IDENTIFICATION OF MICROSCOPY CELL IMAGES BY USING CONVOLUTIONAL NEURAL NETWORK APPLICATION

NORFATIN FARISYA BINTI AJIS

RESEARCH PROJECT SUBMITTED TO THE

FACULTY OF BIOMEDICAL ENGINEERING

UNIVERSITY OF MALAYA,

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF BIOMEDICAL ENGINEERING

2020

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: NORFATIN FARISYA BINTI AJIS

Registration/Matric No: KQB170013

Name of Degree: MASTER OF BIOMEDICAL ENGINEERING

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):
IDENTIFICATION OF MICROSCOPY CELL IMAGES BY USING CONVOLUTIONAL NEURAL NETWORK APPLICATION

Field of Study:
BIOMEDICAL ENGINEERING, MACHINE LEARNING

I do solemnly and sincerely declare that:
1) I am the sole author/writer of this Work;
2) This Work is original;
3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:
Name:
Designation:

IDENTIFICATION OF MICROSCOPY CELL IMAGES BY USING CONVOLUTIONAL NEURAL NETWORK APPLICATION

ABSTRACT

Breast cancer has been the major factor of cancer death and the second main cause of women's deaths in the world. The false positive results of this cancer cell detection during the screening test leads to false treatment and emotional disturbance of the patients. Thus, breast cancer cell lines (MCF7) is used as the microscopy image samples together with the Human Bone Osteosarcoma Epithelial Cells (U2OS), and Human Hepatocyte as control to study the effectiveness of convolutional neural network (CNN) as a method of image recognition. The objectives of this study are to determine the ability of convolutional neural network in the classification of MCF7, U2OS, and human hepatocyte and also to compare the accuracy of convolutional neural network in the detection of microscopic cells by using Resnet-50 architecture model and self-training. The first procedure is a normal CNN training without using any CNN architecture model and later added with Resnet-50 model as a transfer learning to compare the results of efficiency for both method. Both results found that it can detect the image samples according to the cell type but the training accuracy gives different percentage, where the image training accuracy for Resnet-50 gives 100% accuracy result and 79.57% accuracy for self-training. Increasing the number of image samples and image augmentation has been done to improve the result of self-training but still, both methods are not comparable because the self-training is not completely done under certain circumstances. This is a small scale work where the findings may not contribute much to the medical research, but more to exposure of machine learning application.

Keywords: convolutional neural network, MCF7, cell recognition, machine learning, Resnet-50

PENGENALPASTIAN SEL-SEL MIKROSKOPI DENGAN MENGGUNAKAN APLIKASI RANGKAIAN SARAF KONVOLUSIONAL

ABSTRAK

Kanser payudara merupakan faktor utama yang menyumbang kepada kematian pesakit kanser dan menjadi faktor kedua tertinggi terhadap kematian wanita di dunia. Kesalahan mendiagnosis pesakit kanser payudara boleh menyebabkan pemberian rawatan yang salah dan juga gangguan emosi terhadap pesakit. Jadi, sel kanser payudara (MCF7) telah digunakan sebagai sampel kajian bersama-sama dengan sel epitelium tulang belulang manusia (U2OS) dan hepatosit manusia. Objektif kajian ini adalah untuk menentukan kebolehan rangkaian saraf konvolusional dalam mengklasifikasi MCF7, U2OS, dan hepatosit manusia mengikut kumpulannya serta untuk membandingkan ketepatan rangkaian saraf konvolusional dalam mengenalpasti sel-sel mikroskopi dengan menggunakan model Resnet-50 dan juga latihan rangkaian tanpa model. Kajian ini menggunakan model Resnet-50 untuk membandingkan hasil efisiensi sekiranya dilatih tanpa model CNN. Keputusan menunjukan bahawa CNN berjaya mengklasifikasi sampel gambar mengikut kumpulan masing-masing tetapi berbeza peratus ketepatan dengan merujuk kepada 100% efisiensi bagi Resnet-50 dan 79.57% bagi latihan rangkaian tanpa model. Walaupun jumlah sampel gambar telah ditingkatkan dan augmentasi gambar telah dilakukan, hasil bagi latihan rangkaian tanpa model tidak dapat dibandingkan dengan hasil yang ditunjukkan oleh Resnet-50 disebabkan oleh beberapa faktor yang dinyatakan dalam perbincangan kajian. Walau bagaimanapun, kajian berskala kecil ini mungkin tidak dapat menyumbang kepada kajian perubatan, namun sebagai pendedahan kepada aplikasi pembelajaran mesin.

Katakunci*: rangkaian saraf konvolusional, MCF7, pengenalpastian sel, pembelajaran mesin, Resnet-50*

ACKNOWLEDGEMENTS

Foremost, I sincerely want to express my gratitude to my supervisor, Dr. Wan Safwani binti Wan Kamarul Zaman for giving me the chance to enrol my final project under her supervision. I really appreciate her concern, especially when she suggested to change my project topic into something than can be done without clashing with my working schedule. Even though this topic is something that I have never experienced before, her patience and guidance helped me to keep going in times when I want to give up.

Besides my supervisor, I want to thank Dr. Wan Safwani's students, Ms. Siti Nurainie Tukimin for helping me with getting image resources and information about the lab work, and also Mr. Azrul Azizi Ahmad Zulkarnain for helping me to settle down with the command error in Matlab. Thank you for answering my questions even in unreasonable times.

Last but not least, I really thankful for the support and cherish from my parents, Mr. Ajis bin Tamrin and Mrs. Rosliza binti Khalid. They have always motivate me from discontinuing my study and keep going to finish my master study in times when I lose hope. Thank you for always reminding me for the times and money that we have spent together from my first semester.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

CNN	Convolutional neural network
MCF7	Breast cancer cell line
U2OS	Human bone osteosarcoma epithelial cells
ResNet	Residual network
ML	Machine learning
ANN	Artificial neural network
BN	Bayesian network
DT	Decision tree
SVM	Support vector machine
RNN	Recurrent neural network
VGG net	Visual graphic group net
ReLU	Rectified linear unit
ATCC	American type culture collection
DMSO	Dimethyl sulfoxide
imds	Image datastore
RReLU	Randomized rectified linear unit
LReLU	Leaky rectified linear unit
PReLU	Parametric rectified linear unit
%	Percentage

CHAPTER 1: INTRODUCTION

Image recognition and identification by using convolutional neural network (CNN) has getting popular and readily available for its application, especially in medical and biological field for screening purpose (Sommer & Gerlich, 2013). The main objective of machine learning as mentioned by Kourou et al., (2015) is to produce an algorithm for subsequent classification, diagnose and detection. From the development of conventional machine learning, it branches into deep learning, which creates many layers of neural networks to acquire a high degree of presumptions in a data given (Fu & Menzies, 2017). According to Indolia et al. (2018), CNN is a type of deep learning application which is broadly used in resolving complicated problems and it is always referred to as ConvNet.

Zhou (2018) has also said that our body have a complicated image identification system which is able to differentiate and group the objects by ourselves. So, CNN is one of the methods which copy our neural network, especially by the usage of the neurons to receive and transfer information (Zhou, 2018). It is a method from deep learning approach, which is specially used for the utilization of computer imaging that requires image identification and categorization which includes cells, cancer, face, voice, handwriting, language and also useful in business analysis ((Dargan et al., 2019; Gogul & Kumar, 2017; Indolia et al., 2018; Zhou, 2018). Kusumoto and Yuasa (2019) also have proposed that convolutional neural network (CNN) is a new advanced method from deep learning, which caused significant results in stem cell research.

The application of machine learning for the purpose of cancer cell detection has readily been studied widely across cancer research for these past years, in which it can be applied for daily clinical practices and medical discoveries (Chen et al., 2019;

Kourou et al., 2015). Thus, this study is significant in exploring the cancer detection which needs lengthy works and more time consuming when it was done classically by the experts. Training can be done with or without using a pre-trained CNN architecture model. However, training the image sample by using any CNN model is known as transfer learning because the network has been trained and can be readily used for other image training (Gogul & Kumar, 2017). This work used ResNet-50 as pre-trained model because a lot of previous research shows its effectiveness, accuracy, and simplicity during the training process (Alom et al., 2019).

Breast cancer cell line (MCF7) is chosen as one of the sample for this study because it is ranked as the fifth placed that causes death and was also identified as the major factor of cancer deaths in females and second main cause of women’s deaths in the world (Asri et al., 2016; Talari et al. 2019; Yue et al., 2018). The current constraint clinical method leads to emotional disturbance to the patients which is caused by the false positive results during the screening test (Talari et al., 2019). By accessing the accuracy and working procedure of CNN, this may increase the confidence level of using this method for diagnosing the presence of cancer cells. The decision of choosing human Bone Osteosarcoma Epithelial Cells (U2OS) and Human Hepatocyte during the training were done randomly as they are grouped in the same microscopy images and act as control, which have no effect to the result. So, recently, the researchers have pay attention to deep learning approach, such as diagnosis of skin cancer by Munir et al. (2019) and microscopy of normal cells by Akogo and Palmer (2019). Resnet-50 model has been added to the procedure because it is the most favoured CNN architecture model which has been used to classify the images (Uchida et al., 2016).

Thus, the objectives of this study are:

1. To determine the ability of convolutional neural network in the classification of MCF7, U2OS, and human hepatocyte.

2. To access the accuracy of the developed Resnet model in the detection of MCF7, U2OS, and human hepatocyte.

CHAPTER 2: LITERATURE REVIEW

2.1 Machine Learning

The term machine learning (ML) itself was acquired from the idea of human knowledge which collects knowledge from learning and incorporated those knowledge into a machine for a self-machine learning from previous events (Indolia et al., 2018). Kusumoto and Yuasa (2019) define machine learning as an algorithm tool that study from enormous dataset to outline a pattern and classification. It has three methods of learning, which are learning under guidance, without guidance and semi-guidance learning (Indolia et al., 2018). The learning software will be able to identify the image details through a technology of image identification which applies artificial intelligence by using a recent ML technique (Han & Li, 2015). General main goal of machine learning is to learn from a set of training data to generate an accurate projection on the next enormous data samples (Bishop, 2006; de Ridder et al., 2013; Hastie et al., 2005; Sommer & Gerlich, 2013; Tarca et al., 2007).

The main objective of machine learning as mentioned by Kourou et al. (2015) is to produce an algorithm for subsequent classification, diagnose and detection. Despite the diagnoses and diseases progression detection, the approach of the ML also has greatly impacted the pharmaceutical discovery and developing genomic data model (Shah et al., 2019). These techniques are referred as artificial intelligence, where several branches of the systems need to be validated by further research, in which it can be applied for daily clinical practices and medical discoveries (Chen et al., 2019). The improved technologies is crucial than the biomarkers due to its late detection of cancer development (Tukimin et al., 2019).

2.1.1 Application of Machine Learning

The application of machine learning (ML) for the purpose of cancer cell detection has readily been studied widely across cancer research for these past years, as such Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Decision Trees (DTs) and Support Vector Machines (SVMs) (Kourou et al., 2015). Besides, several influential methods have been proposed and used by Asri et al. (2016) and Yue et al. (2018) to develop the machine learning, such as the ANNs, Naive Bayes (NB), DTs, SVM and k-Nearest Neighbors (k-NN). Both K-NN and SVM are the enhanced version of conventional categorization technique, and currently, the SVM shows excellent results for this task (Hu et al., 2015). However, Kusumoto and Yuasa (2019) proposed that convolutional neural network (CNN) is a new advanced method from deep learning, which caused significant results in stem cell research.

By comparing SVM with convolutional neural network (CNN) in image identification, CNN requires higher consumption of times from hours to weeks to train the network, but it promises satisfaction to the performance results of matrices as compared to SVM (Fu & Menzies, 2017). It has also proven through an experimental analysis, which was mentioned by Fu and Menzies (2017), where deep learning harvest 96.5% precision, while 80% precision for SVM from a data acquired from an Android app. CNN has been used as a morphology-based recognizing system to assign automation of cell identification, differentiation and their growth stages without molecular tagging (Hu et al., 2015; Kusumoto & Yuasa, 2019). The classification of accurate stages of MCF7 is important to determine the best and correct treatment for breast cancer patients (Yue et al., 2018).

Machine learning is divided into supervised and unsupervised machine learning (Sommer & Gerlich, 2013). Supervised machine learning has been used widely in various biological fields for the screening purpose (Sommer & Gerlich, 2013). Supervised learning can be understood as learning from an external huge set of image data which need to be trained by artificial intelligent method to come out with image identification prediction from its available previous learning data (Dargan et al., 2019; Gogul & Kumar, 2017). It is divided into two problem categories, which are the classification problem and regression problem (Joss & Muller, 2019). Classification relates to the work of recognizing and distinguishes the observed data into particular named group or features (discrete data), while regression would come out with numerical or continuous data to figure out the analytical illustration which can relate the observed variables (Joss & Muller, 2019).

