29 (2020) 9140–9151. `doi:10.1109/TIP.2020.3023615`. 50, 54

[421] Z. Chen, V. Calhoun, Effect of spatial smoothing on task fMRI ICA and functional connectivity, Frontiers in Neuroscience 12 (FEB) (2018) 1–10. `doi:10.3389/fnins.2018.00015`. 50, 54

[422] S. Moran, P. Marza, S. McDonagh, S. Parisot, G. Slabaugh, DeepLPF: Deep local parametric filters for image enhancement, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 12823–12832. `doi:10.1109/CVPR42600.2020.01284`. 50, 57

[423] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, J. Jia, Underexposed photo enhancement using deep illumination estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 6842–6850. `doi:10.1109/CVPR.2019.00701`. 50

[424] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, R. Cong, Zero-reference deep curve estimation for low-light image enhancement, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 1777–1786. `doi:10.1109/CVPR42600.2020.00185`. 50

[425] J. Park, J.-Y. Lee, D. Yoo, I. S. Kweon, Distort-and-recover: Color enhancement using deep reinforcement learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 5928–5936. `doi:10.1109/CVPR.2018.00621`. 50

[426] C. Wang, H. Song, L. Chen, Q. Li, J. Yang, X. T. Hu, L. Zhang, Automatic liver segmentation using multi-plane integrated fully convolutional neural networks, in: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2018, pp. 1–6. `doi:10.1109/BIBM.2018.8621257`. 50, 58, 59, 60, 61, 65, 67, 72, 77, 80

[427] Y. Wang, J. Guo, H. Gao, H. Yue, Uiecˆ 2-net: Cnn-based underwater image enhancement using two color space, Signal Processing: Image Communication 96 (2021) 116250. `doi:10.1016/j.image.2021.116250`. 50, 57

[428] J. Wang, X. H. Han, Y. Xu, L. Lin, H. Hu, C. Jin, Y. W. Chen, Sparse Codebook Model of Local Structures for Retrieval of Focal Liver Lesions Using Multiphase Medical Images, International Journal of Biomedical Imaging 2017 (2017). `doi:10.1155/2017/1413297`. 50

[429] C. Chen, Q. Chen, J. Xu, V. Koltun, Learning to see in the dark, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 3291–3300. `doi:10.1109/CVPR.2018.00347`. 50

[430] A. Jamadandi, U. Mudenagudi, Exemplar-based underwater image enhancement augmented by wavelet corrected transforms, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 11–17. 50, 57

[431] T. Vu, C. Van Nguyen, T. X. Pham, T. M. Luu, C. D. Yoo, Fast and efficient image quality enhancement via desubpixel convolutional neural networks, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0. 50

[432] M. Jahidul Islam, P. Luo, J. Sattar, Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception, arXiv (2020). `arXiv:2002.01155v1`. 50, 54

[433] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, L. Shao, Learning enriched features for real image restoration and enhancement, in: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16, Springer, 2020, pp. 492–511. `doi:10.1007/978-3-030-58595-2_30`. 50

[434] H. Zeng, J. Cai, L. Li, Z. Cao, L. Zhang, Learning image-adaptive 3D lookup tables for high performance photo enhancement in real-time, IEEE Transactions on Pattern Analysis and Machine Intelligence (2020) 1–1`doi:10.1109/TPAMI.2020.3026740`. 50, 55, 57, 58, 59, 60

[435] H.-U. Kim, Y. J. Koh, C.-S. Kim, PieNet: Personalized image enhancement network, in: European Conference on Computer Vision, Springer, 2020, pp. 374–390. `doi:10.1007/978-3-030-58577-8_23`. 50

[436] J. Lee, T. Yoon, B. H. Lee, Post-Processing method for image reconstruction enhancement in integrating-bucket-based full-field Optical Coherence Tomography, Applied Sciences 10 (3) (2020) 830. `doi:10.3390/app10030830`. 50

[437] T. Son, J. Kang, N. Kim, S. Cho, S. Kwak, URIE: Universal Image Enhancement for Visual Recognition in the Wild, in: European Conference on Computer Vision, Springer, 2020, pp. 749–765. `doi:10.1007/978-3-030-58545-7_43`. 50

[438] S. Moran, S. McDonagh, G. Slabaugh, Curl: Neural curve layers for global image enhancement, in: International Conference on Pattern Recognition (ICPR), IEEE, 2021, pp. 9796–9803. `doi:10.1109/ICPR48806.2021.9412677`. 50

[439] J. Wang, W. Tan, X. Niu, B. Yan, RDGAN: Retinex decomposition based adversarial learning for low-light enhancement, in: IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2019, pp. 1186–1191. `doi:10.1109/ICME.2019.00207`. 50, 54

[440] Z. Zhao, Z. Liu, M. Larson, Adversarial color enhancement: Generating unrestricted adversarial images by optimizing a color filter, arXiv (2020). `arXiv:2002.01008v3`. 50

[441] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, Vol. 15, 2015, pp. 234–241. `doi:10.1007/978-3-319-24574-4_28`. 53, 65, 67, 68, 81

[442] D. Eigen, C. Puhrsch, R. Fergus, Depth map prediction from a single image using a multi-scale deep network, in: International Conference on Neural Information Processing Systems, 2014, p. 23662374. 53

[443] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, Z. Wang, EnlightenGAN: Deep light enhancement without paired supervision, IEEE Transactions on Image Processing 30 (8) (2021) 2340–2349. `doi:10.1109/TIP.2021.3051462`. 54

[444] S. Saladi, N. Amutha Prabha, Analysis of denoising filters on MRI brain images, International Journal of Imaging Systems and Technology 27 (3) (2017) 201–208. `doi:10.1002/ima.22225`. 55

[445] H. Rabbouch, F. Saâdaoui, A wavelet-assisted subband denoising for tomographic image reconstruction, Journal of Visual Communication and Image Representation 55 (2018) 115–130. `doi:10.1016/j.jvcir.2018.05.004`. 55

