T.C.

ANTALYA BILIM UNIVERSITY

INSTITUTE OF POSTGRADUATE EDUCATION

DISSERTATION MASTER'S PROGRAM OF ELECTRICAL AND COMPUTER ENGINEERING

Breast Cancer Classification Using deep Neural Network

DISSERTATION

Prepared By

M. Uzair Javed

Student ID- 181212024

ANTALYA – 2020

T.C.

ANTALYA BILIM UNIVERSITY

INSTITUTE OF POSTGRADUATE EDUCATION

DISSERTATION MASTER'S PROGRAM OF ELECTRICAL AND COMPUTER ENGINEERING

Breast Cancer Classification Using deep Neural Network

DISSERTATION

Prepared By

M. Uzair Javed

Student ID- 181212024

Thesis Advisor

Asst. Prof. Dr. Shahram Taheri

ANTALYA – 2020

ANTALYA BİLİM UNIVERSİTY

INSTITUTE OF POST-GRADUATE EDUCATION

M. Uzair Javed, a M.Sc. student of Antalya Bılım University, Institute of Post Graduate Education, Electrical, and Computer Engineering owning student ID 181212024, successfully defended the thesis/dissertation entitled “Breast Cancer Classification Using Deep Neural Network”, which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Academic Tittle, Name-Surname, Signature

Thesis Advisor **:** Prof. Dr. Shahram Taheri, ……

Jury Member **:** Asst. Prof. Dr. Yusuf ÖZTÜRK, ……..

Jury Member **:** Asst. Prof., Dr. Hüseyin Gökhan AKÇAY ……..…..

Director of The Institute **:** ……..., ……

Date of Submission: --28/10/2020

Date of Defence: --28/10/2020

ABSTRACT

Breast Cancer Classification Using Deep Neural Network

Over the last few decades, cases of breast cancer have increased enormously. It is the second most popular cause of deaths in women in both developed and undeveloped countries. 8 out of 100 women face this popular and dangerous disease in their life period. The only way to cure this disease is to detect breast cancer at early stages. Delay in identifying breast cancer leads to an increase in the death rate. An appropriate data representation determines the performance of classification systems. In this work, we have done some classification on Breast Cancer histopathological images from publically available Break-His dataset using machine learning and deep learning techniques. We proposed different classifiers in our work, Resnet-101, Resnet-18, and Densenet-201, etc as Cnn and multiple handcrafted features like LBP, HOG, and MPT for more accurate classification of breast cancer images. Deep learning extract and organizes features from data. We have organized and prepared a competitive comparison of these different implementations by evaluating their accuracy using deep learning and machine learning techniques. We also organized competitive results for handcrafted feature extractors and matched CNN and handcrafted features extractor accuracies with recent work done. We have achieved some enormous results using these different techniques. We brought in light that how deep networks and CNN are taking the place of handcrafted feature extractors in different image classifications.

Keywords: Convolutional Neural Networks, Deep Learning, Deep Neural Networks, Local Binary Patterns, Resnet-101, Densenet-201, Computer-Aided Diagnosis Systems

DEDICATION AND ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisors, Prof. Dr. Shahram Taheri for his guidance, constant encouragement, endless support, and great help throughout this journey. I am glad to have work with them and extremely grateful for the knowledge and experience. Their dedication and close follow up with their students is something to be admired and truly appreciated. I have been able to learn and acquire under their mentorship. I am very grateful to all the professors who taught me and helped me in learning a lot on this during this Master program.

To the greatest persons in my life: my mother, my father, and my brother. I dedicate this work to you and I hope I can always make you very proud and happy. Your limitless, undying love and daily calls mean the world to me.

CONTENTS

INSTITUTE OF POSTGRADUATE EDUCATION

ELECTRICAL AND COMPUTER ENGINEERING

PROGRAM WITH THESIS

ACADEMIC DECLARATION

I hereby declare that this master's thesis titled "Breast Cancer Classification Using Deep Neural Network" has been written by myself under the academic rules and ethical conduct of the Antalya Bilim University.

I also declare that the work attached to this declaration complies with the university requirements and is my work.

I also declare that all materials used in this thesis consist of the mentioned resources in the reference list. I verify all these with my honor.

28/10/2020

Muhammad Uzair Javed

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

CNN	Convolutional Neural Network
DNN	Deep Neural Network
ANN	Artificial Neural Networks
CAD	Computer Aided Diagnosis System
SVM	Support Vector Machine
RNN	Recurrent Neural Network
WHO	World Health Organization
MIB	Magnification Independent Binary Classification
MSB	Magnification Specific Binary Classification
MIM	Magnification Independent Multi Classification
MSM	Magnification Independent Multi Classification
RBM	Small Boltzmann machines
RGB	Red Green Blue
PAD	Presentation Attack Detection
PCA	Patient-Controlled Analgesia
LDA	Linear Discriminate Analysis
SFM	Structure-from Motion
MVS	Multi-View Stereo

ReLU	Rectified Linear Unit
LBP	Local Binary Patterns
HOG	Histogram of Gradients
CLBP	Complete Local Binary Patterns
POEM	Patterns of Oriented Edge Magnitudes
GDP	Gradient-Based Patterns
GLTP	Generalized Local Ternary Patterns
IWBC	Improved Weber local descriptor
LAP	Local Anisotropic Pattern
LDiP	Local Directional Pattern
LDN	Local Directional Number Pattern (LDN)
LTP	Local Ternary Pattern (LTP)
BSIF	Binarized Statistical Image Features
BPPC	Binary Pattern of Phase Congruence
LDTP	Local Directional Ternary Pattern
LFD	Local feature descriptor
LGBPHS	Local Gabor Binary Pattern Histogram Sequence
LGiP	Local Gradient Increasing Patterns
LGP	Local Gradient Patters
MBC	Monogenic Binary Coding
MBP	Monogenic Binary Patterns

MRELBP	Median Robust Extended Local Binary Patterns
MTP	Median Ternary Pattern
MWLD	Weber Local Descriptor
PT	Phyllodes Tumor
TA	Tubular Adenona
LC	Lobular Carcinoma
MC	Mucinous Carcinoma
PC	Papillary Carcinoma
ILA	Image Level Accuracy
PLA	Patient Level Accuracy
ILSVRC	Large Scale Visual Recognition Challenge

CHAPTER ONE

1. Introduction

Recent studies show that Breast cancer cases have increased enormously due to a lack of early detection [1] of the disease. Cancer has evolved as one of the most leading and massive health problems in the whole world. According to the world's renowned organization (WHO), 8.3 million deaths occurred due to breast cancer in 2012 and it is expected that cases will reach up-to 27 million by 2030 [2]. No of the Breast cancer cases are drastically increasing among women besides skin and lung cancer from recent years. The major cause of breast cancer by a typo or a mutation in a cell. These mutations can cause reckless cell division. It is expected that early detection of cancer can help in reducing the death rate and limiting it to reach a chronic level. Machine learning and deep learning techniques are being used worldwide by professionals to detect cancer at earlier stages. Throughout the world, digital mammography is a major diagnostic system and is recommended for Breast cancer detection by professionals. CAD systems are used worldwide for detecting different diseases including Breast cancer, lung cancer with the accurate decisions [3]. CAD systems are built in such a way that they provide a perfect decisions by examining the patient's conditions and help professionals to make the right decision and diagnose different stages of the disease.

Deep learning is dominantly getting used from speech recognition, face detection, tumors detection, and iris recognition to do different AI-related systems. For over the last 20 years [4] enormous no of new techniques and algorithms are developed using deep learning techniques.AI and Deep learning altogether have gained the attention of researchers for the last few years. Deep learning uses deep neural networks (DNN), a convolutional neural network with recurrent neural networks (RNN), and for supervised and un-supervised learning it uses Long-Short term memory. Recently breasts cancer deaths are kept under control due to enhancement and development in the detection and treatment of mammography.

Deep learning is the computer science method used for the classification and detection of different cancers at early stages. Classification methods associated with deep learning most rely upon methods in the field of medicine. The most efficient and powerful way to cope with complex

classification and detection of different diseases is by deep learning classification algorithms. CNN (Convolutional Neural Network): is class of deep learning and applied for visualizing digital images [5].Cnn is based on multiple perceptron concepts because they are fully connected like input layers are connected to all hidden layers etc.

Over time by time amount of cancer cases among women is increasing drastically. The major diagnosis model used frequently now a days around the world is Digital Mammography. Main cause of breast cancer can be family disease, hormones or may be reproductive factors [6]. According to the report of WHO over more than millions of women are newly diagnosed with breast cancer every year and also many of them may die because professionals are unable to detect breast cancer at early stages.

Machine learning is composed of four different phases, collecting data, picking the suitable model, training the model and testing the model. Deep learning techniques have been used from so long time for breast cancer tumors and other malignant. In our study we proposed different neural classifiers for detecting breast cancer tumors mainly benign tumors and malignant tumors in early stages with more accurate results. [7] Malignant tumors increases and spreads to nearby cells that can reach to other parts while benign tumors does not expand. Purpose of this study is to develop effective machine learning approaches for cancer classification in a dataset by these different classifier and to achieve better accuracy rate and results.

1.1. Contributions

We provide an overview of the research contributions in this thesis towards the problem of breast cancer detection and classification at the early stages. At first, we provided the basic methods that are followed by researchers in their previous work, respectively what methods and techniques they used in their work. In our work, we focused on Magnification independent binary classification (MIB) and Magnification specific binary classification (MSB), used different techniques for tackling data imbalance and stain normalization, and followed techniques of old related work with different modifications. These steps and techniques helped in achieving some promising results in our respective work. Next, we described different CNN's and handcrafted features that are used for detection and classification. We thoroughly covered most of the CNN's and Feature descriptors for developing a better understanding related to the classification of different tumors. We described a thorough study of the dataset including the pros and cons and

also the problems that we faced using dataset. Lastly, we provide extensive evaluations to confirm the performance of the proposed methods compared to the state-of-the-art methods, along with crucial insights to the biological findings.

1.2. Thesis Organization

This thesis is organized in such a way that literature review and the previous researches and basic old methodologies are described in chapter 2.Chapter 3 contains post-applied approaches and also major approaches that we applied and can be applied in the future. Chapter 4 contains information on the dataset that we used for our work and chapter 5 consists of experiments that we performed using multiple approaches and our results that how they differ from other work. In this chapter, we presented a conclusion about which approaches for better for what kind of image classifications.

CHAPTER TWO

2. Literature Review

A huge number of techniques are developed for breast cancer CAD systems. The accuracy of these techniques has been evaluated using the Break-His data set and UCI repository dataset. Neural network's idea is based on neurons that how they are interconnected with each other. Their working principle follows the working of biological neural networks. Dendrites are the input and output channels of biological neural networks. An ANN forms a highly connected network due to its millions of processing units which in results processes a large amount of information. A neural network is composed of 1 input layer and 1 or more hidden layer which is connected to the output layer. Simple deep neural network layers are defined in Fig. 1 that how each input image goes through different layers and how features and features maps are extracted from the input image.

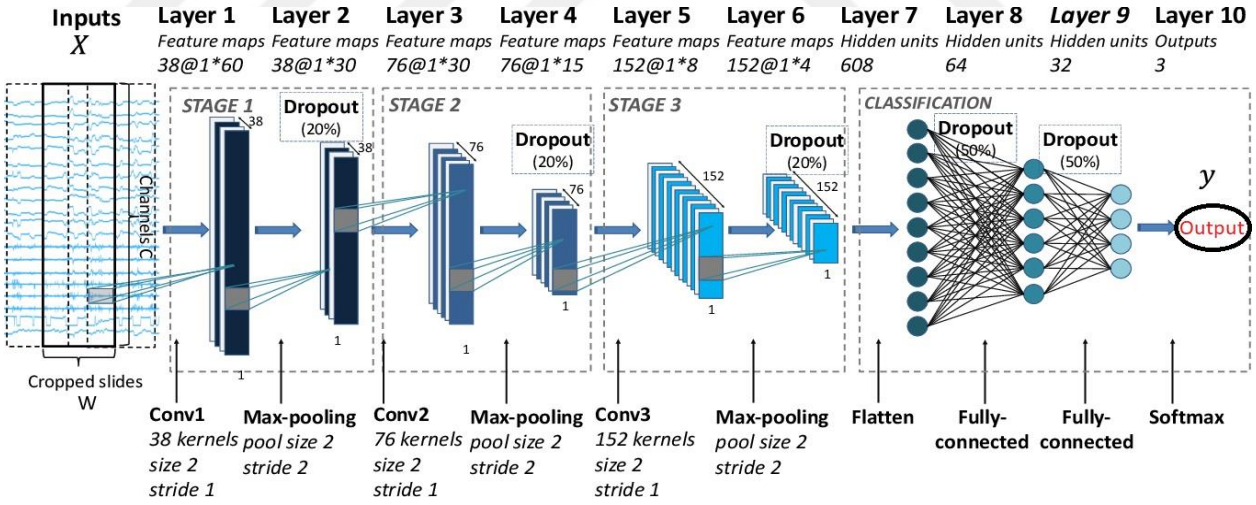

Figure 1: Deep Neural Network Architecture [7]

One of the frequently used learning algorithm for single hidden layer is Extreme machine learning algorithm. It was introduced by Huang ET al. It randomly initializes initial inputs and biases and analyzed calculation of output weights. Few steps are required for training a network. Abbas developed an Emetic pareto-artificial neural network that was composed with pareto-differential evaluation algorithm with local search scheme [7].

Y. Benhammou and S. Tabik in their work achieved the highest accuracy of 85.7% using inception v3 on an image level. [8].In their work median filter was used for and preprocessing and min-max techniques were used for data normalization and none of the data was removed from thedataset. They used 70% for training and 30% for testing in their work for achieving such efficient accuracy [9]. AN expert system was developed by Karabatak and Ince for breast cancer classification in which they used association rules for reducing the dimension of dataset [10, 9]. In this system, they proposed AR1 and AR2 for features reduction. The classification was carried out on both AR1 and AR2 by a convolutional neural networks. They attained an accuracy of 95.6% on AR1 and 97.1% [11] by using 3-fold cross-validation.

Quinlan reached 94.74% accuracy by using C4.5 [12] decision tree method. Abonyi and Szeifert achieved 95.47% accuracy by using the Supervised Fuzzy Clustering (SFC) technique. Pena-Reyes and Sipper used the Fuzzy –GA method and achieved an accuracy of 97.16% [13]. Particle Swarm Optimized Wavelet a new breast cancer classification approach was used by Dheeba, J. ET Aland detected abnormalities in breast mammograms. They achieved 93.671%, 92.105%, and 94.167% for accuracy, specificity, and sensitivity [14].

Meriem AMRANE and Saliva OUKID used machine learning algorithms, Naive Bayes (NB) classifier, and k-nearest neighbor (KNN) for classification of breast cancer images [15]. They evaluated accuracy using Cross-validation algorithms. Their main area of interest was to classify binary cancers that are benign and malignant. The data set they used is the UCI Repository dataset provided by the University of California. This dataset is comprised of 699 clinical cases. In their work, they first divided their dataset into two parts, training, and testing. By using KNN they achieved a high accuracy of 97.11% [15] that was quite precise and by using Nearest Neighbor they achieved an accuracy of 96.19%.

Jouni et al proposed a model that was comprised of BPNN and multilayered perceptron networks [13] along with artificial neural networks. This model is used to classify tumors whether they are benign or malignant. They also used Wisconsin Breast Cancer Dataset (WBCD) from the UCI repository for breast cancer classification and the classifier included weight factors and bias values as well.

Mohamed NEMISSI, Halima SALAH [16] proposed a classification system that was based on hidden neural networks and the proposed system is trained by extreme machine learning

algorithms. Wisconsin Breast Cancer Dataset (WBCD) from the UCI repository was used for benign and malignant tumors classification. Their work mainly focuses on the activation function of hidden neurons and for their optimization, they used genetic algorithms.

Ali Raad, Ali Kalakech, and Mohammad Ayache upon using Wisconsin Diagnostic and Prognostic Breast Cancer datasets achieved 97% [14] success by using RBF neural network. Ahmed M. Abdel-Zaher , Ayman M. Eldeib working on image classifications developed a more improved CAD system by using a deep belief network unsupervised path which was further followed by back propagation supervised [15] path.

In this study, we applied deep learning techniques on our publically available BreakHis dataset [9] by using 70% data for training and 30% data for testing using deep learning techniques. We used CNN classifiers and handcrafted feature descriptors for the classification of our data set. Our proposed classifiers achieved great accuracies with some précised results.

2.1. Deep learning Approach using Pre-trained neural network

Image classification is a particular process that takes some input as an image and gives output as a class. Image classification is challenging mainly microscopic images from the histopathologic section because they have issues like a large amounts of intra-class variability, rich geometric structures, and complex textures. Deep learning explores results from input data rather than hand-crafted features.