While the unsupervised machine learning has grouped the similar feature of data samples into a cluster and accelerate the data extraction by decreasing the complication of data (Sommer & Gerlich, 2013). This method helps in the research of new and undefined phenotype which have been favourably defined the phenotypic identification of drug performance (Sommer & Gerlich, 2013). The unsupervised learning also means the lack or non-existence of the target data (Dargan et al., 2019).

2.2 Deep Learning

Deep learning is part of machine learning (Fu & Menzies, 2017). From the development of conventional machine learning, it branches into deep learning, which creates many layers of neural networks to acquire a high degree of presumptions in a data given (Chen et al., 2019; Fu & Menzies, 2017). Deep learning is an approach which avoids characteristic selection by human favors, where the selection will be extracted from raw materials (Indolia et al., 2018). It has gotten its place in various fields which has arose and improvised from machine learning application from the aspect of better supervision, saving time and budget and also more effective in solving complex data (Dargan et al., 2019). Zhou (2018) has also said that our body have a complicated image identification system which is able to differentiate and group the objects by ourselves. So, CNN is one of the methods which copy our neural network, especially by the usage of the neurons to receive and transfer information (Zhou, 2018). It is a method from deep learning approach, which is specially used for the utilization of computer imaging that requires image identification and categorization (Gogul & Kumar, 2017; Zhou, 2018).

It serves as a method for diverse application, especially in object identification, which includes cells, cancer, face, voice, handwriting, language and also useful in business analysis (Dargan et al., 2019; Indolia et al., 2018; Khan et al., 2018; Liang, 2017). There are three types of neural networks that use the principal of deep learning, namely the convolutional neural network (CNN), artificial neural neural network (ANN), and recurrent neural network (RNN) (Fu & Menzies, 2017). However, Dargan et al. (2019) provides more examples of deep learning architecture, such as deep stacking network (DSN), auto-encoder (AE) and restricted Boltzmann machine (RBM). For this study, CNN serves as a good application at studying and identifying objects' features effectively, such as the language expression, identification of image, voice, and face (Indolia et al., 2018; Zhou, 2018). This method can decrease the amount of parameters that can be exercised to enhance the algorithm backward propagation (Han & Li, 2015).

The basic technique of deep learning includes the phase of training and testing, where a particular characteristic is developed and trained from a given sample data in a training phase before it can be used in the testing phase for any similar traits or characteristics (Dargan et al., 2019). ANN is also part of deep learning but ANN is not fit for image identification because it cannot bear large size of image which leads to over-fitting (Gogul & Kumar, 2017). If ANN needs to be used, it may need larger processor to fit the input weights (Gogul & Kumar, 2017). Thus, the effectiveness of a particular deep learning approach or application is depending on the size of data (Dargan et al., 2019).

The main factors of that contribute to the deep learning techniques are the non-linear processing at its numerous layers and also the unsupervised and supervised learning method as mentioned in the machine learning technique (Dargan et al., 2019). The non-linear processing can be explained by the passing of extracted information from the current layer to the next layer accordingly (Dargan et al., 2019). The next layer, which acts as the receiver will receive the information as the input and then, it will be passed out to the higher level as the output (Dargan et al., 2019).

As can be seen in **Figure 2.1** and **Figure 2.2**, the main difference between ANN and CNN is the connection of layers and the neurons, where the neurons in every layer of ANN is linked to each other, while in CNN, the fully-linked neurons only happen at the terminal layer of CNN (Gogul & Kumar, 2017). In other words, there is only small part of regions in each layer which are linked together to gather information (Gogul & Kumar, 2017). The layers added in between the input and output layers are known as the hidden layers, where else CNN does not only consisted of hidden layers in between the input and output layer, but also the convolutional layers and pooling layers (Dargan et al., 2019).

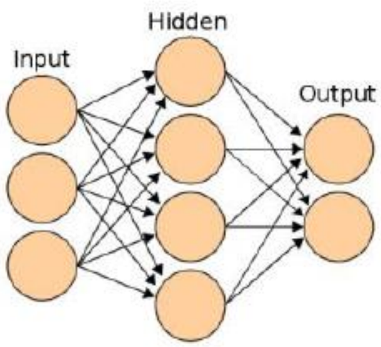

Figure 2.1: Concept of Artificial Neural Network (Gogul & Kumar, 2017)

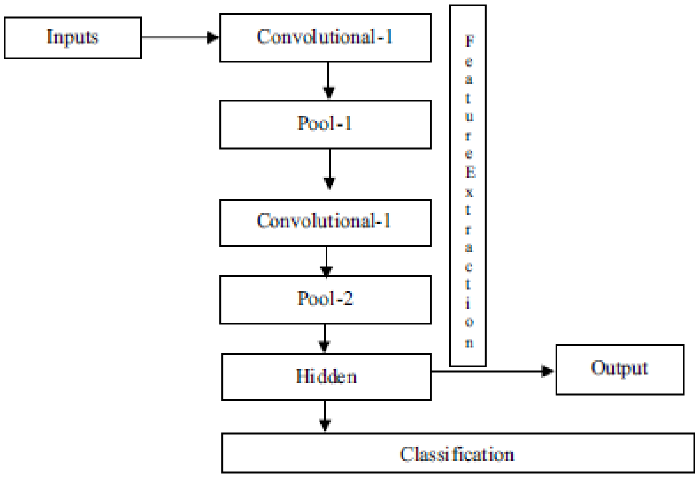

Figure 2.2: Concept of Convolutional Neural Network (Dargan et al., 2019)

Deep learning can be differentiated from machine learning in term of its algorithm, where deep learning will build its own algorithms in multiple layers to develop the networks and presents its own findings, while machine learning has to be provided with readily available algorithms to continue with the learning process and data presentation (Dargan et al., 2019). If one needs to deal with high volume of sample data, then deep learning would be a better choice than the machine learning (Dargan et al., 2019). However, deep learning needs hardware with rather huge processing ability and time consuming to compute and build the complex network (Dargan et al., 2019). Despite the drawbacks of the deep learning, it can be encountered with its contented results and high precision (Dargan et al., 2019).

2.3 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) was first introduced by LeCun (Dargan et al., 2019) and it is a neural network with hierarchical structure which works on objects' identification by applying important process, for instance the backward propagation and also Gradient Descent (Zhou, 2018). Hierarchical means the information is passed out from one layer to the next layer accordingly as the input and output, which means that it works on numerous layers (Dargan et al., 2019). It can be also referred to as a feedforward method, where the signal is processed directly without any loops or cycles (Dargan et al., 2019; Munir et al., 2019). Han and Li (2015) and Hu et al. (2015) define CNN as the multiple layer of neural network which is particularly built to work on the categorization of 2-dimensional (2-D) image data with great results.

It is a type of deep learning application which is broadly used in resolving complicated problems and it is always referred as ConvNet (Gu et al., 2017; Indolia et al., 2018). CNN is believed to be the first technique of deep learning which has been successfully proven to be powerful in applying the multiple layers ranking in the networks (Han & Li, 2015), especially in case of acquiring a good image quality in medical imaging (Ayas & Ekinci, 2019). This model was initiated by McCulloch and Pitts in 1943 to develop a model which is based on humans' neural network and imitate the reaction of the neurons by using a computer (Zhou, 2018). This is the most used deep learning architecture due to its high precision (Dargan et al., 2019).

2.3.1 Characteristics and Architecture of CNN

As stated by Indolia et al. (2018), CNN has three important features that make it as a classical representation than other tools, which is the application of load sharing to reduce the number of parameters that is required to be trained for smoother results. Next, the progression of learning is assimilated into two stages, which are categorization stage and also trait selection stage. Lastly, the implementation of CNN is much easier than ANN because ANN applies a huge network by using a common model for its operation. Zhou (2018) supported these statements by saying that the load sharing and decreasing in its dimension is the strength of CNN. These features have become the important proof in producing results with high competency and influential in calculating the error and speed rate (Zhou, 2018).

There are four general layers stated by Indolia et al. (2018), which are the Convolution Layer, Pooling Layer, Fully Connected Layer, and Activation Function. While Hu et al. (2015), has classified the above layers into five main layers, which starts with Input Layer and ends with Output Layer that replaces the Activation Function. According to Zhou (2018), the basic of CNN structure starts with convolution layer and ends with output layer. The middle layers which are located between the first and final layer is known as the hidden layers that constitute the pooling layers or sub-sampling layers (Zhou, 2018). The function of the convolution layer is to gain the image details and transferred the details to the hidden layers to be calculated before harvesting the results at the output layers (Zhou, 2018).

A better idea of how these layers carry out their tasks can be seen in **Figure 2.3**. This figure also clearly shows that the neurons in the hidden layers are partially-linked, except at the output layers where the layer is fully-linked. These partially-linked neurons give advantage in decreasing the calculating weight by using load sharing, which is known as kernels (Zhou, 2018). CNN applies a method of developing the image features from the lowest layer to the highest layer (Gogul & Kumar, 2017). For instance, it will gather different range of information at each layer, hierarchically, for the process of building-up the knowledge to fasten the speed of calculation (Gogul & Kumar, 2017; Zhou, 2018).

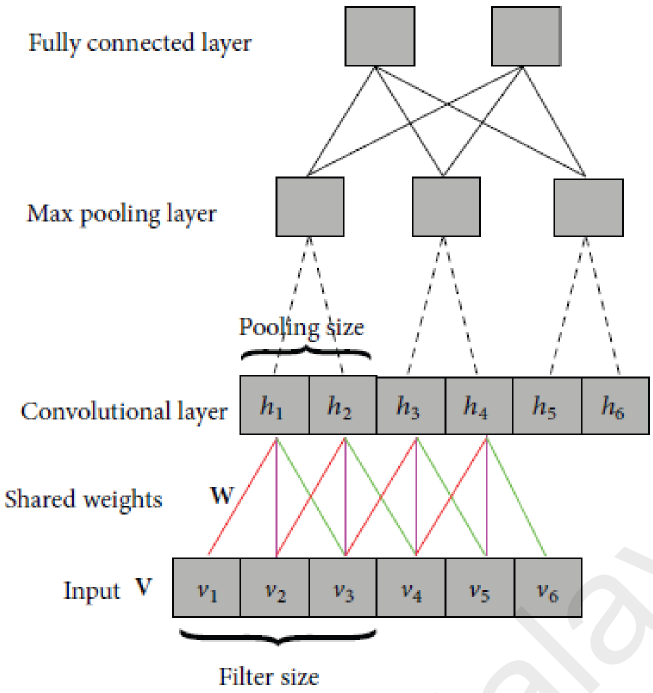

Figure 2.3: Basic Layers of CNN Architecture (Hu et al., 2015)

Figure 2.4 shows general steps of CNN, where the image is further subsample before convoluting and pooling the adjacent pixels. The summarized of adjacent pixels will be replaced in the output prediction layer with summarized features or traits (Munir et al., 2019). The purpose of pooling process give a hand in translating the output images with no changes or minimal changes to the input images (Munir et al., 2019). A study that was carried out by Oei et al. (2019) was proven that CNN method gives a better result than human expert in classifying the cell lines despite its aggressive level.

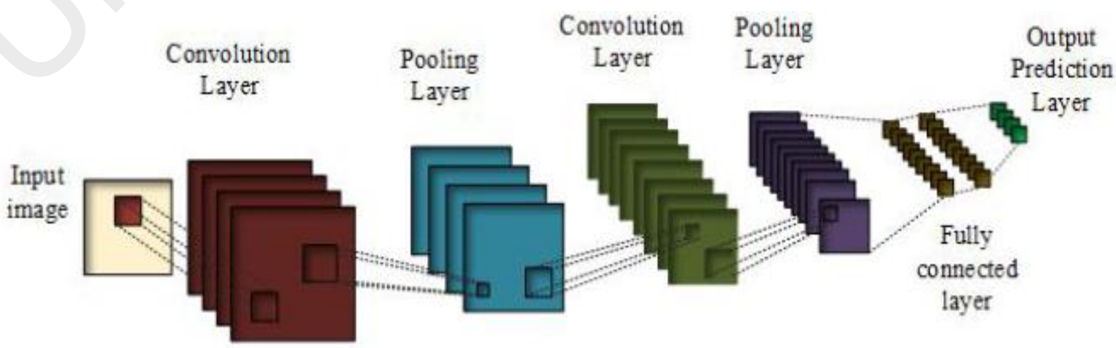

Figure 2.4: **General CNN** (Munir et al., 2019)

There are three types of CNN architectures which were mentioned by Indolia et al. (2018), such as LeNet Architecture, AlexNet Architecture, and GoogleNet Architecture. AlexNet works best with graphics processing unit (GPU) for its computation speed (Dargan et al., 2019). LeNet Architecture is suitable for operating the analysis of image identification and classification due to its capability to study and digest very difficult and challenging dimensional features (Indolia et al., 2018). LeNet was established by LeCun in 1989 for a digital recognition (LeCun et al., 1989; Zhou, 2018). By referring to **Figure 2.5**, LeNet5 comes from layer of five convolutions and the remaining of three connected layers which makes it a total of eight layers altogether (Indolia et al. 2018). Han and Li (2015) also have mentioned that LeNet can be used to solve various image identification problems with satisfied results (Han & Li, 2015).