[446] J. Lu, H. Yang, L. Shen, Y. Zou, Ultrasound image restoration based on a learned dictionary and a higher-order MRF, Computers and Mathematics with Applications 77 (4) (2019) 991–1009. `doi:10.1016/j.camwa.2018.10.031`. 55

[447] I. Amsr, H. E. Msrcr, O. Psnr, F. Lv, F. Lu, J. Wu, C. Lim, MBLLEn: Low-light image/video enhancement using CNNs, British Machine Vision Conference (BMVC) (2019) 1–13. 55, 56, 57, 58, 59, 60, 61

[448] H. Ibrahim, N. S. P. Kong, Brightness preserving dynamic histogram equalization for image contrast enhancement, IEEE Transactions on Consumer Electronics 53 (4) (2007) 1752–1758. `doi:10.1109/TCE.2007.4429280`. 57, 58, 59, 60, 61

[449] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., {TensorFlow}: a system for {Large-Scale} machine learning, in: 12th USENIX symposium on operating systems design and implementation (OSDI 16), 2016, pp. 265–283. 57

[450] I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT press, 2016. 57

[451] Z. Baig, N. Abu-Omar, R. A. Khan, C. Verdiales, R. Frehlick, J. Shaw, F.-X. Wu, Y. Luo, Prognosticating outcome in pancreatic head cancer with the use of a machine learning algorithm, Technology in cancer research & treatment 20 (2021). `doi:10.1177/15330338211050767`. 60, 64, 94

[452] A. Ziaei, H. Yeganeh, K. Faez, S. Sargolzaei, A novel approach for contrast enhancement in biomedical images based on histogram equalization, BioMedical Engineering and Informatics: New Development and the Future - Proceedings of the 1st International Conference on BioMedical Engineering and Informatics, BMEI 2008 1 (2008) 855–858. `doi:10.1109/BMEI.2008.300`. 61

[453] L. Si-Yao, D. Ren, Z. Hu, J. Li, Q. Yin, P. Guo, A concatenated residual network for image deblurring, arXiv (2018). `arXiv:1804.06042`. 61

[454] M. Lysaker, A. Lundervold, X. C. Tai, Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time, IEEE Transactions on Image Processing 12 (12) (2003) 1579–1589. `doi:10.1109/TIP.2003.819229`. 61

[455] Y. Hu, M. Jacob, Image recovery using improved total variation regularization, Proceedings - International Symposium on Biomedical Imaging (2011) 1154–1157`doi:10.1109/ISBI.2011.5872606`. 61

[456] B.-y. Park, K. Byeon, H. Park, FuNP (Fusion of Neuroimaging Preprocessing) Pipelines: A Fully Automated Preprocessing Software for Functional Magnetic Resonance Imaging, Frontiers in Neuroinformatics 13 (February) (2019) 1–14. `doi:10.3389/fninf.2019.00005`. 61

[457] O. Esteban, C. J. Markiewicz, R. W. Blair, C. A. Moodie, A. I. Isik, A. Erramuzpe, J. D. Kent, M. Goncalves, E. DuPre, M. Snyder, H. Oya, S. S. Ghosh, J. Wright, J. Durnez, R. A. Poldrack, K. J. Gorgolewski, fMRIPrep: a robust preprocessing pipeline for functional MRI., Nature methods 16 (1) (2019) 111–116. `doi:10.1038/s41592-018-0235-4`. 61

[458] N. Sharma, A. Ray, K. Shukla, S. Sharma, S. Pradhan, A. Srivastva, L. Aggarwal, Automated medical image segmentation techniques, Journal of Medical Physics 35 (1) (2010) 3. `doi:10.4103/0971-6203.58777`. 64

[459] W. Li, F. Jia, Q. Hu, Automatic segmentation of liver tumor in CT images with deep convolutional neural networks, Journal of Computer and Communications 03 (11) (2015) 146–151. `doi:10.4236/jcc.2015.311023`. 64

[460] G. Chlebus, A. Schenk, J. H. Moltz, H. K. Hahn, H. Meine, Deep learning based automatic liver tumor segmentation in CT with shape-based post-processing, in: International conference on Medical Imaging with Deep Learning, 2018, pp. 1–9. 64, 65, 66, 75, 77, 79

[461] X. Xu, Q. Lu, L. Yang, S. Hu, D. Chen, Y. Hu, Y. Shi, Quantization of fully convolutional networks for accurate biomedical image segmentation, in: 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8300–8308. `doi:10.1109/CVPR.2018.00866`. 64

[462] M. Moghbel, S. Mashohor, R. Mahmud, M. Iqbal Bin Saripan, Automatic liver tumor segmentation on computed tomography for patient treatment planning and monitoring, EXCLI Journal 15 (2016) 406–423. `doi:10.17179/excli2016-402`. 64

[463] S. Zheng, B. Fang, L. Li, M. Gao, Y. Wang, A variational approach to liver segmentation using statistics from multiple sources, Physics in Medicine & Biology 63 (2) (2018) 025024. `doi:10.1088/1361-6560/aaa360`. 64, 75, 77

[464] X. X. Wang, Y. Zheng, L. Gan, X. X. Wang, X. Sang, X. Kong, J. Zhao, Liver segmentation from CT images using a sparse priori statistical shape model (SP-SSM), PLOS ONE 12 (10) (2017) e0185249. `doi:10.1371/journal.pone.0185249`. 64

[465] S. Luo, X. Li, J. Li, Review on the Methods of Automatic Liver Segmentation from Abdominal Images, Journal of Computer and Communications 02 (02) (2014) 1–7. `doi:10.4236/jcc.2014.22001`. 64

[466] L. Huang, M. Weng, H. Shuai, Y. Huang, J. Sun, F. Gao, Automatic Liver Segmentation from CT Images Using Single-Block Linear Detection, BioMed Research International 2016 (2016) 1–11. `doi:10.1155/2016/9420148`. 64, 77

