T.C.

ALTINBAS UNIVERSITY

Institute of Graduate Studies

Information Technologies

USING OF CONVOLUTIONAL NEURAL NETWORK FOR GRAYSCALE IMAGE COLORIZATION

Esra Mohamed ISSA

Master of Science

Supervisor

Asst. Prof. Dr. Oğuz ATA

Istanbul, 2021

USING OF CONVOLUTIONAL NEURAL NETWORK FOR GRAYSCALE IMAGE COLORIZATION

By

ESRA MOHAMED ISSA

Information Technology

Submitted to the Institute of Graduate Studies in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2021

The thesis titled “Using of convolutional neural network for grayscale image colorization” prepared and presented by “Esra ISSA ” was accepted as a Master of Science Thesis in Information Technology Department.

Asst. Prof. Oğuz ATA

Supervisor

Thesis Defense Jury Members:

Asst. Prof. Dr.Oğuz ATA	School of Engineering and Natural Sciences Altinbas University	_____
Asst. Prof. Dr.Cağdas ATILLA	School of Engineering and Natural Sciences Altinbas University	_____
Asst. Prof. Dr.Aytuğ BOYACI	School of Computer Engineering Milli Savunma University	_____

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

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

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

Esra ISSA

ACKNOWLEDGEMENTS

I would like to thanks to my great supervisor Assist Professor Dr. Oğuz ATA to helping me in preparing the thesis. This thesis is wholeheartedly dedicated to me my father, my ideal mom, may God prolong her life to be of my help to my dear family members, my husband how incarce me and give hole support, and my children how

ABSTRACT

USING OF CONVOLUTIONAL NEURAL NETWORK FOR GRAYSCALE IMAGE COLORIZATION

Esra ISSA

Master of Science/ Information Technologies Altınbaş University

Supervisor: Assist Professor Dr. Oğuz ATA

March/2021

Pages:45

The advancement of data technology had motivated more facilitation of human life including security and pattern recognition. Large information can be obtained from the image after proper feature extraction. Thus, the field of image processing has gained extended attention especially after incorporating artificial intelligence (AI). One of the vital problems that image processing is looking after is dealing with ancient images or in other words, the grayscale images which were captured before the introduction of RBG technology. Analyzing grayscale images for more curtail knowledge extraction is possible through colorizing technology where the grayscale image is converted into an RBG image. In this thesis, the development of an accurate image colorization approach is proposed using a convolutional neural network (CNN). Optimization of colorization performance is conducted through tuning of the CNN model. Results have shown that the proposed classifier have been scored of 80.91 Percent accuracy.

Keywords: CNN, Colorization, RBG, Bands, Normalization, Pixel, Neural network.

Table of Contents

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

1. INTRODUCTION

1.1 OVERVIEW

Extraction of information from an image, converting it into other useful formats, or recognizing of image objects are folded under image processing which is part of computer engineering that incorporated plenty of applications related to pattern recognition, personal verification, and security systems. Image colorization is gained wide attention in current days as a path of image processing that keen one converting the grayscale image into a colored one [1]. The process of image colorization is a vital technological operation since it deals with the regeneration (reproduction) of information in an acceptable error rate [2]. A colored image reveals more information than the normal (grayscale) image and hence the idea of colorization is raised for fulfilling such needs. This application has been widely utilized in the coloring of historical images and old movies for appealing more visuality. A colored image consisting of three main channels called red, blue, and gray. Such image format is called tri-channel format or RBG which is representable by a three-dimensional matrix. Coloring of an image involves generating more two scales (channels) namely red and blue channels. The combining of those three channels is resulting from the final color image. The gray-scale image is one-dimensional matric including the gray channel information only with no details about the other blue and red channels. Colorization technology is widely used in ancient studies for coloring old paints situated on the cave walls and others that sketched on leaves [3]. More recently, image colorization is performed using the deep learning approaches where a large number of images in colored form and uncolored form are fed into a pattern recognition paradigm where supervised learning can take place. The algorithm can predict the other red and blue scales on the input grayscale image resulting colored image as accurate as the training phase is. This thesis demonstrates the process of image colorization using a big dataset and deep learning methods. This chapter is providing the motivations with the problem statement of this thesis as well as research objectives and thesis structure.

1.2 PROBLEM STATEMENT

In today's life, images have become vital for various applications such as pattern recognition, security system, face recognition, remote sensing, etc. Due to the popularity of image processing applications, new fields begin to establish using the technology of image processing such as image colorization. The main idea behind image processing is information preserved by each pixel. However, such information is considered as a cornerstone for all applications folded under the image processing e.g. (face recognition). The problem is raised when no enough pixel information in some/specific images making it unfit for processing under all mentioned applications. In other words, the value of information preserved by any image is a function of what each pixel preserved. This problem is manifested in grayscale images where only information about the gray channel is available and no chronic information exists. Due to that, grayscale images are unfit for image processing since the pixels information of the image is represented by one value only (grayscale channel) and no strong correlation can be ensured between any two images in grayscale (assumed for the same event/object). Considering that noise interference with the image could manipulate the information of every channel and the noise impact is different in each channel. The need for multiple channels while processing the image digitally is a must and is the only way to correlated two images with each other in applications like face recognition or pattern recognition. Noise impacting in the signal channel such as grayscale image may corrupt all information preserved in the pixel while the noise impact is lesser while dealing with a multi-channel image. The current state of the art is using the machine learning approaches for image colorization has left the colorization accuracy in dispute due to the uncertainty of the machine learning performance in colorization of image.

1.3 LITERATURE SURVEY

The image processing technology has used widely in any applications such as pattern recognition and biometrical verification, the colored image can be only used for such application as was discussed in above sections, below demonstration involves the examples of colored image processing technology which have been performed in the literature. From the other hand, the need for robust classifier is another corner stone for image processing paradigms, this section demonstrates various applications and types of machine learning and deep learning approaches.

At [1], author stated that colored image of palm can be used for palm recognition which can be used as key feature of robust security enforcement techniques. Palm recognition can yield permanent features from the palm image, such features are constant and resistive to the time effects and age effects. Palm image is recognized in this study using the essential lines and the wrinkles persist on the palm surface. However, the wrinkles and, delta points are one of the richest sources of recognizable features of palm as in Figure 1.1.

In this study, palm recognition was deployed for security authentication in airports and other sensitive places. The designed system was used for inquiring the palm image using scanner (same like this used for paper scanning). One of the steps that used for efficient data inquiry is differentiating between the palm object and the background object. For this, scanner laid is left opened so that the background object will appear black totally.