Besides that, Dargan et al., (2019) added more CNN architecture model which have been used for building convolutional neural network, such as residual network (ResNet) and visual graphic group net (VGG). The current used or architecture model with the lowest error rate is ResNet model, which was built by Kaiming (Alom et al., 2019; Uchida et al., 2016).

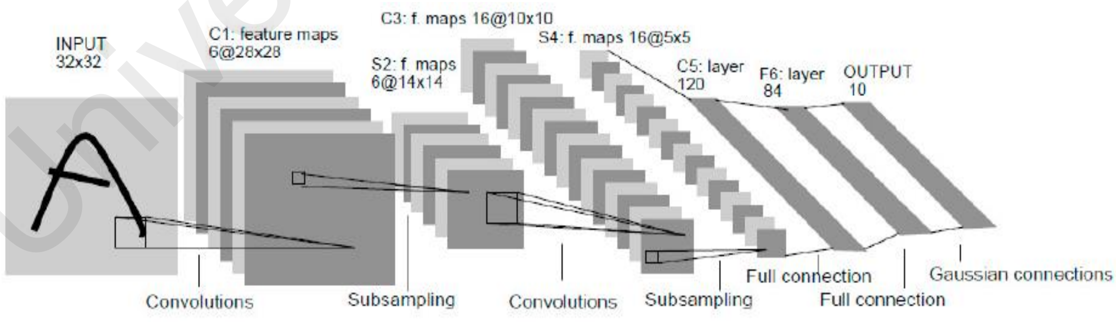

Figure 2.5: **CNN Architecture of LeNet5** (Indolia et al., 2018)

In a network of CNN, a tiny part from the input image will be extracted as the input figures which will be passed by a forward propagation to many series of layers in the network architecture (Han & Li, 2015). For example, a portion of 5 x 5 unit area in a first layer is accepted from a full image of the input layer, which is known as the receptive field (Indolia et al., 2018) or an initial sensing area (Han & Li, 2015). While calculating the extracted unit area from the input image, the result at the feature map can be produced variously with a variety of load sharing from the same input layer. This will produce various characteristics at the resulting feature map (Indolia et al., 2018). However, as mentioned previously by Indolia et al. (2018) about the important features of CNN, the same load sharing is applied for each unit in the plane. And later, the transferred output unit area can be seen exactly at the same position in each feature map of that layer. So, we can say that each layer is connected to each other by image sampling from its preceding layer. Sigmoid activation function is used on the output unit area of the first convolution layer which contains of neurons (Indolia et al., 2018).

The second layer in **Figure 2.5** shows that subsampling is applied which produced the same total output number of feature map as in the first convolution layer. So, there are 6 feature maps contained within the first convolution layer and second subsampling layer, and also 16 feature maps in the third convolution layer and forth subsampling layer. A filter is placed in every layer in the network to compute the significant values of the obtained data (Han & Li, 2015). To be specific, the filter (sometimes referred as a kernel in some paper) is placed in the convolutional layer (Abiyev & Ma'aitah, 2018). For that action here, subsampling is carried out to decrease the complicated values of significant digits because it is said that the significant values at the extracted image location is not critical (Indolia et al., 2018). In other words, subsampling will decrease the dimensions of the data by improving the network strength which is affected by minor disturbance or noise (Han & Li, 2015). Then, a portion of 2 x 2 unit area is now become the input image and the feature map in the next adjacent layer can be seen increasing in its number with decreasing spatial resolution (Indolia et al., 2018).

2.3.2 CNN Algorithms

Munir et al. (2019) has figured out several convolution functions as listed in **Table 2.1.** There are Atrous Convolution Function, Sigmoid Activation Function, Tanh Function (Scaled Sigmoid Function), and Rectified Linear Unit (ReLU) Function. Both the Tanh and ReLU Function is considered as non-linear activation function which carry out computational tasks to guess the probability of the results for predicting the image label and group (Gogul & Kumar, 2017). These algorithms were developed to improve the previous problem, which can only deal with linear models of object categorization in 1969 (Zhou, 2018). The first algorithm for solving the non-linear problems was developed in 1986, which is a back-propagation algorithm and by Rumelhart et al. (1988) which introduced a multiple layer sensor, and also Sigmoid Function where these techniques have become the key resolution to the categorization of the non-linear function (Zhou, 2018).

Table 2.1: Convolution Functions (Munir et al., 2019)

Atrous Convolution Function	y[i] = $\sum_{k=1}^{k} x\,[i + r\,.\,k]w[k]$
Sigmoid Activation Function	$\sigma(x) = \frac{1}{1+e^{-x}}$
Tanh Function	$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$
Scaled Sigmoid Function	tanh(x) = 2sigmoid(2x) – 1
Rectified Linear Unit (ReLU) Function	G(x) = max(0,x)

Activation Function is the extensive model from the traditional ML algorithms which applies sigmoid activation function as can be seen in **Table 2.1** (Munir et al., 2019). In addition to the sigmoid activation function, Rectified Linear Unit (ReLU) has been introduced due to the uncomplicated technique while computing the partial derivative, requires a shorter duration than sigmoid function, and also restricts the dissipation of gradient during the image convolution and pooling (Indolia et al., 2018). However, they have provided way to overcome the ineffectiveness of ReLU, where the neuron can become inactivated if huge gradient is going through the networks. The have tried to resolve this problem by using Leaky ReLU, where if

$x < 0$ is detected, this will operate the function of $f(x) = x$, and if

$x > 0$ is detected, this will operate the function of αx (α = small constant).

Later, Gradient Descent, which is a type of optimizing algorithm is computed to achieve satisfied outputs as the targeted outputs by adjusting the deciding parameters by using error backpropagation (Indolia et al., 2018). This method is carried by calculating the outputs of each layer and identifies the error elements that can be found at the final layer (Indolia et al., 2018). The enhanced network can be achieved by calculating these gradients backpropagated and is repeated until it reaches efficiency (Indolia et al., 2018).

2.3.3 CNN of Previous Studies

Several steps need to be taken before obtaining a quality image for cancer diagnosis, which are the removal of noises during the pre-processing of raw microscopic images, extraction of desired regions from its background during image segmentation, grabbing traits during image post-processing, diagnosis of ABCD-rule (A – asymmetry, B – border, C – color, D – diameter), obtaining a score value by using a technique of seven-point checklist, the next one is carrying out *Menzies* Technique by observing the presence and absence of particular positive and negative traits and the last one is tracing a local or global patterns on the images (Munir et al., 2019). However, these steps were taken by Munir et al. (2019) to diagnose a skin cancer which may be beneficiary for other cancer analysis too by the mean of deep learning.

A machine learning offer a powerful method for the detection of microscopic image data to improve previous conventional image analysis algorithm method which are more tedious for the biological cell lab works with limited expertise and knowledge in the image mathematical progress (Sommer & Gerlich, 2013). The main phase in machine learning is the training phase, which is a work of building and learning the relationship from a group of data samples by engineering software (Sommer & Gerlich, 2013). This technique screens the morphological images of particular images with the presence of fluorescent markers (Sommer & Gerlich, 2013). Sommer and Gerlich (2013), have explained the basic steps for the pipeline (as in **Figure 2.6**), which has used chromatin marker for human HeLa cells.

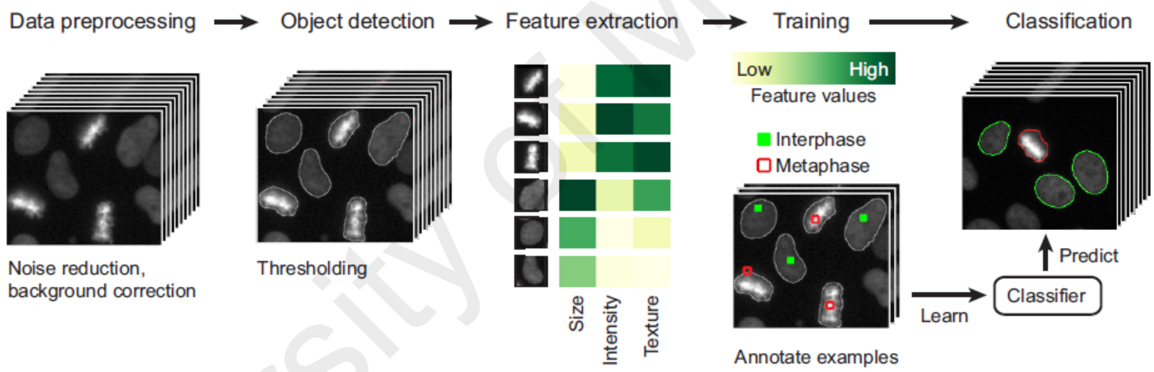

Figure 2.6: A machine pipeline for microscopy data analysis (Sommer & Gerlich, 2013)

Identifying microscopy images of cells also has been carried out by Akogo and Palmer (2019) by using CNN. CNN was a successful method to identify handwriting and optical characters in 1980s which was in the era it was initially been introduced (Akogo & Palmer, 2019). In recent years, it is also recognized as an important factor of overcoming various types of problems, especially in the image recognition and segmentation (Akogo & Palmer, 2019; Chen et al., 2014; Redmon et al., 2015; Ren et al., 2015) and medical image analysis (Akogo & Palmer, 2019; Albarqouni et al., 2016; Esteva et al., 2017; van Grinsven et al., 2016; Yang et al., 2017). This model was elected for their ability to recognize and differentiate classes of breast cancer cell lines (Akogo & Palmer, 2019).

Gogul and Kumar (2017) have carried out a study about the identification of flower variants by using CNN method of deep learning which produce high precision. The extraction processes were accomplished by using a Transfer Learning method from images which have been taken by using a cell phone, and a union of CNN with Transfer Learning method was successfully harvest high precision of image identification (Gogul & Kumar, 2017). They have suggested several approaches, such as the Random Forest or Logistic Regression to harvest a better precision by reducing the usage of hardware to calculate the training measure of CNN (Gogul & Kumar, 2017).

A study of distinguishing between the cancer lines and normal cell lines by using nuclear mechano-morphometric biomarkers for machine learning has been carried out by Radhakrishnan et al. (2017) which can be seen in **Figure 2.7**. They have adapted a CNN method which is proven to give high accuracy and good explainable properties. **Figure 2.7(a)** shows the identification and extraction of nuclei images from several cell lines by using linear models and CNN machine learning, while **Figure 2.7(b)** shows several nuclear morphometric images from different types of cell lines.

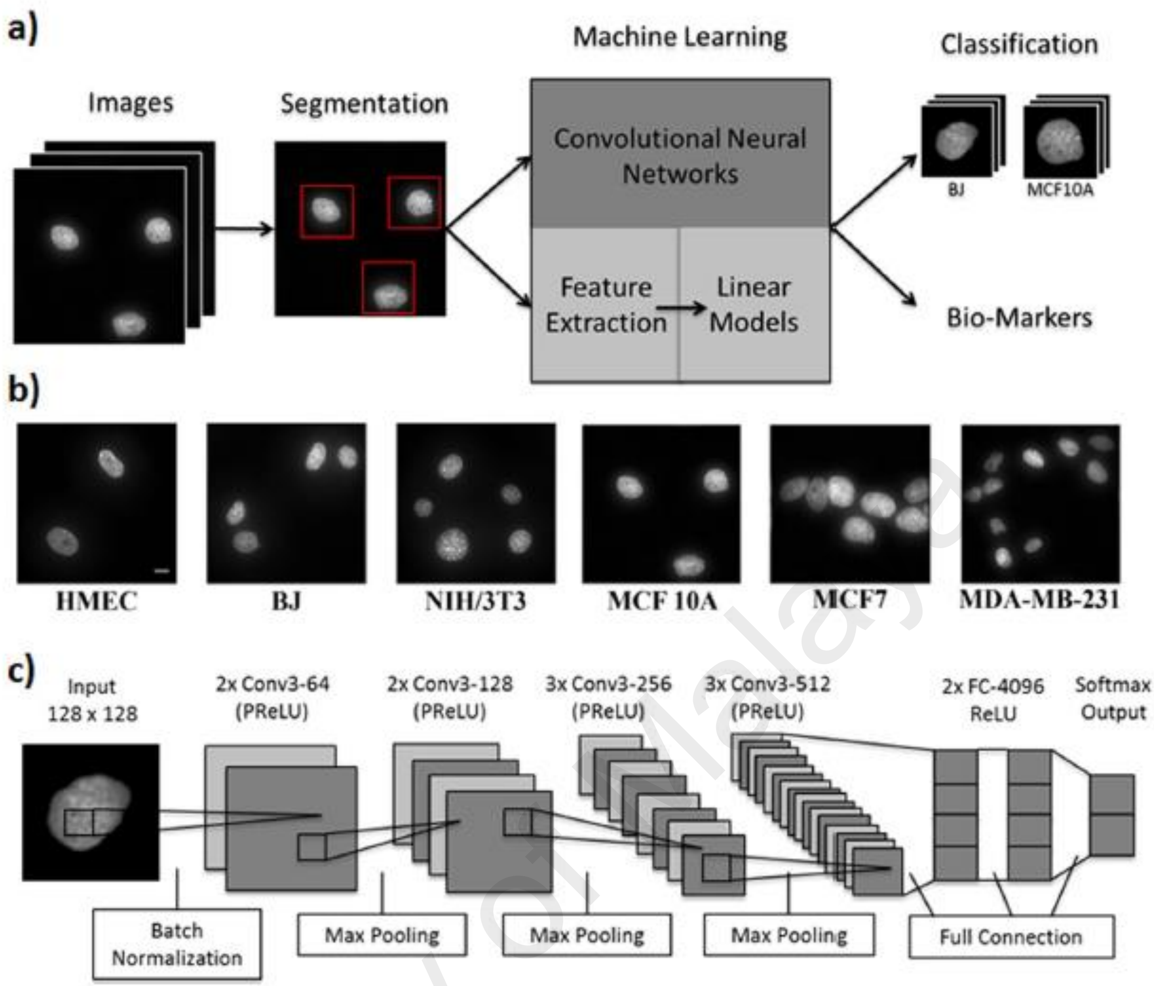

Figure 2.7: A machine pipeline which combine single-cell nuclear imaging

(Radhakrishnan et al., 2017)

