T.C.

ALTINBAS UNIVERSITY

Institute of Graduate Studies

Information Technologies

AUTOMATED LIVER TUMOR SEGMENTATION AND CLASSIFICATION USING AI AND FULLY CONVOLUTIONAL NEURAL NETWORK (CNN)

Idrees Ahmed Dhahir DHAHIR

Master of Science

Supervisor

Assoc. Prof. Dr. Yasa EKŞİOĞLU ÖZOK

Co-Supervisor

Assoc. Prof. Dr. Adil Deniz DURU

Istanbul, 2021

AUTOMATED LIVER TUMOR SEGMENTATION AND CLASSIFICATION USING AI AND FULLY CONVOLUTIONAL NEURAL NETWORK (CNN)

by

Idrees Ahmed Dhahir DHAHIR

Information Technologies

Submitted to the Institute of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2020

The thesis titled "AUTOMATED LIVER TUMOR SEGMENTATION AND CLASSIFICATION USING AI AND FULLY CONVOLUTIONAL NEURAL NETWORK (CNN)" prepared and presented by Idrees Ahmed Dhahir DHAHIRwas accepted as a Master of Science Thesis in Information Technologies.

Assoc. Prof. Dr. Adil Deniz DURU Co-Supervisor	Assoc. Prof. Dr. Yasa EKŞİOĞLU ÖZOK Supervisor

Thesis Defense Jury Members:

Assoc. Prof. Dr. Yasa EKŞİOĞLU ÖZOK	School of Engineering and Natural Sciences, Altinbas University	_____
Asst. Prof. Dr. Doğu Çağdaş ATİLLA	School of Engineering and Natural Sciences, Altinbas University	_____
Asst. Prof. Dr. Aydın Tarık ZENGİN	Faculty of Engineering and Natural Sciences, İstanbul Sabahattin Zaim University	_____

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Approval Date of Institute of Graduate Studies: ____/____/____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Idrees Ahmed Dhahir DHAHIR

DEDICATION

I devote and pledge this research work to my supervisor who is salient for guiding me through whole research work as well as my family for always assisting me in my hard time.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Assoc. Prof. Dr. Yasa EKŞİOĞLU ÖZOK for guiding and helping me along the way in writing this dissertation. Discussing my progress, problems, and ideas with my supervisor Assoc. Prof. Dr. Yasa EKŞİOĞLU ÖZOK a couple of times every week helped me tremendously in understanding the logic behind the research. It made me better realize the technical need for this research work.

ABSTRACT

AUTOMATED LIVER TUMOR SEGMENTATION AND CLASSIFICATION USING AI AND FULLY CONVOLUTIONAL NEURAL NETWORK (CNN)

DHAHIR, Idrees Ahmed Dhahir,

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Assoc. Prof. Dr. Yasa EKŞİOĞLU ÖZOK

Date: March/2021

Pages: 53

The aim of this thesis is to design and implement automated liver tumor segmentation and classification using AI and fully convolutional neural network (FCNN). This represents more of a challenge when speaking of advance object recognition systems. A practical example of this issue is the recognition of objects in images. This is a task that humans can perform very well, but Fully Convolutional Network (FCN) systems don't struggle to perform. Visual Geometry Group (VGG-16) pre-trained model was used for the training the dataset because of it trouble-free architecture on very large scale dataset "LiTS dataset" was used in R2019a Matlab. The dataset was split with the ratio of 70% for training and 30% for the testing part. This has prompted fully convolutional neural network to start experimenting with networks architectures as well as new algorithms to train them. This research thesis presents an approach to train networks such as improving their robustness to the recognition of liver tumorous images on R2019a Matlab. This training strategy is then evaluated for designed VGG-16 network architecture. The result of the study was that the training algorithm could improve robustness to different liver lesion sample recognition at the expense of an increase in performance for the performance of images of diseases with high accuracy of segmentation and classification. When the advantages of different types of

architectures were evaluated, it was found that accuracy of liver tumor were around 91.28% based on the liver lesion image with 300 epochs of training taking 28-hours with 10-cross fold validation. The research innovatively combines VGG-16 and the fully automated FCN model for liver tumor segmentation. Secondly, in regards of the label imbalance issue in medical imaging, we integrate object detection techniques into our semantic segmentation framework for the purpose of locating ROIs and performing local segmentations. This lesion localization step imposes an additional level of supervision on training the segmentation networks.

Keywords: Liver, Tumor, Segmentation, Classification, Deep Learning, Optimization, AI, Medical Imaging.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

LIST OF ABBREVIATIONS

AI	:	Artificial Intelligence
ANN	:	Artificial Neural Network
CV	:	Computer Vision
CNN	:	Convolutional Neural Network
DL	:	Deep Learning
FCNN	:	Fully Convolutional Neural Network
GPU	:	Graphic Processing Unit
LTI	:	Liver Tumor Identification
MRI	:	Magnetic Resonance Imaging
ML	:	Machine Learning
ROI	:	Region Of Interest
SVM	:	Support Vector Machine
VGG	:	Visual Geometry Group

1. INTRODUCTION

The most obvious distinction is the data format. Natural images are commonly 2D RGB images with 3 channels, whereas medical images are 2D grayscale images with only 1 channel. Some medical data are formatted as 3D volumes or even 4D (volumes changing over time). The volumes can be processed directly as a 3D tensor, or sliced into multiple 2D images for further analysis.

The variation in intensity is another important difference between natural and medical images. For natural images, the exact intensity of objects is usually not considered as a determinative feature. For example, a cat should be recognized as a cat, regardless of its color, or the contrast and saturation level of the image itself. Therefore, in natural image processing, edges and gradients are captured as features more often. On the contrary, for medical images, the exact pixel values actually convey the information of the tissues. For instance, in CT scans, the commonly used intensity scale is the Hounsfield scale [1], in which air has a value of -1000, fat has a value between -90 and 130, and other tissues have higher values. In tasks involving anomaly detection, the pathological regions usually differ from the context in terms of pixel intensity.

The impact of orientation is also a major distinction. Objects in natural images usually have canonical orientations. For example, in street view images, pedestrians have their heads on the top and feet at the bottom, vehicles have their wheels below bodies as described in [2]. As showed in Figure 1.1. Nevertheless, objects in medical images can emerge without canonical orientations. Especially, the abnormalities often do not have a certain structure, and there is no predefined alignment between their placements.

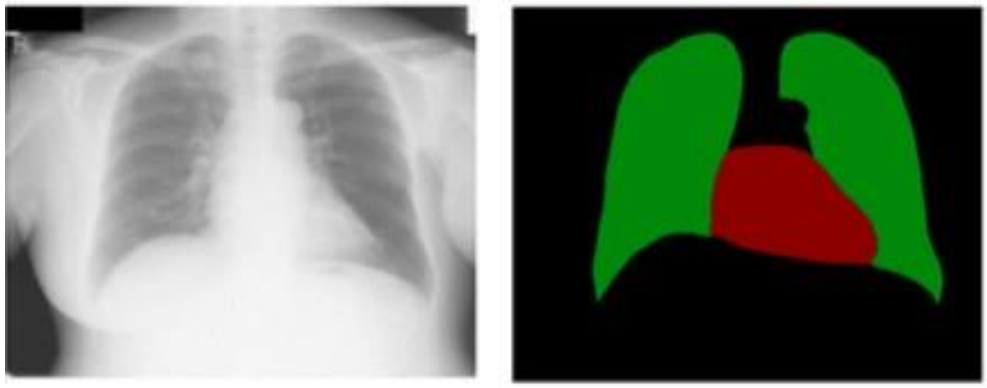


Figure 1.1: Example of organ segmentation in chest X-ray images [2].

Medical imaging semantic segmentation is often associated with the label imbalance problem. For example, in tasks involving segmenting tumors, the size of lesions can be very tiny, and abnormal cases may only take up a small portion of the dataset as described in [3]. In this scenario, there are much less tumor samples compared to non-tumor ones, which increases the difficulty of capturing the features of tumors.

1.1 PROBLEM STATEMENT

There are many related issues of automated liver tumor classification. The most widely recognized ones are:

- **Liver Tumor Localization:** Means to decide the picture position of each single location of liver tumors.
- **Liver Tumor Identification:** Plans to recognize the nearness and area of highlights, for example, the area where the tumor lesions are located compared to non-tumor ones would be used with the Liver Tumor Identification (LTI).
- **Liver Tumor Recognizable Proof:** Thinks about an info picture against a database and reports a match, assuming any.
- **Liver Tumor Verification Issue:** Checks the case of the normalization of liver tumor in an information picture.
- **Liver Tumor Articulation Acknowledgment:** Concerns recognizing the full of states (To check if the lines on tumor lesions changes over time in different images) of different people.

Obviously, automated liver tumor lesion classification is the finalized phase in any mechanized framework which takes care of the above issues.

1.2 MOTIVATION

In spite of the distinctions mentioned above, the approaches to solving semantic segmentation problems of natural and medical images are generally the same. Before deep learning became predominant for computer vision, most successful models for semantic segmentation were

classifiers based on hand crafted features. These features were basically derived from the texture of the image, and obtained by applying multiple filters. TextonBoost [4], texton forests [4] and other frameworks based on decision trees and random forests were the standard approaches to semantic segmentation problems in the past.

The first attempt of applying deep learning models in computer vision problems was to tackle image classification problems using convolutional neural networks (CNNs). In AlexNet [5] won the ImageNet competition [5], and improved the classification accuracy by a large scale. In the following years, a number of CNNs with various structures were proposed, including VGG [6], GoogLeNet [6], and ResNet [6] etc., which further updated the state- of-the-art performance subsequently.

Deep learning is characterized for its automatic feature extraction, thus combines feature engineering and traditional machine learning. Upon their success in image classification tasks, CNNs were soon applied to the field of semantic segmentation, replacing the old mode of designing hand engineered features in conjunction with decision forests. The fully convolutional network (FCN) was the first end-to-end deep learning model that can perform dense prediction, and its segmentation accuracy was revolutionary compared to traditional methods. Motivated by FCN, many networks adopting similar structures were proposed afterwards, such as SegNet, RefineNet, and PSPNet [7]. These networks use an encoder-decoder structure, and apply convolutional layers in the base networks trained for image classification tasks as feature extractors[8].

Variations of FCNN have also become the backbone model of medical image semantic segmentation. U-Net [9] fuses the multiscale learned features in the decoding process and incorporates low level features into the inference in order to achieve a deeper supervision, which is considered critical to improving the segmentation accuracy of medical images [9]. Following U-Net, V-Net [10] was proposed to process 3D medical image volumes directly, and it introduced the Dice loss function to tackle the label imbalance problem. Moreover, the attention mechanism was integrated into U-Net in [11], which allowed the model to implicitly learn to focus on relevant regions for pancreas segmentation problems.

Even though having greatly enhanced the segmentation accuracy and accelerated the learning process, models built upon FCN still have limitations. As summarized in [12], FCN models perform dense prediction by independently classifying each pixel, through decoding the features extracted by the encoder. However, this independent per-pixel inference scheme sometimes disobeys common visual characteristics. For example, linear classifiers like FCN can predict different labels for similar pixels in the body of the same object. Moreover, in FCN models, the spatial information loss caused by dimension reduction in the encoding process often leads to the loss of segmentation precision at boundaries between different objects. Conditional random fields (CRFs) are a type of undirected graphical models, that allow prior knowledge of a good segmentation output to be included in the modelling. By integrating CRFs into the segmentation framework, the interactions between pixels can be explicitly modelled, serving as a complement to the linear classifier. Hence, CRFs are typically applied as a post-processing tool to enhance the segmentation output produced by FCN models.

CRFs were widely used in image semantic segmentation even before the era of deep learning. Initially, simpler models like grid graph CRFs [12, 13] were favored, as there were only efficient inference algorithms for them at that time. Once the connections became denser, the complexity of the algorithm would rapidly grow. In 2016, researchers [14] popularized the fully connected CRF model by adopting an approximating inference algorithm accompanied with fast high-dimensional filtering [15], allowing for more efficient message passing. The linear classifiers like FCN can predict different labels for similar pixels in the body of the same object. From then on, it became the standard CRF model for most segmentation frameworks.

The first attempt of combining CNNs and CRFs was made in [15]. By using the per-pixel labelling distribution predicted by a FCN model to compute the unary potentials in a fully connected CRF, it showed that CRFs could significantly improve the segmentation accuracy, compared to other deep learning models without post-processing. Researcher in [16] further consolidated CRFs and CNNs into a system that could be trained end-to-end, which enabled the CRF parameters to be learned by the algorithm, delivering additional improvements to the performance. Up to now, the mode of combining CNNs and CRFs has become one of the most popular approaches to semantic

segmentation problems, and this applies to medical data as well. Researchers also used FCN models and 3D fully connected CRFs on liver lesion segmentations, respectively.

1.3 DESIGN GOALS

Our system is designed to maximize user experience. Moreover, our system differs from existing systems in the following ways.

Advanced Algorithm: Many existing systems uses old and obsolete techniques of machine learning and data mining however this research is full automated by the deep learning. We feel this limits the usability of the application and aim to keep the prediction of the system entirely marker free.

Real-time: Our system should run smoothly on average modern computers with a dedicated graphics card. The system should also recognize liver tumors at a high frame rate, nearing real-time speeds. Our desired frame rate 30Hz; we achieved a frame rate around 10Hz. In our design and implementation, one driving goal is to squeeze every millisecond as possible.

No calibration: Once trained, our system should require no further calibration before working in a new environment for liver tumor classification.