At [2], biometrical features are being used to enhance the security methods where those features cannot be cheated or penetrated as easily as passwords are; mentioned by the authors. However, biometrical features are unique for every person. One of the biometrical features is eye print, this have been used widely in personal identification systems. [3]

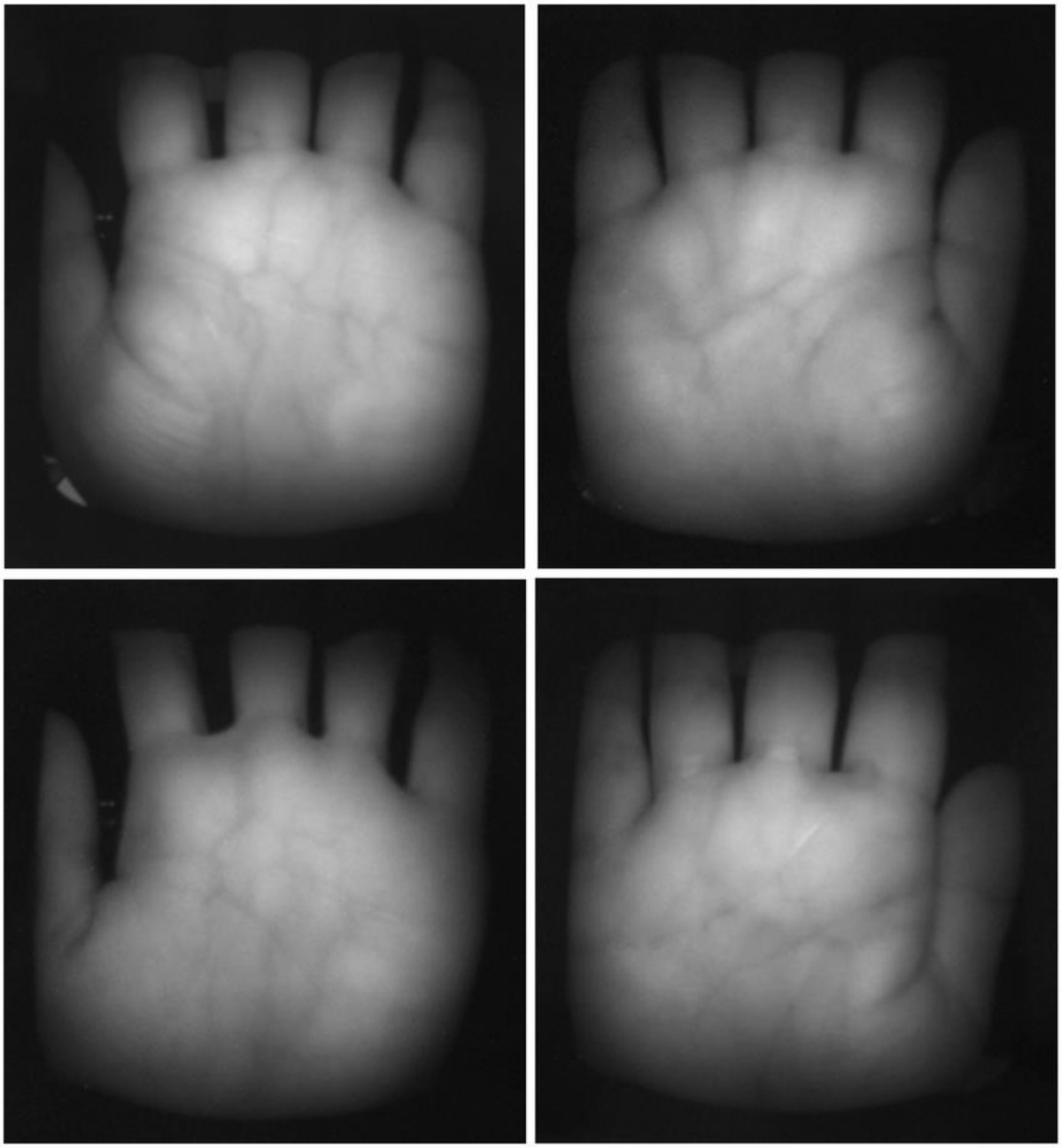

Figure 0.1: Image colorization (palm) for features extraction.

The features from eye can be detected and then applied into object classifiers for implementing of eye recognition system. Those system are being used in sensitive security missions where passwords and other features stands not enough. Author mentioned that cost of adaptation this technology is high more likely, eye image inquiry is to be taken using special camera which add more cost to the project. From the other hand, eye can be impacted

by the age problems where shape of eye ball can be varying according to the eye health and person age.

At [4], another approach of biometrical recognition is made over this study by utilizing finger print. Such biometrical features from fingers alike thumb is used for personal identification. Features from fingers are being extracted from the thumb image, this image is taken from the finger print scanner. [5] [6] [7] Dataset of two hundred images are gathered from the two hundred candidates. Each image represents a finger print of the right handed thumb with green color background. Features alike circles and curved lines are recognize from each image using the pixel resolution values. Artificial neural network is used for classification of the image data for detection the character using the thumb features.

At [8] [9] [10], occluded image recognition is another approach made in this study in the interest of personal identification. In most of the times, hand image is used for biometrical verification which proved a noticed performance in the security applications. One of the challenges raised in case of occluded hand palm. The idea of recognizing the occluded image is based on some image features alike fingers width and the length and the minimum and maximum points (local maxima and local minima) of the hand. However, this approach can be applied when the biometrical features mentioned above are required for personal identification.

This approach is not applicable if the recognition process is based on other features alike palm veins and lines features. [11] [12]. This study utilized the fact that the hand can be recognized using the geometrical features of the hand itself if the testing palm image is occluded with gloves and the actual training data is not occluded with any object. So, in order to tackle this occlusion, all the dimensions are scaled down by some factor to recover the previous dimensions of the not occluded image.

At [13] [14] [15], this study aimed to simulated the hand signs (Hamburg sings) that used for communication between the mute and def people. However, the approach is developed to detect the hand movement by a computer based program and translated them into virtual (avatar) object that represents the hand signs language. Authors mentioned that hand movement is being detected using camera device from KINECT company which designed especially for hand movement detection.

In order to provide the Matlab with the required hand movement image, another program is used provide a direct connectivity between the KINECT device and Matlab. This device can provide a real time images decomposed into the three bands (green, red and blue). Matlab program is made to detect the edges of each proved image and match the same with reference Hamburg image in order to return the appropriate avatar object [16] [17].

At [18] [19], most of the advanced personal identification systems are based on human body organs biometrical recognition. Biometrical recognition involves face, hand palm, finger print and eye print recognition. Researchers aiming to enhance the performance of those recognition systems by defining new methods of recognitions. Problem raised when the data taken for face are relied for personal identification. This data is being pre-processed and used to train machine learning paradigm.