2.1.1. Deep Neural Networks

Deep neural networks (DNN) mainly called artificial neural networks have multiple layers between input and output layers. They find precise mathematical manipulation in both linear and non-linear relationships and convert input layers to output. DNN is also called Convolutional Networks. In DNN computations are performed by different layers. Computations for a network of the hidden layer is denoted as:

$$f(x)=f[a^{(L+1)} \quad (h^{(L)} (a^{(L)}](\ldots(h^{(2)} (a^{(2)} (h^{(1)} (a^{(1)} (x))))) \text{ [17]}$$

Each pre-activation function a (l) (x) is typically a linear operation with matrix W (l) and bias b (l), which can be combined into a parameter θ:

$$a^{(l)}(x) = W^{(l)}x + b^{(l)} \text{ [17]}$$

$$a^{(l)}(x) = \theta^{(l)}x \text{ , } l=1 \text{ [17]}$$

$$a^{(l)}\, h^{(l-1)} = \theta^{(l)}\, h^{(l-1)} \text{ [17]}$$

Hidden-layer activation functions h(l)(x) often have the same form at each level, but this is not a requirement.

A deep neural network is the hierarchical architecture of neurons where each neuron is connected to other neurons. Same as neuron this learning method has layers that pass messages from one layer to another and form a complex network that learns with some feedback mechanism. Fig 2 illustrates the input, hidden, and final layers of DNN. [18]

A deep learning algorithm Convolutional neural network also called CovNet takes an input images and assign some sort of weights or biases to this image so that it can differentiate one from another. Covnet has very much less preprocessing as compared to other algorithms. Filters are hand engineered in primitive methods, [18] Covnets are used to avoid hand engineered algorithms because they are pretrained on datasets. Deep learning is comprised of numerous algorithms that follows hierarchal structures. Deep transfer learning are based on CNN. CNN's are pretrained networks and they classify ImageNet. [18]

An algorithm learns the same way as a child learns because a lot of images are to be showed to algorithms for training and testing. Computers 'see' the world in a different way than we human beings do. Their world consists of only values and numbers. Every image represented as 2-dimensional arrays of numbers can be known as pixels. CNN Architecture is composed of three layers. Convolutional layer is composed of extracting features that extracts input features such as endpoints, corner, and edges. Convolution layer extracts features. Pooling Layer Function Reduces the dimensionality of each feature map but retains the most important information. The pooling layer decreases the image resolution which reduces the precision accuracy of the effect (shift and distortion). All the activations in the previous layer are completely connected to the layer. They

act as a classification system and produce the production. Fully connected, each two neurons are connected in each layer.

2.2. The proposed Classification algorithms

The major aim of our work is to make awareness about pre-trained neural networks for breast cancer classification. These pre-trained neural networks include GoogLeNet, densenet201, resnet-101, resnet-18, mobilenet-v2 etc. In Break-His data set we have two classes benign and malignant. The breast histopathologic images are organized in four different magnification factors, i.e., 40x, 100x, 200x, 400x. According to theorems and old research work we know that feed forward network with single input are enough for representing any function. Since AlexNet have 5 convolutional layer and Google net have 22 layers still it is hard and challenging for training a deep network because of vanishing gradient problems and data over-fitting [19].The mean purpose of introducing Resent and other pretrained neural network in deep learning world was to "identify shortcut connections" that can jump one or more layers. Similarly resnet18 is another pretrained neural network that contains 18 layers, resnet-101 includes 101 layers, and similarly Densenet 201 contains 201 layers. Their main concern is to classify millions of objects with much accuracy and with high performance. Transfer learning techniques are widely in used these days. Transfer learning improves the accuracy of classifier and also accelerates the learning process [19].

2.2.1. Hand Crafted Features

Hand crafted features are being used from past of the years for image classification and object detection [20]. With time number of features is increased and has shown some enormous results in object detection and image classifications. Feature based approaches are mainly combination of different features. Most commonly hand crafted features are buildup of 3 major techniques : Local Phase Quantization (LPQ) [20], Local Ternary Pattern (LTP) and Local Binary Patterns (LBP) Histogram of Oriented Gradients (HOG) and Gabor filter .We have used multiple hand craft feature descriptors in our work. Mostly feature descriptors are based on LBP.LBP feature extraction techniques is widely in used by researches for Computer vision technologies such as facial recognition and recognition of the face of expression, recognition of finger-veins, and estimation of human age. Fig. 2 illustrates how different feature maps are extracted from an input image using deep learning principle.

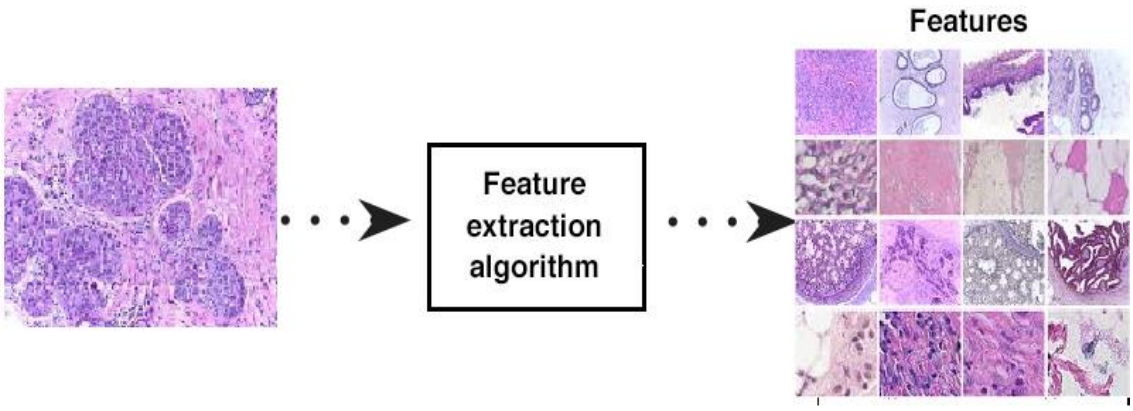

Figure 2: Deep learning Principle

In different applications systems, biometric recognition systems are now commonly used because they are difficult to steal, are highly reliable to recognize and convenient for users. The disparity between people's physical and behavioral properties is based on this recognition process. For example, fingerprints and faces have been used historically for identifying a person. Several biometric features have been used in identification Face and fingerprint features such as iris, veins, ears, or pal-prints [21]. Methods for the face recognition program have been investigated for the presentation attack detection (PAD) [21]. Past researches are divided into two groups of methods of extraction based on non-training and training based feature extraction. As the first level, the sparse low rank, bilinear, differential image model used by Tan et al. to distinguish images from other forms of presentation-attack, derived by varying Gaussian (DoG) [22] or logarithmic total (LTV) approaches. They showed that the real and presented image attacks can be discriminated against using its own tool using the NUAA database. Maatta et al. [23] used three feature extraction methods, Gabor filter, Local Phase Quantization (LPQ) and Local Binary Pattern (LBP), with the NUAA database, to extract image features and identify actual and presentation attack images using vector support machines (SVMs). The classification error was substantially reduced compared to those in Tan et al's study, based on their results.

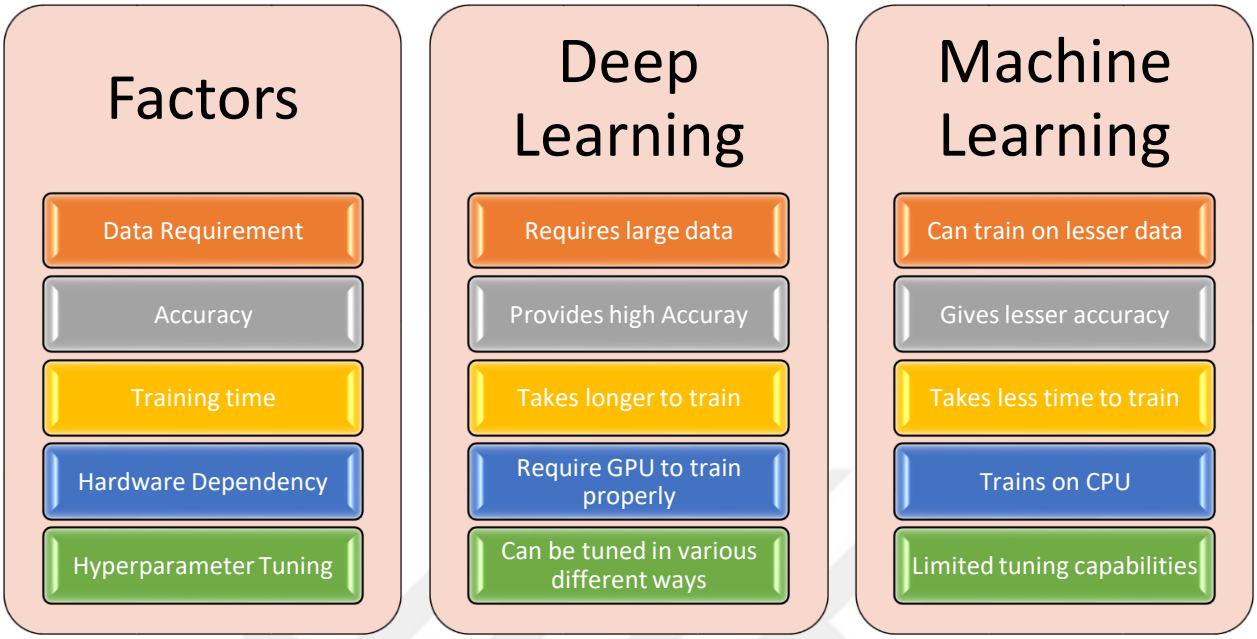

Figure 3: Machine learning vs. Deep Learning

Fig. 3 includes complete difference between deep learning and machine learning that which one is easy to train or which gives more high accuracy etc. The number of features has increased over time to allow researchers to better adapt to the various tasks being dealt with. The features we have developed are based on LBP in our Local Binary Pattern, Combined Local Binary Pattern (CLBP) [23] research. Both these approaches remove the overall complexity, so that a standard computer vision architecture can be better created.

2.2.2. Deep Learning-based Approaches

Deep neural networks mainly called as artificial neural networks have multiple layers between input and output layers. They find precise mathematical manipulation in both linear and non-linear relationships and converts input layers to output. DNN is also called Convolutional Networks. For last 10 years, deep neural learning methods have become increasingly common because of the advanced technologies in language recognition, computer vision and other fields. Our recent success can be attributed to increased data availability, improved hardware and software, and numerous algorithmic breakdowns, which accelerate training and generalize new data. Mostly used learning algorithms for classification and edge detection are shallow architectures (1-3 levels) (KNN, MOG, SVM, KDE, Parzen Kernel regression, perceptron, PCA)

.In DNN computations are performed by different layers. Computations for a network of hidden layer is denoted as:

$$f(x)=f[a^{(L+1)} \quad (h^{(L)} (a^{(L)}](\ldots(h^{(2)} (a^{(2)} (h^{(1)} (a^{(1)} (x))))) [12]$$

Each pre-activation function a (l) (x) is typically a linear operation with matrix W (l) and bias b (l), which can be combined into a parameter θ:

$$a^{(l)} (x)= W^{(l)}x + b^{(l)} [24]$$

$$a^{(l)} (x)= \theta^{(l)}x \text{ , } l=1 [24]$$

$$a^{(l)} h^{(l-1)} = \theta^{(l)} h^{(l-1)} [24]$$

Hidden-layer activation functions h (l) (x) often have the same form at each level, but this is not a requirement.

Deep neural network is neuron hierarchical architecture where each neuron is connected with other neurons. Same as neuron this learning method has layers that passes message from one layer to other and forms a complex network that learns with some feedback mechanism [25].Fig 5 illustrates deep neural network architecture that how it extracts features and converts into different maps the extracted features:

Covnet has very much less preprocessing as compared to other algorithms. While in primitive methods filters are hand-engineered, [25]. Figure 6 illustrates how an input image goes through different stages in traditional machine learning while in deep learning algorithms all these feature are read at single stage.

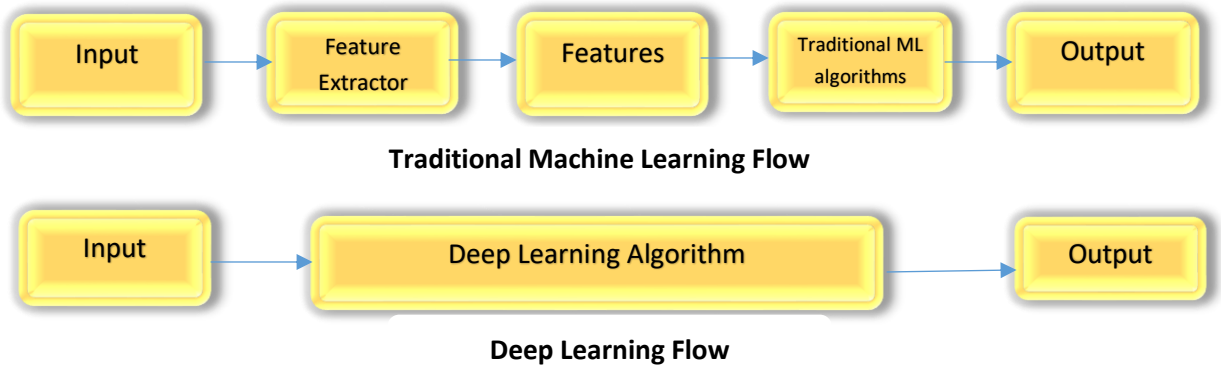

Figure 4: Machine learning flow vs. Deep learning Flow

Fig. 4 illustrates different processing and steps that are followed by machine learning algorithms and Deep learning algorithm. Usually in deep learning we have pre-trained networks so we don't have to follow different parameters while in machine learning we have to create different parameter for better results. Researchers reported positive experiment results at normal levels of two or three (e.g. one or two secret layers), but the more profound training resulted in worse results. Scientists have been reporting for decades of deep neural networks. Deep learning is a new process, which is used more and more in CAD systems and overcomes conventional computer education. Features are usually manually selected during computer education and fully automated in profound learning. Features like textures, colors, and edges can be retrieved during machine learning, while other hierarchical or compositional features can be accessed during the training [25]. Deep supervised network takes too much time for training and also they are difficult to train. CNN is less connected and less educated and have less parameters than deep supervised networks and it is easy to use.

CHAPTER THREE

3. BACKGROUND MATERIALS

3.1. Convolutional Neural Network (CNN)

For last few decades CNN has attained the attention of researchers because they have less parameters and easily trainable as compared to DNN.CNN are the most successful classifiers and tools for mostly accurate classification of breast cancers diagnosis, lung nodule detection, and prostate cancer localization. Deep learning methods are usually classifiable into four categories: CNN-based techniques, small Boltzmann machines (RBMs). The CNN-based approaches have recently gained more and more attention around the world, producing promising results in literature [26]. A typical CNN framework consists of three types of layer: Convolution layer, pooling layer and fully connected layer. Complete CNN architecture is illustrated in Fig. 5 is listed below with all respective layers:

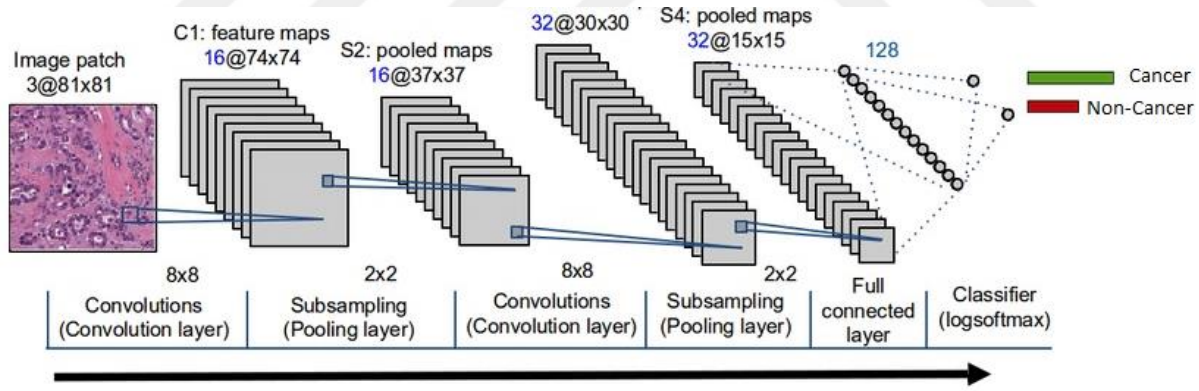

Figure 5: CNN Architecture [26]

Discussing about principles of CNN, when an image is given to a CNN it tries in every possible position to match the required features. We make a filter by calculating match across the whole image and this math work done or obtained filter is convolution. Pooling layer extracts important information from large images and shrink images into small size after obtaining important information. A pooling layer is simply a pooling of an image or a series of images. The output has equal numbers of images, but each pixel is smaller. Taking an 8 megapixel image down to a 2

megapixel image makes it downstream much simpler for life. Rectified Linear Unit or ReLU is small but most important layer in CNN's. [26]

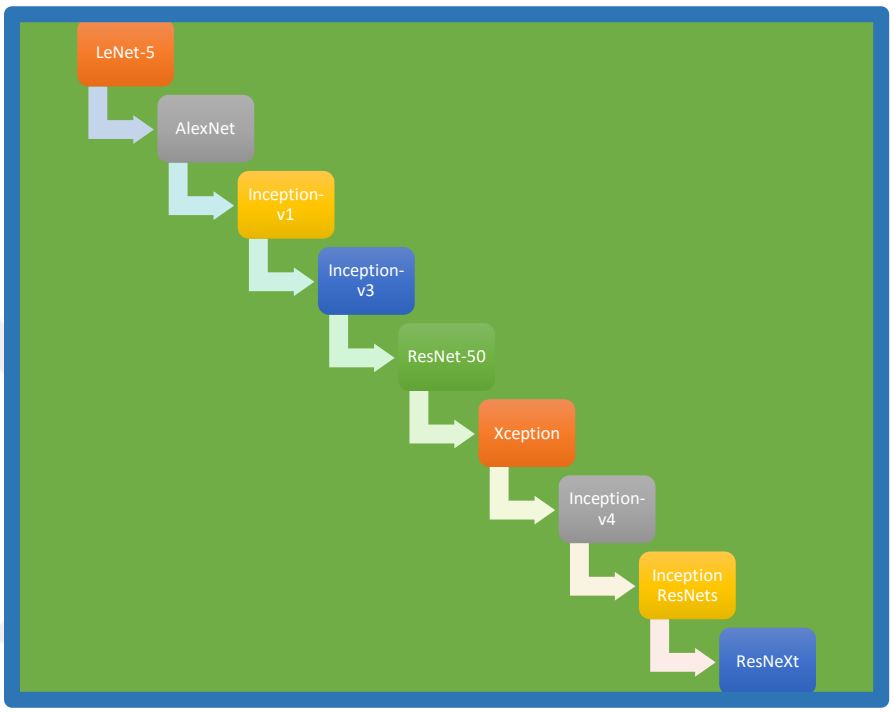

Figure 6: CNN architectures over a time from 1998-2019

Fig. 6 describes the formation of different CNN architectures over a time. The use of CNN-based image-based methods is distinct from natural image-based methods. On the other hand, for training and testing of CNNs, a large labeled dataset, such as the ImageNet, is needed. Rather of having RGB channels, [26] medical pictures are typically grayscale. However, medical image databases on a wide scale are not always available due to the extensive labeling research and the need for expert expertise.