[467] M. S. Fasihi, W. B. Mikhael, Overview of Current Biomedical Image Segmentation Methods, Proceedings - 2016 International Conference on Computational Science and Computational Intelligence, CSCI 2016 (2017) 803–808 doi:10.1109/CSCI.2016.0156. 64

[468] M. Jayanthi, Comparative study of different techniques used for medical image segmentation of liver from abdominal CT scan, Proceedings of the 2016 IEEE International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2016 (2016) 1462–1465 doi:10.1109/WiSPNET.2016.7566379. 64

[469] Y. Guo, Y. Liu, T. Georgiou, M. S. Lew, A review of semantic segmentation using deep neural networks, International Journal of Multimedia Information Retrieval 7 (2) (2018) 87–93. doi:10.1007/s13735-017-0141-z. 64

[470] A. M. Mharib, A. R. Ramli, S. Mashohor, R. B. Mahmood, Survey on liver CT image segmentation methods (2012). doi:10.1007/s10462-011-9220-3. 64

[471] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, P.-A. Heng, H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation From CT Volumes, IEEE Transactions on Medical Imaging 37 (12) (2018) 2663–2674. doi:10.1109/TMI.2018.2845918. 65, 66, 67, 71, 72, 75, 77, 79, 80

[472] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Clarkson, D. C. Barratt, Automatic multi-organ segmentation on abdominal CT with dense V-networks, IEEE Transactions on Medical Imaging 37 (8) (2018) 1822–1834. doi:10.1109/TMI.2018.2806309. 65, 72, 80

[473] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: Learning dense volumetric segmentation from sparse annotation, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, pp. 424–432. doi:10.1007/978-3-319-46723-8_49. 65

[474] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Medical Image Analysis 36 (2017) 61–78. doi:10.1016/j.media.2016.10.004. 65, 66, 80

[475] J. Hai, K. Qiao, J. Chen, H. Tan, J. Xu, L. Zeng, D. Shi, B. Yan, Fully convolutional densenet with multiscale context for automated breast tumor segmentation, Journal of Healthcare Engineering 2019 (2019). doi:10.1155/2019/8415485. 65, 66, 80

[476] F. Yu, V. Koltun, Multi-scale context aggregation by dilated convolutions, International Conference on Learning Representations (ICLP) (2015). arXiv:1511.07122. 65, 80

[477] L. Zhong, T. Li, H. Shu, C. Huang, J. M. Johnson, D. F. Schomer, H.-L. Liu, Q. Feng, W. Yang, H. Zhu, 2WM: Tumor segmentation and tract statistics for assessing white matter integrity with applications to glioblastoma patients, NeuroImage 223 (2020) 117368. doi:10.1016/j.neuroimage.2020.117368. 65

[478] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, H. Larochelle, Brain tumor segmentation with Deep Neural Networks, Medical Image Analysis 35 (2017) 18–31. doi:10.1016/j.media.2016.05.004. 65, 67

[479] H. R. Roth, H. Oda, X. Zhou, N. Shimizu, Y. Yang, Y. Hayashi, M. Oda, M. Fujiwara, K. Misawa, K. Mori, An application of cascaded 3D fully convolutional networks for medical image segmentation, Computerized Medical Imaging and Graphics 66 (2018) 90–99. doi:10.1016/j.compmedimag.2018.03.001. 66, 67

[480] M. Drozdzal, G. Chartrand, E. Vorontsov, M. Shakeri, L. Di Jorio, A. Tang, A. Romero, Y. Bengio, C. Pal, S. Kadoury, Learning normalized inputs for iterative estimation in medical image segmentation, Medical Image Analysis 44 (2018) 1–13. doi:10.1016/j.media.2017.11.005. 66

[481] S. Jgou, M. Drozdzal, D. Vazquez, A. Romero, Y. Bengio, The one hundred layers tiramisu: Fully convolutional denseNets for semantic segmentation, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1175–1183. `doi:10.1109/CVPRW.2017.156`. 66, 80

[482] S. Lyu, E. P. Simoncelli, Nonlinear image representation using divisive normalization, in: 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2008, pp. 1–8. `doi:10.1109/CVPR.2008.4587821`. 68

[483] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net: Fully convolutional neural networks for volumetric medical image segmentation, 2016 Fourth International Conference on 3D Vision (3DV) (2016) 565–571`doi:10.1109/3DV.2016.79`. 69, 80

[484] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification, in: 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, 2015, pp. 1026–1034. `doi:10.1109/ICCV.2015.123`. 70

[485] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, in: 3rd International Conference on Learning Representations, ICLR - Conference Track Proceedings, 2015, pp. 1–15. `arXiv:1412.6980`. 70

[486] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells, F. A. Jolesz, R. Kikinis, Statistical validation of image segmentation quality based on a spatial overlap index, Academic Radiology 11 (2) (2004) 178–189. `doi:10.1016/S1076-6332(03)00671-8`. 72

[487] F. Chung, H. Delingette, Regional appearance modeling based on the clustering of intensity profiles, Computer Vision and Image Understanding 117 (6) (2013) 705–717. `doi:10.1016/j.cviu.2013.01.011`. 77

[488] G. Li, X. Chen, F. Shi, W. Zhu, J. Tian, D. Xiang, Automatic liver segmentation based on shape constraints and deformable graph cut in CT images, IEEE Transactions on Image Processing 24 (12) (2015) 5315–5329. `doi:10.1109/TIP.2015.2481326`. 77

[489] Y. Yuan, Y. W. Chen, C. Dong, H. Yu, Z. Zhu, Hybrid method combining superpixel, random walk and active contour model for fast and accurate liver segmentation, Computerized Medical Imaging and Graphics 70 (2018) 119–134. `doi:10.1016/j.compmedimag.2018.08.012`. 77