Face features data reliability cannot be garneted as permanent features to recognize the face. However, face is one of the human parts that impacted directly by the age effects. By growing in age, face features changes and this is obvious as face during the childhood is changing while entering the youngster age and similarly while entering old age [20].

Face will be occluded with wrinkles and rashes due to growing in age so the data of the face recognition system is to be upgraded periodically depending on the person age. Considering a big face recognition system involving a large number of candidates of different age groups; the updating of such data is not an easy task. The updating of face image data will be required both preprocessing and training stage updating which will increase the complexity and cost of the system [21] [22].

At [23] [24] [25], another approach is made to inquire the hand image while hand is moving. The detection of hand object while hand is not is stable mode (capturing the hand image while person speaking and using hand action) is remained challengeable. One of the proposed approaches is using gloves that able to detect the hand while movement using infrared sensors. Author mentioned that, in order to capture the moving hand image, gloves can be wore in the hand, this glove are contained of infrared sources placed in between the fingers and the reflection of infrared light is being sensed using special camera placed at the finger top. One algorithm called as TTRPP is being used to capture the hand contour. A computer program is made using this algorithm (TTRPP) for detection of hand contour in real-time, the program is made using JAVA language [26] [27].

At [28] [29] [30], this approach has stated that palm features are one of the best and reliable features for personal identification systems as compared to those of face and fingers impression. The reason is that finger dimensions can be changes and other factors such as wounds or finger wrinkles can occlude the features of fingers. Similar concept is applicable with face recognition, such features are highly affected by the age and hence it can not be treated as reliable features for biometrical recognition. However, palm is consisting of features that resistive for age impact and hence it has a good chance to be reliable of personal identification systems.

At [31] [32], palm image is being processed using four different masks where each mask will produce a different extension of the palm image. However, each extension provided by the mask (filter) will be used to generate particular features. Features in each extension of those four will be generated by taking the standard deviation and variance of each extension. Those features are being used to recognize the palm.

Author is also mentioned that before all these must be robust procedure to crop the region of interest from the palm image. It was also stated in this research that region of interest should be captured with enough accuracy so that the features extraction will be more reliable. Region of interest can be cropped using many methods alike radius method and centroid method. The main condition behind successful region of interest cropping is that: region of interest should be covering the same region (part) of the hand in all the dataset images and that is practically difficult [33] [34].

At [35] [36], author mentioned that palm can be recognized using several parameters related to the dimensions of palm and its geometry. Hand palm can be recognized with those features after performing of two starting stages namely image capturing and preprocessing. Author emphasized that image capturing must be performed with adequate device that ensures the required resolution as well as the uniformity amongst all the captured images. The images should preserve same content of the palm more likely, same palm orientation which make the palms appearing in similar snaps are required.

Authors used the following features to recognize every palm: widths of the all five fingers in the hand (which yields five features) and length of each finger in the hand (which yield five features as well). eventually ten statistical features are detected from each palm which can be used for recognition process.

At [37] , a new approach is proposed to recognize the palm using the features of palm blood veins. Authors stated that the recognizable features from the blood veins are robust than those features of face. The reason behind that is blood veins under the skin of the palm can yield a feature that very resistive to the time and age (not impacted by the age of human) so, it is considered as permanent features.

This study is applied on the old images that included palm print on the tree leaf and kept inside those old caves in the area of Thailand. This technology is applied using a hardware (microchip) that involves the recognition program. Authors mentioned that deep learning algorithm con outperform for palm recognition after good features extraction. However, performance metrics alike, training time, level of the recognition accuracy are evaluated and shown good results.

At [38], one of the proposed methods to extract the features from the palm image especially those images printed on ancient manuscripts e.g. tree leaf and caves; this method relays on the knowledge obtained from those images using spectrum domain analysis. However, spectrum domain information can be yielded from the pixel domain by using Fourier analysis. The frequency domain information is considered as good source of features from the faded images or old historical images which always can tells more about the objects associated in this image more than the spatial domain.

Soon after features extraction, authors stated that machine learning approaches such as Principle Component Analysis is used for palm image classification. Authors stated that the method in their proposal are resulted image classification with better performance than other available algorithms.

At [39], authors of this study emphasized that palm image is the richest available biometrical image that may provide features that not impacted by the natural effects such as hair growing (palm is not witnessing any growing of hair unlike the face area) as well as not affected by the skin color fade by the sun light. Hence, palm is one of the finest biometrical features sources that can participate the personal verification systems.

Within the palm recognition context itself, many technologies are used to extract the palm features. One of those technologies are the blood veins features. Author mentioned that blood veins features is mainly depending on the camera type and normal camera cannot be used for enquiring such images includes a veins feature. Veins features are extractable using infrared rays reflection from the blood cells (hemoglobin). So, first thing is to provide a camera which can produced a digital image with infrared band.

At [40], another study is presented here which proposed using different kind of features combinations as multi-dimensional features vector. This vector involves features alike widths of all palm fingers and length pf all palm fingers as well as the skin color density (shad level). Those features are taken from each image and hence features matrix is formulated. The features matrix will be used for training a deep learning paradigm such as convolutional neural networks to predict the character base on the palm features.

Convolutional neural network is trained using seventy five percent of the features available in features matrix. Performance is being measured using the percentage of correct decisions during the neural network testing phase.

1.4 RESEARCH OBJECTIVES

From the above problem statement, the need for a strong colorization approach is crucial for the success of image processing and for allowing the primitive images to be reused in the current technology of image processing. Potential information from historical images and other basic images can be extracted with the help of the image colorization technique. Thus, in this thesis, the following objectives are to be fulfilled:

a> To establish a robust image colorization paradigm that capable to intake the grayscale images and predicting their missing channels (scales) e.g. (blue and red).

b> To implement the colonization system so that images decolorization can be taking place for preparing the gray-scale of each image and then to colorize the said image. However, the original colored image (before decolorization) can be used for verification by testing the performance of the colorization system.

c> To develop an accurate convolutional neural network (CNN) based image colorization technique. The performance of the said CNN to be optimized by optimizing each layer of the classifier.

d> To examine the performance of the proposed classifier using accuracy measures as well as the performance metrics measures such as mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), and time.

1.5 THESIS ORGANIZATION

This thesis is consisting of four technical sections named as chapters, it is commenced by chapter one, the Introduction and terminated by chapter four, the Conclusion and Discussions. The chapter including chapter one and chapter four and the in-between chapters can be illustrated briefly as hereinafter.