Robust and accurate: Our system should have an accurate estimation of where the tumor is located and classify them with the use of low false positive. Moreover, the system should be insensitive to various background, user's location, tumor position and other noise.

Arbitrary Tumor Identification: Our system should be able to easily incorporate new types of tumor identification. With sufficient training, our system should support many complex tumor identifications such as those seen in the inner chest area of patient.

1.4 AIM OF CONTRIBUTION

We summarize our contributions as follows:

The main contributions of this thesis is 10-fold. Firstly, we innovatively combine VGG-16 and the fully automated FCN model, and propose a new type of energy function for liver tumor

segmentation. We also solve for the exact solution with applying a new technology. Secondly, in regards of the label imbalance issue in medical imaging, we integrate object detection techniques into our semantic segmentation framework for the purpose of locating ROIs and performing local segmentations. This lesion localization step imposes an additional level of supervision on training the segmentation networks. The results of the liver tumor segmentation task show that the proposed formulated automated CNN can effectively enhance the segmentation accuracy. The results of the liver tumor segmentation task strengthen this enhancement, and show that the overall performance produced by combining the formulated CNN post-processing and lesion localization is better than conventional methods. This weighted information is the basis for the second tree. The concept here is to enhance the first image projections. The main challenge of the research work was to manage dataset as medical dataset is not publicly accessible in our country and dataset that we have used is not big enough. To arrange this, finding the hyperplane which is good at separating the different classes. It is a double straight classification non-probabilistic that frequently controls how-ever with an auxiliary non-straight and mathematical description and a robust program for the algorithm. Although the limitation, survey managed to solve four case studies and showed the best outputs in tree based structured deep learning and machine learning algorithms. Therefore, deep learning-based algorithms are our new model.

1.5 STRUCTURE OF THESIS

The research structure presented in this thesis is organized as follows:

Chapter 1 introduce the domain and provides problem statement for the research and draws the significance of research that contributes to the advancement of liver tumor classification technology. The aim of the study and the technical approaches are also presented in this chapter.

Chapter 2 provides an overview of liver tumor classification operation and a comprehensive literature review on relevant aspects of liver spectrum access in current liver tumor classification network technology and its deployment scenarios. This chapter focuses on the challenges and solution branches of the access mechanisms. The required branches of a liver tumor classification network, such as the network architecture, CNN modeling, sensing, and access methods, are thoroughly reviewed along with the current challenges and future trends.

Chapter 3 shows the impacts of liver lesion identification and verification methods on the design of the CNN operation through a capacity measurement of the CNN method opportunity. The underlying obstacles of the sensing-throughput trade-off, issue is synthesized with the analysis of the liver ROI and access probability. To achieve the highest possible liver tumor classification results, a novel sensing mechanism is proposed by differentiating the target liver lesions. And provides the design aspects of the CNN based methodology in order to achieve the highest possible throughput. The achievable throughput by CNN is formulated based on the proposed automated CNN mechanism. To obtain maximum throughput under the constraint of protection, an optimization is conducted between the sensing period and the detection probability of the liver tumor segmentation mechanism. This optimization shows the significance of the liver classification mechanism for designing.

Chapter 4 provides the simulation-based results for designing automated liver classification. The CNN based framework of the identification and verification mechanism is described, with consideration of practical liver tumor lesion identification scenario. The derivation of the detection sensitivity configuration according to the contention parameters is provided in this chapter. The achievable throughput is formulated with the automated CNN mechanism.

Chapter 5presents a complete discussion of performed experiments for a liver tumor classification network with existing techniques in literature review. The automated liver tumor classification framework exhibited in chapter 4 is exploited with the contribution of embedding the classification and doing comparison of accuracy with existing literature.

Chapter 6 presents the conclusions of this thesis along with the essentials of the tasks accomplished and the technical contributions made. In addition, recommendations are made for the further improvement of the access strategies in liver tumor classification technology.

2. RELATED WORK

2.1 LITERATURE REVIEW

Analyzing liver regions various aspects of the image are used in previous research. The most common type of image imaging in Magnetic Resonance Imaging (MRI) describes the location patterns. Text features typically consist of Gray Level Co-occurrence Matrix (GLCM) measurements, gradient histograms and features calculated using the Discrete Wavelet Transform (DWT) [17]. Text measurements are often used in isolating tissue or other abnormalities in body parts. Common methods of machine learning include Multi-Layer Perceptron's (MLP), Support Vector Machines (SVM) and Artificial Neural Networks (ANN) [18].

The author in the article [19] performed brain tumor classification using dwt image features fed into the neural network. Image dimensions are reduced by key analysis (PCA) and understanding of the stages of progression as Figure 2.1 from [20] describes the progression of liver disease. They achieve high segmentation accuracy with a neural network consisting of a single hidden layer and a low number of nodes.

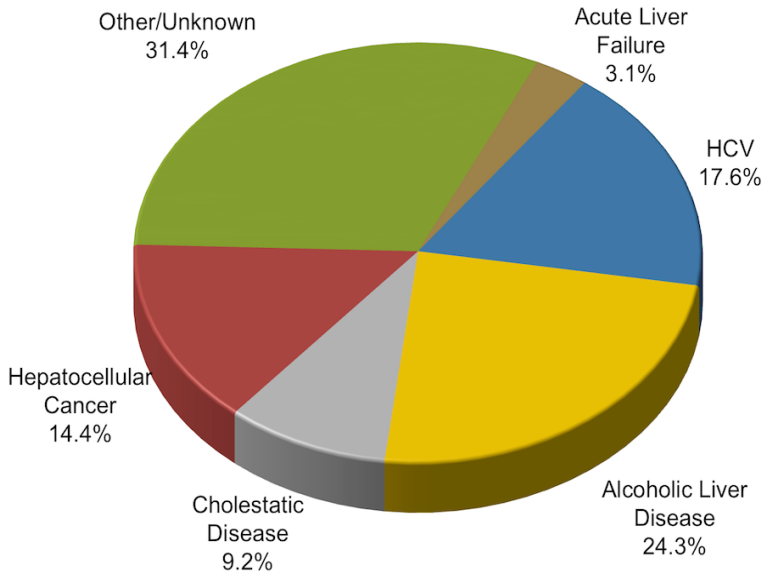


Figure 2.1: Progression of liver diseases and failures with different types [20].

The author in the article [21] conducts research by diagnosing cirrhosis liver using gradient magnitude histograms and GLCM tissue analysis in the subset ROI of the MR image. The learning approach was used to integrate informal learning k and gained sensitivity and clarity of 72% and 60% respectively.

The author in the article [22] demonstrates that the analysis of tissue formation and detection of liver lesions on MRI scans using GLCM, DWT and gray gradient level integration are important techniques that achieve good results. They compared several segregation mechanisms including Neural Networks, Support Vector Machines and k-Nearest Neighbor functioning [23].

The authors in the article [24] used image features based on histograms of DWT level and gray to locate liver regions. Appropriate features are obtained using a genetic algorithm inspired by the law of evolution by liver anatomy shown in Figure 2.2 [25]. The characteristic vector used is inserted into the neural network and they have a pixel segment accuracy of 86.2%.

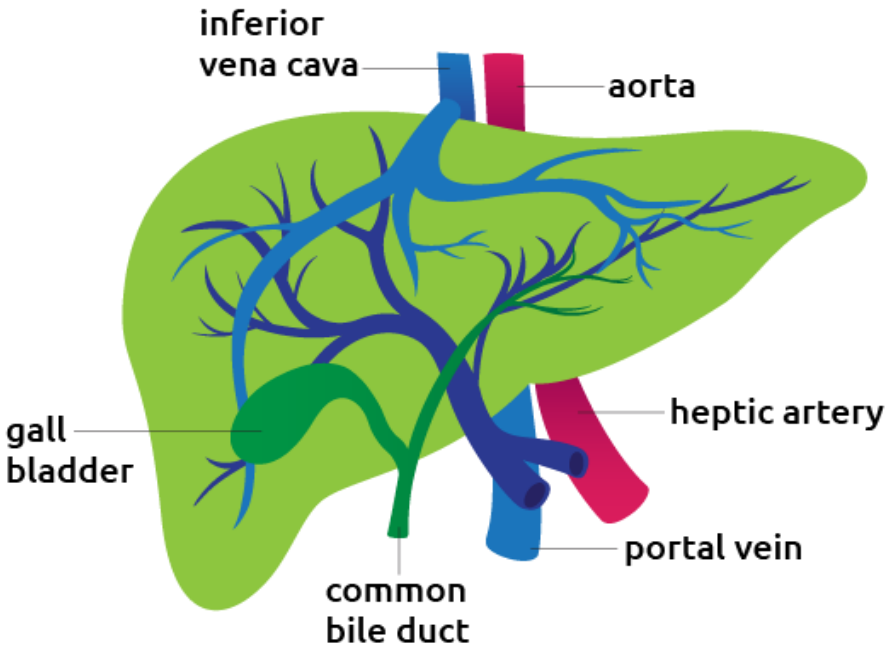


Figure 2.2: Anatomy of normal human liver with arteries [25].

In this article [27], the authors discussed the current methods and suggested improved technology based on the deep learning based (Extreme Gradient Boost), which combined significant characteristics of the F-scores and evaluated four pre-processing scenarios. In addition, we provided machine training methods for anticipating chronic renal disease with clinical information. Four techniques of master teaching are explored including Support Vector Regression (SVR), logistic Repressor (LR), Ada-Boost, Gradient Boosting Tree and Deep Learning Regression. The components are made from MRI dataset of chronic brain disease and the results of these models are compared to determine the best regression model for the prediction. From this four preprocessing cases, replacing missing values with mean values of each column and choosing important features was most logical as it allows to train with more data without dropping. However, XG-Boost gave the best outcomes in all four cases where it obtained 98% accuracy in case one where nulled valued are dropped, 98.75% testing accuracy for both case two and three where null values were replaced with minimum and maximum values of each column and it scores 100% accuracy in case four where null values are replaced with mean values as mentioned in [28]. It is implemented by dividing the information into two or more subsets that support info factor estimates. The most divided areas can be seen by an estimated function or a discordant foundation.

2.2 LIVER REGION OF INTEREST

Defining Region of Interest (ROI) is the process of separating the desired region or object of interest from the whole background for the further process of feature extraction. ROI in medical images can be defined using both manual or semi-automatic method and automatic method. Manual or semi-automatic method requires user interaction while fully automatic methods is free of user intervention as described in [29]. Generally, a rough image segmentation is performed for defining ROI in medical images. Some of the common methods for defining ROI are thresholding, region growing and boundary tracking, classifier methods, deformable models and atlas-guided methods etc.

Thresholding determines an intensity value either locally or globally to separate the desired object from background. It usually separates the pixels with certain range of intensity as one class and remaining as other class. Though it is simple and a powerful technique, it has its own limitation of inability to take in account of spatial characteristics of image, sensitivity to noise and intensity

inhomogeneity [30]. Region growing extracts the connected regions of image based on predefined condition: intensity, image feature etc. [31]. It begins from an initial point called seed point and stops on reaching the stopping criterion. Atlas guided method segments organs using information about the anatomy of interest termed as atlas of anatomy [32]. In this process, the atlas images are transformed using linear, non-linear or a combination of multiple linear or non-linear transformations to real images and segmentation is concerned with determining resemblance between two objects of similar objects. This method is only appropriate when the structures are consistent over slices [33]. Classification is the process of separating a given object into different classes. The classification step classifies the extracted and selected features obtained from previous step into different classes. The classification can be both supervised and unsupervised as described in [34]. In supervised method, the features sets are classified into predefined classes while in unsupervised method, the feature sets are assigned to unknown classes with stages of liver damage is displayed in Figure 2.3. A classifier always requires a training and testing data to train and evaluate the performance of classifier. The data for training in medical images are usually obtained from one or more experts who has assigned labels to a set of objects.

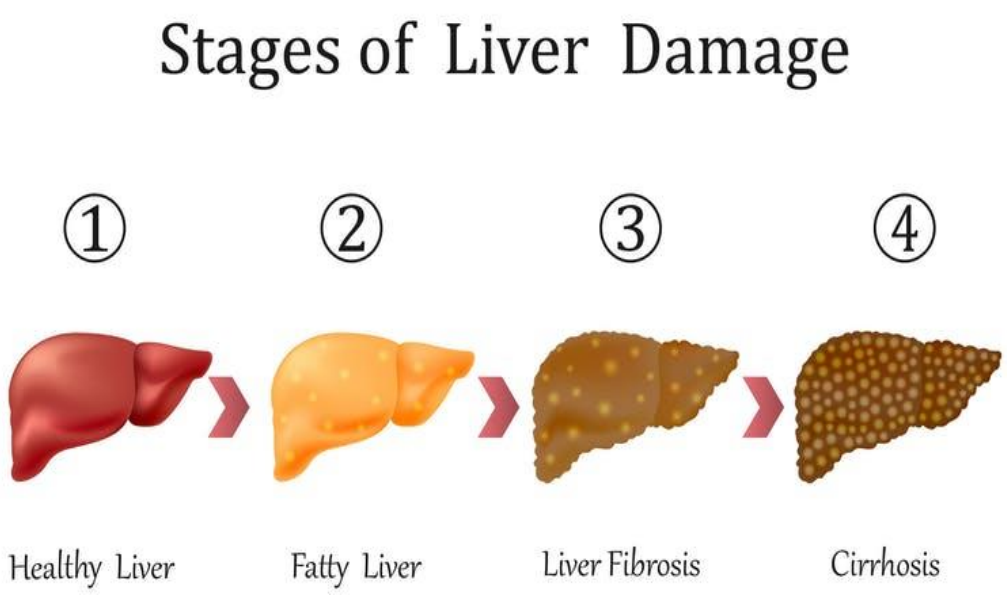


Figure 2.3: Different stages of liver damage from healthy liver to the cirrhosis liver [35].