2.4 MATLAB

Matlab is a computer-based system and world convenient scientific tool to execute computer algorithms which is coded in a particular programming language (Elsayed & Yousef, 2019). The word MATLAB is derived from MATrix LABoratory, which serve as a simple approach to matrix software (Houcque, 2005). It was first established in 1970's by Cleve Moler which is equipped with Eigen system package (EISPACK) and linear system package (LINPACK) with high presentation language for a practical computing (Houcque, 2005). It was not formally distributed before 1984 until Mathworks Inc. took place the responsibility to promote and continue the growth of MATLAB (Houcque, 2005). After LINPACK and EISPACK were introduced, Linear

Algebra Package (LAPACK) was launched to replace LINPACK and EISPACK (Houcque, 2005). This tool undergoes an excessive growth for its ability to correlate an enormous amount of experimental and synthetic figures (Joss & Muller, 2019). While openCV is another tool in computer algorithm coding using (C++)-based library but with faster execution speed (Elsayed & Yousef, 2019). Based on a research carried out by Elsayed and Yousef (2019), it was proven that openCV performs better in term of speed behaviour in most of the group recognition experiments (Elsayed & Yousef, 2019).

However, one of the experiments favour MATLAB over openCV for its speed when changes were made by several factor, such as sample size, the training set, the number combination of dimentionality and the frequency of observation (Elsayed & Yousef, 2019). In addition, according to Houcque (2005), the programming has been modernized with polished data configuration, visualisation, assist the programming for object-designated materials and also provide tools for built-in editing and debugging setting. Despite the speed factor, MATLAB is undoubtedly proven to be convenient for research purpose with its heavy toolboxes and library, together with successful presentation data analysis and visualization (Elsayed & Yousef, 2019). This tool is not difficult to learn and easy to fix the flaws if there is any (Houcque, 2005). Its function can easily be described as an enumerator and a programming language which carry out very fast calculation, especially for matrices (Houcque, 2005). MATLAB works great in a way of combining an enumeration tasks with a visual plotting and it also does have feature of object-designated elements (Houcque, 2005).

In comparison with C++ language, this tool performs faster than MATLAB when it comes to expressing the language and it does not fit with other tools for language commands due to its specified MATLAB operations (Houcque, 2005). In addition, MATLAB does not aim at a common programming language because it is modelled for research-based computing and does not fit for other usage (Houcque, 2005). Besides its great computing application, MATLAB competes with other programming tools, such as Mathematica, Rlab, Scilab, and GNU Octave (Houcque, 2005). Ones will have to favour their programming tools according to the tools' speciality, as MATLAB has specialty over linear algebra and calculation interpretation, while Mathematica works best with symbolic operation (Houcque, 2005).

2.5 Cell Classification

Differentiating a normal cell with a cancer cells can be carried out by observing the structures of actin cytoskeleton, which will give new features about the changes of the malignant cells (Oei et al., 2019). This will be useful as a supplementary marker of analysis (Oei et al., 2019). Oei et al. (2019) has also proposed a similar method as previous review about using CNN to classify the cells by looking at the microscopy images of the actin-labelled fluorescence. The cytoplasm in the eukaryotic cells contain cytoskeleton which consisted of microtubule, intermediate filaments and actin filaments (Oei et al., 2019) as illustrated in **Figure 2.8** below.

According to Oei et al. (2019), there is high possibility of cancer growth when modification and changes to the actin filaments are observed. Actin filaments are known for retaining the cell shape and movement and these properties undergo a modification in cancer cells to spread and move to adjacent tissues to form another tumor at different tissue region by the mean of lymph and blood vessels (Oei et al., 2019). Currently, the study of actin filaments behaviour has not been covered extensively by machine learning due the shortage of huge dataset to assist the application. Thus, the observations of modification for these actin filaments by the mean of machine learning application will contribute to the idea of biological cell behaviour.

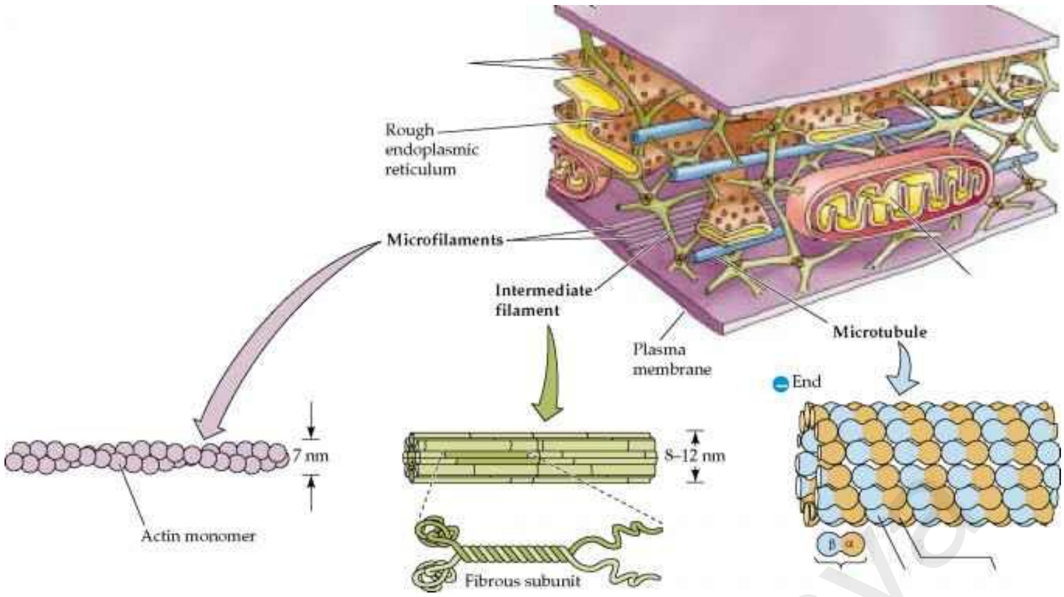

Figure 2.8: A cytoskeleton subunit which consisted of actin filaments (microfilaments), intermediate filaments and microtubule. Source: Google Image (https://www.78stepshealth.us/plasma-membrane/the-cytoskeleton.html)

Categorizing a particular cell with high accuracy is highly challenging but it really helps in medical analysis to develop ways of disease hindrance and setting up a suitable treatment to the disease (Oei et al., 2019). Machine learning comes in handy in the medical images and tissue specimens diagnostic and also cell categorization which is based on microscopic cell images (Oei et al., 2019). The shape and arrangement of actin filaments in the cells give an important potential data about the biological cell behaviour (Oei et al., 2019). However, it needs fluorescence element tools to tag the cells, where else the image acquired were not used any fluorescent markers and therefore, the training of the image was carried without the fluorescent labelling image.

2.5.1 Breast Cancer Cell Line (MCF7)

In 2017, there were at least 40, 000 deaths caused by breast cancer reported in American women and it could be higher in those developing countries with poor diagnosed and medical equipment (Akogo & Palmer, 2019). This is the type of cancer which is frequently happened among women and caused serious effect to the citizen's health and population (Comsa et al., 2015). The latest data has reported that 20% of men and 16.7% of women discover their cancer development during their lifespan, while 12.5% of men and 9.1% of women could not survive from the illness (Talari et al., 2019). Breast cancer that was diagnosed among women is ranked as the fifth placed that causes death and was also identified as the major factor of cancer deaths in females (Talari et al., 2019). So, over 40 years, the scientists have developed a cell line, which is called MCF-7 to raise in its laboratory cancer research (Comsa et al., 2015). MCF-7 gained its name from a research which was carried out at the Michigan Cancer Foundation (MCF) by a group of researcher which is led by Dr. Soule in 1973 (Comsa et al., 2015). The current leading constraints in diagnosing the breast cancer can be overcome by limiting the false positive result from the screening test, dealing with patients' discomfort, be it emotional or physical attributes and also their financial needs (Talari et al., 2019).

In history, the MCF-7 were extracted from a pleural effusion of a patient (69 years old) with metastatic disease which has undergo a surgery to remove her breast (mastectomy) in 7 years back before MCF-7 culture started (Comsa et al., 2015). The research of MCF-7 were further be enhanced by the application of estrogen to stimulate its growth and it was also found that its growth can be inhibited by anti-estrogens tamoxifen (Comsa et al., 2015). Before tamoxifen was used, synthetic estrogen diethylstilbestrol was applied with triple times longer to the disease since it is hormone-responsive (Comsa et al., 2015). MCF-7 is the popular cell line for breast cancer that has been cultivated over the years globally (Comsa et al., 2015). This cell line is lowly aggressive, has poor metastatic ability and also is not invasive (Comsa et al., 2015).

The number of chromosomes present in MCF-7 cells can be varied from 60 to 140 which show a chromosomal abnormality number (aneuploidy) in a haploid set (Comsa et al., 2015). A different culture state and technique will have an effect on the karyotypic distinction which is caused by the selective pressure (Comsa et al., 2015).

These conditions also contribute to the upraised level of genetic fragility in the cell line (Comsa et al., 2015). Comsa et al., (2015) stated that there is a fragment of stem cells which stimulate the production of clonal variation and those variants in the cell line will encounter differentiation at two levels, which are the RNA expression and genomic level. The development of human MCF-7 line is influenced by estrogen receptor (ER)-positive, progesterone receptor (PR)-positive and also plasma membrane-associated growth factor receptors (Comsa et al., 2015). Both the *in vivo* and *in vitro* method were carried out to mimic the condition as in *in vivo* for a whole representation of interactivity between the cancer cell line with human metabolism processes to get a proper assessment than the basic cell culture (Comsa et al., 2015).

CHAPTER 3: METHODOLOGY

3.1 Culturing of MCF7

MCF is a breast cancer cell line, which stands for Michigan Cancer Foundation. The guidelines for handling this cell is based on MCF7 form American Type Culture Collection (ATCC® HTB-22™). While receiving the frozen cells, they have to be placed quickly at a temperature lower than -130°C and liquid nitrogen would be an ideal storage to keep the cells alive. The frozen vials need to be handled with full protection, which includes the wearing of gloves, safety clothes, and a fully-covered face mask. This precaution step is important to overcome the vessels explosion or blowing off of the vessels' cap which is expected to result from the leakage of the vessels that force the liquid nitrogen to fill into the vessels when they are submerged in the liquid nitrogen. This explosion is the results of the conversion of liquid nitrogen into its gaseous state. Thus, before submerging the vials into the liquid nitrogen, it is important to do the screening or inspection on any broken or leakage vials.

A 75 cm^2 flask was used for sub-culturing. Before starting the sub-culturing, the presence of floating cells need to be observed. If there are any, the floating cells can be taken out and be placed at first and second sub-culture as mentioned below:

The culture medium is transferred to a centrifuge tube. The cell layer is washed thoroughly with 0.25% Trypsin and 0.53 mM EDTA solution to ensure that it is clean from all of the serum traces. It has to be cleaned because it contains trypsin inhibitor. Trypsin-EDTA solution with 2.0 to 3.0 mL was added to flask for the cells to be examined by an inverted microscope. This can be done in a range of 5 to 15 minutes where the cell layers were observed to be evenly dispersed. It is important to prevent agitation or shaking of the flask which may cause the cells to clump together. If the cells are hardly being found to detach from the cells, they can be put at 37°C to fasten the dispersal rate.

Now, the aspirated cells and the growth medium of 6.0 to 8.0 mL can be softly pipetted to a centrifuge tube. Then, the cell suspension can be taken into the centrifuge tube together with the cultured medium from step 1 to be centrifuged for 5 to 10 minutes at 125 x g. later, the supernatant can be discarded. When the supernatant has been discarded, the remaining cell pellet can be resuspend in a newly growth medium. A portion of the cell suspension can be added to a new culture vessel. The cultures need to be incubated at 37°C, while the growth medium need to be renewed 2 to3 times weekly.