3.2. Neural Network

Over the past 10 years, a methodology that is called "deep learning", has succeeded in the most powerful artificial intelligence systems [12] — for example speech sensors on smartphones, or Google's new automated translators. Deep learning is currently a modern concept for a neural networks technique that has been used and not used for more than 70 years. In 1944, two researchers at Chicago's McCullough and Walter Pitts proposed neural networks [25] and, in 1952,

moved to MIT, the director of the most widely recognized Cognitive Science Department. Neural networks were a key research field in both the neuroscience and the informatics until 1969, [27] when the MIT mathematician Marvin Minsky and Seymour Papert, co-directing the new MIT Artificial Intelligence Laboratory a year later, destroyed them, according to computer science reports. The strategy then made a comeback in the 1980s, was again eclipsed in the first decade of the new century, and returned, largely driven by the rising computing power in the graphical chips. A simple neural network with 2 input layered and several hidden layers is shown in fig 9:

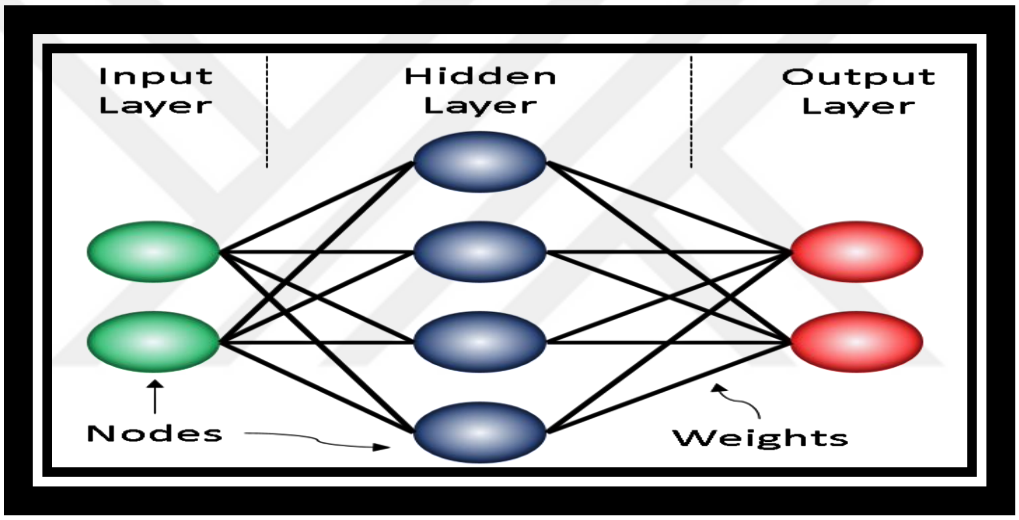

Figure 7: Simple Neural Network

A simple Neural Network with input layers, hidden layers and output layers is illustrated in Fig. 7. Neural networks are a way to teach machine learning, by analyzing training examples, a computer may do some research. The examples have typically been labelled manually beforehand. For example, an object recognition system may be used in the pictures, which consistently correspond with particular labels, by feeding thousands of marked pictures of vehicles, homes and coffee cups, and so on. Incoming connections will be assigned a number known as a "weight" by nodes. After a neural net has been formed, all its weights and thresholds are set on a random basis. The training data is transmitted to the lower levels – the input layer – and the following levels are reached, multiplying and adding together in dynamic ways before it eventually enters the output layer, dramatically transformed. The threshold and weights are continuously modified during preparation, until preparation data with the same labels reliably produce identical outputs.

Most famous pre-trained neural networks include GoogLeNet, densenet201, resnet-101, resnet-18 [27], MobileNet -v2 etc.

3.2.1. Resent

Resnet-1o1 is a pre-trained neural network and it consists of 101 layers. This architecture can be loaded on more than million images and provides outstanding results with great accuracy.Resnet-101 can classify images into huge no of objects e.g. dogs , face , gender , cancer etc. It takes an images input of 224-by-224.Similarly resnet18 is another pre-trained neural network that contains 18 layers and takes an input size image of 224-by-224 [28]. It I almost similar to resnet-101 as it also takes millions of images and has the ability to classify more than 1000 objects. Due to the notorious problem of vanishing gradient, deep networks are difficult to train. Resent is one of the first neural network classifier that skips one more layer for overcoming vanishing gradient problem.

Experiments show, however, that the Highway Network is nothing better than ResNet, a strange thing, given that the Highway network solution area contains ResNet and should therefore perform at least as well as ResNet. This means that such "gradient highways" are more critical than a broader variety of solutions. A drop down architecture also known as highway architecture of VGG-19 and Resnet-152 along with feed forward neural networks is shown in Fig. 8:

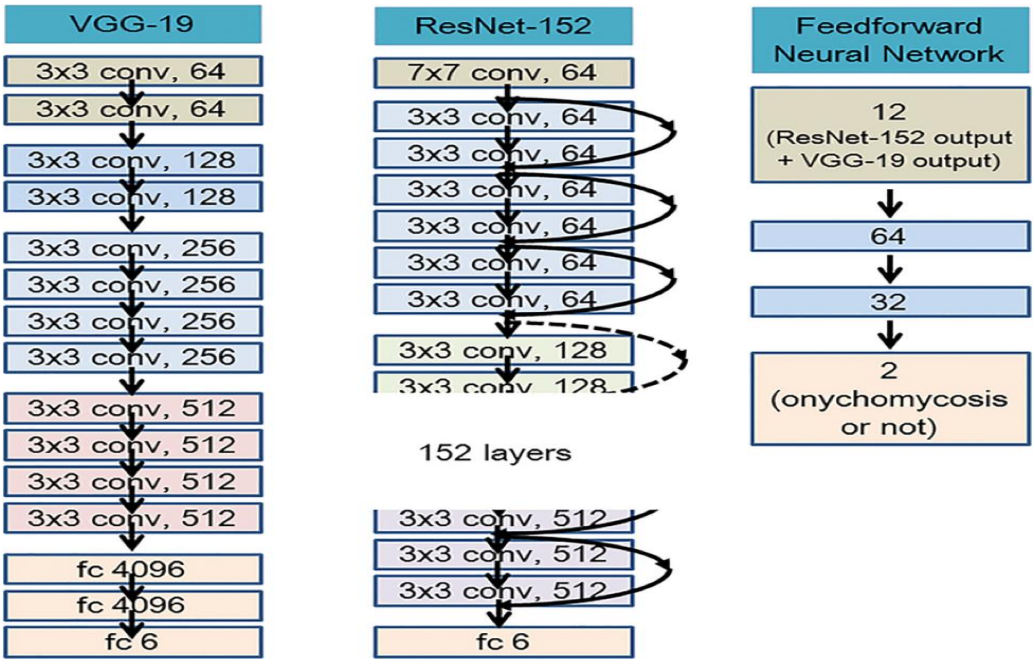

Figure 8: Resent Highway Structure [28]

Resent is mostly in study and in attention of different researchers these days and few new variants of ResNet are also introduced in few years. ResNext is the new variant of Resent that is proposed by Xie at el [23].

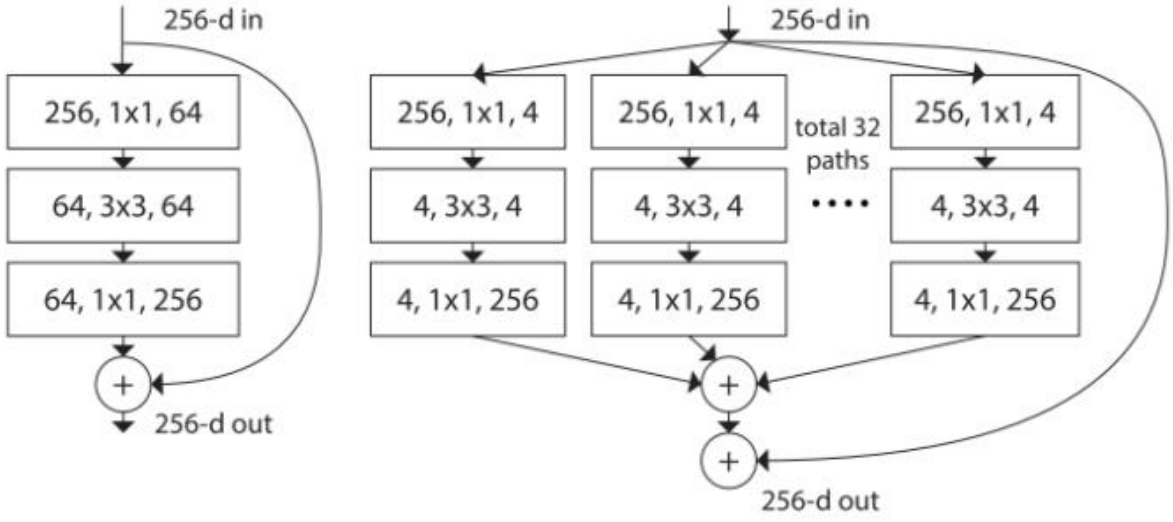

Figure 9: Blocks of ResNext [28]

Fig. 9 shows that how features set are extracted from input image through these blocks. This structure mostly look like same as inception module because it also follows the same split-transform-merge paradigm as inception module but in this variant output is carried out by merging different paths and then adding them. Each path in inception is different from other while in this variant architecture follows same path topology.

3.2.2. Densenet

Densenet also known as shortcut connection architecture—was proposed by Huang et al, this structure connects all the layers directly with each other. One of its well-known variant .Densenet-201 is also a renowned pre-trained neural network with having 201 deep layers. It also takes an input image of size 224-by-224 [28]. It can also classify millions of objects with much accuracy and with high performance. Each layer's input consists of feature maps of each earlier layer, which are then transferred to each layer. The characteristics of the maps are applied to depth. About this architecture authors say that it ensures feature reuse and makes the network highly parameter-efficient. One easy understanding is that the identity mapping output has been applied

to the next block, which may prevent information flow if the two-layer function map has very different distributions. Concatenated feature maps can therefore retain them all and increase the performance variance, which promotes reuse of the features. Densenet architecture with its layers architecture in listed down:

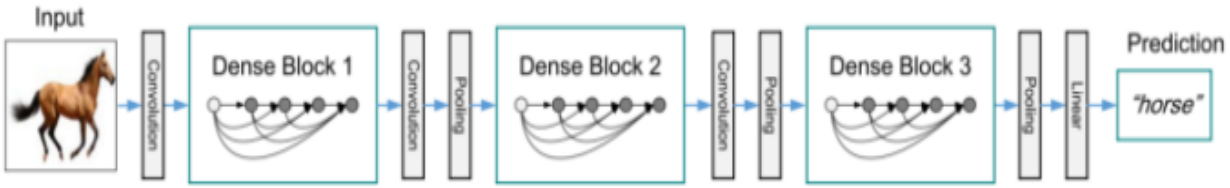

Figure 10: Layers Structure of Densenet [29]

Fig. 10 illustrates Densenet architecture that how can image moves through different layers and block and how features are extracted through these layers and at the end output is predicted in reference to extracted features. Due to this model we know that al' th layer is to have input maps of k * (l-1) + k 0 where the number of input image channels is k 0. The researcher used a growth rate(k) of hyper parameter and a 1x1 convolution bottleneck layer to reduce the number of maps until the expensive convolution (3x 3) [4] is reached, to prevent the network from becoming too big. The complete layer to layer architecture is reflected in the Fig. 11 below:

Layers	Output Size	DenseNet-121($k = 32$)	DenseNet-169($k = 32$)	DenseNet-201($k = 32$)	DenseNet-161($k = 48$)
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	[1 × 1 conv, 3 × 3 conv] × 24	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 48	[1 × 1 conv, 3 × 3 conv] × 36
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	[1 × 1 conv, 3 × 3 conv] × 16	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 24
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

Figure 11: Complete Flow Densenet Architecture [29]

3.2.3. Googlenet (inception_v1)

In the research paper "Going Deeper with Convolutions", Google Net (or the Inception V1) was proposed by the research team at Google in 2014 [30]. GoogLeNet won the image classification challenge at ILSVRC 2014. Compared to previous winners AlexNet (ILSVRC 2012 winner) and ZF-Net (ILSVRC 2013 winner), [30] the error rate was significantly lower than that of VGG (2014 runner up) as a result of the error rate. 1x1 convolutions and global average pooling techniques are used by this architecture.

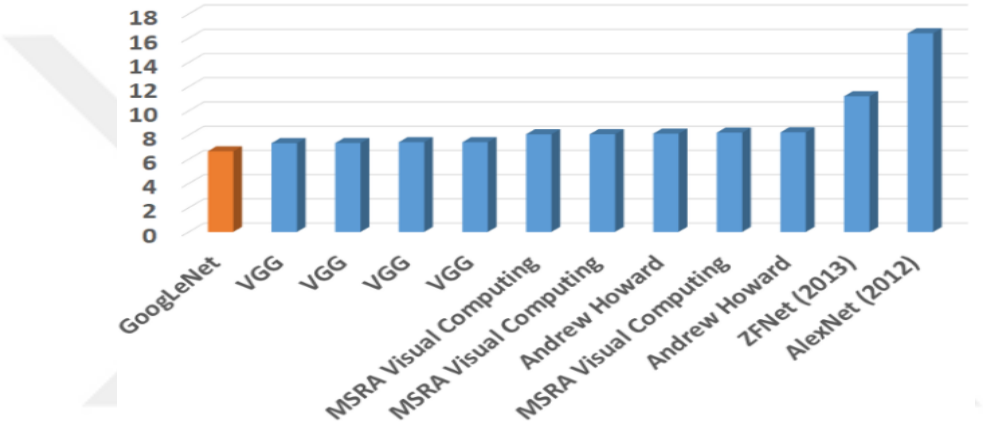

Figure 12: Error rate in ILSVRC 2014

According to ILSVRC 2014 report Error rate among different CNN's is illustrated in Fig. 12. 1x1 convolutional architecture used by GoogLeNet was introduced by NIN and this architecture uses RELU with 1x1 convolution. 1 to 1 convolution in GoogLeNet is used as a dimension reduction module which in turn reduces computations. The depth and width can be improved by the reduction of bottleneck computation. A bottleneck GoogLeNet architecture is shown in Fig. 13

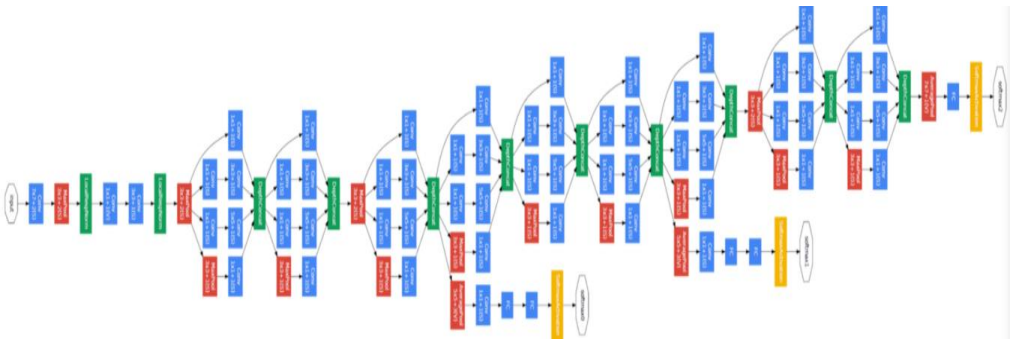

Figure 13: GoogLeNet Architecture [30]

This architecture is comprised of total of 22 layers. Compared to previous AlexNet, ZFNet, and VGGNet, it is already a very deep model. So we can see that multiple initial modules are wired together to go deeper. (But not too deep in contrast to ResNet created afterwards.) (There are some intermediate branches of softmax in the middle.Softmax branches consist of 5x5 pooling layers having 3 stride and 1x1 convolutional with 128 filters with 1024 fully connected layers.