2.3 EXISTING PREDICTIVE FEATURES

A large number of research has been done for liver and liver tumor segmentation based on semi-automatic, automatic and manual segmentation techniques [36]. While manual segmentation focuses on different segments or sectors or hemi-liver or vessels segmenting liver, semi-automatic and automatic technique are focused on different computer based algorithm to facilitate segmentation of liver or tumor from medical images with more or less or null user interaction [37]. There is always active participation of expert or doctor required to delineate liver or tumor from medical images in either of manual (with complete involvement) or semi-automatic method (partial involvement). The following sections provides an overview of the literature or related work done in the field of liver segmentation and tumor segmentation with the ultrasound of liver is given in Figure 2.4 [38].

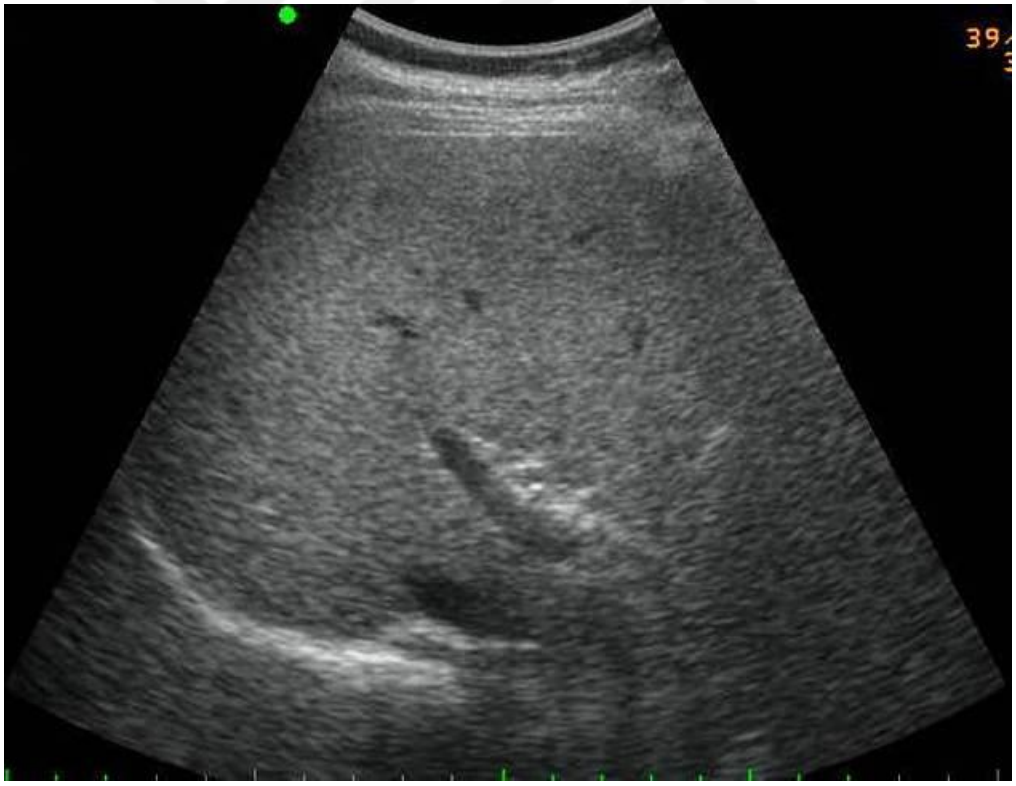


Figure 2.4: Ultrasound of human healthy liver with no preliminary diseases [38].

A general Computer Aided Diagnosis (CAD) System uses computerized analysis to detect and characterize anomaly from medical images. These analyses are based on certain techniques or

methods and helps radiologists or doctors to detect lesions, access the extent of disease and make diagnostic decision. CAD system have found a lot of applications in various imaging modalities like CT, MRI and US to address various diagnostic problems [39]. Image processing, computer vision, pattern recognition, mathematics, Artificial Intelligence (AI) technology and statistics are integrated into CAD system to assist medical personnel in decision making. A general overview of a CAD system for liver and tumor segmentation as described in [40]. Spatial domain approach and Spectral domain approach are two of the methods available for removing noise in present medical images with the percentage change in European age with mortality rate is given in Figure 2.5 [41].

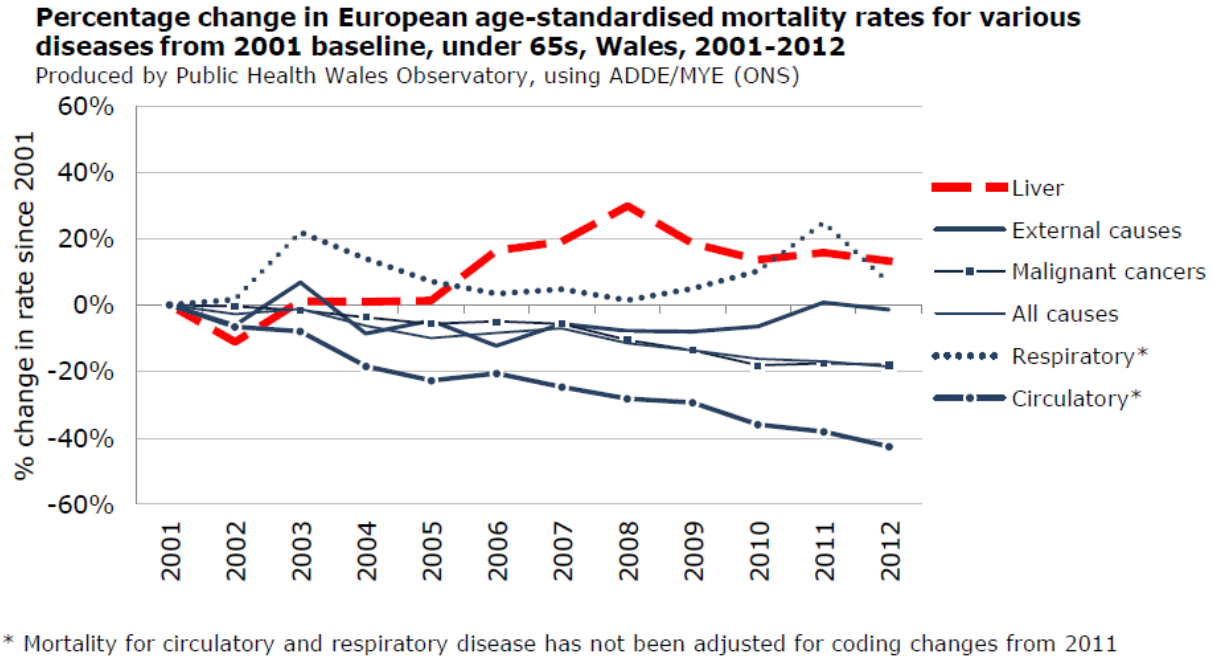


Figure 2.5: The percentage of mortality for liver respiratory diseases from 2001 to 2012 [41]

Wavelet shrinkage thresholds the wavelet coefficients thereby suppressing noise coefficient while retaining image feature coefficients. It is generally performed using either hard thresholding or soft thresholding and requires priori knowledge of wavelet coefficient statistics. Weiner filtering is one of the spectral domain filter type applied on wavelet domain and performs better than wavelet thresholding [42].

There have been several semi-automatic and manual technique to segment the tumor in liver. But, all these methods require user-interaction, and are therefore susceptible to user-error, biasness of individual to select feature and latency. Thus, there is a need of fully automatic segmentation technique which can segment both liver and tumor in a single run. Such technique is very helpful in assisting doctor or radiologist without user interference to reduce reading time, increase detection sensitivity, improved diagnosis accuracy and early detection of tumor as described in [43]. These methods also can be used after proper testing in places where there is no expert in liver imaging.

Research in recent years has been towards fully automatic method that can make accurate and timely prediction of liver tumor saving lot of effort with the temporal trends in recording of features. Transplants from liver conditions for men and women with overall analysis for different age groups for men and women with overall rating is given in Figure 2.6 [44]. The benefit of fully automatic method is they evolve over time through its output and integration of ranging conditions and inputs. Many researchers have emerged lately supports this proposition.

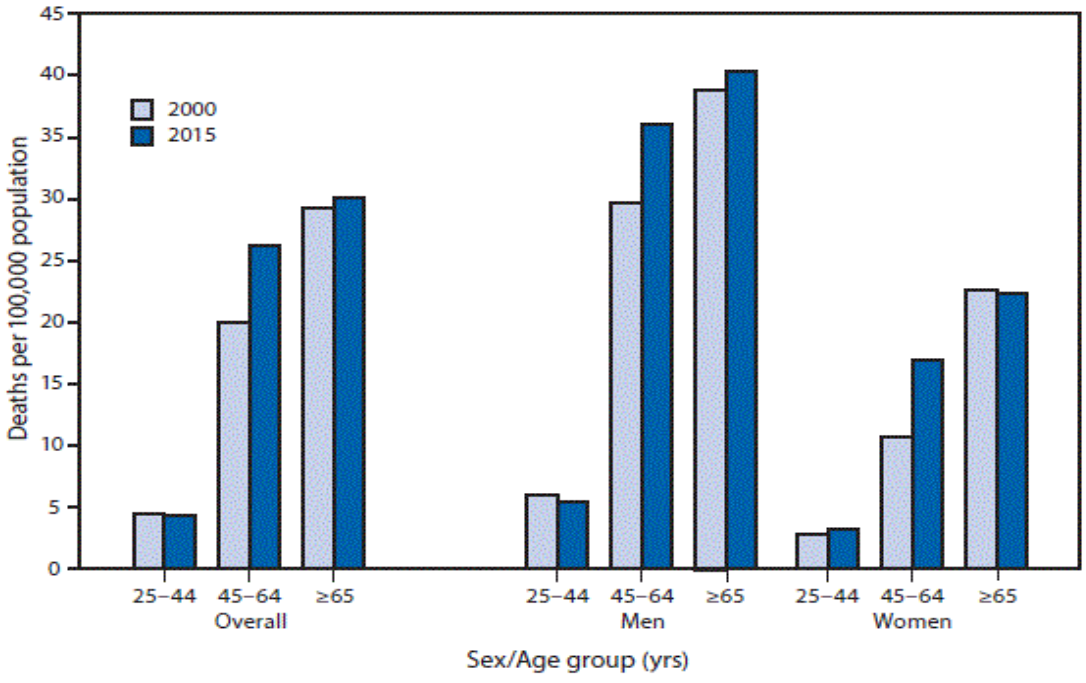


Figure 2.6: Temporal trends in recording of features. Transplants from liver conditions for men and women with overall analysis for different age groups [44]

Though the segmentation done by experts are reliable, this requires lots of time and effort. In addition, such expert who can make accurate and fine segmentation are very few and are in limited access to the people of developing countries where the issue of liver cancer is more prominent. These issues create a need of computer aided or computer assisted system that utilizes a better, accurate and effective algorithm for segmentation, to determine tumor size, shape and location [45].

2.4 LIVER AS AN OBJECT RECOGNITION

After applying the activation function, the output signal is passed on to the neurons in the next layer. By repeating this calculation for every neuron, the signals are processed throughout the layers of the network, which yields more and more abstract, internal representations of the original input data. The last network layer generates the prediction output (e. g. a distribution of class probabilities). In general, all network parameters are initialized randomly. Therefore, the network is not able to make any meaningful predictions in the beginning. Training the network via supervised learning means to perform an initial prediction (forward pass), evaluate the prediction performance, tune the neuron's weights and biases (backward pass), and repeat the whole process iteratively until the network learns to predict the desired output for a specific input. researchers found that active devices require the user to interact physically with the interface device, which requires some learning and teaching, and it's not the natural way for communication using series. Active liver imaging devices use sensors to give very accurate range measurements. This makes these devices expensive as mentioned in [13]. Active devices have high accuracy and low complexity, but they can be uncomfortable since it needs physical contact between the user and the device. While passive devices are computationally demanding, they don't require physical contact, and they are friendly compared to active devices. Passive based approaches rely on tracking the natural user movements without requiring users to wear or hold any sensor. A famous example of passive based techniques is camera-based imaging algorithms. When using camera-based approaches, the liver needs to be in a lightened environment. In camera-based devices, the analyzing time motion is captured using video cameras. Then this video is processed frame by frame and decomposed into a set of features. This kind of modulation generates a signal that has a frequency that increases or decreases linearly with time. This is usually called chirp, if the

frequency is increasing it is called image un-scale if the frequency is decreasing it is called imaging scale. Using this type of signal in series recognition applications has many advantages. Firstly, they utilize channel bandwidth. Secondly, they allow high resolution on the time axis, which makes them superb for ranging applications. The researchers performed the deep learning algorithms differ greatly by identifying weak learner's inadequacies throughout the two algorithms. While the Deep learning model recognizes the weaknesses with elevated weight information points, the increase in gradients works by the same with loss feature gradients. The loss function is a measure of how good the coefficients of the model fit the information at the base. A logical interpretation of loss function is what we try to optimize. For instance, if the sales price is predicted by regression, it would be based on the error between true and forecast house prices. the loss functions. Likewise, the loss function would assess how useful our predictive model is for poor loans to be classified if it's our objective to categorize loan defaults as mentioned in [46].