3.1.1 Culture Medium

ATCC-formulated Eagle's Minimum Essential Medium is used as the base medium for this breast cancer cell line, where the catalogue number is No. 30-2003 MCF7 (ATCC® HTB-22™). After that, additional components need to be added to the base medium, such as 0.01 mg/ml human recombinant insulin and 10% fetal bovine serum. It is also nourished with 5% DMSO.

3.2 Image training by Using MATLAB

The images were pre-trained in Matlab version R2019b. This work has used CPU (central processing unit) computation with AMD Radeon HD 6320 Graphics, 2 GB RAM and C++ programming language. The categories of the cells that were trained were consisted of three types of cells, which are the breast cancer cell line (MCF7), Human Bone Osteosarcoma Epithelial Cells (U2OS), and Human Hepatocyte. The images of U2OS and human hepatocytes were obtained randomly from a Broad Institute website database (https://data.broadinstitute.org/bbbc/image_sets.html).

The cells were grouped into three types of file, namely the U2OS (Human Bone Osteosarcoma Epithelial Cells), MCF7 (breast cancer cells), and Hepatocyte (Human

Hepatocyte). The network training consisted of image samples with of 20X magnification. Image Data Store (imds) was started to create in the Matlab R2019b to store the chosen images.

The images need to undergo an image processing before sending the image to CNN. ResNet-50 takes the size of 224 x 224 x 3 as the image input. Both the 224 represents width and height of the image dimension of the image, while 3 represents the depth of the image, which is three colour channels of RGB (red, green, and blue). Augmented image data store was applied to resize and convert any grayscale image to RGB image. Another benefit of applying this work is that it can be used for additional data augmentation when used for network training. So, the images were divided into two categories, which are the training set and the test set. The percentage of images that were taken into training set is 30%, while the remaining percentage can be used for test set or validation. The training data set were resized according to the image size that needs to follow the requirement of input image size of ResNet-50, which is 224x224x3. The same thing was done to the test image set.

Figure 3.1 shows a mathematical model to represent the information flow in a neural network which is modelled by a neuron as can be seen in **Figure 3.2**. The flow of information in **Figure 3.2** represents as magnitude in **Figure 3.1** to carry out the computation of activation function, which will give the classification output. These figures are the overall representation of the neural network computation, starting from the inputs to the output. The neurons compute the summation of weights (W) and inputs (I), and the compare it to the threshold (T). The weights carry magnitude which decides on the amount of input gives to the output.

The summation of these input and weight multiplication represents as f(x) and acts as the activation function that connects the input with the output. CNN undergoes two processes, which are the forward propagation and backward propagation. The function of the forward propagation is to calculate the results at the output by the activation function f(x) as shown in **Figure 3.1**. Depending on the type of the activation function used, T is set between 1 and 0. If the result of the summation is higher than the T, the output result is equal to 1, and vice versa.

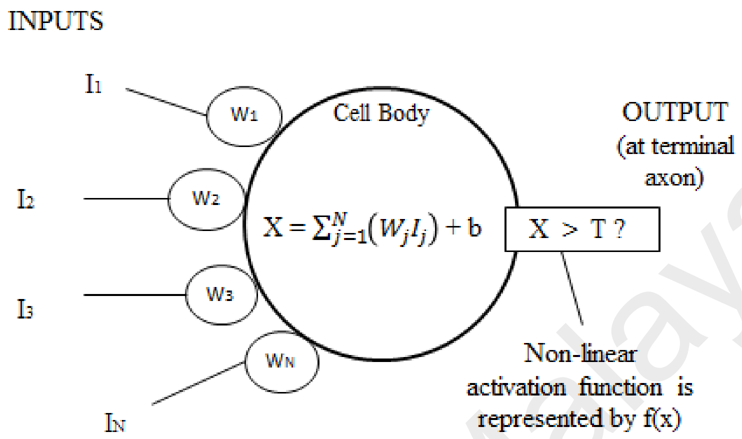

Figure 3.1: Mathematical model representing a neuron model

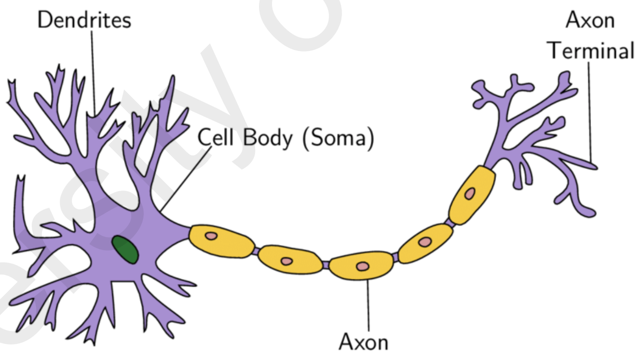

Figure 3.2: A neuron model in a human neural network (Bekolay, 2011)

By referring to **Figure 3.3**, it simplifies the connection of forward and backward propagation in one figure. Let z = x y z, H_1 = hidden layer 1, H_2 = hidden layer 2 and f(z), p(z), q(z) is the function represented at each layer. The backward propagation is the inverse calculation of the forward propagation to learn the design of the input features.

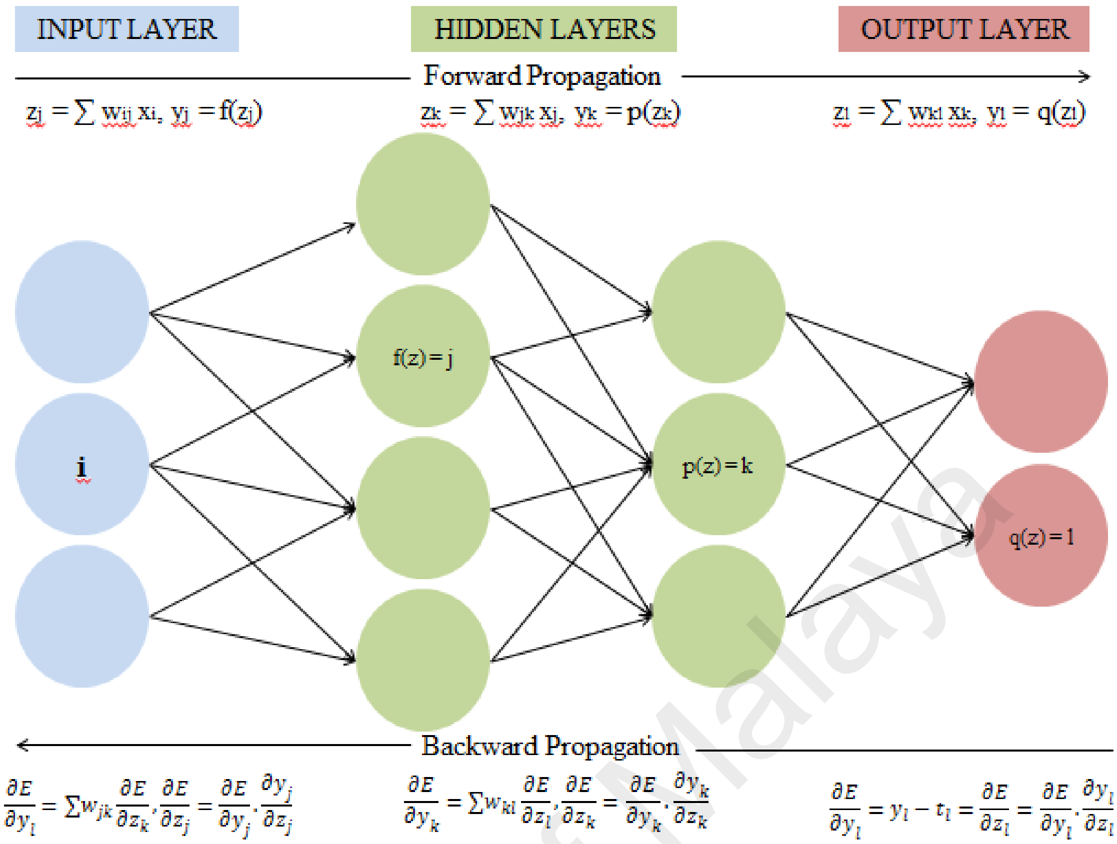

Figure 3.3: General forward and backward propagation of neural network (Liang, 2017)

3.2.1 Convolutional Layer

This is the most crucial layer in a CNN architecture which contains several filters (convolutional filters) that can produce several feature maps. The feature map that was extracted depended on the convolution filter we are using. The feature maps and the filters are linked by weights. The filters of convolutional layer are two dimensional matrices. These filters are determined through training process. If there are 4 convolutional layers, then it will generate 4 feature maps. The difficulty of network can be decreased by sharing of weights (parameter) which happened when there is similar feature map in the network. Parameters in this network refer to the W (weights) and b (biases) as shown in **Figure 3.1** that represents the neurons. This reduction of difficulties can be maintained by ensuring that the number of weights sharing is small.

An example of the connection between the filter and the extracted images can be seen in **Figure 3.4**. As mentioned previously, the dimension of the input image follows the requirement of ResNet-50, which is 224 x 224 x 3. The filter in the convolutional layer is connected to a portion size of the input image and cover the full depth from which the spatial image was taken. The result of the convolution between the input images with the filter produced an output image as a pixel, which is also called as feature maps as mentioned previously. Abiyev and Ma'aitah (2018) have noted that the amount of kernels in the convolutional layer is similar to the depth of the feature maps that is produced. **Formula (3.1)** shows the computation of feature maps (output image) with the filters at the yth convolution layer. The bias is denoted as $B_v^{(y)}$, while the filter is referred as $K_{v,x}^{(y-1)}$. This filter linked the feature map of vth with xth in the same convolutional layer, which is (y-1). Note that the first convolutional layer is $C_v^{(o)} = X_v$ which is simplified after the $C_v^{(y-1)}$.

$$C_v^{(y)} = B_v^{(y)} + \sum K_{v,x}^{(y-1)} . C_x^{(y)} \quad \textbf{(3.1)}$$

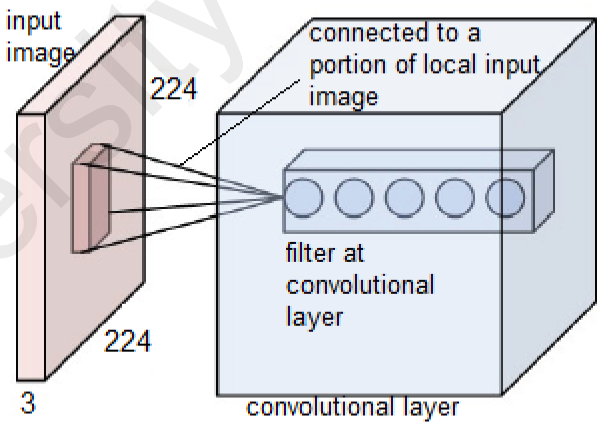

Figure 3.4: Connection of filter in the convolutional layer to the input image

Figure 3.5 shows the simplified architecture of CNN for this work. The size of kernel or filter for the convolution and subsampling steps each represents as m*m and n*n respectively. Higher validation accuracy during the training is linked to the number of feature maps produced, where increasing in the number of feature maps which representing the input images can increase the percentage of accuracy results (Alom et al., 2019).

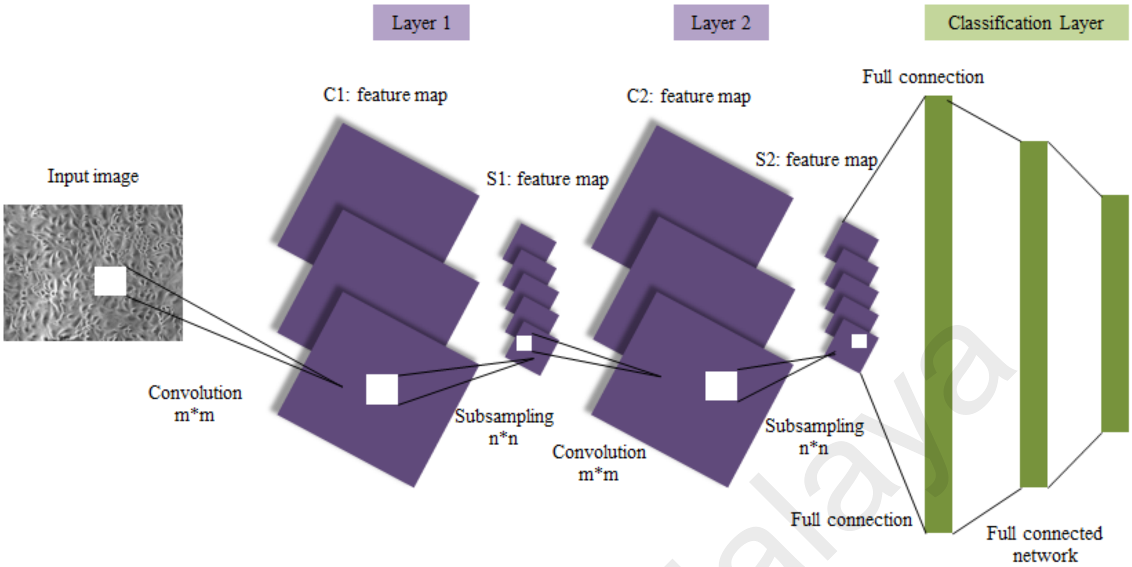

Figure 3.5: Convolutional neural network (CNN) architecture

3.2.2 Pooling Layer

There are two types of pooling layers, which are the maximum pooling and average pooling or also known as mean pooling. The function of the pooling layer is to reduce the size or dimension of the image. The progress of maximum pooling and average pooling has been presented in **Figure 3.6**, where an example of 4x4 input image gives out the image output of 2x2 image size for both pooling layer with 2x2 filter size and stride of 2. This work has applied maximum pooling with 10x10 filter size for training and the same stride. Several filter size have been tested below and above 10x10 filter size, but the results produce longer training period and with low accuracy results. Mean pooling can be calculated by averaging the numbers in the same coloured convolution areas to form into 2x2 average pooling. While, the calculation of maximum pooling is much easier by taking the largest value from each four same coloured convolution areas to form into 2x2 max pooling. Max pooling is chosen over mean pooling because of its ability to produce rapid calculation at the pooling layer (Rawat & Wang, 2017).