Chapter one: The Introduction, which demonstrates the preface of the image colorization approach, its importance and the problem statement of the thesis. The motivations of this work are also included in the problem formulation section. Thesis objectives and organization (structure) are the last two sections of this chapter.

Chapter two: The Methodology, which involves the technical and theoretical approaches that been used for performing this work and detailing the results obtained at the end of the work.

Chapter three: Results and discussions involves the results that represents the performance of the classifiers using the proposed performance metrics.

Chapter four: The Conclusion and Discussions, demonstrates the results discussions and details the points concluded in this work according to the achieved results.

The final sections of this thesis include, the references which were used for implementing this thesis.

2 METHODOLOGY

2.1 PREFACE

Image colorization is a process of allotment of color for each pixel in the crystal image with efforts of artifact minimization in the resultant image. In order to colorize an image, deep knowledge is required of the nature of that image from the contents perspectives more likely, what could the image contains, and also the knowledge of common objects colors. While considering objects' nature, the colorization model must be trained about various objects' color more likely if the color is blue if the object was water and if color is green then the object can be a grace of the garden. Similarly, the colorization model is to be trained to perform the colorization process. The researchers established in this regard are attempting to perform the colorization with a level of accuracy by tackling the artifacts in the resultant image.

Coming to the point of image colorization system modeling, two kinds of approaches are proposed in the published studies namely, the classical approach and the deep learning approach. In a classical approach, one needs to tell the system about the colors of some regions and hence system may propagate this color into the other areas and trying to performing coloring of the picture area-wise. This approach is not as accurate as required and hence, a new approach named a deep learning-based colorization model is implemented.

However, this chapter is detailing the process of image colorization using the deep learning paradigm. The performance of this model is examined using various performance metrics such as mean square error, mean absolute error, root means a square error, time, accuracy, and pixel signal to noise ratio.

The model is designed using the conventional neural network with a large number of hidden layers to enable the learning of each pixel value in the colored image and hence to apply the color by predicting the new pixel value in the crystal image.

2.2 IMAGE DECOMPOSITION

Digital images are populated in the science and engineering domains due to their efficiency in prescribing the natural objects. However, the term digital is called to those images that allotting a value to each dot. Images as a result are contained a large number of dots with different intensities (value) which produce the final tent of the image. Giving it the final look that can be sensed by the naked eye.

With the development of digital systems, new cameras are designed to intake color images instead of old fashion crystallizing images. The digital images are now providing clear colored images with multi-channel.

Images of interest are depending on the cameras can be classified into Red, Blue, Green, and under the infrared channel. A binary image is in turn represents the image with only black and white colors and is representing the pixels with binary value more likely zeros and ones. Figure 3.1 demonstrates the four bands that are aforementioned above along with the united colors image that is termed as a colored image that includes all channels above.

It is a noteworthy statement of the under infrared images channel that the same image can be used with special applications such as remote sensing in satellite image processing which can give the other information that helps to recognize the geological objects from satellite images.

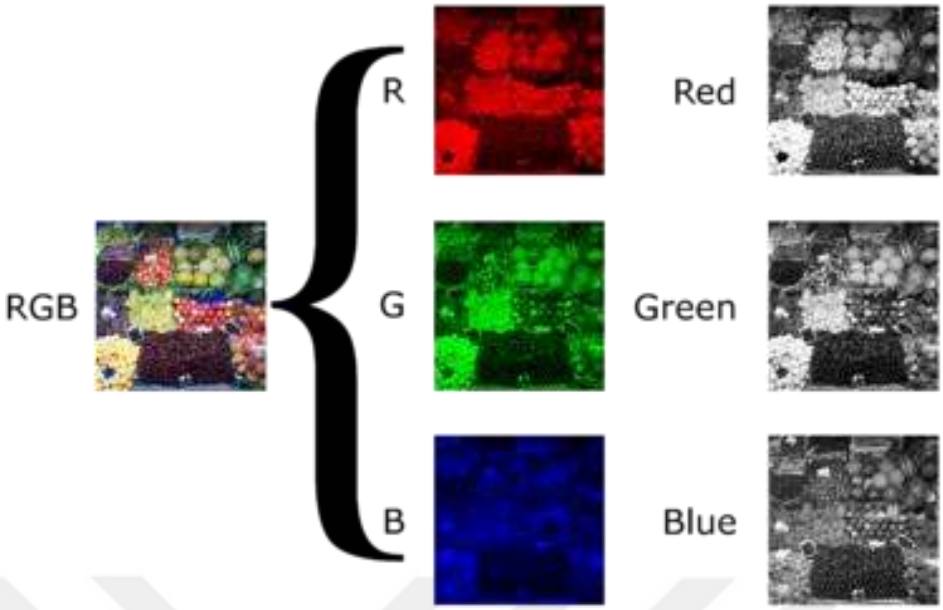

Figure 3.1: Colored and RGB channels of image.

Generally, the main point of image processing is understanding what the images equal in mathematical scale are. The image can always be represented by a combination of their (three-four) channels information that forms the final image colors. The color image can be represented with a three-dimensional matrix with dimensions of x, y, and z.

The very standard image representation is of the binary image which allots every pixel with a value of zero or one depending on the pixel location and intensity and formulates the final shade of the image. If assuming that image s is a binary image of nine pixels, the representation of that can be as in following.

$$s = \begin{matrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{matrix}$$

Whereas, the colored image is a combination of three channels namely, red, blue, and green, and in some cases, the under infrared channel is implemented. So, let the red color is

represented by the matrix R, and blue and green colors are represented by the matrixes B and G respectively.

$$R = \begin{matrix} 1.2 & 0.5 & 1 \\ 1.25 & 0.98 & 1.8 \\ 1.32 & 2.1 & 0.84 \end{matrix}$$

$$B = \begin{matrix} 1.8 & 0.89 & 0.25 \\ 0.112 & 0.285 & 1.69 \\ 1.12 & 1.1 & 1.07 \end{matrix}$$

$$G = \begin{matrix} 1.879 & 0.213 & 0.748 \\ 0.396 & 0.203 & 1.369 \\ 1.102 & 1.279 & 0.258 \end{matrix}$$

So, the combination of the three channels can yield the final.

$$Colour = \begin{matrix} R \\ B \\ G \end{matrix}$$

2.3 NEURAL NETWORK MODELING

With the development of machine learning algorithms, new approaches had implemented to provide supper learning capability and enhanced prediction performance. A neural network is inspired by the human neural system and how does the neuron cells are working. It reflects the same fundamentals of the human learning system. The term artificial is allotted to those kinds of algorithm which represents a smart generation of machine learning technology.