[490] A. Wimmer, J. Hornegger, G. Soza, Implicit active shape model employing boundary classifier, in: 2008 19th International Conference on Pattern Recognition, IEEE, 2008, pp. 1–4. `doi:10.1109/ICPR.2008.4760968`. 77

[491] M. G. Linguraru, W. J. Richbourg, Jianfei Liu, J. M. Watt, V. Pamulapati, Shijun Wang, R. M. Summers, Tumor Burden Analysis on Computed Tomography by Automated Liver and Tumor Segmentation, IEEE Transactions on Medical Imaging 31 (10) (2012) 1965–1976. `doi:10.1109/TMI.2012.2211887`. 77

[492] A. S. Maklad, M. Matsuhiro, H. Suzuki, Y. Kawata, N. Niki, M. Satake, N. Moriyama, T. Utsunomiya, M. Shimada, Blood vessel-based liver segmentation using the portal phase of an abdominal CT dataset, Medical Physics 40 (11) (2013) 113501. `doi:10.1118/1.4823765`. 76, 77

[493] R. Gauriau, R. Cuingnet, R. Prevost, B. Mory, R. Ardon, D. Lesage, I. Bloch, A generic, robust and fully-automatic workflow for 3D CT liver segmentation, in: International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging, Springer, 2013, pp. 241–250. `doi:10.1007/978-3-642-41083-3_27`. 77

[494] S. D. Salman Al-Shaikhli, M. Y. Yang, B. Rosenhahn, 3D automatic liver segmentation using feature-constrained Mahalanobis distance in CT images, Biomedical Engineering / Biomedizinische Technik 61 (4) (2016) 401–412. `doi:10.1515/bmt-2015-0017`. 77

[495] A. H. Foruzan, Y.-W. Chen, Improved segmentation of low-contrast lesions using sigmoid edge model, International Journal of Computer Assisted Radiology and Surgery 11 (7) (2016) 1267–1283. `doi:10.1007/s11548-015-1323-x`. 79

[496] L. Liu, F. X. Wu, J. Wang, Efficient multi-kernel DCNN with pixel dropout for stroke MRI segmentation, Neurocomputing 350 (2019) 117–127. `doi:10.1016/j.neucom.2019.03.049`. 80

[497] F. Feo, R. M. Pascale, Multifocal hepatocellular carcinoma: intrahepatic metastasis or multicentric carcinogenesis?, Annals of translational medicine 3 (1) (2015). `doi:10.3978/j.issn.2305-5839.2014.12.08`. 84

[498] G. Zhang, Z. Peng, C. Yan, J. Wang, J. Luo, H. Luo, A novel liver cancer diagnosis method based on patient similarity network and densegcn, Scientific Reports 12 (1) (2022) 1–10. `doi:10.1038/s41598-022-10441-3`. 84

[499] J. Cao, S. Liu, H. Liu, H. Lu, CS-MRI reconstruction based on analysis dictionary learning and manifold structure regularization, Neural Networks 123 (2020) 217–233. `doi:10.1016/j.neunet.2019.12.010`. 87

[500] H.-T. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir, R. Anil, Z. Haque, L. Hong, V. Jain, X. Liu, H. Shah, Wide & deep learning for recommender systems, in: Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, DLRS 2016, Association for Computing Machinery, New York, NY, USA, 2016, p. 710. `doi:10.1145/2988450.2988454`. 89

Appendix A

Supplementary for deblurring

This section provides supplementary materials for image formation model A.1, image gradient magnitude A.1.1, and additional visual results for the proposed deblurring neural network A.1.2.

A.1 Image Formation Model

Mainly two planes are considered in image formation. One is a real-world plane, and the second one is an image plane in which we try to develop an image. Therefore, a 3D scene in the real world plane is transformed into a 2D image plane. Meanwhile, the object in a real-world scene is assumed to be a source of radiant energy. Suppose the 3D scene is defined in a homogeneous real-world coordination system with projective geometry as $O(x, y, z, w)$, where (x, y, z) are related to 3D coordinates, and w is the projective distance between real-world plane and image plane. There are 3D objects with radiant energy presented by $e_i(x, y, z, w)$, where $i = \{1, 2, \cdots, n\}$ are a number of objects. These 3D objects contain radiant energy assumed to propagate from real-world plane to image plane while intercepting through image formation system based on multiple factors such as projective distance w, aperture or pin-hole size, etc. By geometric transformation function $q(x, y; e_i(x, y, z, w))$ the objects are transformed from $O(x, y, z, w)$ to $O'\left(x', y'\right)$ image plane, eliminating hidden surfaces of objects and generating the radiant intensity for visible surface points. Thus, by the transformation, the radiant energy function for each visible object point becomes [114],

$$I(x, y) = q(x, y; e_i(x, y, z, w)) \tag{A.1}$$

A function that describes the transformation of energy from a real-world plane (x, y) to an image plane $\left(x', y'\right)$ is defined as

$$B\left(x', y'\right) = k\left(x', y', x, y, I(x, y)\right) \tag{A.2}$$

Since the image formation model is linear with the superposition additive property, we can rewrite it as

$$k\left(x', y', x, y, I(x, y)\right) = k\left(x', y', x, y\right) I(x, y) \tag{A.3}$$

Summing the infinitesimal contribution of object points to image plane radiant intensity, the general image formation can be written in the spatial domain as [113],

$$
\begin{aligned}
B(x', y') &= \iint_{-\infty}^{\infty} k\left(x', y', x, y, I(x, y)\right) dx dy \\
&= \iint_{-\infty}^{\infty} k\left(x', y', x, y\right) I(x, y) dx dy
\end{aligned}
\tag{A.4}
$$

where $k\left(x', y', x, y\right)$ is known to be the point-spread function (PSF) that describes radiant energy distribution in the image plane due to a point source of radiant energy located in the object plane. The $k\left(x', y', x, y\right)$ can define energy at any position in both planes; hence it is said to be space-variant PSF. However, space invariant PSF independent of position can be defined by taking the difference of coordinate systems. Therefore, linear space invariant PSF can be expressed using convolution integral as [113, 114],

$$B(x', y') = \iint_{-\infty}^{\infty} k\left(x' - x, y' - y\right) I(x, y) dx dy \tag{A.5}$$