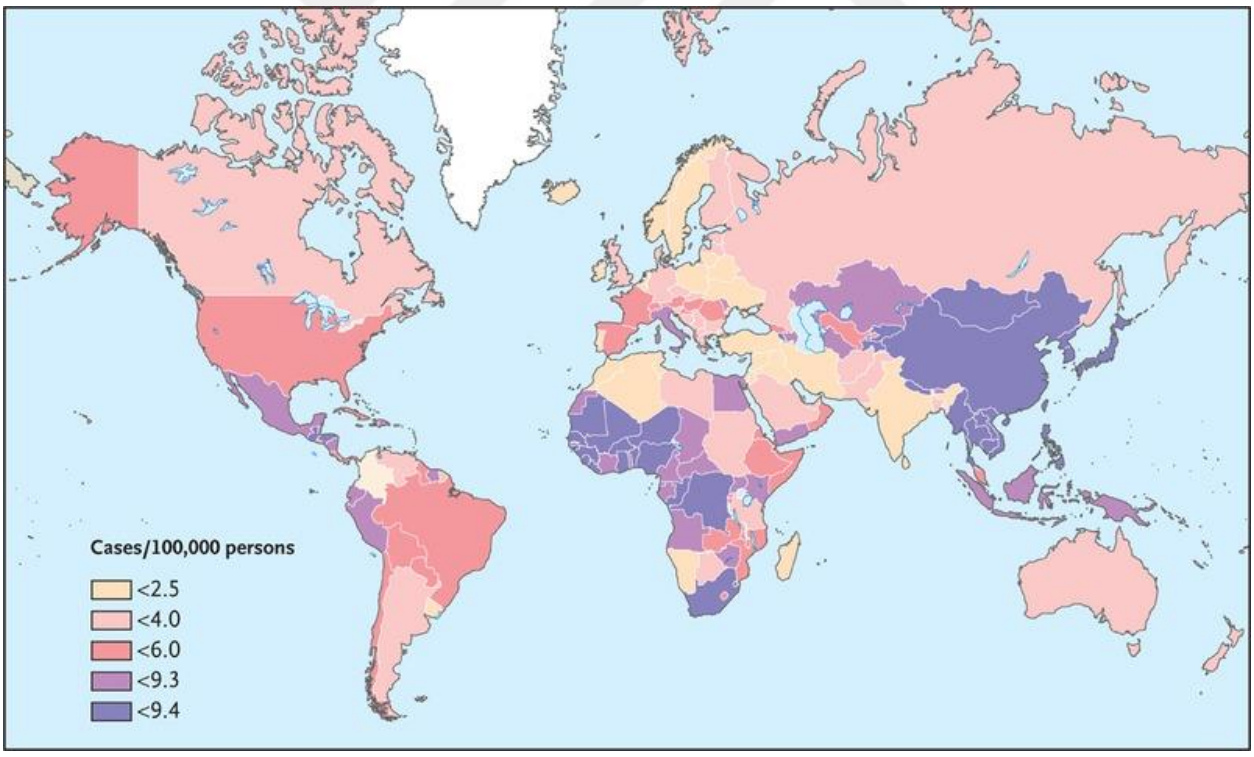


Figure 2.7: Number of hepatocellular liver cancer cases per 0.1 million people in different regions of the world [46].

2.5 LIVER HEALTH DATA PROCESSING

In the analysis of liver data for direct output, the value of each pixel in each piece will change at intervals (0.255) depending on their small area and size as the images are gray in data based on the image in [47]. However, the model does not work under this setting. The reason is that each patient's data is taken at a different time with some different obscure analyzes (because physicians will not care how the image details were processed) as given in [48]. As a result, the distribution of original data varies greatly. Some contain a value equal to the interval (0.255) while others contain a value in (-200,8000). Also, each piece of data for the wound has a different local and local dimension, and it happens that the background is not always the smallest number of lesions (because each piece contains a minus pixel value which is also a vague reason). Technically speaking, all the pieces convert to an interval (0.255) with different insertion methods and this affects many looks. To address those issues, a minimum value for all pieces is hand-tied to 0 (which must be the real back value) and the highest value for the global size of each wound being processed. All pieces of data will switch back to the interval (0,255) according to these two values. So now each piece of data has the same amount of data from the same patient than the same high value.

3. METHODOLOGY

3.1 PROPOSED SEMANTIC SEGMENTATION

Our proposed method consists of a FCN model and the CNN model. The VGG-16 is solved by Data Augmentation (DA). As shown in Figure 3.1, the input image is first passed into the FCN, which produces a coarse segmentation output. Next, we crop the regions enclosing the predicted tumors as ROIs. Then we feed the ROIs, along with the required parameters derived from the input and the FCN predictions, to apply the VGG-16 on the ROIs to perform post-processing. The refined segmentation maps of the ROIs replace the FCN predictions and yield the final output. We explain these steps in detail in the next sections.

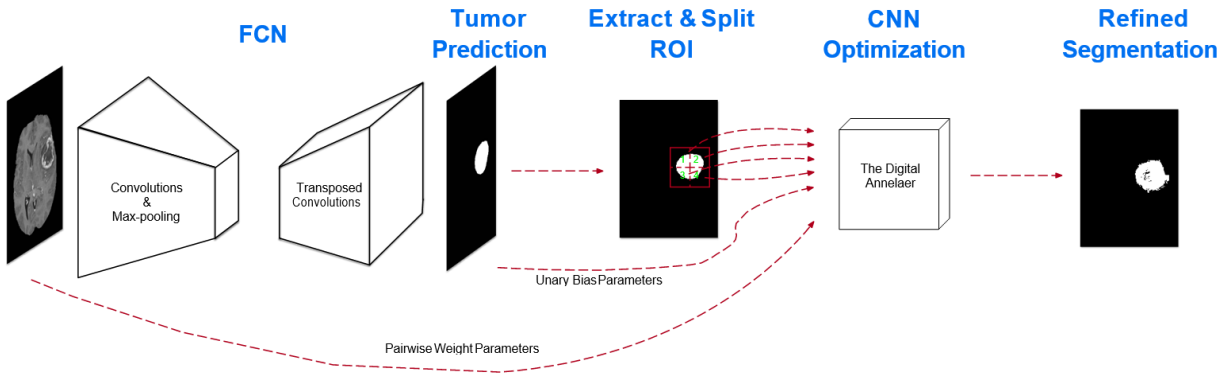


Figure 3.1: Proposed model architecture for liver tumor semantic segmentation [49].

3.2 LIVER FEATURE MAPPING

The FCNs are the backbone of today's almost every semantic segmentation model. The name is self-explanatory, which indicates that the network consists of only convolutional (including the transposed convolutional) layers. CNNs that are used in image classification problems reshape the feature maps extracted by the convolutional layers into 1D and adopt fully connected layers to get the image level prediction (which is a scalar). Note that FCNs retain the shape of the features in 2D by performing only convolutions. This adaption maintains the spatial information in the feed forward phase, and extends CNN inference from 1D to 2D, which promotes image level classifications to pixel level classifications.

The feature maps are further processed by the bottleneck layer which consists of only convolutions. After that, the network up-samples the feature maps by 3-stage transposed convolutions, and finally yields the prediction map that has the same size as the input. The outputs from the last and the second last pooling layers are fused with intermediate up-sampled feature maps by element-wise addition. Note that batch normalizations (BNs) are adopted inside each convolution layer to stabilize the gradients during backpropagation and accelerate the training.

A detailed FCNN block diagram specifying the feature map dimensions is shown in Figure 3.2. The depth of the output prediction in terms of each pixel is 2, which can be perceived as the "score" of the pixel being a tumor pixel or a non-tumor (background) pixel.

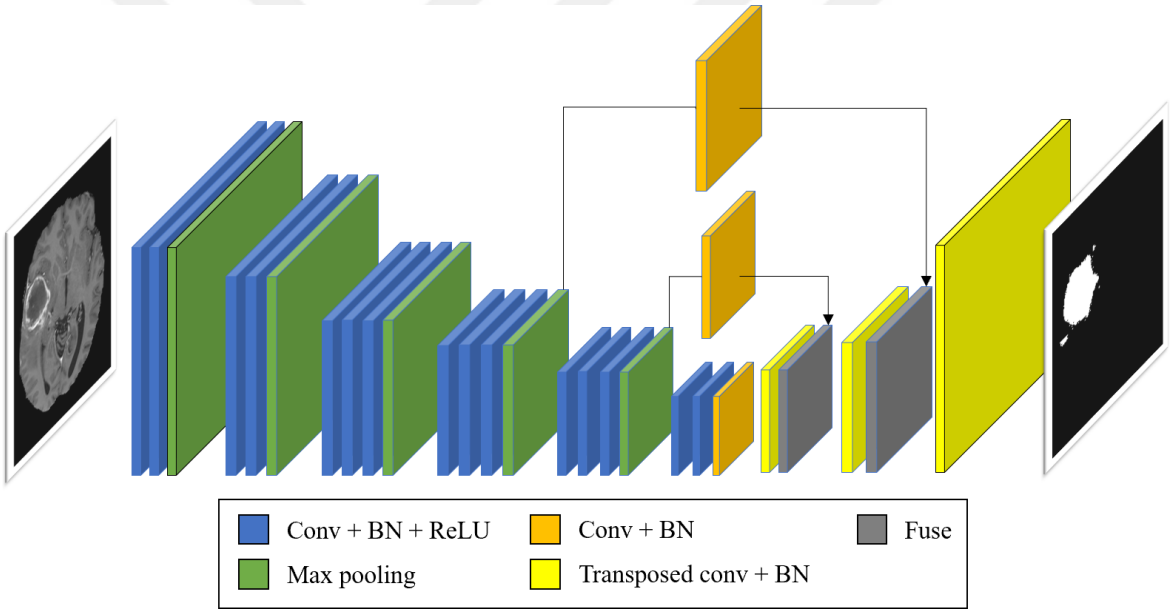


Figure 3.2: A systematic and pictorial representation of architectures using the Fully Connected Convolutional Network (FCN)for liver tumor classification [49].

The first is to consider position of individual points in all directions of liver. However, the simple positions cannot be considered for understanding task because of the variability of size, shape and positions itself in a 3D world. Second, trajectory will have velocity by which the region segmentation is performed, it will have different set of points at each different velocity. The third and most reliable solution for decoding trajectory is to find an angle between subsequent points.

An angle will be independent of scale, size and velocity of the liver region and will be perfect contender to represent the set of points as sequence of angles for tumor classification[50].

3.3 DATASET ACQUISITION

The dataset was obtained from an open-source repository for automated liver tumor segmentation and classification known as LiTS – Liver Tumor Segmentation Challenge (LiTS17) consisting of 16.66 GB in size. We performed tests which demonstrated similar (scalability) results by processing this dataset. A heuristic solution had been proposed such that we divided the total number of scales by half. The first half containing the heavy sized scales images of liver lesions would be parallelized using a row wise parallel for, while the second half containing the light sized scales would be parallelized computing in parallel each scale. That's how the dataset was processed using the multi-cross folds. The dataset can be acquired from [51].

3.4 TRAINING PARAMETERS

A common approach to training CNNs is transfer learning. For example, in semantic segmentation, the weights and biases of the convolution layers of FCN models are often initialized with the parameters from other CNNs that are well trained for image classification tasks, and the target network is fine-tuned based on these transferred parameters rather than random initialization. This is a valid and effective approach for natural image processing, as the low-level features of objects can be shared, regardless of the varying purposes of distinct tasks. Nevertheless, applying transfer learning on medical image processing is considered unwarranted. The features in medical images and natural images are very different from each other, therefore transferring the parameters learned from natural images does not provide a strong base for the training process. Furthermore, there is a lack of comprehensive medical image datasets, resulting in a lack of robust CNNs trained for processing medical images. Because of the considerations above, we decided to train the FCN from scratch. The batch normalization operations enabled training with few samples.

We initialized all parameters with Xavier initialization. We have trained the network with FCN model. We set the initial learning rate to 3×10^{-4} and trained the network for 300 epochs with a batch size of 4. We picked the model with optimal parameters that produced the best validation

accuracy after each epoch of training. The training was performed by one NVIDIA GTX 24 GB graphic processing unit (GPU). LiverNet generates 16 × 16 × 2 = 512 anchors for each input image. For each anchor, the classification layer predicts its probability of enclosing tumors (referred to as the foreground probability), and that of not enclosing tumors (referred to as the background probability); the regression layer computes the horizontal and vertical distances, by which each anchor should move so as to adjust its position to enclose tumors. In comparison with RPN, our LiverNet does not learn the horizontal and vertical scaling factors of each anchor, as keeping the fixed aspect ratio is important for TumorNet's inference in the next stage. We also remove the ROI pooling layer to simplify the network. Via two 1 × 1 convolution operations in parallel, the feature maps in layer Pool 4 are converted to the classification and regression layer. The output of each layer can be viewed as a 3D tensor, which can further be separated into 16 × 16 = 256 vectors with a length of 4. Each vector represents some predictions over the anchor group centered at that grid in the 16 × 16 feature map the training and testing graph mending on the 300 epochs is displayed in Figured 3.3.