Basically, the neural network works in two stages namely: learning stage and the testing stage.

During the learning stage, the neural network studies the data and target and reaches the concept that relating both data at the input and the results (or the expected results). The following stapes can be considered while the training stage.

A neural network is basically constructed of three parts namely input layer, hidden layer, and output layer. The number of hidden layers always differs from single to multiple layers depending on the design requirements. Figure 2.1 demonstrates the construction of a single hidden layer neural network.

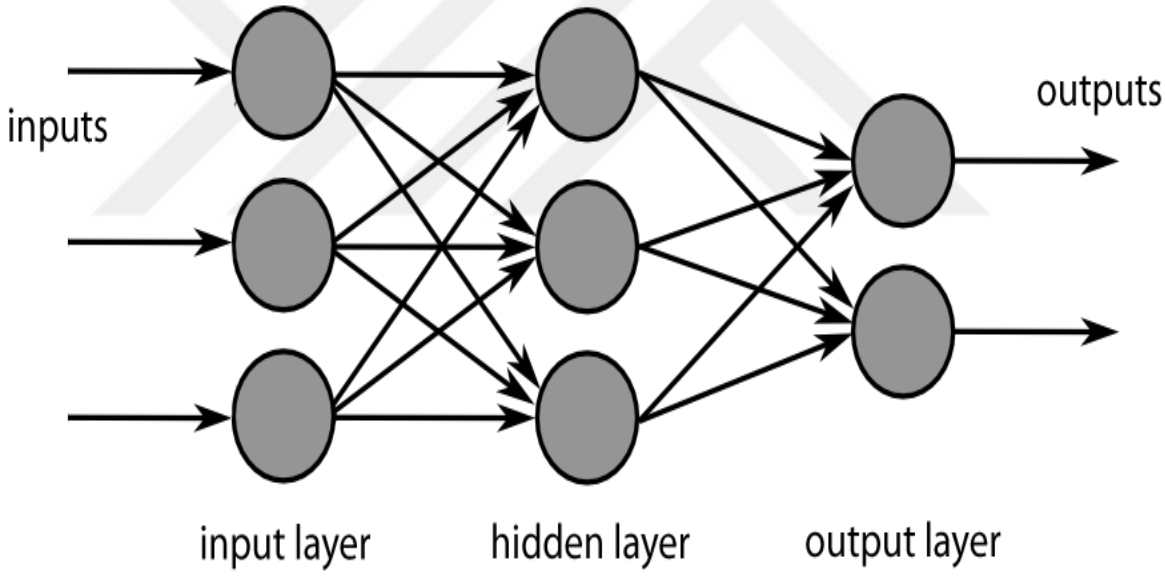

Figure 2.1: Neural network fundamental structure.

Each layer in the neural network is constructed with several (multiple) nodes, nodes number is decided according to the input length and generally, large node number might increase the execution time of the network.

Nodes are connected with each other using weights, the weights are acting like neurons in the human nervous system. These weights are represented by numerical value which represents the number that scales the input when the input passes from a layer to another layer.

Let x[n] is a single-dimensional array of one column and n rows which is passed into the neural network model at the input layer. The output of the hidden layer is expressed by y[n] and can be written as:

$$y[n] = \sum_{n=1}^{n=N} x[n].w_i \quad (1)$$

Whereas, the z[n] can stand for the output of the hidden layer and can be expressed as in the following equation.

$$z[n] = \sum_{n=1}^{n=N} y[n].w_h \quad (2)$$

It is worth saying that three different weight coefficients to be presented on the network namely input layer weight, hidden layer weight, and output layer weight. However, the final output of the network model can be expressed as in the equation below which is entitled as in the following equation.

$$\text{m[n]}.m[n] = \sum_{n=1}^{n=N} z[n].w_o \quad (3)$$

Training stage: a neural network is firstly trained on specific data and the training stage aims to provide the network with all information related to the data structure and nature of this data more likely, how data elements are related to each other and how do they

relate to the target. This process is an essential task in the neural network operation. The aim accuracy of the neural network is depending on how the weight allotment is accurate.

Weight allotment is basically performed to satisfy the equation (4). Weight is the main role player in the process of output generation or in other words, the process of mapping the input to its particular class.

Let the input be x[n], the y[n] is to be the output from this process and the weight general formula is represented by W, and the output after passing from the weight is represented by the following equation:

$$W = \frac{x[n]}{y[n]} \tag{4}$$

Knowing that a weight vector can contain a large number of elements that might be reaching thousands of weights which always depends on the number of dataset elements.

Weight is being allotted to the neural network model using optimization algorithms more likely LM algorithm which acts as standard weight generation algorithm in the neural network model.

2.4 MODEL ESTABLISHMENT

In order to perform the image colourization, model is established using deep learning approach and inspired by conventional neural network. Hence, model is made to perform image colouring by prediction the value of each pixel in the image. Model in firstly trained with large dataset that includes hug number of colored images and however, model is trained with thirty-two category of images in order to inject the neural network with complete knowledge about the object colors so that it can deal with various number of images.

In those thirty-two training stages, thirty-two hidden layer were established and hence each hidden layer is trained with particular data in order to provide the network with complete understanding of diversity of objects. Training of each layer imply allowing the layer to gain popular with images so that network can recognize the image group and hence predict the colour of its pixels accordingly. The training stage is demonstrated in the Figure 2.2.

The process of training will be controller by how well model is doing. So-to-say, training algorithm with is integrated with the neural network model might monitor the performance of the network by evaluating the results accuracy and accordingly the optimization algorithm might regenerate the weight in order to minimize the error in the results until reaching the minimal error in the output. Figure 2.3 is representing the training process of neural network model.

The training process is beginning by generating a random member and allotting those numbers to the weights and hence after evaluating the performance of the machine using the fitness functions. So-to-say, training process night is repeated periodically I during the training stages until reaching the weight values that minimize the fitness function. Fitness function in this case is the mean square error. Figure 2.2 is demonstrating all the process of the said training model.