3.2.4. MobileNet

MobileNet consists of 2 layers. The first layer is referred to as a depth wise convolution, and makes lightweight filtering with a single convolutionary filter. Point-wise convolution which is second layer of MobileNet is 1x1 convolution, that is responsible for compiling new features by computing linear input channel combinations.Mobilenet_v2 consists of two different types of blocks . 1 block for drowsing and other residual block with stride 1. All forms of blocks have 3 layers. The first layer this time is 1x1 convolution with ReLU6 [31]. Depth-wise convolution is second layer . The third layer is another 1x1l, with no non-linearity. It is stated that if ReLU is reused the deep networks would only have the power of a linear classifier on the output domain's non-zero volume portion. MobileNet V2 architecture with different layers and parameters in shown in Fig. 14:

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	

Figure 14: MobileNet V2 architecture [31]

3.3. Transfer learning

The basic principle of transferring learning is to use information gained from activities for which there are many labeled data available in settings where there is only a limited amount of labeled data. It is expensive to construct labeled data, therefore making efficient use of existing

datasets is important. The primary goal in a traditional model of machine learning is to generalize unknown data based on patterns learned from training data. With transfer learning, you seek to kick-start this cycle of generalization by starting from patterns learned for a specific task. Essentially, instead of beginning the learning cycle from a blank sheet (often initialized at random), you start from patterns that were learned to solve a particular problem. In any form of learning transferring learning is important [32]. To be successful in this, humans are not taught every single task or question. Everybody gets into challenges that have never been experienced, and we try to solve problems in an ad-hoc way. The opportunity to learn from a wide variety of perspectives and export 'information' to new worlds is just what learning transfer is all about. In that point, on a logical level, transfer learning and generalization are very close. The key difference is that learning transfer is mostly used to 'transfer information through tasks, rather than generalizing within a single task.

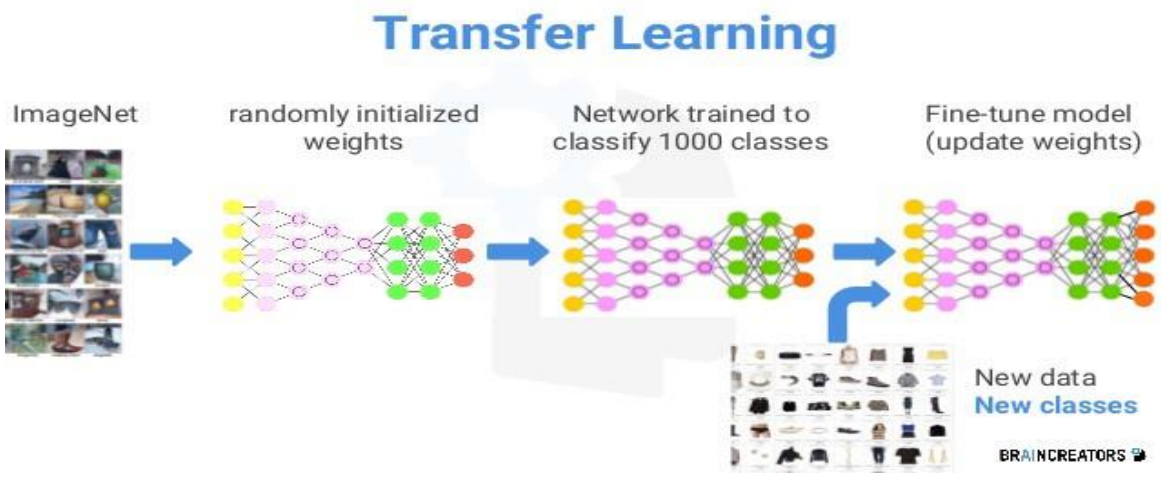

Figure 15: Principle of Transfer Learning [32]

Fig. 15 ensure the complete workflow of Transfer learning that how a new data is classified on fine-tuned pre-trained neural networks for better results. To ensure the advancement of deep learning techniques in a large number of small-data environments, transfer learning is crucial. Deep learning in science is pretty much everywhere, but many real-life situations usually don't have millions of data points labeled for training a model. Deep learning techniques require large data volumes to change the millions of parameters in a neural network. This means you need a lot of (highly expensive) labeled data, particularly in the case of supervised learning. The labeling of

images sounds simple, but expert expertise is needed in Natural Language Processing (NLP), [32] for example, to construct a large labeled dataset. Transfer learning is one way to reduce the necessary size of datasets to provide a viable alternative for neural networks. Other viable options shift towards more probabilistically based models, which are usually best suited for handling small data sets.

As the name suggests, transfer learning involves the ability to transfer information from one domain to another. Transfer learning can be understood at a high level, i.e. NLP model architectures can be reused in sequence prediction issues as many NLP issues can be implicitly reduced to sequence prediction issues. Transfer learning can also be viewed at a low level, where parameters are simply imported from one model into another model (skip-gram, continuous bag-of-words etc.). Transfer learning techniques are widely in used these days. Transfer learning improves the accuracy of classifier and also accelerates the learning process.

3.4. Hand-crafted Feature Descriptors

Matching local image descriptors is an essential move for several computer-vision applications. This function has been used for more than a decade by hand-made descriptors such as SIFT. Several new descriptors obtained from the data were recently introduced and have shown that SIFT is more discriminatory. Matching the local image function is a significant step for many computer vision applications, including image recovery and picture-based localization for Structure-from Motion (SFM) and Multi-View Stereo (MVS) [33]. The overall output in many of these applications is highly dependent upon the quality of the initial matching feature level. A significant part of the computer vision community therefore has an important interest in deciding which local feature descriptors have the most discrimination and the best matching performance. For over a decade, SIFT was possibly the most common feature descriptor for such activities. The ability of neural networks to learn data features higher than previous hand-made representations has recently brought significant changes in computer vision, e.g. recognition and object detection. Therefore, neural networks were implemented in order to create more inclusive representations of local features to the learning question of descriptors. The findings show major changes over conventional handcrafted representations such as SURF, DAISY or SIFT. In Fig. 16 a whole hand engineered architecture is presented that how an input image is classified and feature maps are

extracted though different parameter, then presented to different feature descriptors which further classify for better output prediction.

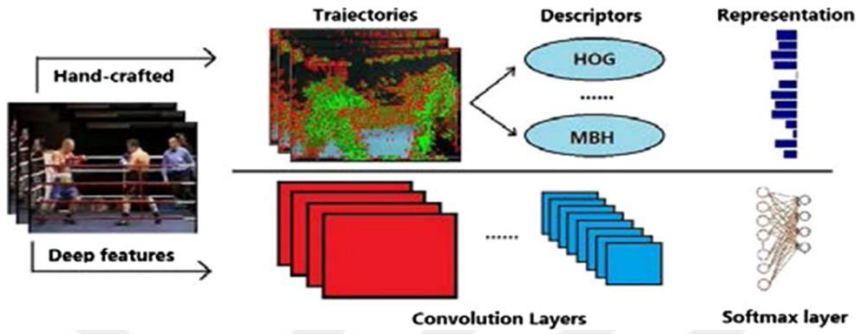

Figure 16: Work Flow of Hand Crafted Feature Classifier [33]

3.4.1. Local Binary Patterns (LBP)

LBP is suggested successful descriptor of texture, which is characteristically local neighborhood texture, was the Local Binary Pattern. LBP encodes image pixels by thresholding the central pixel by using tiny 3 x 3 pixel image regions. If the center pixel has a gc value and the neighborhood pixels a gp strength (p = 0, 1, 7), the pixels are then measured [34].

The proposed LBP by Ojala, mainly due to its small computing complexities and the ability to code fine details, has soon become a common descriptor. The canonical operator LBP [37] shall be calculated at each pixel location, taking into account the values of a small circular neighborhood with a R pixel radius, in the central pixel Ic.

$$\text{LBP}\ (N, R) = \sum_{n=0}^{N-1} s(I_n - I_c)2^n \quad [34]$$

When N is the number of pixels in the neighborhood, R is the radius, and s(x) = 1 if x 0, s(x) = 0 otherwise. The descriptor is the histogram of such binary numbers.

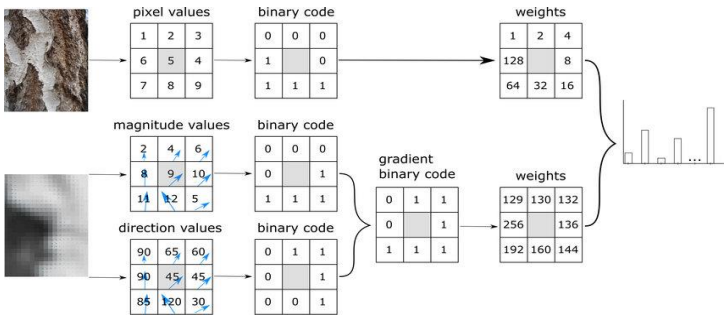

Figure 17: Local Binary Patterns applied on grey scale Images [34]

The image descriptor is displayed as the occurrence frequency (i.e., histogram) for all patterns collected in the image. When the LBP is computed for all pixels, a histogram describes its distribution over the whole image, given the space resolution of W / H pixels:

$$H = \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} f(LBP(i,j),p), p \in [0,N], [34]$$

And

$$f(x,y) = \begin{cases} 1, \text{if} x = y \\ 0, \text{if} x \neq y, \end{cases} [34]$$

Where N shows the LBP pattern's maximum value. A total of 28=256 different patterns are represented by integer values in [0,255] for square 3 to 3 image areas with 8 neighbors. The final histogram is a vector of 256 dimensions that can be used to describe the images. The ability to capture fine grained image data is a benefit of the original LBP operator. The limitation of the descriptor to capture image details at different scales however is limited to a 3 to 3 dimension, In addition to vibration and noise response.

3.4.2. Histograms of Oriented Gradients (HOG feature descriptor)

The histogram of oriented gradients (HOG) defines the incidence of gradient orientation on a histogram basis in localized areas of the image. In recognition task and more specifically pedestrian detection, HOG is proposed for characterizing visual objects. The image input has been decomposed into small squared cells, where each pixel in the cell is measured by its gradient orientation and a histogram is constructed on the basis of each cell's contribution. The histogram may have 9 or 18 bins according to the range of orientation angles used, whether they be 0 ° C to 180 ° C or 0 ° C to 360 ° C [19]. In order to increase the robustness of histograms against light

variations, local histograms are averaged over wider areas called blockage. The L2 norm is implemented in blocks, where the same number of cells exists for each row. The normalized histograms are eventually linked to the final HOG vector which represents the input image. HOG features are good for recognizing and differentiating objects, achieving excellent detection and object recognition efficiency.

3.4.3. Local Ternary Pattern (LTP)

The Local Ternary Patterns (LTP) suggested by Tan and Trigg's 362 are another model that has inspired many others. They use a 3-strength encoding scheme that includes a zero grid to determine the gray-scale difference at the local level. The ternary coding shall be completed by inserting the threshold in the canonical function LBP s(x):

$$s(x) = \begin{cases} 1, x \geq \tau \\ 0, |x| < \tau \\ -1, x \leq -\tau \end{cases} \quad [35]$$

The initial LTP code is divided into the positive and the negative LBP code due to the duration of the ternary coding.

3.4.4. Completed Local Binary Patterns (CLBP)

The CLBP is a texture descriptor that defines local image structures and sets certain limits and fixes them of the original LBP operator, such as noise and rotating sensitivities... The authors suggested that their pixels and Local Difference Sign-Magnitude Transformation (LDSMT) encode small image regions. [29]

Formulae shows the range of P pixels which are also located around one center pixel with intensity gc on a circle of R radius. Local distance between center pixels is given as follows:

$$d_p = s_p \cdot m_p \text{where} \begin{cases} s_p = sign(d_p) \\ m_p = |d_p|. \end{cases} \quad [35]$$

The CLBP follows three key steps to create the feature histogram that represents the entire image. First, with the global value t as a threshold which represents the average gray level of the whole image, the center pixel gc becomes binary:

$$CLBP\_C = T(g_c, t), \text{and} T(g, t) = \begin{cases} 1, \text{if} g \geq t \\ 0, \text{if} g < t. \end{cases} [35]$$

The sSign (S) are encoded by CLBP-Sign (CLBP_S) operator:

$$CLBP\_S = \sum_{p=0}^{P-1} s(g_p - g_c) \times 2^p, t(x) = \begin{cases} 1, \text{if} x \geq 0 \\ 0, \text{if} x < 0, \end{cases} [35]$$

Magnitude (M) is encoded by CLBP-MAGNITUDE (CLBP_M) operator:

$$CLBP\_M = \sum_{p=0}^{P-1} t(m_p, c) \times 2^p, t(i, j) = \begin{cases} 1, \text{if} i \geq j \\ 0, \text{if} i < j, \end{cases} [35]$$

These three feature maps are encoded in order to get required CLBP histogram.

3.4.5. Adjacent Evaluation Local Binary Patterns (AELBP)

Local Binary Pattern for Adjacent Evaluation (AELBP) is an effective LBP operator noise tolerant version. The AELBP user considers the neighboring pixels to be measurement centers around a central pixel to create a collection of adjacent windows with the adjacent pixels. A set of pre-detailed P-Measurement Windows with neighborhood pixels (p=0, P-1) is constructed for the central pixel of the GC-intensity and the P-pixel neighborhood equally spaced on the R-circle. The mean value for each window will be replaced by all of the pixels in the window except the middle pixel.

The central pixel AELBP code is given as:

$$AELBP = \sum_{p=0}^{P-1} s(a_p - g_c) \times 2^p, s(x) = \begin{cases} 1, \text{if} x \geq 0 \\ 0, \text{if} x < 0, \end{cases} [35]$$

The AELBP operator can be combined with other LBP models in order to have more inclusive and reliable features. The combination with the operator for the Texture Descriptor Completed Local Binary Patterns (AECLBP) for Adjacent Evaluation is taken into account here. Therefore the additional LDSMT sign (sp) components and magnitudes (mp) are specified as:

$$s_p = s(a_p - g_c) \text{and} m_p = |a_p - g_c|.$$

AECLBP_M operator is obtained by:

$$AELBP\_M = \sum_{p=0}^{P-1} t(m_p, c) \times 2^p, t(x, c) = \begin{cases} 1, \text{if} x \geq c \\ 0, \text{if} x < c, \end{cases} [35]$$

Where c is average value that is computed over whole image. In comparison, AECLBP C Center Operator is similarly encoded to CLBP C). The three operators are then combined together to build the final histogram for each image.

3.4.6 Patterns of Oriented Edge Magnitudes (POEM)

Magnitudes of Oriented Edge Patterns (POEM) is an image descriptor that encodes each pixel with gradient and local form in the surrounding areas to characterize local object appearance, shape, and form. It builds on the two well-known HOG and LBP characteristics. Similar to the first, the gradients of cells accumulate, using the LBP technique, in order to represent the spatial relationship between the accumulated gradients in the various directions. The gradient picture is determined first to measure the POEM features. The direction of the gradient is then uniformly discreet over a range of guidelines. Furthermore, for each pixel, a gradient orientation magnitude-weighted histogram is obtained from all pixels in predefined spatial regions known as cells. Therefore a vector of n values replaces any pixel location where n is the number of orientations used to construct the histogram [36]. The third step is to apply the LBP technique on each orientation with a 3/3 district on the cumulative gradient majors. Eventually, by concatenating all POEM characteristics at each n orientation, the POEM function histogram is obtained. The POEM descriptor is a strong imaging technology, with outstanding results, which provides outstanding recognition of all competing methods.

3.4.7 Binarized Statistical Image Features (BSIF)

Binarized Statistical Image Features (BSIF) is an LBP and LPQ-inspired technique of texture encoding and histogram image representation. In BSIF, however, binary codes are computed with a linear filters, which are first computed to the image, and subsequently binarized in each pixel. Bit of the code is related to a certain filter, and the number of filters added defines the length of the code. The binary codes are subsequently mapped to integer values between zero and 2^{m-1}, similar to LBP and LPQ, where m is the filter number. Lastly, the picture descriptor is a global histogram constructed from local histograms of BSIF [37] codes computed from a small picture block sequence. The BSIF is a strong texture descriptor with a distinctive function that uses filters learned through natural image statistics rather than heuristic code calculation.