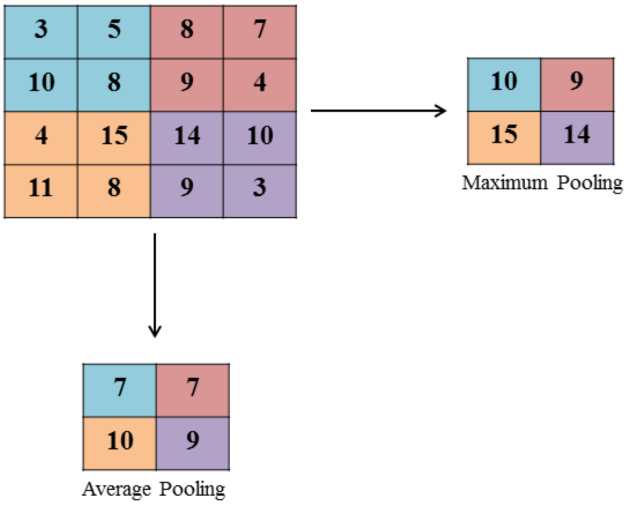

Figure 3.6: Examples of maximum pooling and average pooling

3.2.3 Activation Layer

There is another layer between convolution filter and feature maps, which is the activation function. The activation function in CNN is similar to the activation functions that have been used in neural network. Sigmoid function, tanh function and ReLU are the examples of activation function in neural network, but ReLU is the most popular activation function, especially in case of image recognition works.

3.2.4 Output Layer

The output layer is depended on the number of categories or classes that have been put into testing. This layer is referred to the classification layer where the performance can be best achieved with forward propagation.

3.3 Utilization of ResNet-50 as a pre-trained model for transfer learning

They were trained using two methods. The first method utilized the pre-trained model of residual network-50 (ResNet-50), while the second method does not utilized any pre-trained model. The computation period of both method yielded high amount of time due to the size of the image input layer, which is 224 x 224 (50, 176) and low computer memory. By applying this pre-trained model as a characteristic extractor, the invested period of training can be reduced.

The required installation toolboxes from the Matlab are the Machine Learning Toolbox, Deep Learning Toolbox, Statistics and also ResNet-50 Network. imageDatastore was used to load the image set and arrange them. This is suitable with huge data collections because its efficiency causes the images to be read first before they are loaded into memory. The imds below consisted of the images together with the assigned labels, in which the labels were extracted from the folder of data files. Then, countEachlabel was used to compute the number of picture from each group and the number can be made equally same by using minSetCount. The number of images for U2OS, MCF7, and Hepatocyte are 50, 46, and 864 accordingly before they are made balanced into 46.

ResNet-50 was installed and loaded to the command session. It is one of the pre-trained models which have trained 1000 image categories. The net.Layers(1) below refers to the size of image input, where this model requires the image input size of 224x224x3. Between the first layer and the final layer, there are sequences of convolutional layers, ReLU layers and maximum pooling layers before it enters the final layers, which is the classification layer that consisted of three layers for this case three group of image sample to be classified. However, this model has the ability to resolve 1000 method of classification problem which is defined by fc1000 below. These layers from the input to the final layers make up a total of 177 layers altogether.

The images were split into training set and test set, where 0.3 below indicates that 30% of the image samples were used for the training purpose, while the remaining images were used for the test purpose. The percentage of the image training set has been varied from 0.3 to 0.9 which give the same result of the accuracy and it is set to 0.3 to make it similar when doing the self-training. They are randomized to avoid bias while choosing the image for training.

Image pre-processing utilize the augmentedImageDatastore to keep from resaving the image into 224x224x3 format everytime the image need to be trained. According to Masters and Luschi (2018), the best range of MiniBatchSize is from 2 to 32. The MiniBatchSize below was set to 32 and it can be lowered if the memory size is not enough. Fitcecoc stands for Fit-Class Error Correcting Output Codes which is a powerful function that uses K(K-1)/2 binary support vector machine, where K refers to the number of unique class labels. The extracted features were stored in a variable named trainingLables as below.

3.4 Network Training

The image samples have been tried to be trained and recognized by a simple command instructions provided by MathWorks. Several modifications have been done to the layers and options to obtain higher accuracy or the same accuracy as in Resnet-50 model. The sample images were resized (224x224x3) by using IrfanView software before it is trained in the Matlab2019b. The size of filters as in convolution2dLayer below were resized to 10x10, where else the one in the ResNet50 used 7x7 filter size. By referring to **Figure 3.7**, it shows the filter size, number of filters, stride, maximum pool and average pool accordingly in each layer, which consisted of 177 total number of layers.

Layer Name	Input Size	Resnet-50		Output Size
Conv1	224 x 224 x 3	7 x 7, 64, stride 2		112 x 112
Conv2_x		3 x 3 max pool, stride 2 [1 x 1, 64] [3 x 3, 64] [1 x 1, 256]	X3	56 x 56

Conv3_x	[1 x 1, 128] [3 x 3, 128] [1 x 1, 512]	X4	28 x 28
Conv4_x	[1 x 1, 256] [3 X 3, 256] [1 X 1, 1024]	X6	14 x 14
Conv5_x	[1 x 1, 512] [3 x 3, 512] [1 x 1, 2048]	X3	7 x 7
	Average pool, 1000 fc, softmax		1 x 1
FLOPs	3.8×10^9		

Figure 3.7: Resnet-50 Algorithm Flow

Table 3.1 shows the comparison of parameters in CN layers for optimization. The final resort of optimization was done to the filter size which is 10x10, instead of 5x5. The augmentation of image have been tried which is good at improving the image and preventing the network from using the same image for training by randomizing them. However, these optimizations were not only unsuccessfully produce good result when running the commands, even it consumes a lot of time. The learning rate is set to 10^{x4} and after trying range of learning rate from 0.0 to 1.0, 10^{x4} has found to be the most acceptable value because the is no restrictions in determining the learning rate. While, the number of epochs were increased from 20 to 30 to increase the accuracy result. Both methods give the classification output of 3 classes, which indicates the U2OS, MCF7, and Hepatocyte with softmax layer of fc100.

Table 3.1: Parameters of CNN training

Properties of Layers	ResNet-50	Trained Network
Image input size	224x224x3	224x224x3
Number of image samples	138	138
Number of image training (%)	41 (30%)	41 (30%)
Learning rate	0.0003	0.0004
Activation function layer	ReLU	ReLU
Filter size	7	10
Number of filters	64	20
Stride	2	2
Padding size	[0,0,0,0]	[0,0,0,0]
Softmax layer	fc1000	fc1000
Classification output	3	3

CHAPTER 4: RESULTS AND DISCUSSION

960 numbers of images were used from which, 864, 46, and 50 sets of images came from Human Hepatocyte, Breast Cancer Cell Line (MCF7), and Human Bone Osteosarcoma Epithelial Cells (U2OS) respectively. They were classified in a table from which, each three categories of cells were divided into the same amount of image number, which is 46. Each category were grouped and named as U2OS, MCF7 and Hepatocyte accordingly as shown in **Figure 4.1**. They were trained using two methods. The first method utilized the pre-trained model of residual network-50 (ResNet-50), while the second method does not utilized any pre-trained model. The computation period of both method yielded high amount of time due to the size of the image input layer, which is 224 x 224 (50,176).

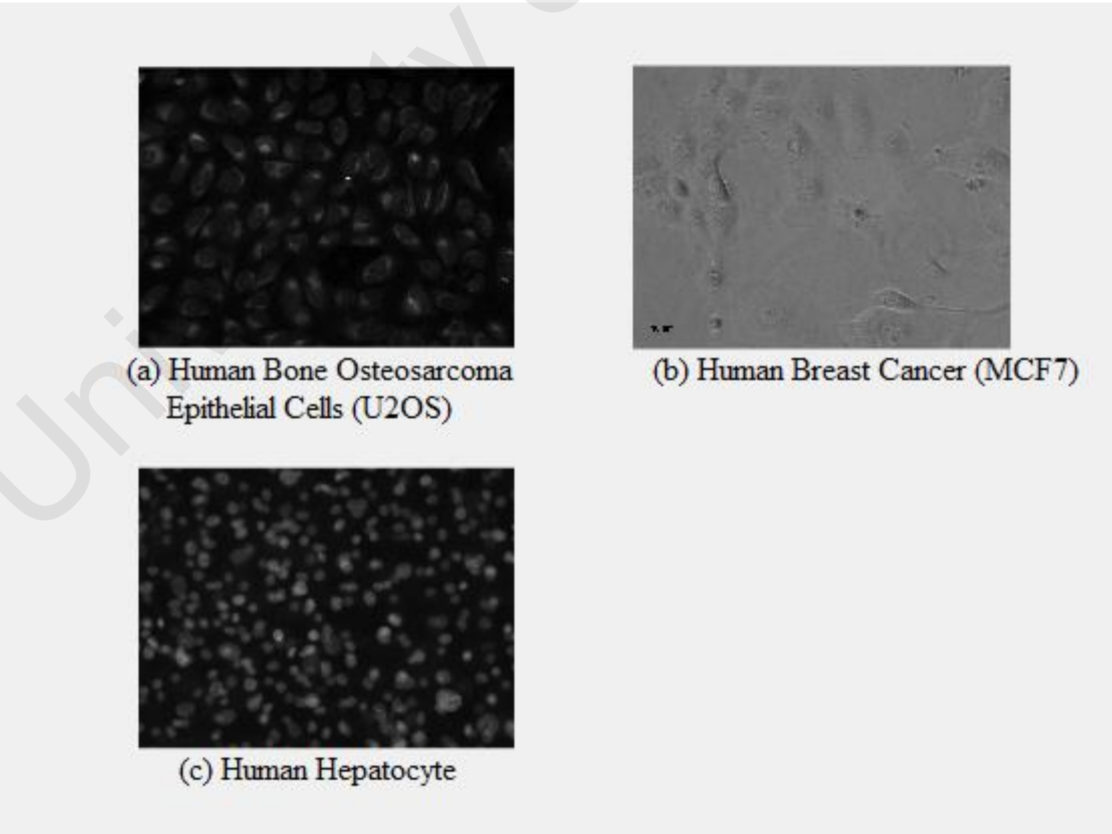

Figure 4.1: Classification of images into three categories

4.1 Pretrained Model of Resnet-50

Resnet-50 is one of the most favoured CNN architecture models which have been used to classify the images (Uchida et al., 2016). It is also one of the most current architecture model used in CNN and it is always referred to as a transfer learning due to the pre-trained model (Uchida et al., 2016). The percentage rate of errors shown by ResNet is significantly the lowest among AlexNet, VGG-16, Clarifia and GoogleNet with 3.57% and thus, it serves with higher precision (Alom et al., 2019; Ravilla, 2019). By referring to **Figure 4.2**, it shows layer by layer of the model in which, the first section or the upper part is the image input layer.