In discrete form, it can be represented by summation as,

$$B(x^{'},y^{'}) = \sum_{x,y=0}^{M-1,N-1} k\left(x^{'}-x, y^{'}-y\right) I\left(x,y\right)$$

$$= k\left(x^{'}, y^{'}\right) \otimes I\left(x,y\right)$$

(A.6)

where $B\left(x^{'}, y^{'}\right)$ is the output or obtained image, $I\left(x,y\right)$ is the input or true image, and the size of the image is $M \times N$ and $\otimes$ is the convolution operator.

A.1.1   Image Gradient Magnitude

The image gradient has been implemented in various forms commonly known as Sobel, Intermediate, Central techniques etc. Image gradient is the change in intensity of an image that can be numerically found by Taylor series approximation using finite-difference.
By Taylor series the finite forward difference approximations,

$$f\left(x+h\right) = f\left(x\right) + hf^{'}\left(x\right) + \frac{h^2 f^{''}\left(x\right)}{2} + \frac{h^3 f^{'''}\left(x\right)}{3!} + 0\left(h^4\right)$$

(A.7)

the finite backward difference approximations,

$$f\left(x-h\right) = f\left(x\right) - hf^{'}\left(x\right) + \frac{h^2 f^{''}\left(x\right)}{2} - \frac{h^3 f^{'''}\left(x\right)}{3!} + 0\left(h^4\right)$$

(A.8)

subtracting (A.8) from (A.7) gives the finite central difference,

$$\frac{f\left(x+h\right) - f\left(x-h\right)}{2h} = f^{'}\left(x\right) + 0\left(h^2\right)$$

(A.9)

Hence image gradients in x and y direction become,

$$I_x^{'} = \frac{I\left(x+1,y\right) - I\left(x-1,y\right)}{2}; \quad I_y^{'} = \frac{I\left(x,y+1\right) - I\left(x,y-1\right)}{2}$$

(A.10)

respectively, where the image gradient magnitude can be computed as,

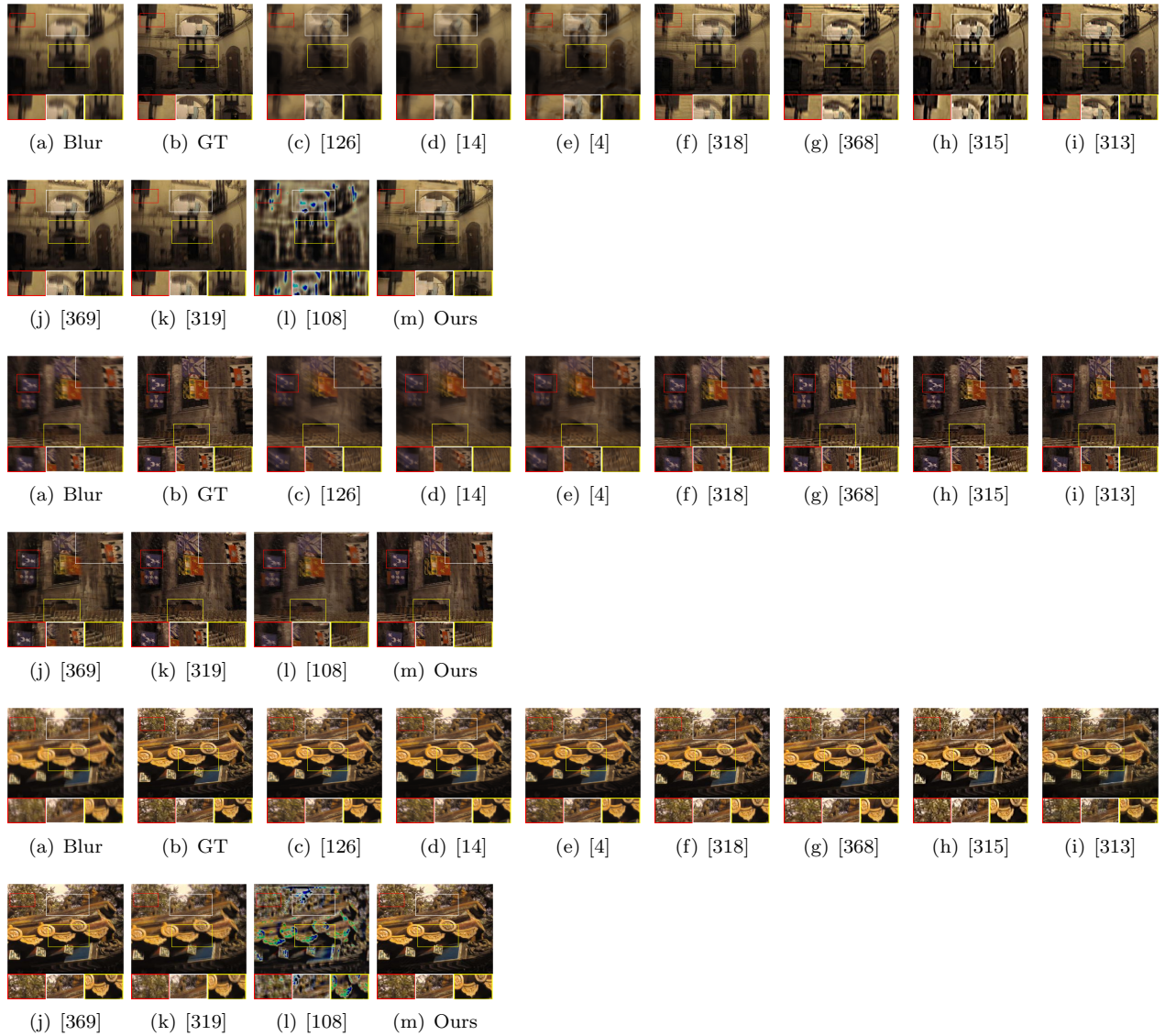$$G = \sqrt{\left(I_x^{'}\right)^2 + \left(I_y^{'}\right)^2}$$

(A.11)

Measuring image gradient assesses how much image has been enhanced or deblurred compared to the original image.