3.4.8. Binary Pattern of Phase Congruence (BPPC)

The suggested descriptor is a multi-scaling local descriptor known as the Binary Pattern of Phase Congruence (BPPC) that can represent a variety of facial image patterns. It is rendered on directed PC images by applying LBP.

3.4.9. Gradient-Based Patterns (GDP)

By quantizing the gradient directional angles, the GDP operator encodes the texture information of a local area to form a binary pattern. The GDP's micro-samples are then used as the feature descriptor for their location and appearance detail. The resulting GDP features retain more information than the grey-level approaches [38] and define local image elements in a more stable manner, as the gradient operator can effectively boost the edge information of an image.

3.4.10. Generalized Local Ternary Patterns (GLTP)

GLTP operator encodes local texture information using three separate types of discrimination, quantizing the gradient magnitude values of a particular neighborhood. The proposed encoding system can differentiate between smooth and highly textured sections and ensure that micro patterns in texture are created that are compatible with local image features (smooth or high-texture). The output of the GLTP function descriptor is evaluated empirically with a Support Vector Machine (SVM). [38]

3.4.11. Improved Weber local descriptor (IWBC)

Improved Weber local descriptor (IWLD) features are encoded by improved weber binary coding. The first encodings are modern Weber magnetic components and orientation components, using standard approaches, including local binary (LBP) and local (LXOP) orientation modules (LCMP) and Local XOR (LXOP) versions.

3.4.12. Local Anisotropic Pattern (LAP)

By mapping visual neurons to image pixels, a binary pattern known as a local anisotropic pattern represent the association between a central pixel and its local neighbors. Experimental results on the texture classification verified the adequacy of the texture classification, invariance rotation, and noise robustness of the proposed LAP.

3.4.13. Local Directional Pattern (LDiP)

A descriptor used for object detection is the Local Directional Pattern (LDP). For each pixel in the image it assigns a code, and the resulting LDP image is split into regions for which a histogram is created. The histogram bins are linked to the final descriptor of all regions [39]. Slightly modification of LDP coding pattern constraints results in the redesign of patterns. DR-LDP's significance for effective object recognition is the compact code generation.

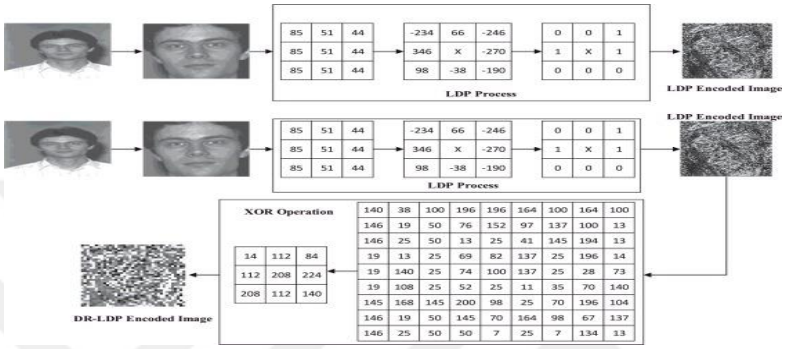

Figure 18: Working of LDiP Operator on Grey Scale Images [39]

Fig. 18 illustrates LDip descriptor feature maps extraction from input grey scale image and then mapping those extracted feature maps for further classification.

3.4.14. Local Directional Number Pattern (LDN)

LDN is a facial descriptor that evaluates local neighborhood directional information for encoding its structure. LDN is primarily a texture descriptor and encodes structural details facial structure and variations in intensity. Calculate therefore the edge responses with the Gaussian neighborhood mask in eight different directions. Then select from all directions the highest positive and negative directions to generate a significant descriptor with related structural patterns in different directions. The first step for the formation of LDN is generation of Gaussian edge responses and they are accomplished through the Gaussian mask. These Gaussian masks are placed over the images of the face, and are combined to create the images of the edge response.

The LDN histogram consists of the coded LDN image and contains different features of potential points, edges, curves, and other facial characteristics of a local texture. The histogram does not encode location information, but encodes event information. The face image is therefore divided into small sections for each section and the histogram is extorted. All histograms are combined, forming a histogram of the LDN.

3.4.15. Local Ternary Pattern (LTP)

Tan and Triggs extended basic LBPs into three encoding valued baptized LTPs with a threshold of extension of two traditionally valued LBP codes (0 and 1) to three (-1, 0 and 1) valued ternary codes.

3.4.16. Local Directional Ternary Pattern (LDTP)

Because of a binary code in the patterns mentioned, many texture-image information is lost. Firstly, it lacks much detail in the neighborhood due to the simplicity of the LBP, LTP and many other LBP-like textures bit string coding strategy. However, these LBP-like methods encode images with pixel intensity shifts, with the micro-level knowledge about edges, spot, and other local features. At the other hand, every instruction is treated equal by the LDP operator so that LDP lacks some guidance details, like the response sign [40]. The ad-hoc nature of the LDP code overlooks the effect of the neighborhood information since the code in the neighborhood contains all the information. In order to prevent these shortcomings, we are explored a new coding scheme that blends LDP's and LTP's 135 values. The local texture modeling proposed is based on local derivative variations and the encryption of contrast information, and inherits relevant characteristics. The suggested descriptor known as LDTP integrates 2 separate types of compass masks: derivative-Gaussian masks to avoid distortion of the noise, make the device resilient for lighting changes.

3.4.17. Local feature descriptor (LFD)

Local Feature Descriptors (LFD) are an increasingly efficient image comparison technique that is currently the best solution to in-plane rotation, distortion, and light changes. However, it is important to remember those assumptions. It is also known as appearance feature. It is still more stable than a pixel feature, but the appearance property still disturb the LFD ability. First, the feature is sensitive to light variations and the ability to explain the appearance depends on its complexity. LFDs overcome the intrusion of the image variations by collecting features through the main areas [40]. That can be seen as a kind of form description extension. However, LFD could not do anything for objects with low appearance complexity. LFD does not collect any useful ball details for this scenario because the ball's differentiation is not the presence of the shape.

3.4.18. Local Gabor Binary Pattern Histogram Sequence (LGBPHS)

A face representation approach focused on non-statistics in which the training procedure for constructing the face model is unnecessary. In the LGBPHS-approach a face image is modeled as a

"histogram sequence", using multiple Gabor mapper and multi-oriented Gabor-filter. Each GMP is transformed into a Local Gabor binary (LGBP) map [41], and each LGBP map is further divided into non-overlapping re. In LGBPHS-approach, the front picture is a "histogram sequence" and Fig. 19 describes how features are extracted from greyscale images LGBPHS feature descriptor:

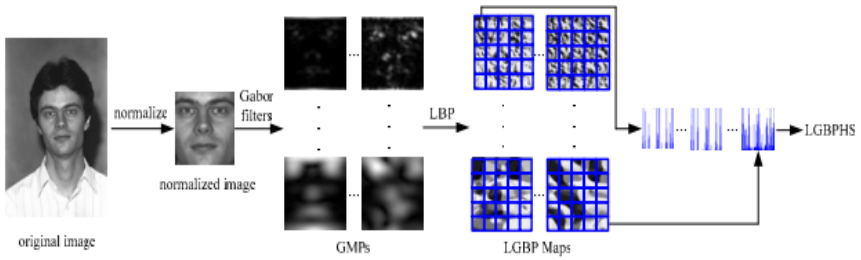

Figure 19: The framework of the proposed LGBPHS face representation approach [41]

3.4.19. Local Gradient Increasing Patterns (LGIP)

It is also one of the most famous for being used as facial recognition feature descriptor. Using eight binary bits, an LGIP feature encodes the intensity-growing trends in eight directions at each pixel, and then a decimal code is assigned to describe the overall trend that is increasing.

3.4.20. Local Gradient Patters (LGP)

It is also one of the most famous for being used as facial recognition feature descriptor. Local pattern is calculated by the local gradient flow through the central pixel in the region of 3x3 pixels from the one side to the other. [42]Two separate two-bit binary patterns, known as the LGP for the pixel, represent the center pixel of the region.

3.2.21. Monogenic Binary Coding (MBC)

In three complementary components, monogenic signal representation decomposes the original signal: amplitude, direction, and phase. In each local region monogenic variation and in each pixel, a monogenic variation is encoded and then the statistic features (e.g. histogram) of the local characteristics extracted are calculated. [43] The local statistical features extracted from the complementary monogenic components are then fused into an effective FR (amplitude, orientation, and phase).

3.4.22. Monogenic Binary Pattern (MBP)

MBP consists of two parts: one is a uniformly encoded magnitude, and another is a quadrant-bit-encoded monogenic orientation. It is also one of the major used facial recognition feature

descriptor. The magnitude is the measurement of local structural energy in the monogenic representation of an image and the conventional LBP operating device is used to encode local energy variation. [44, 45] The LBP operator can encode an 8-bit binary code with a local 3-3 patch. The uniform LBP operator was subsequently suggested, which includes two bitwise transitions at most from 0 to 1 or reverse when the binary chain is regarded as circular. The LBP uses 6 bits to describe information on the local structure without significantly affecting its performance.

3.4.23. Median Robust Extended Local Binary Patterns (MRELBP)

MRELBP compares local image medians instead of brute image intensities in contrast to traditional LBP and many LBP variants. Median Robust Extended Local Binary Pattern are considered as one of the finest texture classification feature descriptor. The local binary model suggested by Ojala ET AL. [46] Is characterized by encoding the differences between the central pixel value and the neighboring points, only taking into account the sign information in order to form a local binary pattern. [46]

3.4.24. Median Ternary Pattern (MTP)

MTP encodes local neighborhood texture information by quantizing the neighbor's intensity values in three different rates around each pixel based on the local median and global threshold value. A facial feature descriptors are to distribute the resulting MTP codes within an image or face patch. In the presence of noise, median filters are efficient. Therefore, a system that uses the median is stronger than noise and therefore produces reliable information. Due to noise or light differences the intensity of a pixel can vary [47]. Deploy a window (threshold ±t) to construct a 3-value code called a ternary pattern instead of using a 2-value number. This reduces the noise effect in nearly uniform regions. Next, the medium strength of the nine pixels is measured for each pixel around a neighbor 3 X 3 [47, 48]. Then the median M within ±t is set to 0, then MTP code comes out as:

$$S_{MTP}(v) = \begin{cases} 1, v > M_c + t \\ 0, M_c - t \leq v \leq M_c + t \\ -1, v < M_c - t \end{cases} [48]$$

Image localization using MTP feature descriptor is shown in Fig. 20 that how features are first divided into patches and then illustrated in localized histograms.

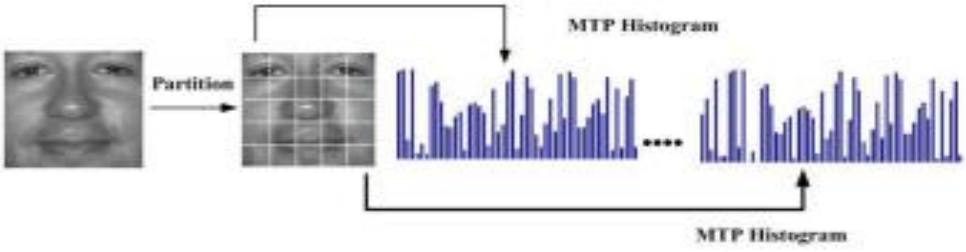

Figure 20: Image localization using MTP [48]

3.4.25. Weber Local Descriptor (mWLD)

Weber Local Descriptor (WLDs), is focused not just on the shift in the stimulus (such as sounds, light) but on the original strength of the stimulus, as a consequence of humans perceiving the pattern. WLD is comprised of two components: differential excitation and orientation. The differential excitation aspect is the function of the relationship between two terms: one is the relative intensity of a current pixel and its neighbors; the other is the current pixel intensity. The orientation part is the current pixel's gradient orientation [49]. For one particular picture, we create a WLD histogram using the two components.

The two WLD differential excitation / orientation components (den) are defined. Following this, we will present how to compute an image (or image region) for a WLD histogram. We use the intensity differences between its neighbors and a current pixel to adjust the current pixel in Differential Excitation. Thus, we hope to find the essential variants in a picture that simulate the perception of the pattern of human beings. In particular, as illustrated, the differential excitation (xc) of the current xc pixel is determined. The variations between its neighbors and the center point are first determined using the filter. Drop down architecture of mWLD [49] feature descriptor including all parameters like filtering, labeling etc is illustrated in Fig. 21:

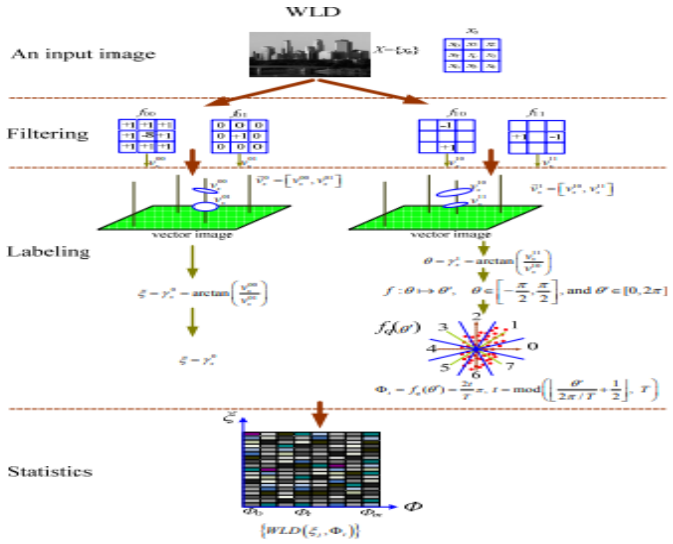

Figure 21: Drop Down WLD Feature extractor Architecture [49]

3.5. Information Fusion

Combining data from different sources and collecting it into secure and streamlined items that are more accurate and closely designed to deduce information than if they had been done by disparate sources. Computer vision integrates information from two or more images into one single image. The process is an image fusion process [50, 51]. The picture resulting is more descriptive than all of the images entered.

Data fusion, awareness and/or criterion fusion, data mining (DM) and the association of data are the four main areas as for fusion of ISHM organic information.

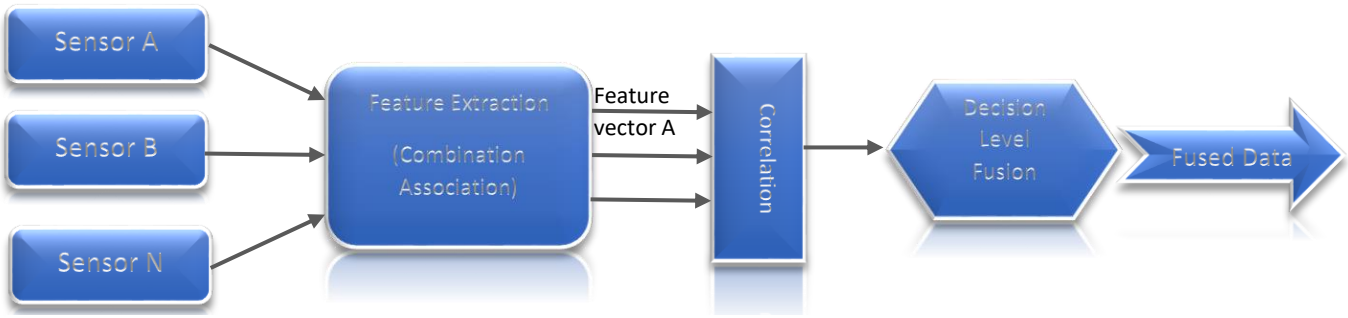

Figure 22: Three levels of fusion

Three levels of Fusion are determined in Fig. 22.Sensing data obtained from various sources are then fused into conceptual knowledge (knowledge laws, knowledge model etc.) of each particular system domain under examination and evaluated in accordance with those knowledge

rules. The fused information is then evaluated. The necessary data is stored on the Information Discovery DM database server.