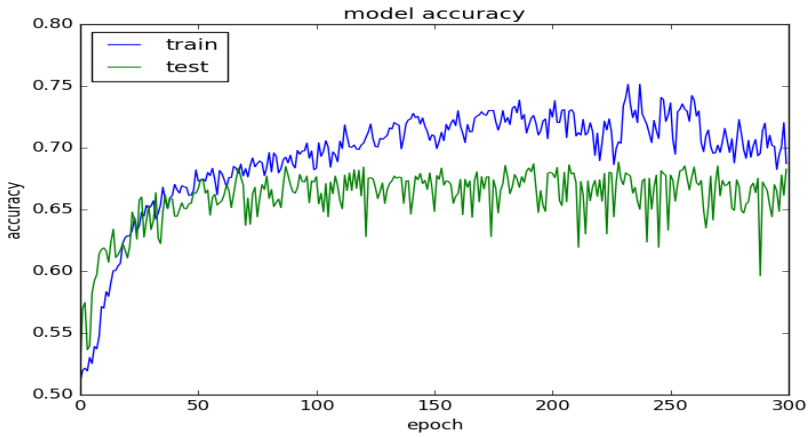


Figure 3.3: The model accuracy of FCN model which is being trained on 300 epochs

3.5 PROPOSED CLASSIFICATION METHOD

We showed that the proposed FCN can enhance the accuracy of the binary liver tumor segmentation problem in the chapter 2. In this chapter, we extend the application of our method to a multi-class liver tumor segmentation problem. The label distribution in medical imaging

semantic segmentation problems are often imbalanced. Our method is proposed to tackle this label imbalance issue.

The objective of the liver tumor classification is to assign each pixel to one of the three classes: liver, tumor, and background. Conventional models that can be trained end-to-end perform a one-time inference on all classes of labels, however it is difficult to train the models to properly learn the features of tumors, as the number of tumor samples is much smaller compared to other classes. The major difference between our method and those models is that we learn the features of tumors and perform segmentations in local regions that may contain tumors rather than on the entire input image. In these local regions, the label imbalance issue is alleviated. We also apply the VGG-16 to refine the tumor segmentation outputs. Our experimental results show that the combination of lesion localization and VGG-16 can boost the classification performance.

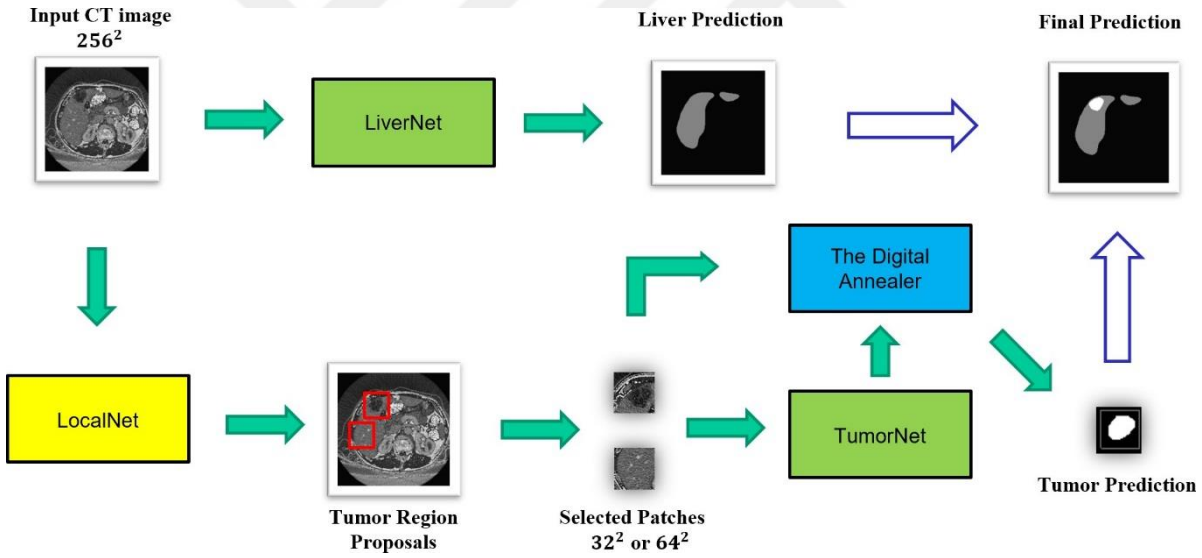


Figure 3.4: A complete flow diagram of proposed model architecture for liver tumor automated classification.

As shown in Figure 3.4, in our model, the input image is fed into two streams. The first stream is LiverNet, which is a FCN model that segments the shape of liver from the background. In the second stream, the input is first passed through LiverNet, which is a CNN that proposes ROIs that potentially contain tumors. These ROIs then become the input to the third network, TumorNet, which is another FCN model. The purpose of TumorNet is to segment tumors within the ROIs.

Next, the output of TumorNet and the information extracted from the ROIs are passed into DA for post-processing following the FCN model. At last, the liver prediction and tumor prediction are consolidated to form a superposition, which is the final predicted segmentation map. Our method divides a ternary semantic segmentation problem into two binary problems.

3.5.1 LiverNet

The block diagram of LiverNet is shown in Figure 3.5. Basically, the structure of LiverNet is to use the FCN model for liver tumor segmentation and classification. The only difference is the dimensions of the feature maps.

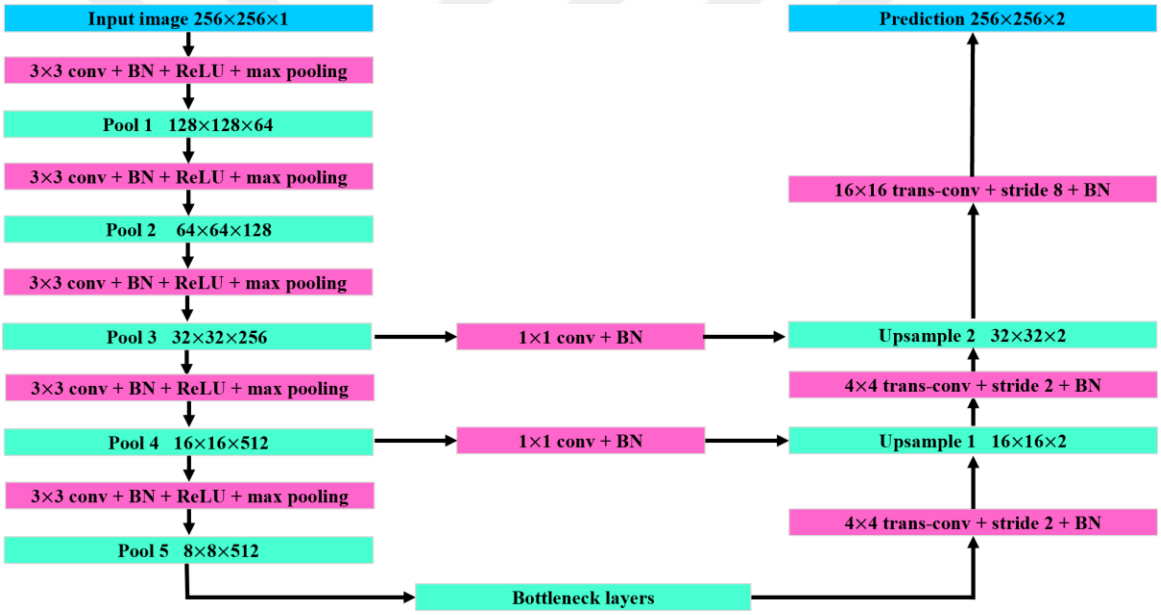


Figure 3.5: LiverNet block diagram.

3.5.2 TumorNet

TumorNet takes the proposed ROIs from LiverNet and performs another binary segmentation. After being trained, TumorNet can process inputs with varied sizes since it is fully convolutional. However, during training, all inputs are resized to 32 × 32 to enhance the efficiency. The structure of TumorNet is shown in Figure 3.6.

TumorNet is a shallow network as the size of inputs are relatively small, which reduces its capabilities of feature extraction. As a result, it is not suitable to use too many pooling operations

in TumorNet, as this might cause excessive loss in information which is averse to the up-sampling process afterwards and limit the segmentation performance. With this consideration, we only down-sample the feature maps once, and adopt multiple dilated convolution operations in parallel for the next layer. The three streams of dilated convolutions process the feature maps with different receptive fields, thus capture both local and global features. This critical step maintains the dimension of the intermediate feature maps and diminishes information loss. These three sources of features are then concatenated and integrated for the up-sampling layers followed next. Same as LiverNet, TumorNet is also appended with a softmax layer at last and is trained by minimizing the cross-entropy loss function.

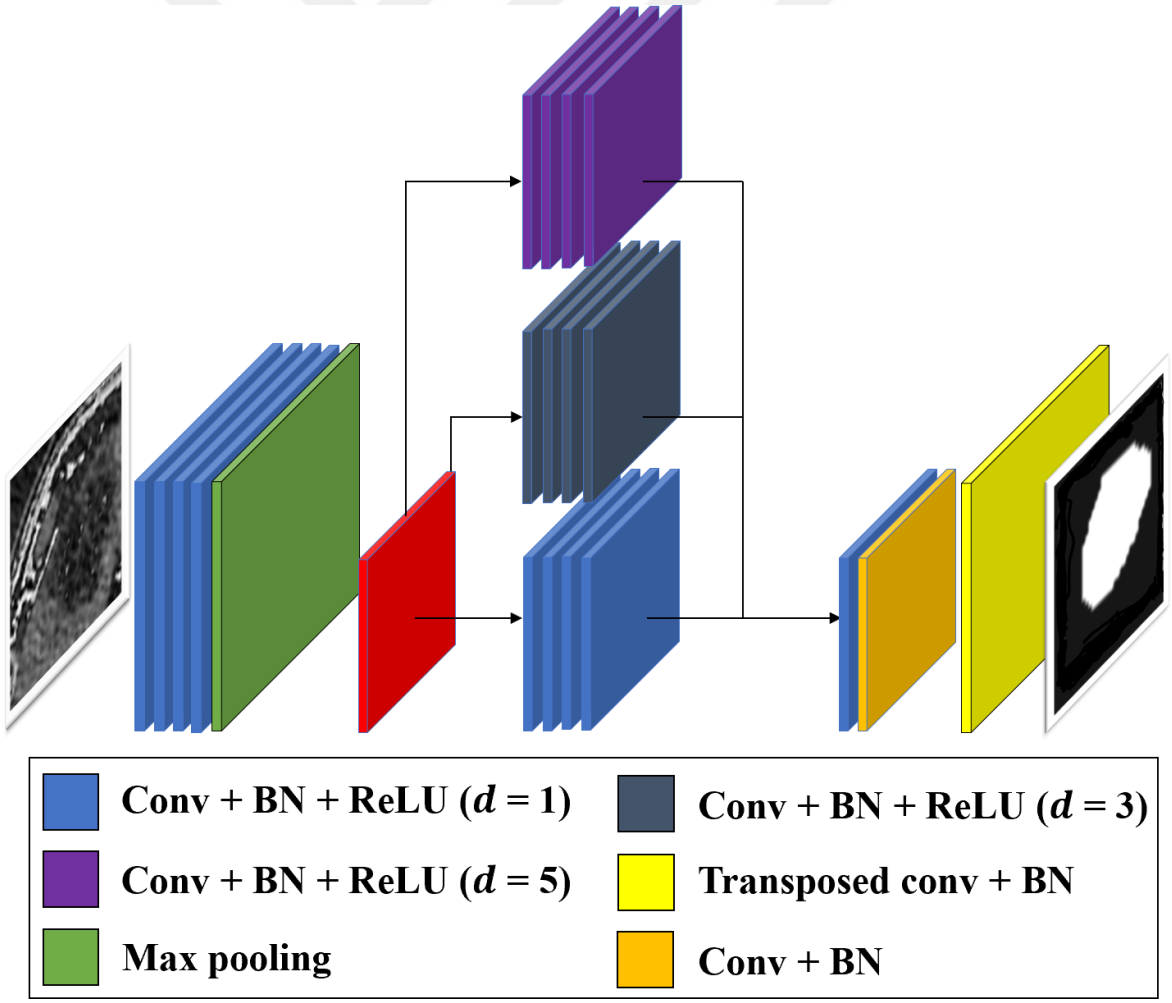
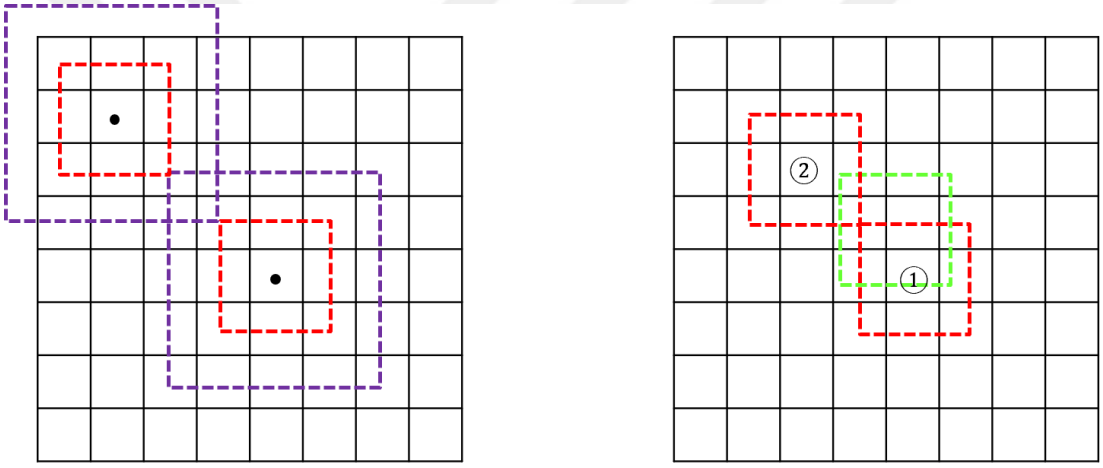


Figure 3.6: TumorNet Structure.