A.1.2 Processed Images

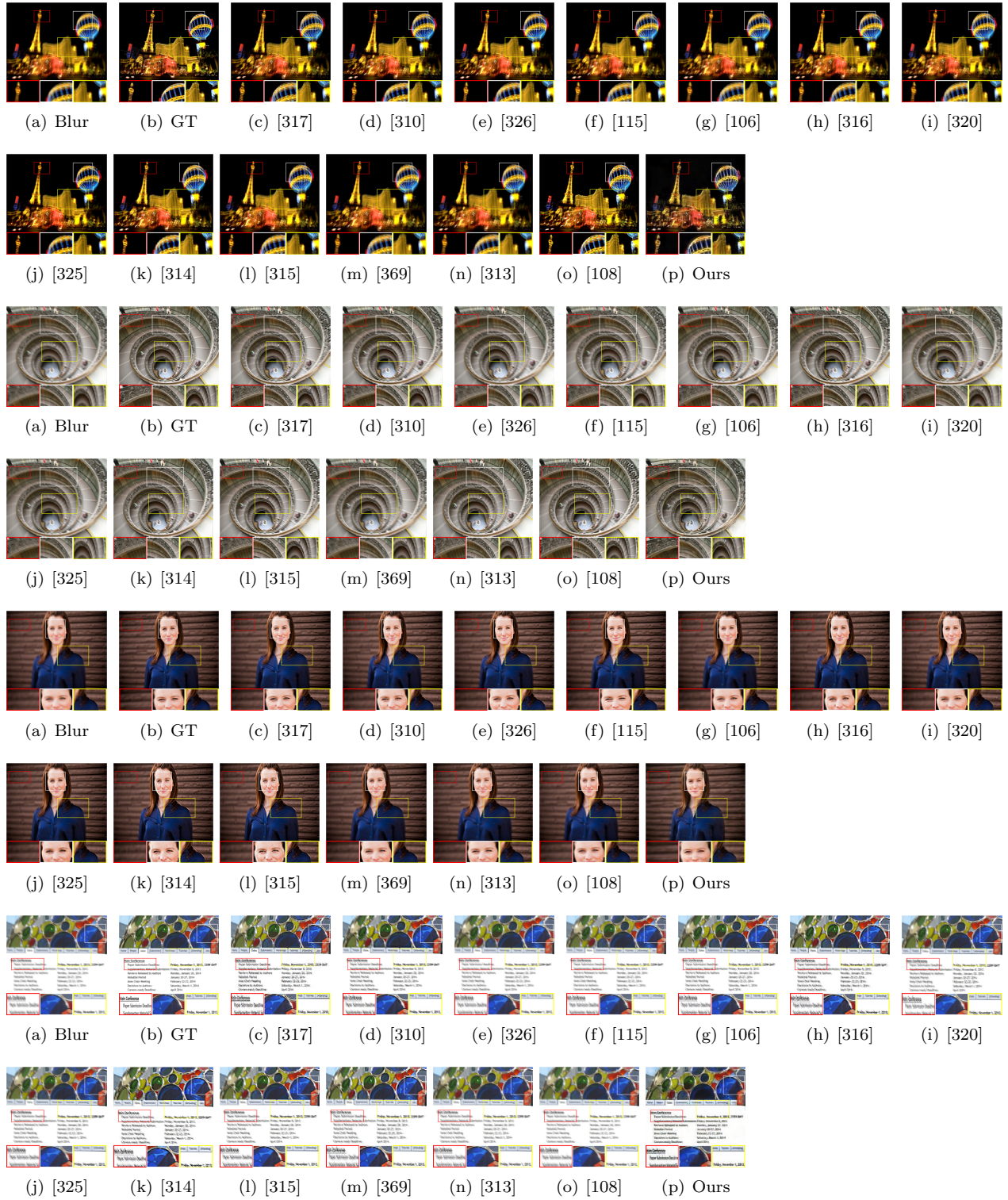The deblurred images for the three datasets GoPro, Köhler, and Lai can be acquired from [361, 5, 6].

(a) Blur   (b) GT   (c) [343]   (d) [14]   (e) [126]   (f) [344]   (g) [4]   (h) [125]   (i) [327]

(j) [129]   (k) [106]   (l) [325]   (m) Ours

(a) Blur   (b) GT   (c) [343]   (d) [127]   (e) [14]   (f) [126]   (g) [344]   (h) [4]   (i) [125]