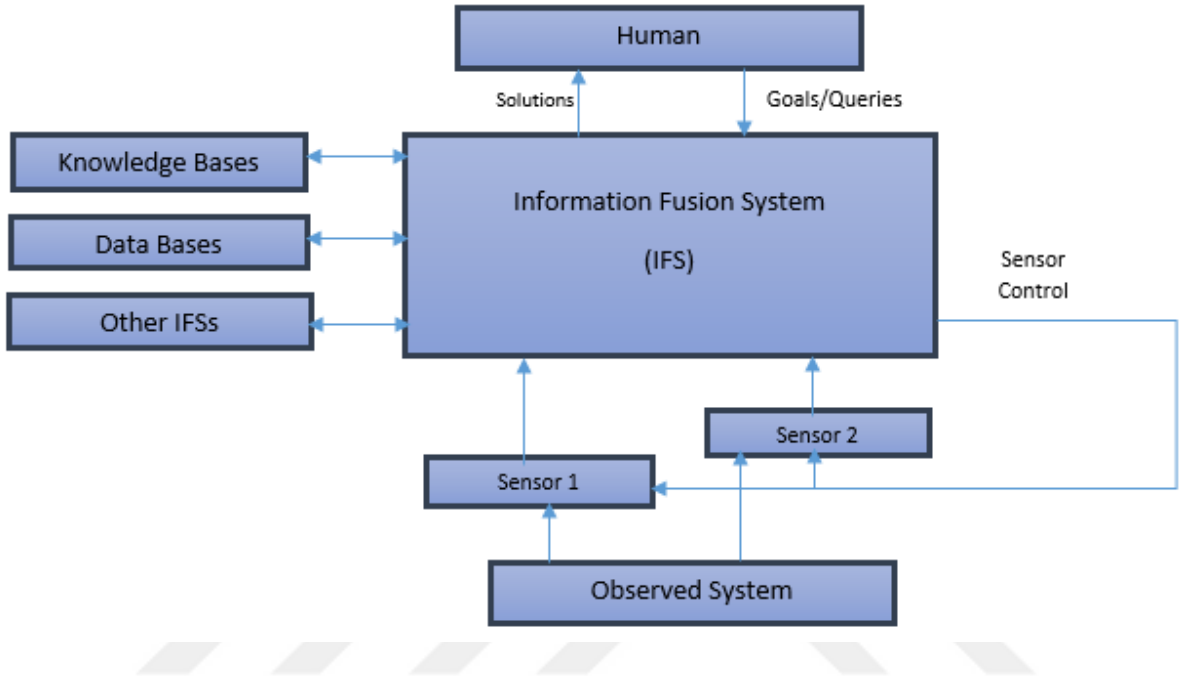

Figure 23: Information fusion systems

Fig. 23 illustrates information fusion systems that how they analyze data and how they perform decisions and other operations. Fusion of information data levels is a processing and analysis involving the use of techniques such as signal filtration, fusion of the feature level information including acquisition of index features by the ISHM [52] functional modules and fusion of information at the decision level including an assessment of the final health management of the system in order to provide decision support to explore Data level information.

Information fusion has three different levels:

3.5.1. Feature Level Fusion

Fusion level fusion consolidates feature sets of multiple biometric sources into a single collection of functions by means of appropriate schemes of standardization, transformation, and reduction. The main advantage of feature level fusion is that the corresponding functional values generated by different bio algorithms are known and a compact set of features is recognized which can improve recognition accuracy [53, 54]. In order to acquire the feature collection, methods of

dimensional reduction are usually used and thus a large number of training data for feature level fusion are expected to be usable.

Fusion of multi-modalities feature sets at the feature level requires combining feature sets. Since the feature set offers richer details than the match or decision on the raw biometric data, better [54] results are anticipated at this stage. Nevertheless, the integration of several approaches at this point is hard in practice because of the fact that the relationship between the properties of the various biometrical systems might not be known (e.g., detailed sets of fingerprints or facial coefficients).

Feature level fusion of hand and facial expression is shown in above image and extracted feature are also shown. Feature level fusion actually extracts the most important feature from an image.

3.5.2. Score-Level Fusion

In score level fusion, the scores are pooled by multiple biometric matches in order to agree on an individual's identity. This consolidation procedure usually results in the generation for the biometric system of a single scalar score. Fusion at this stage is mostly due to the simple access to and analysis of match scores in the biometric methods (compared to the raw biometrical data / feature set derived from the data). At this point fusion methods can be commonly categorized into three types of systems based on density, transformation, and classification. Because of the great trade-off between data ease and better information contents, most multimodal biometric systems incorporate data on the scale [55]. The optimal method for incorporation of biometric data is therefore score-level fusion. Fig. 24 describes how tumors are classified through score-level fusion by moving across different preprocessing and modules.

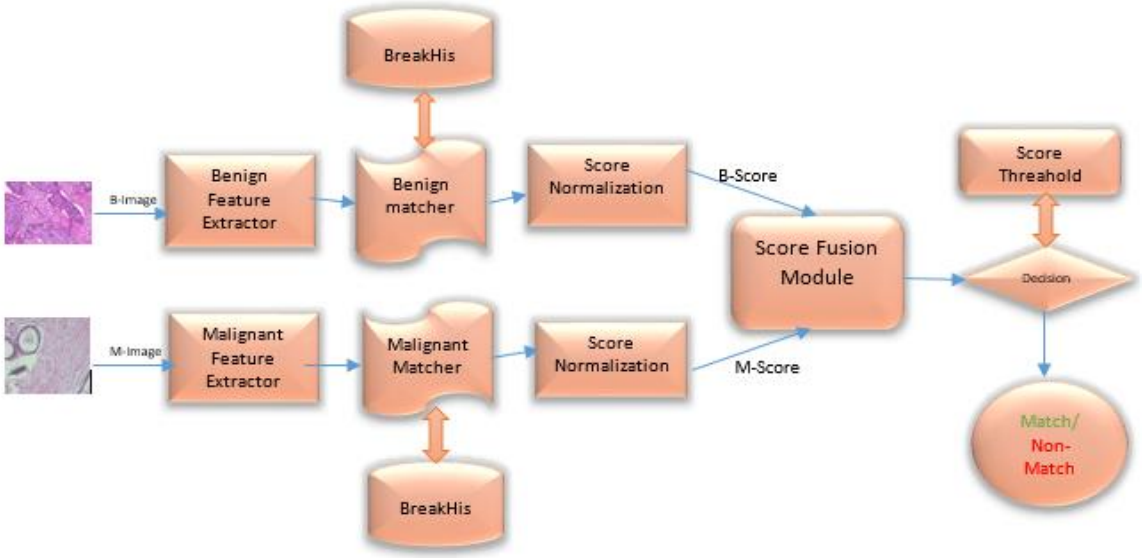

Figure 24: Score level fusion diagram for Benign and Malignant tumors

At score level a new score is produced by using variety of techniques by combining fused score of each modalities which in turn are compare with threshold in decision module. The vector may be defined as impostor or real via a Decision Tree, Vector Support Machine (SVM), or Linear Discriminate Analysis (LDA) process [56]. Therefore, the combination of the individual score produces a single scalar score for the final decision. Compared to the other measures, the combined method combined scores provided a superior performance record. Fusion has at score point two major phases. In the first step, the normalization of the score is accomplished, where determined scores are generated by a certain Si compartment. For example, if a comparator X generates domain scores [1, 50], and a comparator Y generates domain scores [1, 1700], these scores need to be standardized and mapped to common scales. Fusion is second step in score-level.

3.5.3. Decision-Level Fusion

Decision fusion falls within a wider range known to be distributed detection systems, and the method of selecting one hypothesis from several M hypotheses is the choice of N-sensors in the presence of noise and interference. In biometrics, fusion at a decision-level produces a single decision from two traditional impostors or legitimate users, who can or cannot be similar sensors, from multiple biometric sensor choices. Decision-level fusion is also implemented to save bandwidth and to increase accuracy in decision-making. A statistical output pattern for each biometric sensor is required before system-wide optimization with respect to two error levels: false

acceptance rate, imposter admission rate and false rejection rate that rejects the real consumer. The actions are first seen individually in each sensor mode, then these decisions are combined. Numerous sensors typically enhance the detection of complex activities in the real world.

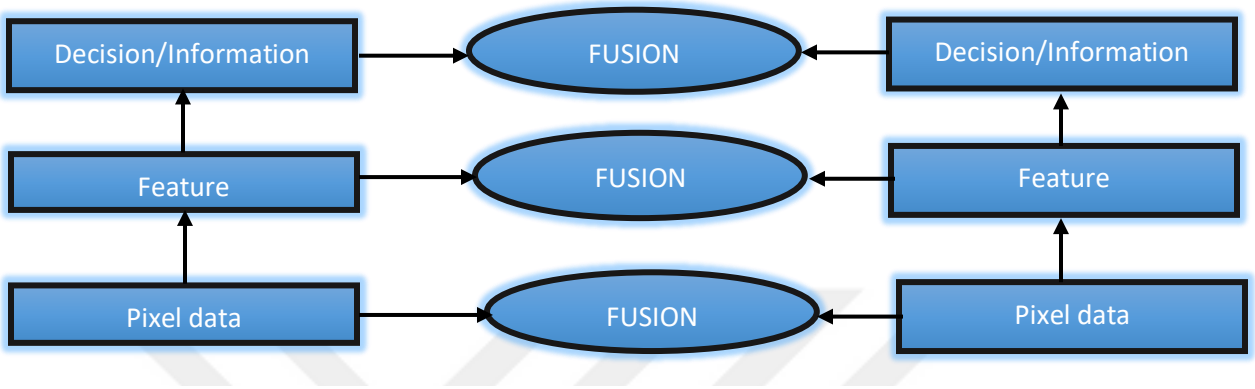

Figure 25: Pixel, feature, and decision level fusion

Fig. 25 illustrates different steps for decision level fusion that how decision is performed on output features of input image. The biometric verification problems can be identified with the match score(s) as a binary hypothesis testing problem. These are H_0: the score(s) indicate an imposter; H_1: the score(s) indicates a real person. The result is H0. Every classifier shall apply a threshold for the match level fusion and send the corresponding decision to the fusion engine. In principle, the threshold can differ across the whole range of matching scores. The null hypothesis is rejected if a match scoring reaches this threshold.

CHAPTER FOUR

4. Dataset

4.1. Break-His dataset

Break-His dataset consist of 2 classes of biopsy tumors, benign and malignant. It is an open source data set available at https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis [24]. The analysis of breast tumors, like most cellular pathologies, relies largely on histopathological slides [4]. During a ROI biopsy procedure, the latter are surgically removed tissues from the patient's breast. The latest edition of BreakHis is composed of 7909 histopathological biopsy images from 82 patients. These images are taken from different magnification factors that are 40X, 100X, 200X, and 400X [11, 17, 1].This dataset is composed of 2,480 benign and 5,429 malignant samples respectively. Data was carried out in 2015 where patient were referred to P&D lab Brazil .Data was anonymized. The board gave permission for study and approved the data and all the patients gave written consents. Malignant tumors increases and spreads to nearby cells that can reach to other parts while benign tumors does not expand.SO these two classes have some further sub-classes. For example dataset contains four histological distinct types of benign breast tumors: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenona (TA); and four malignant tumors (breast cancer): carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC) and papillary carcinoma [57].

Samsung digital color camera SCC-131AN is used by coupling with an OLYUMPUS BX-50 system microscope with a relay lens having a magnification of 3.3X for obtaining digitized images of breast cancer. Images are obtained in RGB channel. Original images are not colored rather they have black textures on left and right and text annotation on top left corner. [1, 58]

In general, multiple samples of breast tissue for each patient with a fine biopsy needle are taken at the operating room. Every sample is then prepared in the following phases: first, formalin fixation, and paraffin embedding, so the tissue structure of the origin and its molecular composition are preserved. Then, using a high precision microtome, sections of 3μm thickness are extracted

from paraffin outcomes. Instead these pieces are placed under microscope on covered glass slides for visualization.

Within raw tissue parts components of interest such as nuclei or cytoplasm are also not clearly visible. Until visualization, an integral process, called tissue staining, occurs. This staining phase aims to accentuate each part under the microscope for a better morphological examination. There are many types of staining, and the most commonly used method is Hematoxyl and Eosin (H&E). In BreakHis, Hematoxylin (H) binds to DNA, [2] coloring the associated structures of interest (i.e. most nuclei) with blue / purple, and Eosin (E) binds to proteins and colors other structures like cytoplasm and stroma with pink.

Magnification	Benign	Malignant	Total
40x	625	1370	1995
100x	644	1437	2081
200x	623	1390	2013
400x	588	1232	1820
total	2480	5429	7909
Patient	24	58	82

Table 1: Image distribution by magnification and classes [1, 30, 58]

Table. 1 is distribution of BreakHis dataset into benign and malignant tumors. Throughout analysis and annotation, pathologists begin by defining ROIs in the lowest slide of magnification level that is 40x, then goes deep into the latter using bit high magnification rates that are 100x and 200x until they reach deep down till 400x . The distribution of BreakHis distribution into four magnification levels is presented in Table for each tumor category and sub-category. Subset of magnifications factor includes precisely 82 patients for the distribution of patients as each patient has images with all magnifications. We found that each sub-set of the magnification component comprises about 24 images per patient following further statistical analysis. On average, 24, 25, 24, and 22 photos are available in the following subsets: 40x, 100x, 200x and 400x.

4.2. Description of dataset

For our work we used Break-His dataset publically available at. From this dataset we took all magnified hypothetical images of breast tumors and trained them on our respectively chosen pre-

trained neural network. Our work is divided into Magnification Specific binary classification and Magnification Independent Binary Classification, mainly Benign and malignant. We used 70% images for training and 30% for test. All images had a pixel size of 700x460 pixels and all are in .png format. We randomly divided dataset in which we assumed 598 images as test images and 1483 images as training and respectively for further four folds.

BreakHis histopathological data set was specifically published to address the absence of previous data sets in terms of clinical information quality and resources. Nevertheless, BreakHis still has a number of common abnormalities that are present in all medical datasets because of this disease nature. The following features are particularly relevant for BreakHis and should be considered in the building of a sturdy CAD system:

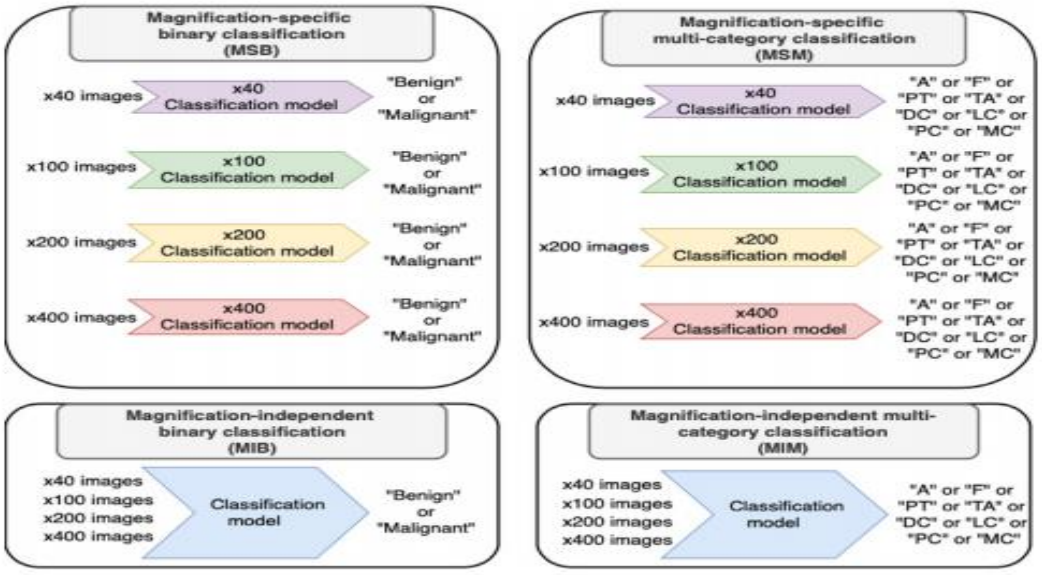

Figure 26: An illustration of each reformulation in Break-His dataset [1, 30]

Above Fig. 26 determines how dataset can be classified in 4 different categories that are MSB, MIB, MSM, and MIM [37]. Data Imbalance is one the major abnormalities and it occurs at different level in Break-his dataset. The malignant / benign class IR ratios are 0.41 at the level of the patient and 0.45 at the level of the image. There is also an inconsistent distribution between image sub-categories and patient level. This question of data imbalance could bias a CAD system

selecting the majority class image and patient level during binary and multi-category classification tasks.

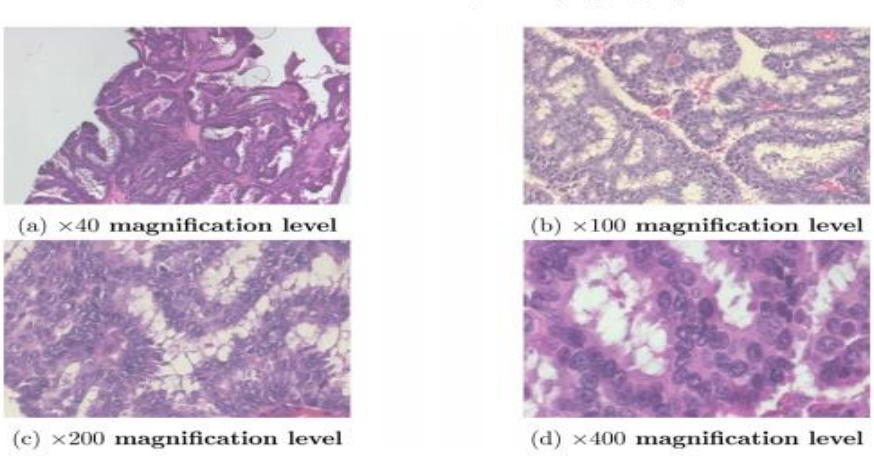

Figure 27: The same malignant PC tissue captured with different optical magnification levels (×40, ×100, ×200, and × 400), taken from the patient with ID: 9146 in Break-His dataset. [1, 40]

Fig. 27 consists of 4 images from BreakHis dataset of all 4 types of magnification from a single patient. Label noise, the breast tumor slowly spreads in the breast tissue from one area to another. This special case can confuse the CAD model with a multi-class classification process which seeks to discover the discriminatory characteristics of both malignant subclasses.