The input image is first processed by 4 consecutive convolution and max pooling blocks, producing a group of feature maps with a dimension of 16 × 16 (ignoring depth). We denote this feature map group as layer Pool 4. For each grid in layer Pool 4, LiverNet creates two anchors with different sizes that are centered at it. The input image with size 256 × 256 is spatially downsampled by a factor of 16 after it reaches layer Pool 4. Therefore, the positions and sizes of the anchors can be mapped back to the original scale by multiplying the same factor, given that each grid in layer Pool 4 corresponding to a 16 × 16 block in the input image. Figure shows two example groups of anchors on a partial feature map. The red and purple anchors have size 32 × 32 and 64 × 64 in the original scale, respectively. The purpose of generating multi-sized anchors is to match with tumors of different volumes. Note that the scope of some anchors residing on the periphery is out of the image.



(a) Anchor groups (b) Positive and negative anchors

(a) The anchor group helps the scanning. (b) The positive and negative anchors are for location finding of tumors.

Figure 3.7: Anchors in the Liver Net.

Training LiverNet requires ground truth (GT) anchor boxes from the training dataset. Before training, we first scanned through the label maps of all images and created bounding boxes to enclose the tumors. It was found that the size of the largest tumor was smaller than 64×64, therefore depending on the size of a given tumor, we used a FCNN with size 32 × 32 or 64×64 to enclose it. For each training image, there may exist multiple GT boxes.

3.6 PRUNING AND OVERFITTING

We studied the effect of pruning on the accuracy of prediction. To study this, we pruned the FCNN trained with 2000 features and three layers for different number of training images. Then we calculated the test set accuracy of the pruned models. It turned out that pruning the tree showed almost no difference in accuracy compared to the fully connected layers. We hypothesize the reason for this is that very high up in the tree the splitting nodes already confidently divide the training points into their class. However, the FCNN algorithm continues to exhaust the features until the entire layer is constructed. By stopping the prediction early in the prediction, we do not suffer in accuracy, and improve the speed of prediction. Finally, we studied how much the model overfits to the training samples by calculating the training accuracy and comparing that with the test accuracy. We fixed 2000 features and three layers and varied the number of training samples. We show evidence of extreme overfitting. The training error is consistently very low (mean 1.06%) and is much lower than the testing accuracy.

3.7 EXPERIENCE AND LIMITATIONS

We learned several lessons in building a practical deep learning and computer vision-based system for liver tumor classification and segmentation.

3.7.1 Rapid optimization.

Although our system is very time-sensitive, we found that it is not necessary to do optimization on one component while other components are not ready. We found it is best to develop quickly and do not profiling to discover the bottleneck and optimize it. This saves us a lot of time in making unnecessary optimization.

3.7.2 Using pipeline rather than events

In the testing of our system, we have incurred many concurrency bugs due to the event-driven programming framework. We solve this by not using locks but adding the processing unit to a pipeline, thus leaving us free concurrency bugs.

3.7.3 Test accuracy is not enough

We found that using the test accuracy is not a good evaluation of the actual system. Sometimes the system performs very well in the test data set, but poorly in practice. The reason behind this is that the environment we evaluate changes all the time: camera position, backgrounds, people's clothes, etc. Therefore, we decide to choose some important parameters based on our experience in the actual environment instead of mere test accuracy. For example, although the experiments indicated the number of layers made big difference, we found to this to be the case in actual tests with varying backgrounds. Therefore, we used a model trained with multiple layers rather than one in fully connected CNN.

3.7.4 Bundle the model with feature extraction

In our system, we separated feature extraction with the model. This turned out to be error-prone. It would happen sometimes that the trained model has the wrong feature extraction (e.g., the offset pairs are not correct), and things could get even worse when the real-time prediction system mistakenly uses the wrong feature extraction method. Our immediate solution is to be extremely careful to this. However, if we were to build the system again, we would have bundled the model with feature extraction to avoid human errors.

3.7.5 Limitations

Our current system is not without limitations. First, the real-time prediction component cannot achieve a frame rate of 30Hz but only 10Hz on average. Second, the FCNN model takes a long time to train, especially with many training samples. Often times, training required around 24 hours or more. Third, the system must be retrained with more training samples every time the user wants to add a new gesture.

4. RECOGNITION RESULTS

A common approach to training FCNNs is transfer learning. For example, in semantic segmentation, the weights and biases of the convolution layers of FCNN models are often initialized with the parameters from other FCNNs that are well trained for image classification tasks, and the target network is fine-tuned based on these transferred parameters rather than random initialization. This is a valid and effective approach for natural image processing, as the low level features of objects can be shared, regardless of the varying purposes of distinct tasks. Nevertheless, applying transfer learning on medical image processing is considered unwarranted. The features in medical images and natural images are very different from each other, therefore transferring the parameters learned from natural images does not provide a strong base for the training process. Furthermore, there is a lack of comprehensive medical image datasets, resulting in a lack of robust FCNNs trained for processing medical images. Because of the considerations above, we decided to train the FCNN from scratch. The batch normalization operations enabled training with few samples.

To facilitate fair comparisons between the different methods, the recognition accuracy after a fixed amount of training steps and the best validation accuracy on the test set, recorded every 200th update step, are consulted. For the former approach, the last training epoch is only taken if its corresponding training loss is within a 50% range of the best recorded training loss, otherwise training continues until this condition is met. This prevents that the training is stopped after a particular detrimental parameter correction update. Both values are averaged over five independent runs to avoid statistical outliers. The respective sample standard deviation is also presented. The minimum validation error gives an indication on the capacity of the method, while the limit on training time considers real-life scenarios where no validation set is available. As the last validation accuracy can heavily differ depending on the number of training epochs selected, the best validation accuracy seems to be the more meaningful and thus is used as the main recognition accuracy metric. Unfortunately, the relatively small original dataset size, compared to the high number of classes and the large sequence variability, does not allow a further split of the training set into a validation set. This would be necessary for the standard approach of finding the best training epoch based on the minimum validation error before testing it on the test set. The proposed

method of using the test set for validation and evaluation and thus is the adequate choice for the particular 3D liver lesions classification action sequence dataset. In the following, the terms test and validation accuracy are used interchangeably.

The fully connected CNN based recognition classifier design parameters improved its average maximum validation accuracy when it was trained on the augmented training sets. For the model with the best generator loss an average maximum validation accuracy of 91.29% was achieved, while the model with the best inception accuracy. The minimum generator loss model therefore outperformed the model with best inception accuracy, which asserts the impression that it is more suitable for the selection of the best model, the maximum number of frames gets clipped to smaller numbers for the created sequences is represented as input image from the dataset, the synthetic samples become less helpful for improving the hot spectral for classification is given in Figure 4.2. The Figure 4.3 depicts the HSV image processing in the liver tumor segmentation and classification pipeline where the Figure 4.4 presents the axial image processing in the liver tumor segmentation and classification pipeline. The visible tumor spectral imaging in the liver tumor segmentation and classification pipeline is given in Figure 4.5. The segmented image and classified image in the liver tumor segmentation and classification pipeline are given in Figure 4.6 and Figure 4.7 respectively. The parametric input image, mask true and the prediction of automated liver tumor classification is presented in Figure 4.8, Figure 4.9, Figure 4.10, Figure 4.11, and Figure 4.12.

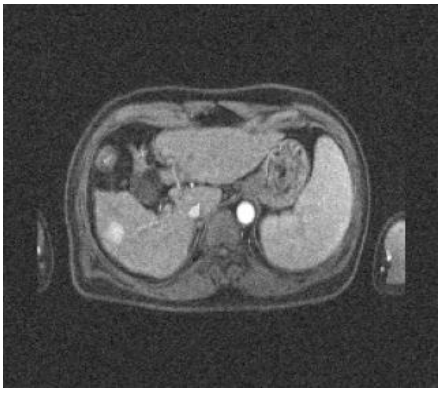


Figure 4.1: The input image in the liver tumor segmentation and classification pipeline.

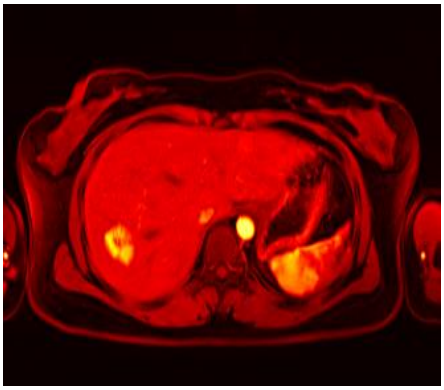


Figure 4.2: The hot spectral image processing in the liver tumor segmentation and classification pipeline.

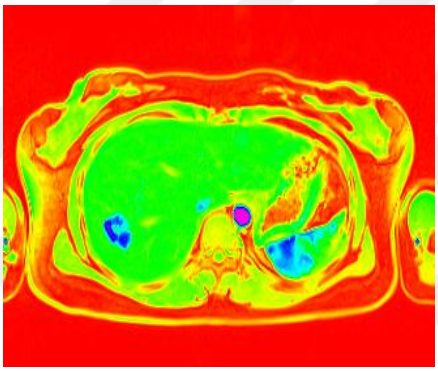


Figure 4.3: The HSV image processing in the liver tumor segmentation and classification pipeline.

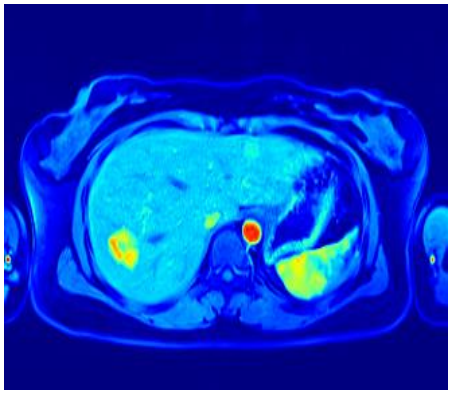


Figure 4.4: The axial image processing in the liver tumor segmentation and classification pipeline.

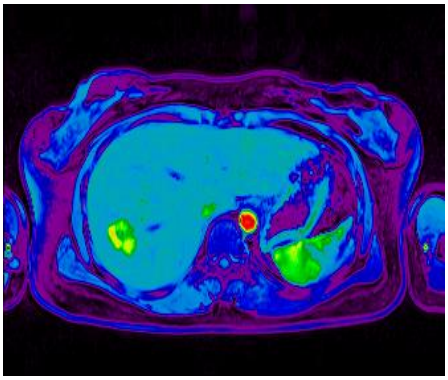


Figure 4.5: The visible tumor spectral imaging in the liver tumor segmentation and classification pipeline.

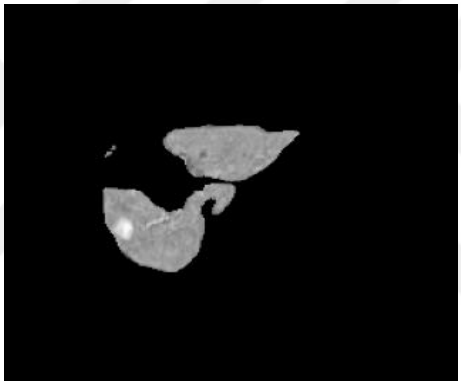


Figure 4.6: The segmented image in the liver tumor segmentation and classification pipeline.

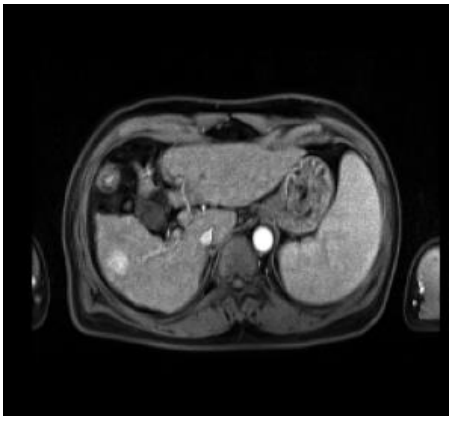


Figure 4.7: The classified image in the liver tumor segmentation and classification pipeline.

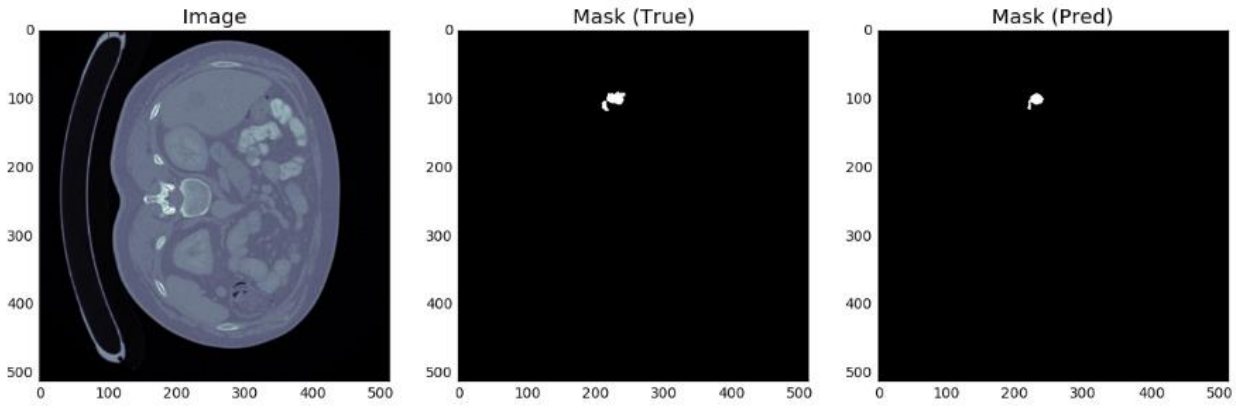


Figure 4.8: The parametric input image, mask true and the prediction of automated liver tumor classification.