(j) [327]   (k) [129]   (l) [106]   (m) [325]   (n) Ours
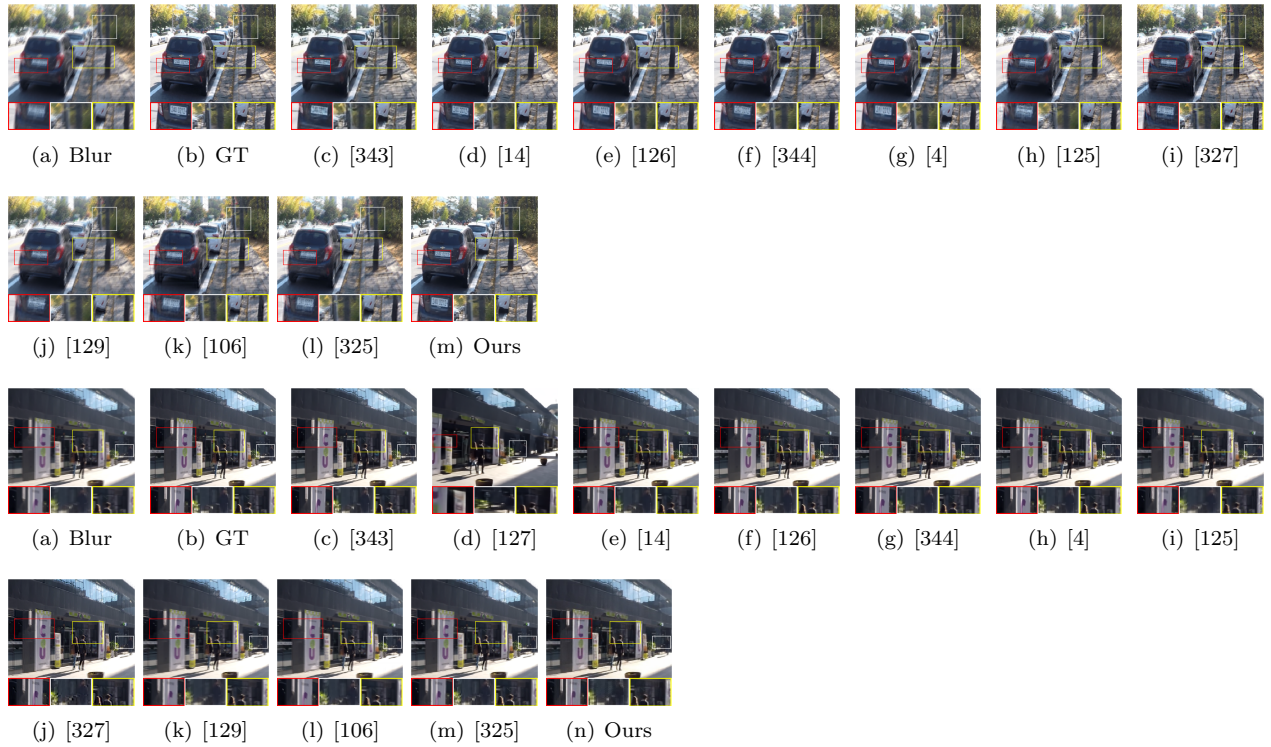
Figure A.1: Visual deblur results of various studies for GoPro [14] dataset. GT represents ground truth image.

(a) Blur   (b) GT   (c) [126]   (d) [14]   (e) [4]   (f) [318]   (g) [368]   (h) [315]   (i) [313]

(j) [369]   (k) [319]   (l) [108]   (m) Ours

(a) Blur   (b) GT   (c) [126]   (d) [14]   (e) [4]   (f) [318]   (g) [368]   (h) [315]   (i) [313]

(j) [369]   (k) [319]   (l) [108]   (m) Ours

(a) Blur   (b) GT   (c) [126]   (d) [14]   (e) [4]   (f) [318]   (g) [368]   (h) [315]   (i) [313]

(j) [369]   (k) [319]   (l) [108]   (m) Ours

Figure A.2: Visual deblur results of various studies for Köhler [5] dataset. GT represents ground truth image.

(a) Blur (b) GT (c) [317] (d) [310] (e) [326] (f) [115] (g) [106] (h) [316] (i) [320]

(j) [325] (k) [314] (l) [315] (m) [369] (n) [313] (o) [108] (p) Ours

(a) Blur (b) GT (c) [317] (d) [310] (e) [326] (f) [115] (g) [106] (h) [316] (i) [320]

(j) [325] (k) [314] (l) [315] (m) [369] (n) [313] (o) [108] (p) Ours

(a) Blur (b) GT (c) [317] (d) [310] (e) [326] (f) [115] (g) [106] (h) [316] (i) [320]

(j) [325] (k) [314] (l) [315] (m) [369] (n) [313] (o) [108] (p) Ours

(a) Blur (b) GT (c) [317] (d) [310] (e) [326] (f) [115] (g) [106] (h) [316] (i) [320]

(j) [325] (k) [314] (l) [315] (m) [369] (n) [313] (o) [108] (p) Ours

Figure A.3: Visual deblur results of various studies for Lai et al.[6] dataset. GT represents ground truth image.

Appendix B

List of Publications

1. **Rayyan Azam Khan**, Yigang Luo, Fang-Xiang Wu, Machine learning based liver disease diagnosis: A systematic review, *Neurocomputing* 468 (2022) 492509. DOI:10.1016/j.neucom.2021.08.138.

2. **Rayyan Azam Khan**, Yigang Luo, Fang-Xiang Wu, Multi-scale GAN with residual image learning for removing heterogeneous blur, *IET Image Processing* 16 (9) (2022) 24122431. DOI:10.1049/ipr2.12497.

3. **Rayyan Azam Khan**, Yigang Luo, Fang-Xiang Wu, Multi-level GAN based enhanced CT scans for liver cancer diagnosis, *Biomedical Signal Processing and Control* 81 (2023) 104450. DOI:10.1016/j.bspc.2022.104450.

4. **Rayyan Azam Khan**, Yigang Luo, Fang-Xiang Wu, RMS-UNet: Residual multi-scale UNet for liver and lesion segmentation, *Artificial Intelligence in Medicine* 124 (2022) 102231. DOI:10.1016/j.artmed.2021.102231.

5. **Rayyan Azam Khan**, Burbridge Brent, Yigang Luo, Fang-Xiang Wu, A multi-modal deep neural network for multi-class liver cancer diagnosis, *Neural Networks* (2022) (Under Review).

6. Baig Zarrukh*, Nawaf Abu-Omar*, **Rayyan Khan***, Carlos Verdiales, Ryan Frehlick, John Shaw, Fang-Xiang Wu, and Yigang Luo., Prognosticating outcome in pancreatic head cancer with the use of a machine learning algorithm, *Technology in cancer research & treatment* 20, (*Authors contributed equally) (2021). DOI:10.1177/15330338211050767.