MAIN ATAGORY		**BENIGN**				**TOTAL BENIGN**	**MALIGNANT**				**TOTAL MAIGNANT**	**TOTAL OF BOTH**
SUB CATAGORY		A	F	TA	PT		DC	LC	MC	PC		
Number of images at each magnification level	40X	114	253	109	149	598	864	156	205	145	1370	1968
	100X	113	260	121	150	614	903	170	222	142	1437	2051
	200X	111	264	108	140	594	896	163	196	135	1390	1984
	400X	106	237	115	130	562	788	137	169	138	1232	1794

Table 2: Dataset Division into multiple Classes [57]

Table. 2 shows the division of dataset into Multiclass tumors of both benign and malignant. Covering all the 8 sub categories of both benign and malignant collectively.

CHAPTER FIVE

5. Results and Experiments

The major aim of our work is to make awareness about pre-trained neural networks and hand crafted features extractors for breast cancer classification. These pre-trained neural networks include GoogLeNet, densenet201, ResNet-101, ResNet-18, MobileNet -v2 etc. The breast histopathologic images are obtained in four different magnification factors, i.e., 40x, 100x, 200x, 400x.In our proposed system we took all magnified factors histopathalogic breast images [1] and classified them on pretrained neural networks and also used these images or handcrafted features on image level obtained relative accuracies. We achieved some enormous accuracy with low false rate using these classifiers. According to theorems and old research work we know that feed forward network with single input are enough for representing any function. Since AlexNet have 5 convolutional layer and Google net have 22 layers still it is hard and challenging for training a deep network because of vanishing gradient problems and data over-fitting [59, 60].The mean purpose of introducing ResNet and other pretrained neural network in deep learning world was to "identify shortcut connections" that can jump one or more layers.

In order to counteract traditional methods, some researchers have been able to specifically pick the world's most relevant features, both by confiding extraction functions and classification tasks to profound learning models. These studies tried successively, starting from exploring various models, to examine specific learning methods and adaptation, to develop their approach to deep learning. In this section we will organize the research in different deep learning aspects in accordance with their contributions. BreakHis authors were the first researchers to test a profound learning CAD method for this dataset by entrusting the extraction and classification tasks to a CNN model, initiated by LeNet evaluation, but its findings were smaller than that stated by its conventional counterpart in the past. As a fairly deeper network, they therefore settled on AlexNet. We compared in our research the performance of different CNN models, CaffeNet, a version of AlexNet, GoogLeNet, and ResNet-101, in order to consider the CNN appropriate for this classification function. Results showed Resent efficiency and the need for data augmentation, the finishing of all layers, large WSIs instead of small patches, and the combination of different amplification-specific models with ensemble learning. The results demonstrated in our research,

we investigated the performances of another CNN called Inception-v3 which is a GoogLeNet version, and our findings indicate that this CNN is more effective than previous works.

5.1. Image Level Accuracy

Image level accuracy is standard image accuracy that is taken into account by just giving images to classifier and obtaining the results given as: $ILA = I_{COT} / I_{TOT}$ [36]

Total no of images are illustrated as I_{TOT} and for total no of correctly classified images I_{COT} is mentioned.

Normally, a malignant case is considered positive during cancer diagnosis, while a benign case is viewed as negative. Clinically more significant is a CAD system's response to positive (malignant) events. The first two metrics, ILA and PLA, are based on the typical precision measured for both positive and negative predictions. [60]

The F1score is often referred to as the F1 metric or Fscore to help highlight the sensitivity of a CAD device to (positive) malignant cases as they demonstrate strong interest in this form of medical diagnosis. The harmonic mean between sensitivity (also known as recall) and precision is the F1-score:

$$F1score = 2 \times Recall \times Precision / Recall + Precision$$

Where: Precision = TruePositives /TruePositives + FalsePositives

Recall = Sensitivity = TruePositives / TruePositives + FalseNegatives [1, 60, 61]

Many break-His CAD systems were developed over the last three years. We suggested a new taxonomy to review all this research, which organizes Break-His work in four separate classes based on his accepted reformulation of the question of classification. Break-his data collection for the construction of CAD systems, their strength and vulnerabilities are the key inspiration behind the proposed taxonomy. However, this categorization has allowed us to define and evaluate from both clinical and functional perspectives, as we know, the most suitable reformulation for this issue.

5.2. Achievements and Results using CNN

As we trained our proposed framework on multiple pre-trained neural network architecture (DCNN) and obtained results are also compared with existing techniques. We followed the same 5-fold experimental settings recommended by dataset owners and also used by different researchers in their work .We began our work by examining the performance of other research works. We evaluated the performance of architecture based on overall prediction accuracies.

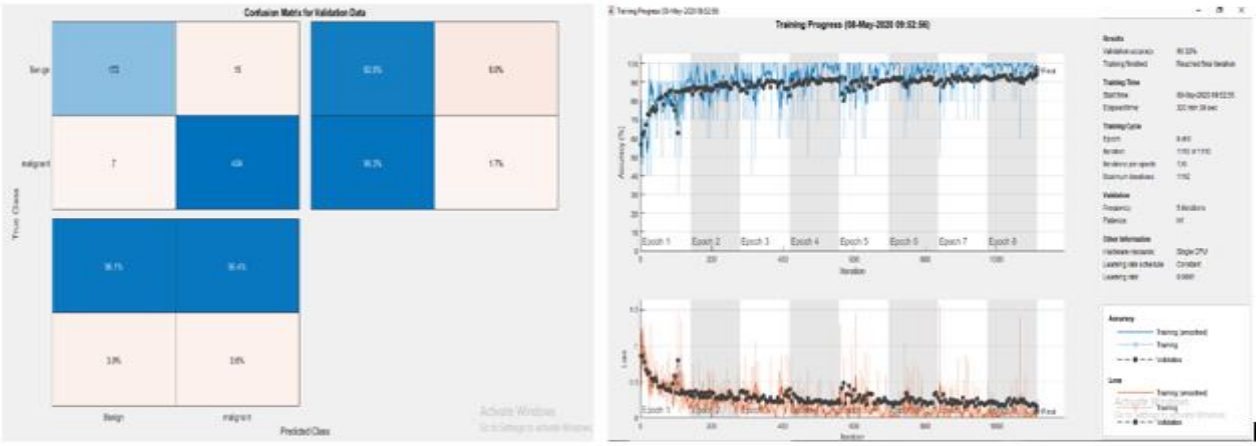

Figure 28: Sample Training Progress of Resnet-18

In the Fig. 28, Confusion matrix of resnet-18 describes that 92% of images were classified as true positive class and 98.3% were classified as true negative class.AS compared to Resnet-101 figure below, we achieved higher average accuracy.

In Table. 3 we used respective data for all magnification factors with various pretrained neural networks for magnification Independent binary classification and achived some promising results.Such as:

Patch/Slide	Feature Extractor	Classifier	Transfer learning	Train/test	Metrics	Results
WSI	Resnet101		ImageNet	70%/30%	ILA	89.49±3.49
WSI	RESNET152		ImageNet	70%/30%	ILA	91.37±3.65
WSI	Densenet201		ImageNet	70%/30%	ILA	**92.13±1.44**
WSI	Densenet169		ImageNet	70%/30%	ILA	91.12±2.22
WSI	xception		ImageNet	70%/30%	ILA	88.11±3.13
WSI	Inception_v3		ImageNet	70%/30%	ILA	85.99±2.60
WSI	inceptionResNet_v2		ImageNet	70%/30%	ILA	87.10±0.77

Table 3: Accuracies of systems over MSB using CNN with Transfer Learning

Similarly in Table. 4 we used respective data for all magnification factors with various pretrained neural networks for magnification specific binary classification (MSB) and achived some promising results.Such as:

Patch/Slide	Feature Extractor	Classifier	Transfer learning	Train/test	Metrics	Results			
						40x	100x	200x	400x
WSI	Resnet101		ImageNet	70%/30%	ILA	88.74± 2.29	90.38± 4.16	94.50± 2.77	88.87± 2.91
WSI	ResNet50		ImageNet	70%/30%	ILA	89.62 ± 2.27	90.63± 4.25	92.17± 1.99	90.00± 1.31
WSI	RESNET152		ImageNet	70%/30%	ILA	90.45 ± 2.36	91.74± 3.81	93.62± 2.04	90.99± 1.04
WSI	Densenet201		ImageNet	70%/30%	ILA	91.12± 1.11	91.75± 4.51	93.25± 2.31	91.24± 3.09
WSI	Densenet169		ImageNet	70%/30%	ILA	88.89± 3.12	90.37± 3.01	92.00± 2.04	90.37± 1.43
WSI	Densenet121		ImageNet	70%/30%	ILA	92.26± 1.81	91.12± 2.97	91.62± 3.49	88.25± 3.28
WSI	Vgg 19		ImageNet	70%/30%	ILA	87.84± 1.32	86.77± 3.18	88.00± 3.48	84.79± 1.75
WSI	xception		ImageNet	70%/30%	ILA	84.18± 1.90	83.35± 1.41	89.05± 3.18	86.00± 1.29
WSI	Inception_v3		ImageNet	70%/30%	ILA	86.41± 2.47	86.21± 2.36	84.09± 1.74	83.25± 1.73
WSI	inceptionResNet_v2		ImageNet	70%/30%	ILA	85.63± 2.49	84.86± 2.31	84.59± 1.37	85.10± 2.22
WSI	mobilenet		ImageNet	70%/30%	ILA	85.25± 3.55	87.49± 2.30	87.12± 1.74	85.17± 2.45

Table 4: Accuracies of systems over MSB using CNN with Transfer Learning

For the first time, we worked on the ResNet-18, Densenet 201, Xception pretrained ImageNet for the binary grading task. The same fine-tuned ResNet-18 and respective with a multi-category rating output layer was subsequently retrained. The second process consisted of two sections, one devoted to benign subclasses and the second to malignant subclasses. Each instance has been given to its corresponding module for sub-class identification after defining their main class in the binary classification process. A picture was classified correctly if both phases were classified correctly. The decision was made at two points for both. t, the decision was reached by combining with the majority voting rule all of the extracted patches at image. In comparison, the change brought about by the normalization of stain and rise in data is not very important to a certain degree. In our binary classification we carried out correctly recognized images of cancer as benign and malignant as no

cancer. Using a pre-trained for each model is a basis of the CNN model [38], and since the chosen model has been pre-trained in a problem with 1000 classification, for each one of its two remaining Binary-Category models we have replaced the last output layer with the binary output softmax model and a 4-output layer. We used WSIs rather than extracted patches during the preparation and we also extracted patches for few experiments. We followed a dual-stage fine tuning strategy to eventually change these three CNNs. First, we retrained the final fully linked layers, then retrained the entire CNN.

5.3. Achievements and Results using Hand crafted features

First Break-His CAD systems have adopted a conventional two-stage approach to derive handcraft characteristics from the images and then use them to train a standalone classification. Some works evaluated numerous descriptors during the extraction process in order to pick the most appropriate ones, while others used a single descriptor. Within this section, all these works will be presented with their image descriptors adopted.

In our work we carried out operations on various feature descriptors like local binary pattern (LBP) along with its variants that is Completed local binary pattern, Histogram of Oriented Gradients (HOG) used for extracting features from image data, Patterns of Oriented Edge Magnitudes, Binarized Statistical Image Features, Binary Pattern of Phase Congruence (BPPC) [17] [32],Gradient-Based Patterns (GDP),GLTP, IWBC, Local Directional Pattern (LDiP), LDN, Local Ternary Pattern (LTP),Local Directional Ternary Pattern (LDTP),Local feature descriptor (LFD),Local Gabor Binary Pattern Histogram Sequence (LGBPHS),Local Gradient Increasing Patterns (LGIP),Local Gradient Patters (LGP),Monogenic Binary Coding (MBC),Monogenic Binary Pattern (MBP),Median Ternary Pattern (MTP),Weber Local Descriptor (mWLD),Pyramid Histogram of Oriented Gradients (PHOG).

So we have been training a 1NN classification [13] with the fractal dimension of each picture to determine the effectiveness of the fractal dimension as the only descriptor. The results show that the use of fractal dimensions as a single descriptor is more acceptable when classifying 40 images with many self-like features, but useless in higher images with fewer self-like features. Another part of this work consists in multiple experiments, each classifying one benign and one malignant. [38] Results obtained from various hand engineered features have been regarded as comparatively appropriate yet highly unreliable preliminary tests. In addition, the major limitation of these typical

approaches is that the consistency of the model depends on the features extracted while it is a difficult task to achieve a highly representative consistency. One of the key difficulties is the correct choice of descriptor, and even when various descriptors combine to increase their discriminatory strength, [60, 61] or turn it into the most suitable, the results obtained between different magnifications remain fairly small and inconsistent.

In Table. 5 we used hand engineered feature extractors and the results are not promising but still we came to know about the working architectures of different hand engineered algorithms. We performed different operations at image level and achieved low accuracies.

Image magnifier	Cell size	LBP	HOG	CLBP
40x	[4 4]	55.78%	39.97%	47.44%
	[8 8]	58.64%	56.89%	49.14%
	[16 16]	62.47%	40.44%	48.28%
	[32 32]	65.56%	43.56%	58.99%
	[64 64]	**66.89%**	57.78%	63.24%
100x	[4 4]	56.32%	58.90%	50.94%
	[8 8]	63.49%	60.29%	54.84%
	[16 16]	69.91%	64.31%	58.76%
	[32 32]	**74.19%**	**67.98%**	61.99%
	[64 64]	**76.69%**	70.19%	**67.74%**
200x	[4 4]	48.47%	52.65%	49.14%
	[8 8]	60.17%	49.99%	51.24%
	[16 16]	64.29%	55.99%	59.21%
	[32 32]	68.80%	50.42%	64.44%
	[64 64]	**73.82%**	64.42%	**67.23%**
400x	[4 4]	44.54%	49%	51.29%
	[8 8]	49.89%	54.31%	56.83%
	[16 16]	53.95%	57.96%	60.94%
	[32 32]	59.36%	61.13%	65.27%
	[64 64]	64.44%	**65.55%**	69.72%

Table 5: LBP, HOG, and CLBP Results

Furthermore we tested our images on 22 feature descriptors famous for carrying out feature and facial expressions from image data in Table. 6. Their illustrations are listed as:

Classifier	40x	100x	200x	400x
GDP	63.31%	61.29%	60.19%	59.91%
GDP_2	56.95%	59.98%	57.59%	61.27%
GLTP	58.14%	51.27%	54.07%	53.41%
IWBC	**66.27%**	58%	56.11%	59.86%
LAP	62.40%	59.01%	56.85%	57.49%
LBP	56.66%	55.21%	56.85%	59.27%
LDiP	64.50%	57.23%	56.67%	54.16%
LDN	59.32%	61.25%	59.07%	54.88%
LDTP	59.72%	62.31%	61.11%	57.64%
LFD	61.24%	51.09%	56.48%	59.91%

LGIP	**66.57%**	59.99%	59.44%	55.24%
LGP	**65.24%**	58.47%	59.44%	60.87%
LGTrP	55.18%	62.20%	60.56%	53.26%
LMP	65.09%	53.91%	58.89%	59.44%
LPQ	61.98%	59.06%	58.33%	61.07%
LTeP	56.95%	58.94%	56.30%	51.39%
LTrP	64.95%	**64.49%**	62.78%	59.93%
MBC	60.21%	61.73%	60.56%	57.41%
MBP	50.30%	54.39%	52.04%	58.04%
MRELBP	53.99%	61.12%	60.19%	57.61%
MTP	57.99%	59.48%	58.52%	64.48%
mWLD	55.29%	53.38%	54.44%	59.93%
PHOG	58.89%	58.37%	60.37%	61.28%

Table 6: Hand Crafted Feature Extractors Results

We listed the results of following 22 feature descriptors in the form of chart in Fig. 29:

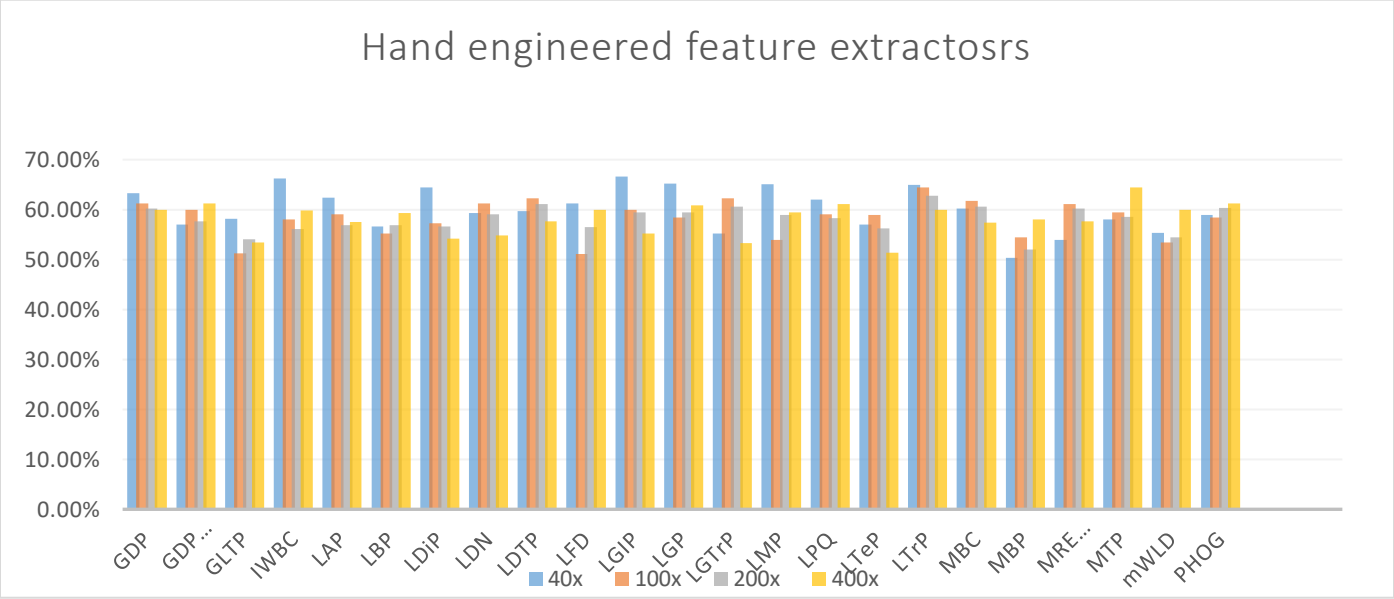

Figure 29: Hand engineered Features Accuracy difference Plot

We summarized our whole work done in all above tables. Relating our work with previously work done on same dataset, Fabio A and Luiz S dis the classification of Break-His histopathological images using convolutional neural network. They extracted the patches of training images for CNN and combined these patches for final classification. They reported the test in two ways. They computed recognition rate at image level. So her we compared our results in Table 7:

Publications	Preprocessing	Feature Extractor/Classifier	Transfer Learning	Results			
				40x	100x	200x	400x
Fabio Alexandre Spanhol [2]	Res(350x230) - SMI	EMAX(AlexNet)	ImageNet	85.6+4.8	83.5+4.8	84.6+4.2	86.1+6.2
Kausik Das [1]	Res(370x230) Dab(Rot,Flip)	GoogLeNet	ImageNet	93.82	93.38	93.67	93.48
E.Deniz [1]	None	AlexNet	ImageNet	90.69	90.46	90.64	90.96
K.Das [1]	none	NDCNN	none	82+2.8	86.2+4.6	84.6+3	81+2.4
E.M.Nijad [40]	Res(224x224) SMI	VG-19	ImageNet	80.00	80.00	80.00	80.00
F.A.Spanhol [1]		Caffenet|LR	ImageNet	84.6+2.9	84.8+4.8	84.2+4.7	81.6+4.7
Proposed	None	ResNet101	ImageNet	88.74± 2.29	90.38± 4.16	93.25 ± 2.77	88.87 ± 2.91
Proposed	None	Densenet-201	ImageNet	91.12± 1.11	91.75± 4.51	94.50 ± 2.31	91.24 ± 3.09
Proposed	None	Densenet-169	ImageNet	90.45 ± 2.36	91.74± 3.81	93.62 ± 2.04	90.99 ± 1.04
Proposed	None	Densenet-121	ImageNet	92.26± 1.81	91.12± 2.97	91.62 ± 3.49	88.25 ± 3.28
Proposed	None	VGG-19	ImageNet	87.84± 1.32	86.77± 3.18	88.00 ± 3.48	84.79 ± 1.75

Table 7: Comparison with State of art work

CHAPTER SIX

6. Conclusion

In our work we conducted experiments on Break-His dataset by using DCNN and we also applied variety of hand-crafted features for better understanding about the working of both CNN and handcrafted features. We showed that how we can train and test tumors images on existing CNN architectures. We achieved enormous and efficient results in accordance with the work done on same dataset using different approaches. For better performance of classifier we implemented some data augmentation techniques .Future work can be done with new and different CNN architectures and more accurate results can be obtained by new deep learning and better stain normalization techniques. In case of hand crafted features we combined those using different techniques yet we were unable to achieve good results and accuracies for our desired breast cancer dataset. Our results differs from other research workers and our results shows the performance of CNN's and hand crafted features. More work can be performed by combining CNN with different handcrafted feature extractors. Accuracy can be improved by removing noises and stain from dataset and performing different preprocessing works.

References

[1] F. Spanhol, L. E. S. d. Oliveira, C. Petitjean and L. Heutte, "Breast Cancer Histopathological Image Classification using Convolutional Neural Networks," in *International Joint Conference on Neural Networks (IJCNN 2016)*, Vancouver, Canada, July,2016.

[2] F.Paulin and S. A, "Classification of Breast cancer by comparing Back propagation training algorithms," *International Journal on Computer Science and Engineering,* pp. 327-332, 2011.

[3] A. A. Samah, M. F. A. Fauzi and S. Mansor, "Classification of benign and malignant tumors in histopathology images," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Kuching, Malaysia, 1 december 2017.

[4] S. Akbar, M. Peikari, S. Salama, S. Nofech-Mozes and A. Martel, "Transitioning Between Convolutional and Fully Connected Layers in Neural Networks," in *International Workshop on Deep Learning in Medical Image Analysis*, 2017.

[5] S. Cascianelli, R. Bello-Cerezo, F. Bianconi, M. L. Fravolini, M. Belal, B. Palumbo and J. N. Kather, "Dimensionality Reduction Strategies for CNN-Based Classification of Histopathological Images," *Intelligent Interactive Multimedia Systems and Services 2017,* vol. 76, pp. 21-30, May 2017.

[6] K. Sekaran, S. P. Ramalingam and C. M. P.V.S.S.R., Breast Cancer Classification Using Machine Learning, Knowledge Computing and Its Applications, February 2018.

[7] V. Gupta and A. Bhavsar, "An Integrated Multi-scale Model for Breast Cancer Histopathological Image Classification with Joint Colour-Texture Features," in *International Conference on Computer Analysis of Images and Patterns*, Mumbai, 2017.

[8] A. T. Azar and S. A. El-Said, "Performance analysis of support vector machines classifiers in breast cancer mammography recognition," *Neural Computing and Applications volume ,* vol. 24, pp. 1163-1177.

[9] A. M. Abdel-Zaher and A. M. Eldeib, Breast Cancer Classification Using Deep Belief Networks, Expert Systems With Applications, 2015.

[10] M. A.Al-antari, M. A.Al-masni, Mun-TaekChoi, Seung-MooHan and Tae-SeongKim, "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification," *International Journal of Medical Informatics,* vol. 117, pp. 44-54, 2018.

[11] Y. Benhammoua, B. Achchabb, F. Herreraa and S. Tabik, "BreakHis based breast cancer automatic diagnosis using deep learning:," *Neurocomputing,* vol. 375, pp. 9-24, 2019.

[12] F. Gorunescu and S. Belciug, "Evolutionary strategy to develop learning-based decision systems. Application to breast cancer and liver fibrosis stadialization," *Journal of Biomedical Informatics,* vol. 49, pp. 112-118, 2014.

[13] S. Kwok, "Multiclass Classification of Breast Cancer in Whole-Slide Images," in *International Conference Image Analysis and Recognition*, 2018.

[14] A.-A. Nahid and Y. Kong, "Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network," School of Engineering, Macquarie University, Sydney,Australia, 16 January 2018.

[15] L. Nanni, S. Ghidoni and S. Brahnam, "Handcrafted vs. non-handcrafted features for computer vision classification," *Pattern Recognition,* vol. 71, pp. 158-172, November 2017.

[16] S. H. Kassani, P. H. Kassani and M. J. Wesolowski, "Breast Cancer Diagnosis with Transfer Learning and Global Pooling," 26 sep,2019.

[17] M. AMRANE, S. OUKID, I. GAGAOUA and T. ENSAR, "Breast Cancer Classification Using Machine Learning," in *Conference: 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, 2018.

[18] H. Li, S. Zhuang, D.-a. Li, J. Zhao and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning," *Biomedical Signal Processing and Control,* vol. 51, pp. 347-354, 2019.

[19] A. Mert, Niyazi Kılıç, E. Bilgili and A. Akan, "Breast Cancer Detection with Reduced Feature Set," 2014.

[20] F. Ghaznavi, A. Evans and A. M. a. M. Feldman, "Digital Imaging in Pathology: Whole-Slide Imaging and Beyond," *Annual Review of Pathology: Mechanisms of Disease,* vol. 8, pp. 331-359, 15, 2012.

[21] Ahmed, M.Abdel-Zaher, Ayman and M.Eldeib, "Breast cancer classification using deep belief networks," *Expert Systems with Applications,* vol. 46, pp. 139-144, March 2016.

[22] N. Zeng, Z. Wang, H. Zhang, W. Liu and F. E. Alsaadi, "Deep Belief Networks for Quantitative Analysis of a Gold Immunochromatographic Strip," *Cognitive Computation,* vol. 8, pp. 684-692, 30 April 2016.

[23] A. Tharwat, "Linear vs. quadratic discriminant analysis classifier: a tutorial," *International Journal of Applied Pattern Recognition (IJAPR),* vol. 3, 2016.

[24] C. Wang, J. Shi, Q. Zhang and S. Ying, "Histopathological image classification with bilinear convolutional neural networks," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, South Korea, 14 September 2017.

[25] K. B. Nahato, K. N. Harichandran and K. Arputharaj, "Knowledge Mining from Clinical Datasets Using Rough Sets and Backpropagation Neural Network," *Computational and Mathematical Methods in Medicine,* 2015.

[26] M. NEMISSI, H. SALAH and H. SERIDI, "Breast cancer diagnosis using an enhanced Extreme Learning Machine based-Neural Network," in *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)*, Guelma, Algeria,, 2018.

[27] F. I, S. A, G. Al, S. M, E. C, P. A and A. P, "Automatic classification of tissue malignancy for breast carcinoma diagnosis.," *Computers in biology and medicine,* vol. 96, pp. 41-51, 2018.

[28] K. Das, S. P. K. Karri, A. G. Roy, J. Chatterjee and D. Sheet, "Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification," *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017),* 19 June 2017.

[29] H. A.Abbass, "An evolutionary artificial neural networks approach for breast cancer diagnosis," *Artificial Intelligence in Medicine,* vol. 25, pp. 265-281, 2002.

[30] W. Zhi, H. W. F. Yueng, Z. Chen, S. M. Zandavi and Z. L. Y. Chung, "Using Transfer Learning with Convolutional Neural Networks to Diagnose Breast Cancer from Histopathological Images," *Neural Information Processing,* vol. 10637, pp. 669-676, 2017.

[31] J. Sun and A. Binder, "Comparison of deep learning architectures for H&E histopathology images," in *2017 IEEE Conference on Big Data and Analytics (ICBDA)*, Kuching, Malaysia, 08 February 2018.

[32] P.J.Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Expert Systems with Applications,* vol. 117, pp. 103-111, 1 March 2019.

[33] Y. Jiang, L. Chen, H. Zhang and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," vol. 10870, pp. 116-121, 29 march 2019.

[34] J. L. Schonberger, H. Hardmeier, T. Sattler and M. Pollefeys, "Comparative Evaluation of Hand-Crafted and Learned Local Features," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 09 November 2017.

[35] R. Venkatesan, P. S. Chandakkar and B. Li, "Simpler Non-Parametric Methods Provide as Good or Better Results to Multiple-Instance Learning," in *IEEE*, Santiago, Chile, 18 February 2016.

[36] M. Sun, T. X. Han, M.-C. Liu and A. Khodayari-Rostamabad, "Multiple Instance Learning Convolutional Neural Networks for object recognition," in *IEEE*, Cancun, Mexico, 24 April 2017.

[37] D. Sanchez-Morillo, J. González, M. García-Rojo and J. Ortega, "Classification of Breast Cancer Histopathological Images Using KAZE Features," *Bioinformatics and Biomedical Engineering ,* vol. 10814, pp. 276-286, 2018.

[38] S.-J. Lee, T. Chen, L. Yu and C.-H. Lai, "Image Classification Based on the Boost Convolutional Neural Network," *IEEE Access,* vol. 6, pp. 12755 - 12768, 25 January 2018.

[39] E. khadiri, A.Chahia, Y. merabet, Y.Ruichek and R.Touahnia, "Local directional ternary pattern: A New texture descriptor for texture classification," *Computer Vision and Image Understanding,* vol. 169, pp. 14-27, April 2018.

[40] E. M. Nejad, L. S. Affendey, R. B. Latip and I. B. Ishak, "Classification of Histopathology Images of Breast into Benign and Malignant using a Single-layer Convolutional Neural Network," *Proceedings of the International Conference on Imaging, Signal Processing and Communication,* pp. 50-53, july 2017.

[41] S.-J. Lee, T. Chen, L. Yu and C.-H. Lai, "Image Classification Based on the Boost Convolutional Neural Network," *IEEE Access ,* vol. 6, no. 25 January 2018, pp. 12755 - 12768, 2018.

[42] M. Guray and A. A. Sahin, "Benign Breast Diseases: Classification, Diagnosis, and Management," *The Oncologist,* vol. 11, no. 5, pp. 435-449, 2006.

[43] F. Perronnin, J. Sánchez and T. Mensink, "Improving the Fisher Kernel for Large-Scale Image Classification," in *European Conference on Computer Vision*, 2010.

[44] Z. Gao, L. Wang, L. Zhou and J. Zhang, "HEp-2 Cell Image Classification With Deep Convolutional Neural Networks," *IEEE Journal of Biomedical and Health Informatics ,* vol. 21, no. 2, pp. 416-428, 2017.

[45] Y. Song, H. Chang, Y. Gao, S. Liu, D. Zhang, J. Yao, W. Chrzanowski and W. Cai, "Feature learning with component selective encoding for histopathology image classification," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington , DC USA, 2018.

[46] M. Sun, Missouri, T. X. Han, M.-C. Liu and A. Khodayari-Rostamabad, "Multiple Instance Learning Convolutional Neural Networks for object recognition," in *International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 2016.

[47] Z. W. H. Z. W. L. &. F. E. A. Nianyin Zeng, "Deep Belief Networks for Quantitative Analysis of a Gold Immunochromatographic Strip," in *Cognitive Computation*, 2016.

[48] J. B. Tenenbaum, V. d. Silva and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science,* vol. 290, no. 5500, pp. 2319-2323, 2000.

[49] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li and A. M.Dobaie, "Facial expression recognition via learning deep sparse autoencoders," *Neurocomputing,* vol. 273, pp. 643-649, 2018.

[50] S. K. Zhou, H. Greenspan and D. Shen, "Deep Learning for Medical Image Analysis," *Journal of Pathalogy Informatics,* p. 443, 2018.

[51] M. Bakator and D. Radosav, "Deep Learning and Medical Diagnosis: A Review of Literature," *Multimodal Technologies and interaction,* 2018.

[52] A. Maier, C. Syben, T. Lasser and C. Riess, "A gentle introduction to deep learning in medical image processing," *Zeitschrift für Medizinische Physik,* vol. 29, no. 2, pp. 86-101, 2019.

[53] D. Shen, G. Wu and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annual Review of Biomedical Engineering,* vol. 19, pp. 221-248, 2017.

[54] M. H. Hesamian, W. Jia, X. He and P. Kennedy, "Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges," *Journal of Digital Imaging ,* vol. 32, pp. 582-596, 2019.

[55] G.-S. Fu, Y. Levin-Schwartz, Q.-H. Lin and D. Zhang, "Machine Learning for Medical Imaging," *Journal Of Health Care Engineering,* 2019.

[56] M. A. Mazurowski, M. Buda, A. Saha and M. R. Bashir, "Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI," *JMRI,* vol. 49, no. 4, pp. 939-954, 2019.

[57] R. Barroso-Sousa and O. Metzger-Filho, "Differences between invasive lobular and invasive ductal carcinoma of the breast: results and therapeutic implications," *Therapeutic Advances in Medical Oncology,* 2016.

[58] A. Tharwat, "Linear vs. quadratic discriminant analysis classifier: a tutorial," *International Journal of Applied Pattern Recognition (IJAPR),* vol. 3, no. 2, 2016.

[59] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *2011 International Conference on Computer Vision*, Barcelona, 2011.

[60] L. P. Coelho, A. Ahmed, A. Arnold, J. Kangas, A.-S. Sheikh, E. P. Xing, W. W. CohenRobert and F. Murphy, "Structured Literature Image Finder: Extracting Information from Text and Images in Biomedical Literature," *Linking Literature, Information, and Knowledge for Biology,* vol. 6004.

[61] R. M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics,* Vols. SMC-3, no. 6, pp. 610-621.