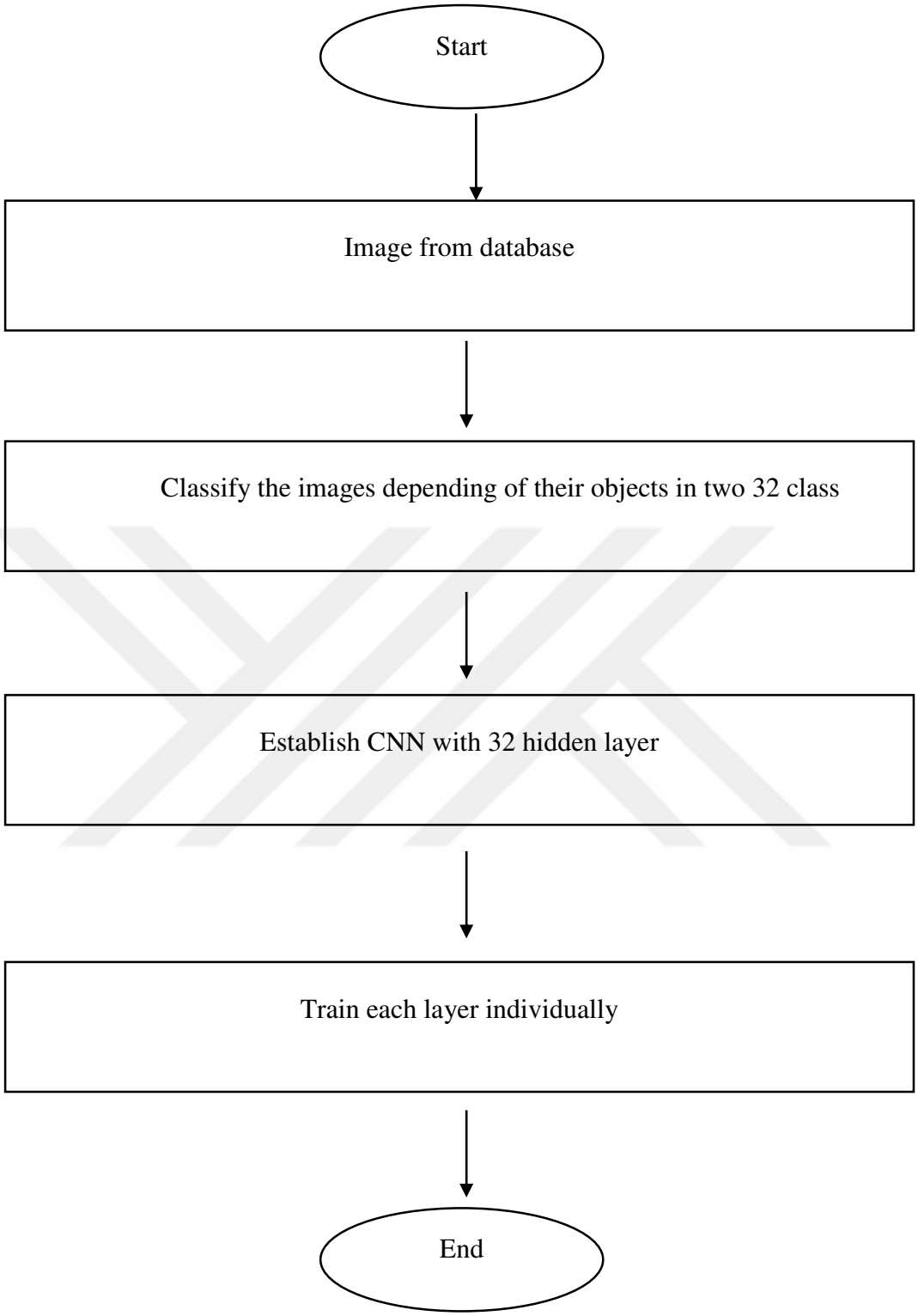

Figure 2.2: A demonstration of CNN training processes

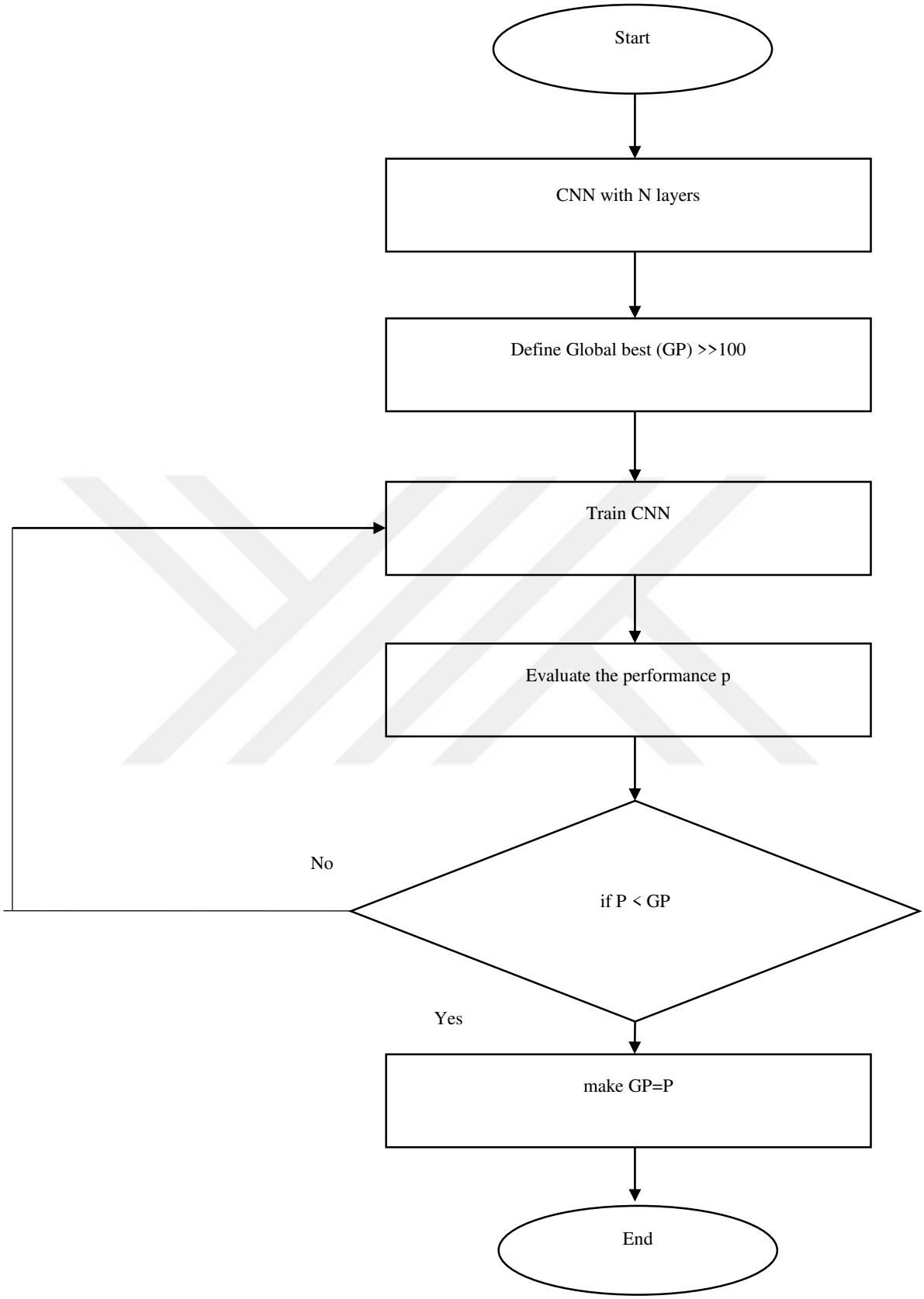

Figure 2.3: CNN training model procedure.