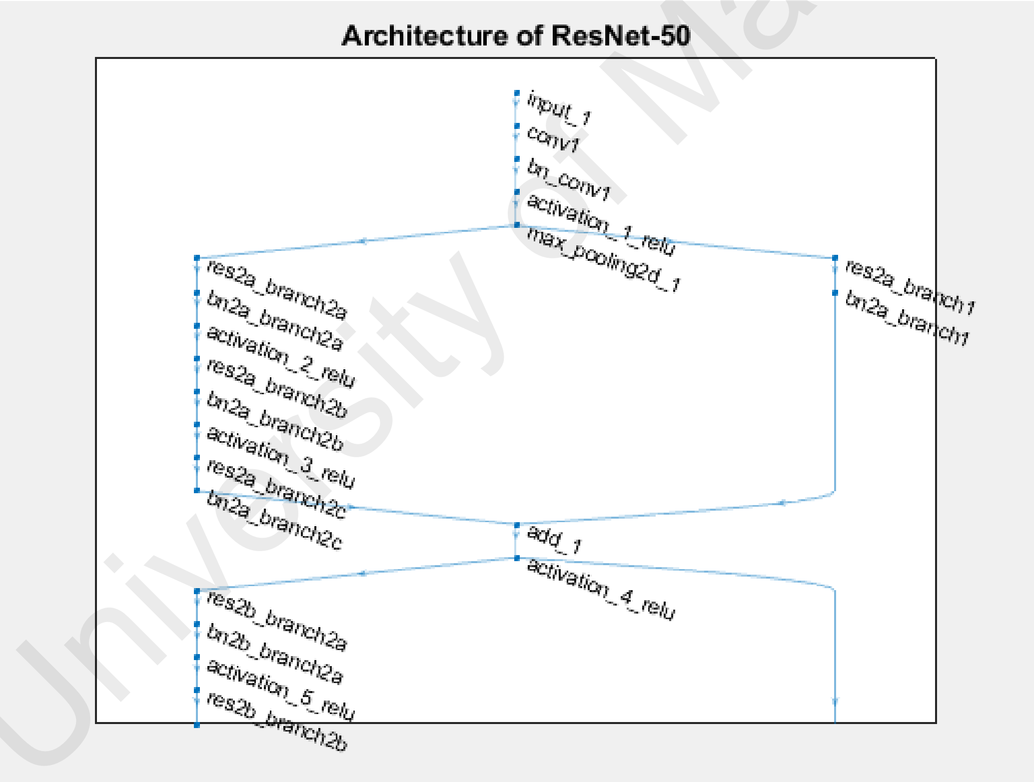

Figure 4.2: Architecture of Resnet-50

The result of the cell classification can be seen in **Figure 4.3**, where CNN network training of pre-trained model was successfully classify the sample images into its respective group or class. The result shows that the sample image of MCF7 that were taken randomly from the internet was classified into its MCF group.

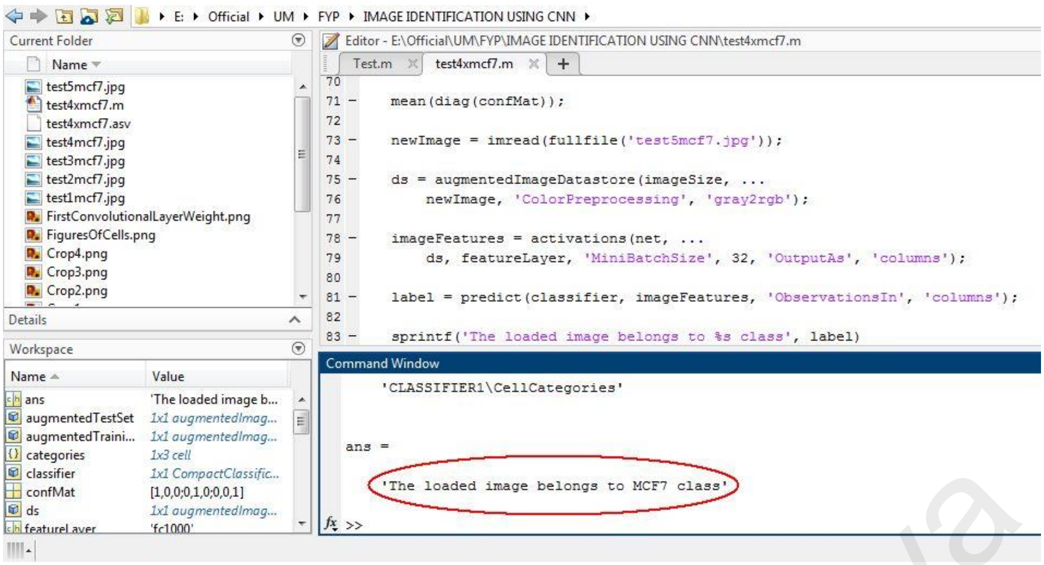

Figure 4.3: The result of classification

4.2 Network Training

The images can also be trained without using any pre-trained model but it will take longer time for the training process to configure the layers, size and number of filters, and the learning rate to get a result with high accuracy. Accuracy can be defined as the percentage of precisely predicted group over the total predicted group (Kannojia & Jaiswal, 2018). By referring to **Figure 4.4** and **Figure 4.5**, the differences of accuracy of training validation happened when there is a change in the validation frequency. The aim is to increase the percentage of validation accuracy for the training network. So, after increasing the validation from 30 to 50, the percentage increases from 41.94% to 79.57%.

The validation accuracy of the samples can be best obtained by increasing the number of image samples and reduce overfitting (NarasingaRao et al., 2018; Uchida et al., 2016). They have used 90% of the total samples for training validation. The maximum number of samples that can be provided for this training is 46 and the optimum number of training samples that can produce higher accuracy at 79.57% is 30% or 0.3 from the total number of samples. After trying to increase the number of image samples with 0.5 and 0.9 for training, it is not only highly time consuming, but the accuracy results drop to 34.78% (with the training period of 3 hours and 32 minutes) and 33.33% (with the training period of 8 hours and 57 minutes) respectively as can be seen in **Figure 4.6** and **Figure 4.7**.

The study that was carried out by Uchida et al. (2016) showed of at least 89% accuracy by using thousands of image datasets, where else this study only consumed of not more than 138 image samples. The best accuracy that can be achieved after several training is 79.57% with 30% of images from thee total image dataset. In addition, the usage of CPU memory is limited, where any addition in the number of image samples to the network training resulted to the error during training. These may be the reasons to the drawbacks and limitations that were obtained during CNN training. The usage of GPU or larger RAM of CPU may be needed in the future to harvest larger image dataset which is to overcome the effect of low validation accuracy during network training.

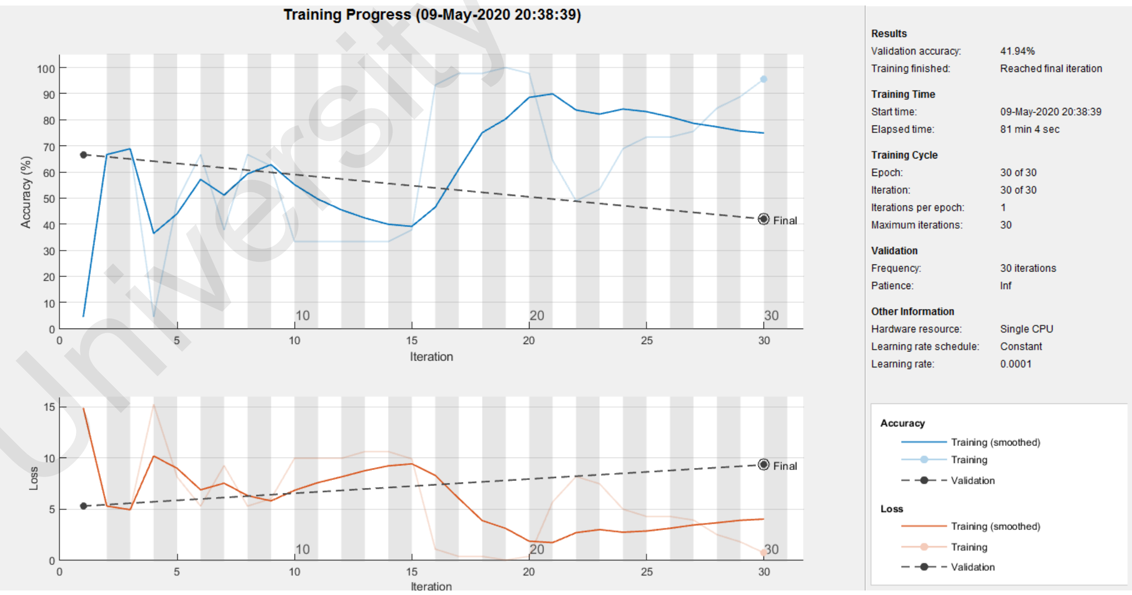

Figure 4.4: Training validation for 0.3 training samples with 41.94% accuracy

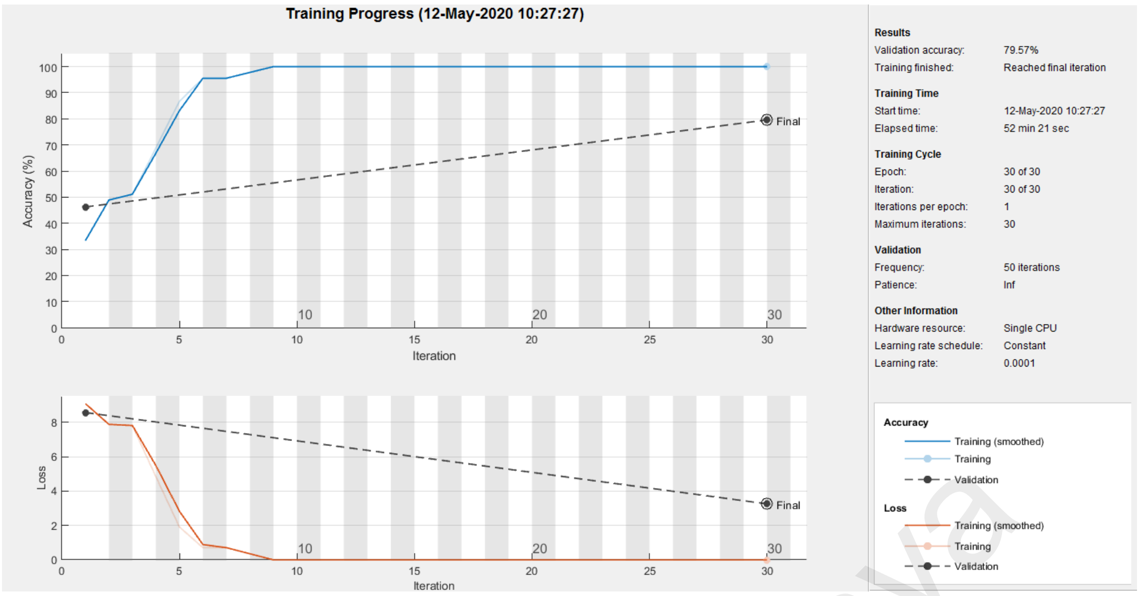

Figure 4.5: Training validation for 0.3 training samples with 79.57% accuracy

Figure 4.6: Training validation for 0.5 training samples with 34.78% accuracy

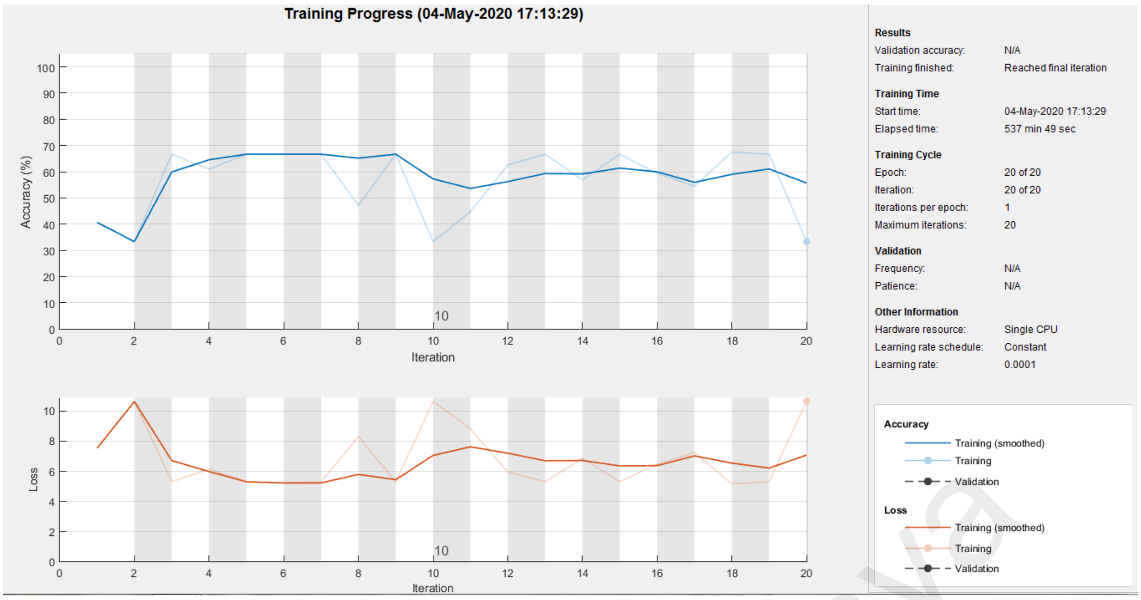

Figure 4.7: Training validation for 0.9 training samples with 33.33% accuracy

The application of ReLU activation function gives advantage to the training speed than the tanh and sigmoid activation function (Gu et al., 2017). It is notably used in many works due to its efficiency and accuracy in training a particular network without using a pre-trained model (Gu et al., 2017). Gu et al. (2017) even have suggested the improvised method of using ReLU by replacing it with Leaky ReLU (LReLU), Randomized ReLU (RReLU) or Parametric ReLU (PReLU). The normal ReLU, G(x) = max(0,x) may drag the training period and cause the continuous inactivity of the units from the start because the weights were not altered by the gradient-based optimization. These limitations were caused by the persistent zero gradients that existed in the units. Where else, LReLU incorporates parameter unit from 0 to 1 and a small additional of parameter numbers in PReLU and RReLU which greatly helps to enhance the accuracy performance and widely used in solving the image recognition works.