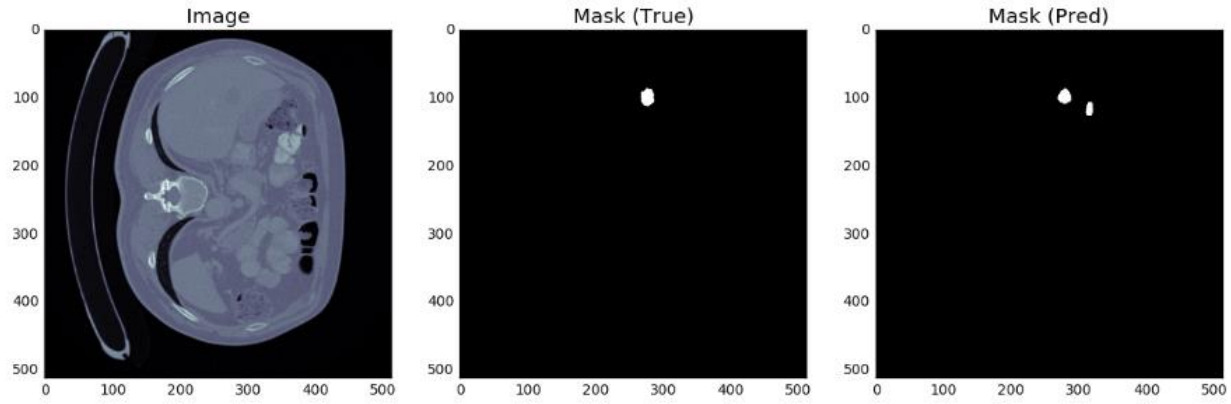


Figure 4.9: The parametric input image, mask true and the prediction of automated liver tumor classification.

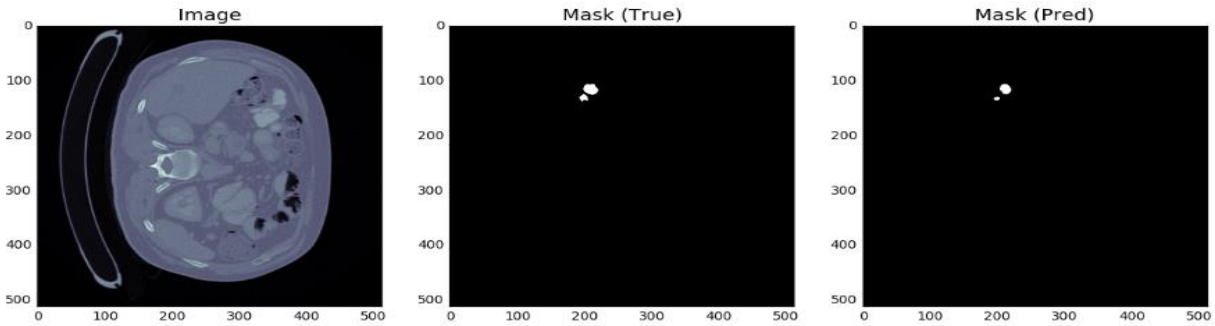


Figure 4.10: The parametric input image, mask true and the prediction of automated liver tumor classification.

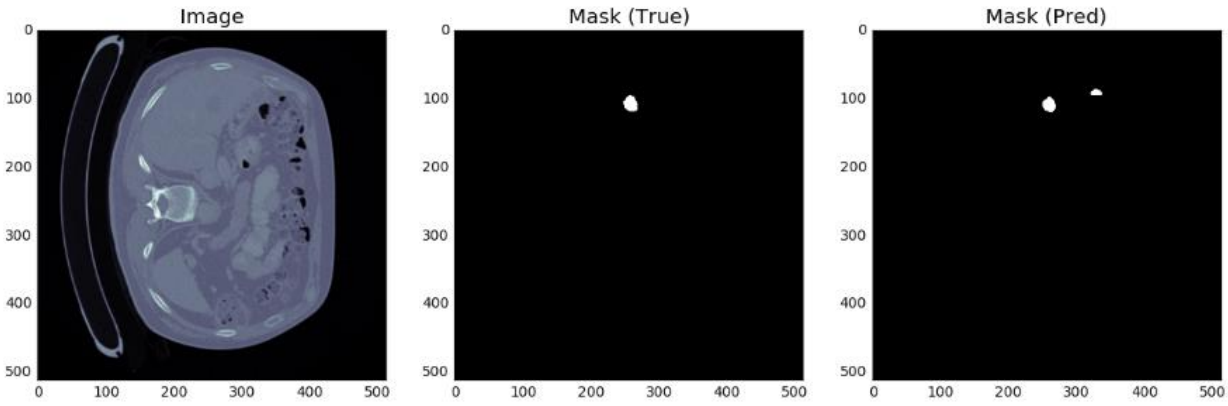


Figure 4.11: The parametric input image, mask true and the prediction of automated liver tumor classification.

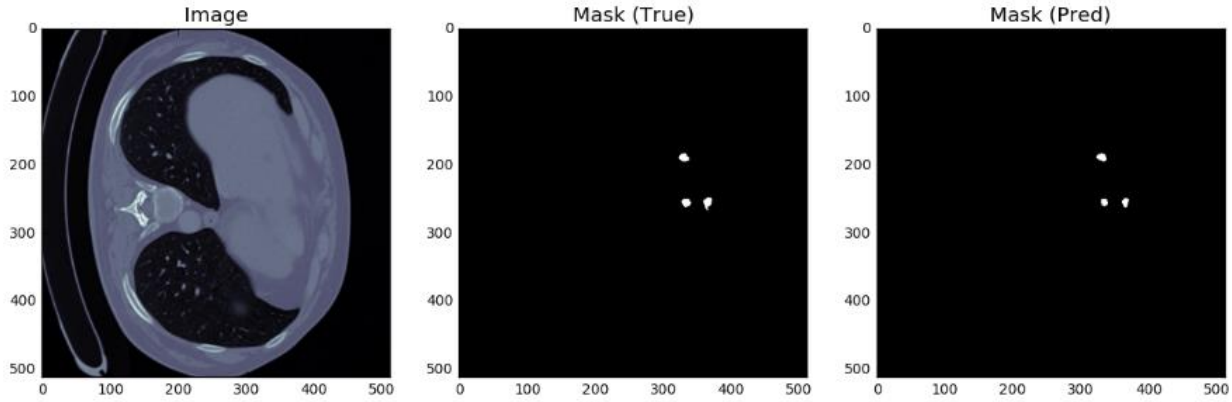


Figure 4.12: The parametric input image, mask true and the prediction of automated liver tumor classification.

In per-pixel liver lesion classification, there are many parameters involved in capturing the training images, extracting features, and training the FCNN classifier. To study the contribution of these parameters, we systematically varied them and studied the accuracy of the resulting classifier. To study the accuracy, we set aside a collection of images as the test set and trained on a different set of images. We didn't use cross-validation due to the time constraints in training a FCNN classifier; it would have taken considerably longer time to test the cross-validation accuracy and would have prevented us from running as many tests as we did. We collected images of liver tumor lesion from the dataset. We refer to each image that we took as a "training sample". Since there are over 300000 pixels in each image, we sampled only 2000 of them so that we can train on more diverse images.

From each image, we sampled 1000 background points, and 1000 points from the liver lesions to ensure we trained evenly on the different types of classes.

Finally, instead of changing the augmentation architecture, it is worth to experiment with alternative augmentation strategies. For this, it was decided to only select synthetic samples for augmentation that were correctly classified by the inception network. The correct identification could be seen as a sign of quality. However, this approach led to a reduction in the average maximum accuracy (91.29%). Once more, the FCNN inception network seem to be a suitable choice for the evaluation of the synthetic liver lesion samples' augmentation capacities as it favors data which is similar to the original training set and thus unlikely to contain new information content.

The different augmentation methods can also be used in conjunction. They can be employed to each generate a new set of samples. These can then be combined to augment the original training set. The results for combining FCNN augmentation methods. The tests were conducted as the alternative augmentation methods were effective for this subset in the previous experiments. For each augmentation method, 60 new samples were independently generated and appended to the real dataset. The improvement that was achieved with the multi-layer FCNN was also observable when it was used together with the alternatively generated sequences. The combination of sampling from random sequence segments and FCNN augmentation was even slightly more effective than the multi-layer FCNN on its own. Interestingly, there was a strong boost in the best validation accuracy when all three alternative methods were applied together. The resulting score is in the same region as for the multi-layer FCNN. Potentially, the classifier was able to learn new information from the three combined transformations that can all be concurrently exercised by the multi-layer FCNN. Finally, using all augmentation methods did not lead to a further boost. Overall, it appears that the multi-layer FCNN data augmentation comprises most of the positive augmentation effects by itself.

5. DISCUSSION

Focusing on the average best validation error, two statistics seem to correlate with its absolute improvement. On one hand, the improvement seems to shrink with the number of classes and thus the number of original samples in the training set. As the number of selected classes is increased from seven to ten, the difference becomes smaller and even negative. However, even though relatively small, the classification accuracy is higher when the data augmentation is used for twenty classes as well as the whole dataset.

On the other hand, the FCNN augmentation becomes more effective for subsets with larger maximum sequence length. While there was a similar amount of original training samples in the three five-layer subsets, the maximum sequence length strongly varied. The largest absolute improvement was achieved for the FCNN where the maximum sequence length was the highest. While the maximum length was equal in the experiments, where a shrinking improvement was observable with increasing training set size, the maximum sequence lengths are significantly larger when the experiments are repeated on the classes. In these cases, the FCNN augmentation has positive effects again. As the maximum number of frames gets clipped to smaller numbers for the created sequences, the synthetic samples become less helpful for improving the classification accuracy on the whole dataset. The classification scores are all smaller than the original baseline when the synthetic sequences have a maximum length of only 100, 200 or 500 frames.

These high-level observations do not allow definite conclusions and it seems likely that the effectiveness of the FCNN augmentation depends on many other factors, foremost the shape, type, and variations of the selected liver lesion. That is why, it is difficult to determine a definite limit for which the FCNN augmentation method starts. However, it is intuitive that the FCNN augmentation is more advantageous on smaller sets which can be more easily enriched with new information. While the visual quality of the synthetic samples is negatively affected by the smaller training set, the generated sequences appear to carry enough information to have a beneficial impact. Moreover, a classifier's proneness to overfitting rises with decreasing amount of training samples, increasing the need for any regularizing input.

Table 5.1: Comparison of proposed technique with existing in terms of performance.

ARTICLE	TECHNIQUE	ACCURACY
[50]	Support Vector Machine (SVM)	88.87%
[51]	Decision Tree and k-Nearest Neighbor Classifier	73.20%
Proposed	Fully Convolutional Neural Network (FCNN)	91.29%

Model based approaches start with a set of hypotheses of the final classification based on rules of the object being classified. For instance, in the case of liver tumor classification a hypothesis can be a particular orientation of the joints. An advantage here is that the hypotheses can be generated from an infinite space of possible classifications. The main disadvantage of model-based techniques is that they are often computationally expensive. In addition, model-based approaches tend to be very complex. An example of a model-based technique can be found in [36]. A possible explanation for the observations on the maximum length of the synthetic sequences could come from the class sequence length distribution. While 95% of all original sequences are shorter than 200 frames, there are outlier sequences that are much longer and skew the arithmetic mean class sequence length. As the length of the synthetic samples is based on an FCNN distribution that makes use of these statistic, the generator might tend to produce longer sequences that help the classifier to better generalize to these rarer inputs. When clipped, this effect cannot be fully exercised by the generated samples. The smaller the number of training samples, the smaller is the number of original outliers, which increases the positive augmentation impact. Interestingly, the maximum validation accuracy is better for the multi-layer FCNN when the synthetic samples are clipped to 100 instead of 200 frames. In this case, the data augmentation emphasizes the inputs that tend to be shorter than the average. Nevertheless, these effects might be rather random statistic correlations and it is unlikely that the positive impact comes from the varying length of the samples alone.

Average inter-class, intra-class and centroid distances for the original, generated and combined datasets. The intra-class distances for the best generator loss indicate that the static FCNN have learned to produce a wider range of different frames for each class compared to their best inception

accuracy counterparts. While the average inter-class distances are also bigger than for the best inception accuracy models, they are still much smaller than for the original training distribution. The FCNN models for liver tumor classification that are very different to each other and thus cause larger average inter- and intra-class distances.

The base multi-layer FCNN had a positive effect on almost all groups of selected fully connected layers as well as on the complete dataset when used for data augmentation. This is a strong proof that the proposed FCNN augmentation method is suitable for improving the classification on liver tumor lesions action sequences. Modifications in the size of the critic architecture led to worse results. While the FCNN-augmentation scored a higher average best validation accuracy on the complete dataset, it performed worse when tested, showed stability during training and generated sequences of higher visual quality. Thus, without the added gradient penalty term, the FCNN appears to be the superior and more stable architectural solution.