The testing stage will begin by applying test data at the input of the neural network model in order to perform the colorization process. The neural network will study the input data and accordingly will predict the value of each pixel. However, this model is designed to intake the binary images and colored images at any time of execution. However, the colored images are converted first into a binary image, and then it is used for the colorization process.

2.5 PERFORMANCE METRICS

To evaluate the performance of the system after the testing, performance is being studied to evaluate how accurate the system is. The following performance metrics are used for testing the system performance. If supposing the input image is of (ten by ten) pixels x[n];

$$x[n,m] = \begin{bmatrix} 1.2 & \cdots & 2.14 \\ \vdots & \ddots & \vdots \\ 5.13 & \cdots & 7.12 \end{bmatrix} \tag{5}$$

Let x[n,m] is the target image that by any chance the neural network needs to predict.

However, the actual prediction of the neural network is seemed to be p[n,m] which is represented by the following matrix.

$$p[n,m] = \begin{bmatrix} 1.2 & \cdots & 2.58 \\ \vdots & \ddots & \vdots \\ 5.74 & \cdots & 7.12 \end{bmatrix} \tag{6}$$

The performance matrices are directly depending upon the error matrix and that is however can be obtained using the formula of actual minus real. Let the error matrix is to be e[n,m].

$$E[n,m] = x[n,m] - p[n,m] = \begin{bmatrix} 0 & \cdots & 0.44 \\ \vdots & \ddots & \vdots \\ 0.61 & \cdots & 0 \end{bmatrix} \quad (7)$$

The performance metrics will be evaluated using the following formulas in the below sub-sections.

2.5.1 MSE

Mean square error is calculated by applying the equation below; known that the N represents the total number of errors in the error matrix. That number is recognized by calculating non zero elements in the error matrix.

$$MSE = \sum_{n=1}^{N} \left(\frac{E[n]}{N}\right)^2 \quad (8)$$

2.5.2 MAE

same way, another metric can be identified called mean absolute error which is given by the below equation.

$$MAE = \sum_{n=1}^{N} |E[n]/N| \quad (9)$$

2.5.3 RMSE

The root mean square error can be evaluated by using the square root of the mean absolute error function as in hereafter;

$$RMSE = \sqrt[2]{\sum_{n=1}^{N} (\frac{E[n]}{N})\text{^}2} \quad (10)$$

2.5.4 Accuracy

The accuracy is representing the amount of correct predictions with respect to the total amount of prediction multiplied by 100 percent, the same is given by the following function.

$$Acccuracy = \frac{number\ of\ correctly\ predicted\ pixels}{total\ number\ of\ pixels} * 100\% \qquad (11)$$

2.5.5 Others

1. Time: represents the amount of time in seconds that the simulation (process of testing only) lasts for each image.

2. PNR: which represents the pixel signal to noise ration which is calculated by detection the amount of white Gaussian noise in the image and hence relating that to the pixel value.

3 RESULTS AND DISCUSSION

3.1 RESULTS OUTLINE

As the system is established using a conventional neural network as a tool to predict the new pixel value to convert that pixel into the color shade. The model was made with thirty-two hidden layers and all are trained for best model tuning. However, seven hundred images are used in the testing phase where the model is set to predict the color valued of the images. It is noteworthy to mention that all images in the database are with the same pixel dimensions and the system is performable with any image after resizing it into the dimensions used while training.

Hence, the model is tested with all those images, and the performance of color prediction is monitored using six performance metrics namely: mean square error, mean absolute error, accuracy, time, and pixel to noise ratio.

The model has executed three times and the results of the aforementioned metrics are recorded and tabled in Table 3.1.

Table 3.1: The performance metrics results of the model (Pixel-SNR).

Simulation trails	**PSNR**
run 1	31.13
run 2	31.19
run 3	31.09

Table 3.2: The performance metrics results of the model (MSE).

Simulation trails	**MSE**
run 1	53.77
run 2	53.1
run 3	54.29

Table 3.3: The performance metrics results of the model (MAE).

Simulation trails	**MAE**
run 1	124.12
run 2	124.04
run 3	124.09

Table 3.4: The performance metrics results of the model (RMSE).

Simulation trails	**RMSE**
run 1	7.2
run 2	7.16
run 3	7.24

Table 3.5: The performance metrics results of the model (Accuracy).

Simulation trails	**ACCURACY**
run 1	80.49%
run 2	80.45%
run 3	80.91%

Table 3.6: The performance metrics results of the model (Time in Seconds).

Simulation trails	**TIME**
run 1	1.033
run 2	0.84
run 3	0.86

Accordingly, results are graphically represented in Figures 3.1 through 3.5. The mean square error is shown minimum at the second rum of the model as per Figure 3.1. Similarly, the mean absolute error and root mean square error as demonstrated in Figures 3.2 and 3.3.