137

Appendix C

Copyrights Permissions

CCC
RightsLink

Multi-scale GAN with residual image learning for removing heterogeneous blur

Author: Fang-Xiang Wu, Yigang Luo, Rayyan Azam Khan
Publication: IET Image Processing
Publisher: John Wiley and Sons
Date: Apr 21, 2022

© 2022 The Authors. IET Image Processing published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

Order Completed

Thank you for your order.

This Agreement between Mr. Rayyan Khan ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

🖨 Printable Details

License Number	5333471471327
License date	Jun 21, 2022

☑ Licensed Content

Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	IET Image Processing
Licensed Content Title	Multi-scale GAN with residual image learning for removing heterogeneous blur
Licensed Content Author	Fang-Xiang Wu, Yigang Luo, Rayyan Azam Khan
Licensed Content Date	Apr 21, 2022
Licensed Content Volume	16
Licensed Content Issue	9
Licensed Content Pages	20

📋 Order Details

Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No

📄 About Your Work

Title	Deep Learning-based diagnostic system for malignant liver detection
Institution name	University of Saskatchewan
Expected presentation date	Sep 2022

📁 Additional Data

Order reference number	GAN1

📍 Requestor

Mr. Rayyan Khan

📑 Tax Details

Publisher Tax ID	EU826007151

$ Price

Total	0.00 USD

Would you like to purchase the full text of this article? If so, please continue on to the content ordering system located here: **Purchase PDF**
If you click on the buttons below or close this window, you will not be able to return to the content ordering system.

Total: 0.00 USD

CLOSE WINDOW   **ORDER MORE**

Multi-level GAN based enhanced CT scans for liver cancer diagnosis

Author: Rayyan Azam Khan,Yigang Luo,Fang-Xiang Wu

Publication: Biomedical Signal Processing and Control

Publisher: Elsevier

Date: March 2023

© 2022 Elsevier Ltd. All rights reserved.

Journal Author Rights

BACK   **CLOSE WINDOW**

RMS-UNet: Residual multi-scale UNet for liver and lesion segmentation

Author: Rayyan Azam Khan,Yigang Luo,Fang-Xiang Wu

Publication: Artificial Intelligence in Medicine

Publisher: Elsevier

Date: February 2022

© 2022 Elsevier B.V. All rights reserved.

Journal Author Rights

BACK   **CLOSE WINDOW**

Special Requests > Special Request Details

Technology in cancer research & treatment

Article: Prognosticating Outcome in Pancreatic Head Cancer With the use of a Machine Learning Algorithm

GENERAL INFORMATION

Request ID	600085187	**Request Date**	21 Jun 2022
Request Status	Denied	**Deny Reason**	Permission may not be needed for this request

ALL DETAILS

ISSN:	1533-0346	**Publisher:**	SAGE Publications, Inc
Type of Use:	Republish in a thesis/dissertation	**Portion:**	Chapter/article

LICENSED CONTENT

Publication Title	Technology in cancer research & treatment	**Rightsholder**	Sage Publications Inc. Journals
Article Title	Prognosticating Outcome in Pancreatic Head Cancer With the use of a Machine Learning Algorithm	**Publication Type**	Journal
		Start Page	153303382110507
		Volume	20
Date	01/01/2002		
Language	English		
Country	United States of America		

REQUEST DETAILS

Page range(s)	1-8	**Distribution**	Worldwide
Total number of pages	8	**Translation**	Original language of publication
Format (select all that apply)	Print, Electronic	**Copies for the disabled?**	Yes
Who will republish the content?	Author of requested content	**Minor editing privileges?**	Yes
Duration of Use	Life of current edition	**Incidental promotional use?**	Yes
Lifetime Unit Quantity	Up to 499	**Currency**	CAD
Rights Requested	Main product and any product related to main product		

NEW WORK DETAILS

Title	Deep Learning-based diagnostic system for malignant liver detection	**Institution name**	University of Saskatchewan
Instructor name	FangXiang Wu	**Expected presentation date**	2022-09-15

ADDITIONAL DETAILS

The requesting person / organization to appear on the license	Rayyan Khan

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Deep Learning-based diagnostic system for malignant liver detection	**Title of the article/chapter the portion is from**	Prognosticating Outcome in Pancreatic Head Cancer With the use of a Machine Learning Algorithm
Editor of portion(s)	Baig, Zarrukh; Abu-Omar, Nawaf; Khan, Rayyan; Verdiales, Carlos; Frehlick, Ryan; Shaw, John; Wu, Fang-Xiang; Luo, Yigang	**Author of portion(s)**	Baig, Zarrukh; Abu-Omar, Nawaf; Khan, Rayyan; Verdiales, Carlos; Frehlick, Ryan; Shaw, John; Wu, Fang-Xiang; Luo, Yigang
Volume of serial or monograph	20		
Page or page range of portion	153303382110507	**Publication date of portion**	2021-11-04

COMMENTS

Add Comment / Attachment

21 Jun 2022 12:30:37 PM, by Mary Ann Price

Thank you for your order. This use is covered under SAGE's journal author reuse policy (https://us.sagepub.com/en-us/nam/journal-author-archiving-policies-and-re-use) which allows you to include the published version of your article (version 3) in your dissertation or thesis, which may be posted in an Institutional Repository or database. Please note that this permission does not cover any 3rd party material that may be found within the work.
In addition, your article is is published under a CC-BY Creative Common License. This allows you to copy and redistribute the material in any medium or format without permission or a fee. However, you will need to cite the source, link to the SAGE article, cite the license, and link to the CC-BY license as well. As permission is pre-approved, we will cancel this order. If you require formal written permission, please place an order via our Customer Permission Portal (https://jira.sagepub.com/servicedesk/customer/portal/9)

Kind regards,
The SAGE Permission Team

View Less