6. CONCLUSION

6.1 CONCLUSION

Liver segmentation and tumor segmentation from abdominal CT scan image is a key task for many clinical applications. Automatic methods for such problem are of the great essence as they reduce time, effort, provide quality assistance to experts as well as early determination and diagnosis of the tumor. To be of practical use, such application of segmenting liver area and tumor area should be very accurate, precise, and verified using metrics popular in medical image processing so that a radiologist can perform segmentation quickly without help from technician. The segmentation of liver and tumor is a very challenging task due to complexity of liver surface, variation in liver and tumor size throughout the CT image slices, ambiguity in boundaries of alike intensity tissues and nearby organs. Many types of research and works have been done previously to develop an automatic and semi-automatic method for liver and tumor segmentation. But, there are only a few methods that do the segmentation of both liver and tumor using the same technique or single approach and provides better and accurate result. Most of the method used practically in clinical applications are the semi-automatic method which requires more or less user interaction for initial contour defining or parameter adjustments, etc. The methods so far that has been developed on previous work are mostly based on statistical method. The drawbacks of such methods are a lot of research work requirement in statistics, and the process of determining information on statistics are very sensitive. Such statistical model and most of the other approach aforementioned requires more research for expansion to some other organs or tissues. The proposed general segmentation algorithm is the latest concept based on training texture features using machine learning algorithm. Texture features are more informative than most of the other features, and the use of machine learning algorithm is in every application with its influence and application to be more increased in future. We took the concept and applied it in the medical field for liver and tumor segmentation with Gabor feature and three different machine learning algorithms. In this thesis, the used LiTS – Liver Tumor Segmentation Challenge (LiTS17) dataset and is publicly available to conduct research which shows the authenticity of data. The different forenamed metrics were used to show the good performance of the proposed method for the individual classifier and segmentation algorithm. The proposed method implementation also provides a perspective for this method to be

practically used with some improvements on the same fundamentals. FCN itself was very good classifier with classification accuracy above 91.28% for classification. We concluded that the presented method takes a significant step in providing insight about the performance, implementation, and comparison of DL algorithm, and it also showed the future with texture feature and DL in the field of medical image processing for automatic liver and tumor classification and segmentation.

6.2 FUTURE WORK

Based on a high value of classifier accuracy and FCN, this idea can be implemented and tested with a large variety of dataset as it has succeeded for this particular set of raw images. Moreover, the technique used in this thesis does not require any post- processing to refine the output. In future, a large number of a dataset can be added to perform analysis provided enough processing speed. The proposed FCN based method can also be tested with some other features such as statistical. In addition, this same approach can be used for different organs like kidney, lungs, brain, etc., and for different medical image formats MRI, ultrasound, etc.

REFERENCES

[1] Salem, M.A.-M.; Atef, A.; Salah, A.; Shams, M. Recent survey on medical image segmentation. Comput. Vis. Concepts Methodol. Tools Appl. 2018, 129–169.

[2] Horn, Z.C.; Auret, L.; Mccoy, J.T.; Aldrich, C.; Herbst, B.M. Performance of convolutional neural networks for feature extraction in froth flotation sensing. IFAC Pap. Line 2017, 50, 13–18.

[3] Li, W.; Jia, F.; Hu, Q. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. J. Comput. Commun. 2015, 3, 146–151.

[4] Li, X.; Chen, H.; Qi, X.; Dou, Q.; Fu, C.; Heng, P. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes. IEEE Trans. Med. Imaging 2018, 37, 2663–2674.

[5] Rajeshwaran, M.; Ahila, A. Segmentation of Liver Cancer Using SVM Techniques. Int. J. Comput. Commun. Inf. Syst. (IJCCIS) 2014, 6, 78–80.

[6] El-Regaily, S.A.; Salem, M.A.M.; Aziz, M.H.A.; Roushdy, M.I. Lung nodule segmentation and detection in computed tomography. In Proceedings of the 8th IEEE International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 5–7 December 2017; pp. 72–78.

[7] Yuan, Y. Hierarchical convolutional-Deconvolutional neural networks for automatic liver and tumor segmentation. arXiv 2017, arXiv:abs/1710.04540.

[8] Yang, D.; Xu, D.; Zhou, S.K.; Georgescu, B.; Chen, M.; Grbic, S.; Metaxas, D.N.; Comaniciu, D. Automatic liver segmentation using an adversarial image-To-Image network. arXiv 2017, arXiv:abs/1707.08037.

[9] Shafaey, M.A.; Salem, M.A.-M.; Ebied, H.M.; Al-Berry, M.N.; Tolba, M.F. Deep Learning for Satellite Image Classification. In Proceedings of the 3rd International Conference on Intelligent Systems and Informatics (AISI2018), Cairo, Egypt, 1–3 September 2018.

[10] Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. arXiv 2014, arXiv:abs/1409.3215.

[11] Ke, Q.; Zhang, J.; Wei, W.; Damaševˇıcius, R.; Wozniak, M. Adaptive Independent Subspace Analysis of Brain Magnetic Resonance Imaging Data. IEEE Access 2019, 7, 12252–12261.

[12] Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 2481–2495.

[13] Ahmed, A.H.; Salem, M.A. Mammogram-Based cancer detection using deep convolutional neural networks. In Proceedings of the 2018 13th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 18–19 December 2018; pp. 694–699.

[14] Lu, F.; Wu, F.; Hu, P.; Peng, Z.; Kong, D. Automatic 3d liver location and segmentation via convolutional neural networks and graph cut. Int. J. Comput. Assist Radiol. Surg. 2017, 12, 171–182.

[15] Wang, P.; Qian, Y.; Soong, F.K.; He, L.; Zhao, H. Part-Of-Speech tagging with bidirectional long short-Term memory recurrent neural network. arXiv 2015, arXiv:abs/1510.06168.

[16] Abdelhalim, A.; Salem, M.A. Intelligent Organization of Multiuser Photo Galleries Using Sub-Event Detection. In Proceedings of the 12th International Conference on Computer Engineering & Systems (ICCES), Cairo, Egypt, 19–20 December 2017; Available online: https://ieeexplore.ieee.org/document/347/.

[17] Elsayed, O.A.; Marzouky, N.A.M.; Atefz, E.; Mohammed, A.-S. Abnormal Action detection in video surveillance. In Proceedings of the 9th IEEE International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 8–10 December 2019.

[18] El-Masry, M.; Fakhr, M.; Salem, M. Action Recognition by Discriminative EdgeBoxes. IET Comput. Vis. Dec. 2017. Available online: http://digital-library.theiet.org/content/journals/10.1049/iet-cvi.2017.0335

[19] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. arXiv 2015, arXiv:abs/1505.04597.

[20] Available Online: https://www.hepatitisc.uw.edu/go/management-cirrhosis-related-complications/liver-transplantation-referral/core-concept/all

[21] Ke, Q.; Zhang, J.; Wei, W.; Połap, D.; Wo ´zniak, M.; Ko´smider, L.; Damaševˇıcius, R. A Neuro-Heuristic approach for recognition of lung diseases from X-Ray images. Expert Syst. Appl. 2019, 126, 218–232.

[22] Noh, H.; Hong, S.; Han, B. Learning Deconvolution Network for Semantic Segmentation. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1520–1528.

[23] Vadera, S.; Jone, J. Couinaud classification of hepatic segments. Available online: https://radiopaedia.org/ articles/couinaud-classification-of-hepatic-segments.

[24] Golan, R.; Jacob, C.; Denzinger, J. Lung nodule detection in CT images using deep convolutional neural networks. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 243–250.

[25] Available Online: https://transplantliving.org/organ-facts/liver/

[26] El-Regaily, S.A.; Salem, M.A.; Abdel Aziz, M.H.A.; Roushdy, M.I. Survey of Computer Aided Detection Systems for Lung Cancer in Computed Tomography. Current. Med. Imaging 2018, 14, 3–18.

[27] Perez, L.; Wang, J. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning. arXiv 2017, arXiv:abs/1712.04621.

[28] Sun, C.; Guo, S.; Zhang, H.; Li, J.; Chen, M.; Ma, S.; Jin, L.; Liu, X.; Li, X.; Qian, X. Automatic segmentation of Liver tumors from multiphase Contrast-Enhanced CT images based on FCN. Artificial Intell. Med. 2017, 83, 58–66.

[29] Badura, P.; Wieclawe, W. Calibrating level set approach by granular computing in computed tomography abdominal organs segmentation. Appl. Soft Comput. 2016, 49, 887–900.

[30] Shi, C.; Cheng, Y.; Liu, F.; Wang, Y.; Bai, J.; Tamura, S. A hierarchical local Region-Based sparse shape composition for liver segmentation in CT scans. Pattern Recognit. 2015.

[31] Hesamian, M.H.; Jia, W.; He, X.; Kennedy, P. Deep Learning Techniques for Medical Image Segmentation: Achievements and challenges. J. Digit. Imaging 2019.

[32] Liu, H.; Jiang, H.; Zheng, R. The Hybrid Feature Selection Algorithm Based on Maximum Minimum Backward Selection Search Strategy for Liver Tissue Pathological Image Classification. Comput. Math. Methods Med. 2016, 2016, 7369137.

[33] Ding, J.; Fu, L. A Hybrid Feature Selection Algorithm Based on Information Gain and Sequential Forward Floating Search. J. Intell. Comput. 2018, 9, 93–101.

[34] Hassan, A.; Abou-Taleb, A.S.; Mohamed, O.A.; Hassan, A.A. Hybrid Feature Selection approach of ensemble multiple Filter methods and wrapper method for Improving the Classification Accuracy of Microarray Data Set. Int. J. Comput. Sci. Inf. Technol. Secur. 2013, 3, 185–190.

[35] Available Online: https://theconversation.com/how-to-grow-human-mini-livers-in-the-lab-to-help-solve-liver-disease-121297

[36] Chuang, L.Y.; Ke, C.H.; Yang, C.H. A Hybrid Both Filter and Wrapper Feature Selection Method for Microarray Classification. arXiv 2016, arXiv:1612.08669

[37] Zhou, N.; Wang, L. Processing Bio-medical Data with Class-Dependent Feature Selection. In Advances in Neural Networks Computational Intelligence for ICT; Bassis, S., Esposito, A., Morabito, F.C., Pasero, E., Eds.; Springer: Cham, Switzerland, 2016; Volume 54; pp. 303–310.

[38] Available Online: https://radiopaedia.org/cases/fatty-liver-grade-i-1

[39] Christ, P.F.; Elshaer, M.E.A.; Ettlinger, F.; Tatavarty, S.; Bickel, M.; Bilic, P.; Rempfler, M.; Armbruster, M.; Hofmann, F.; D'Anastasi, M.; et al. Automatic liver and lesion segmentation

in CT using cascaded fully convolutional neural networks and 3D conditional random fields. arXiv 2016, arXiv:abs/1610.02177.

[40] Christ, P.F.; Ettlinger, F.; Grün, F.; Elshaer, M.E.A.; Lipkova, J.; Schlecht, S.; Ahmaddy, F.; Tatavarty, S.; Bickel, M.; Bilic, P.; et al. Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks. arXiv 2017, arXiv:abs/1702.05970.

[41] Available Online: https://seneddresearch.blog/2015/05/11/liver-disease-in-wales/

[42] Wei, W.; Zhou, B.; Połap, D.; Wo´zniak, M. A regional adaptive variational PDE model for computed tomography image reconstruction. Pattern Recognit. 2019, 92, 64–81.

[43] Whitcomb, E.; Choi, W.T.; Jerome, K.R.; Cook, L.; Landis, C.; Ahn, J.; Westerhoff, M. Biopsy Specimens from Allograft Liver Contain Histologic Features of Hepatitis C Virus Infection After Virus Eradication. Clin. Gastroenterol. Hepatol. 2017, 15, 1279–1285.

[44] Available Online: https://www.cdc.gov/mmwr/volumes/66/wr/mm6638a9.html

[45] Bellver, M.; Maninis, K.-K.; Pont-Tuset, J.; Nieto, X.G.; Torres, J.; Gool, L.V. Detection-Aided liver lesion segmentation using deep learning. Machine Learning 4 Health Workshop - NIPS. arXiv 2017, arXiv:abs/1711.11069.

[46] Ben-Cohen, A.; Klang, E.; Amitai, M.M.; Goldberger, J.; Greenspan, H. Anatomical data augmentation for CNN based pixel-wise classification. In Proceedings of the IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 1096–1099.

[47] Available Online: https://www.nejm.org/doi/full/10.1056/nejmra1001683

[48] Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. CA Cancer J. Clin. 2017, 61, 60–90.

[49] K. Xia, H. Yin, P. Qian, Y. Jiang and S. Wang, "Liver Semantic Segmentation Algorithm Based on Improved Deep Adversarial Networks in Combination of Weighted Loss Function

on Abdominal CT Images," in IEEE Access, vol. 7, pp. 96349-96358, 2019, doi: 10.1109/ACCESS.2019.2929270.

[50] Xu, Sendren Sheng-Dong & Chang, Chun-Chao & Su, Chien-Tien & Phu, Pham. (2019). Classification of Liver Diseases Based on Ultrasound Image Texture Features. Applied Sciences. 9. 342. 10.3390/app9020342.

[51] Ribeiro, Ricardo & Marinho, Rui & Velosa, José & Ramalho, Fernando & Sanches, João. (2011). Chronic liver disease staging classification based on ultrasound, clinical and laboratorial data. Proceedings - International Symposium on Biomedical Imaging. 707 - 710. 10.1109/ISBI.2011.5872504.