Increasing the number of training sample will also increase the period of training as shown in **Table 4.1** which corresponds to **Figure 4.8**. The pre-trained Resnet-50 does not consume longer time as compared to the self-training without the pre-trained model.

In addition, even the training period is significantly increase up to 32269 seconds, the accuracy of sample training does not improve and it keeps decreasing as the number of training sample increased. This may be due to overfitting where the computation progress and the memory of the computer are not bearable. According to Gour et al. (2019), this training needs massive computational strength for the whole process to complete. Image augmentation can improve overfitting and have been tried in the training, but still, the training period too longer while the accuracy keep decreasing (Gour et al., 2019; NarasingaRao et al., 2018).

Table 4.1: Period of training

Number of training sample (%)	Time (seconds)		Accuracy of training (%)	
	Resnet-50	No training model	Resnet-50	No training model
30	2700	3141	100.00	79.57
50	3120	12779	100.00	34.78
90	4200	32269	100.00	33.33

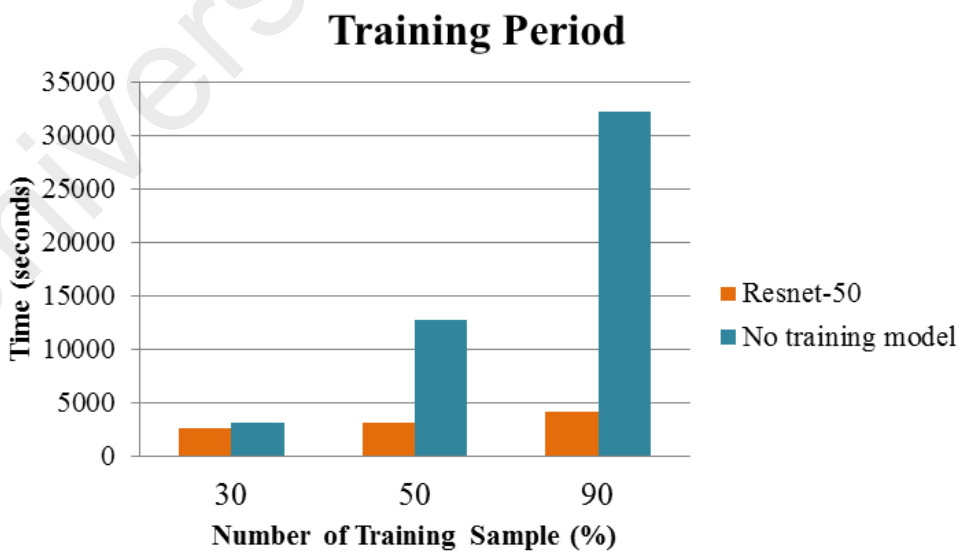

Figure 4.8: Evaluation of training period over the sampling data

Wu and Gu (2015) also stated that the two problem that will be faced during training the CNN network are the computation period and overfitting. They have suggested to use GPU processor to enhance the power and speed of the computation. Overfitting is an event where the training set greatly matches with the model until it becomes hard to derive new images in the training set (NarasingaRao et al., 2018). Thus, it tends to identify particular images in the training set rather than identify its common pattern, which leads to a greater accuracy in the training set than the accuracy in the validation set (NarasingaRao et al., 2018).

CHAPTER 5: CONCLUSION

Based on the results of network training for both methods, the classifiers have been tested by using available images in the internet and also personal storage. The first objective is achieved by showing the successfully classified tested images into their own categories. However, the accuracy of the two classifiers are not the same, where the CNN for pre-trained Resnet-50 is more reliable with 100% training accuracy as compared to self-training results. The training work can be improved by using GPU computation, provided with more RAM memory to smoothen the whole process. When the computer is capable to do so, the number of sample images to be trained can be increased with no difficulties in optimizing the parameters in the network layers to achieve great accuracy.

In conclusion, Resnet-50 serves as an effective tool to help in building up the CNN network. There were many successful products of CNN were born in the industries, such as biometric, face recognition tagging in facebook, keyword search in the internet and more. The main limitations that were faced in this work are the shortage of microcellular images and the computer memory or processor used to compute the function. Future works will need to improve the highlighted cases to explore more available parameters for this network. Increasing image samples with improved computer's ability, such as using GPU computed version will help to achieve the objectives. Besides, instead of using only ResNet-50 model, there are more pre-trained model can be added to the study to improve overfitting, such as VGG-16, GoogleNet, and AlexNet to compare their performance.

REFERENCES

Abiyev, R. H., & Ma'aitah, M. K. S. (2018). Deep convolutional neural networks for chest diseases detection. *Journal of healthcare engineering*, *2018*.

Akogo, D. A., & Palmer, X. L. (2019). End-to-end learning via a convolutional neural network for cancer cell line classification. *Journal of Industry-university Collaboration*, *1*(1).

Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., & Navab, N. (2016). AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Transactions on Medical Imaging*, *35*(5), 1313-1321.

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Essen, B. C. V., Awwal, A. A. S., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics*, *8*(3), 292.

Asri, H., Mousannif, H., Moatassime, H. A., & Noel, T. (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Computer Science*, *83*(2016), 1064-1069.

Ayas, S., & Ekinci, M. (2019). Microscopic image super resolution using deep convolution neural networks. *Multimedia Tools and Applications*, 1-19.

Bekolay, T. (2011). *Learning in large-scale spiking neural networks* (Master's thesis, University of Waterloo).

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York, NY: Springer.

Chen, D., Liu, S., Kingsbury, P., Sohn, S., Storlie, C. B., Habermann, E. B., Naessens, J. M., Larson, D. W., & Liu, H. (2019). Deep learning and alternative learning strategies for retrospective real-world clinical data. *NPJ digital medicine*, *2*(1), 1-5.

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected CRFs. *arXiv preprint arXiv: 1412.7062*

Comsa, S., Cimpean, A. M., & Raica, M. (2015). The story of MCF-7 breast cancer cell line: 40 years of experience iin research. *Anticancer Research*, *35*, 3147-3154.

Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2019). A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. *Archives of Computational Methods in Engineering*, 1-22.

De Ridder, D., de Ridder, J., & Reinders, M. J. (2013). Pattern recognition in bioinformatics. *Brief. Bioinform., 14*, 633-647.

Elsayed, A. A., & Yousef, W. A. (2019). Matlab vs. openCV: A comparative study of different machine learning algorithm. *arXiv preprint arXiv:1905.01213*(2019).

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist- level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115-118.

Fu, W., & Menzies, T. (2017, August). Easy over hard: A case study on deep learning. In *Proceedings of the 2017 11th joint meeting on foundations of software engineering* (pp. 49-60).

Gogul, I., & Kumar, V. S. (2017, March). Flower species recognition system using convolution neural networks and transfer learning. In *2017, Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)* (pp. 1-6). IEEE.

Gour, M., Jain, S., & Sunil Kumar, T. (2020). Residual learning based CNN for breast cancer histopathological image classification. *International Journal of Imaging Systems and Technology*.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*, 354-377.

Han, X, & Li, Y. (2015). The application of convolution neural networks in handwritten numeral recognition. *International Journal of Database Theory and Appplication, 8*(3), 367-376.

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). *The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York*, NY: Springer.

Houcque, D. (2005). Introduction to Matlab for engineering students. *Northwestern University*, 1-64.

Hu, W., Huang, Y., Wei, L., Zhang, F., & Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors, 2015*.

Image Sets. Annotated biological image sets for testing and validating. Broad Bioimage Benchmark Collection. *Broad Institute*, obtained from (https://data.broadinstitute.org/bbbc/image_sets.html).

Indolia, S., Goswami, A. K., Mishra, A. P., & Asopa, P. (2018). Conceptual understanding of convolutional neural network – A deep learning approach. *Procedia Computer Science, 132*, 679-688.

Joss, L., & Muller, E. A. (2019). Machine learning for fluid property correlations: classroom examples with MATLAB. *Journal of Chemical Education*, *96*(4), 697-703.

Kannojia, S. P., & Jaiswal, G. (2018). Effects of Varying Resolution on Performance of CNN based Image Classification: An Experimental Study. *International Journal of Computer Sciences and Engineering*, *6*(9), 451-56p.

Khan, R. U., Zhang, X., Kumar, R., & Aboagye, E. O. (2018, March). Evaluating the performance of resnet model based on image recognition. In *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence* (pp. 86-90).

Kourou, K., Exarchos, T. P., Exarcos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal, 13*(2015), 8-17.

Kusumoto, D., & Yuasa, S. (2019). The application of convolutional neural network to sem cell biology. *Inflammation and Regeneration, 39*(1), 14.

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Comput, 1*(4), 541-551.

Liang, Z. (2017). Automatic Image Recognition of Rapid Malaria Emergency Diagnosis: A Deep Neural Network Approach.

Masters, D., & Luschi, C. (2018). Revisiting small batch training for deep neural networks. *arXiv preprint arXiv:1804.07612*.

MCF7 (ATCC® HTB-22™). www.atcc.org

Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer diagnosis using deep learning: A bibliographic review. *Cancers, 11*(1235).

NarasingaRao, M. R., Prasad, V. V., Teja, P. S., Zindavali, M., & Reddy, O. P. (2018). A survey on prevention of overfitting in convolution neural networks using machine learning techniques. *International Journal of Engineering and Technology (UAE), 7*(2.32), 177-180.

Oei, R. W., Hou, G., Liu, F., Zhong, J., Zhang, J., An, Z., Xu, L., & Yang, Y. (2019). Convolutional neural network for cell classification using microscope images of intracellular actin networks. *PLoS one*, *14*(3).

Radhakrishnan, A., Damodaran, K., Soylemezoglu, A. C., Uhler, C., & Shivashankar, G. V. (2017). Machine learning for nuclear mechano-morphometric biomarkers in cancer diagnostic. *Scientific Reports*, *7*(1), 1-13.

Ravilla, A. K. D. (2019). CNN based image classification model. *International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8*(1S), 2278-3075.

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, *29*(9), 2352-2449.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779-788).

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems* (pp. 91-99).

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive Modeling, 5*(3), 1.

Shah, P., Kendall, F., Khozin, S., Goosen, R., Hu, J., Laramie, J., Ringel, M., & Schork, N. (2019). Artificial intelligence and machine learning in clinical development: a translational perspective. *npj Digital Medicine, 2*(69).

Sommer, C., & Gerlich, D. W. (2013). Machine learning in cell biology – teaching computers to recognize phenotypes. *Journal of Cell Science, 126*(24), 5529-5539.

Talari, A. C. S., Rehman, S., & Rehman, I. U. (2019). Advancing cancer diagnostics with artificial intelligence and spectroscopy: identifying chemical changes associated with breast cancer. *Expert Review of Molecular Diagnostics, 19*(10), 929-940.

Tarca, A. L., Carey, V. J., Chen, X. W., Romero, R., & Draghici, S. (2007). Machine learning and its applications to biology. *PLOS Comput. Biol., 3*, e116.

Tukimin, S. N., Karman. S. B., Ahmad, M. Y., & Zaman, W. S. W. K. (2019). Polarized light-based cancer cell detection techniques: a review. *IEEE Sensors Journal, 19*(20), 9010-9025.

Uchida, S., Ide, S., Iwana, B. K., & Zhu, A. (2016, October). A further step to perfect accuracy by training cnn with larger data. In *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (pp. 405-410). IEEE.

van Grinsven, M. J. J. P., van Ginneken, B., Hoyng, C. B., Theelen, T., & Sanchez, C. I. (2016). Fast convolutional neural network training using selective data sampling application to hemorrhage detection in color fundus images. *IEEE Transactions on Medical Imaging, 35*(5), 1273-1284.

Vedaldi, A., & Lenc, K. (2015, October). Matconvnet: Convolutional neural networks for matlab. In *Proceedings of the 23rd ACM international conference on Multimedia* (pp. 689-692).

Wu, H., & Gu, X. (2015). Towards dropout training for convolutional neural networks. *Neural Networks, 71*, 1-10.

Yang, L., Zhang, Y., Chen, J., Zhang, S., & Chen, D. Z. (2017). Suggestive annotation: A deep active learning framework for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-assisted Intervention* (pp. 399-407). Springer, Cham.

Yue, W., Wang, Z., Chen, H., Payne, A., & Liu, X. (2018). Machine learning with applications in breast cancer diagnosis and prognosis. *Designs, 2*(13).

Zhou, X. (2018, April). Understanding the convolutional neural networks with gradient descent and backpropagation. In *Journal of Physics: Conference Series* (Vol. 1004, No.1, p. 012028). IOP Publishing.