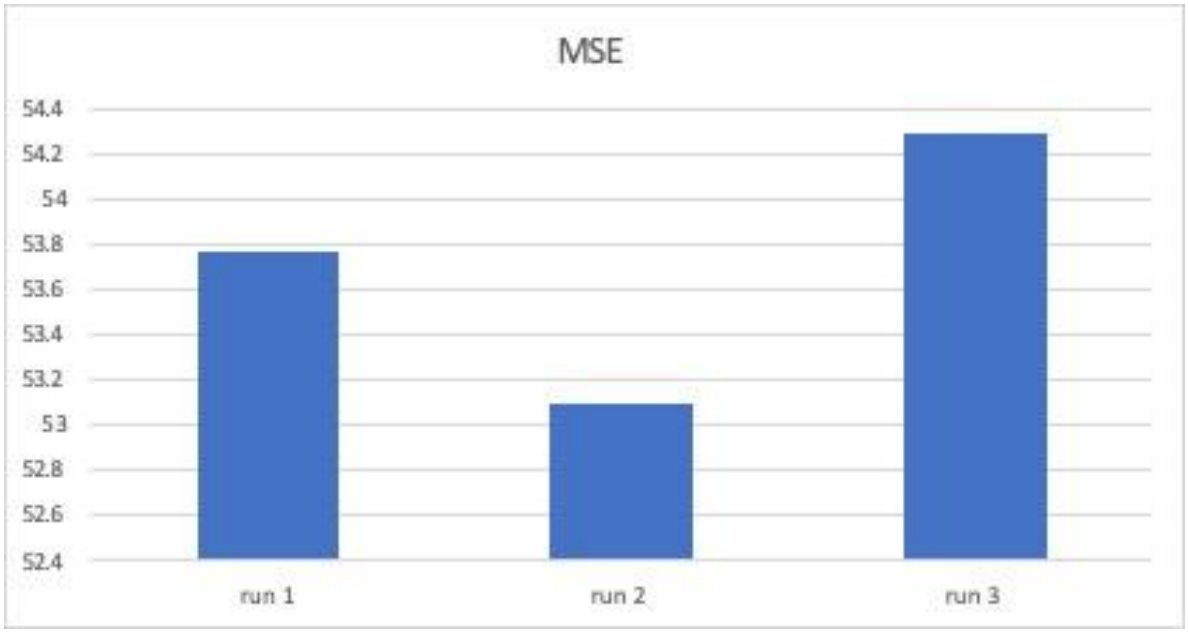

Figure 3.1: Mean square error demonstration for all the runs.

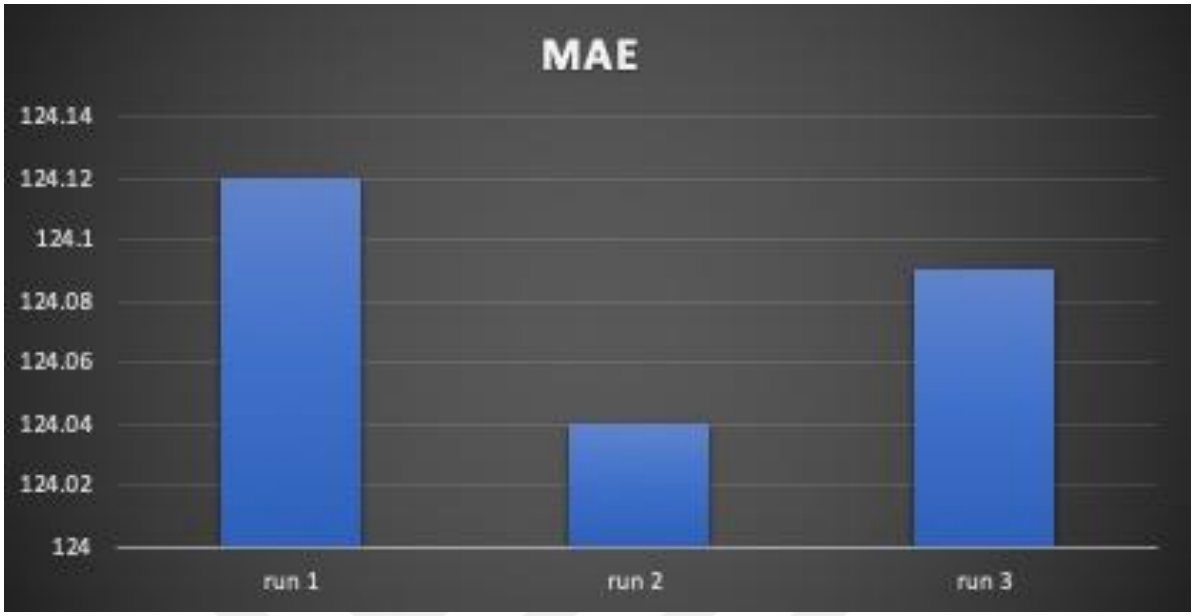

Figure 3.2: Mean absolute error demonstration for all the runs.

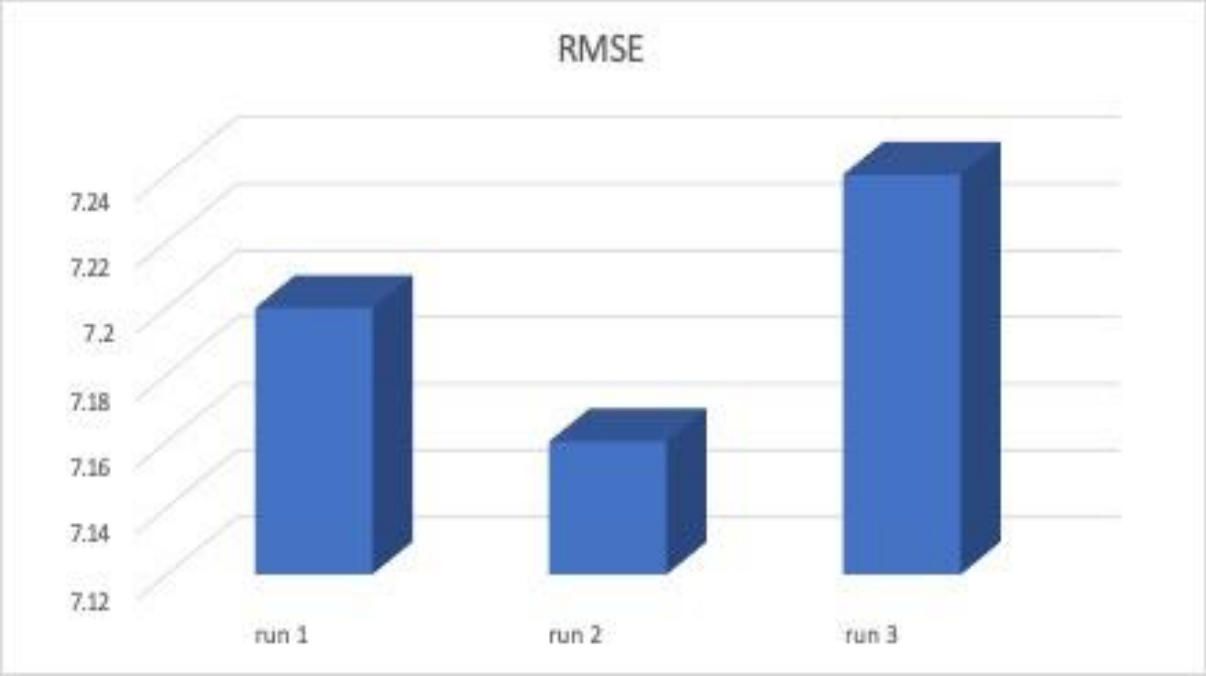

Figure 3.3: Root mean square error demonstration for all the runs.

The accuracy measure is monitored for the whole model testing time during the three stages and hence it is found maximum at the third run of the model as demonstrated in Figure 3.4.

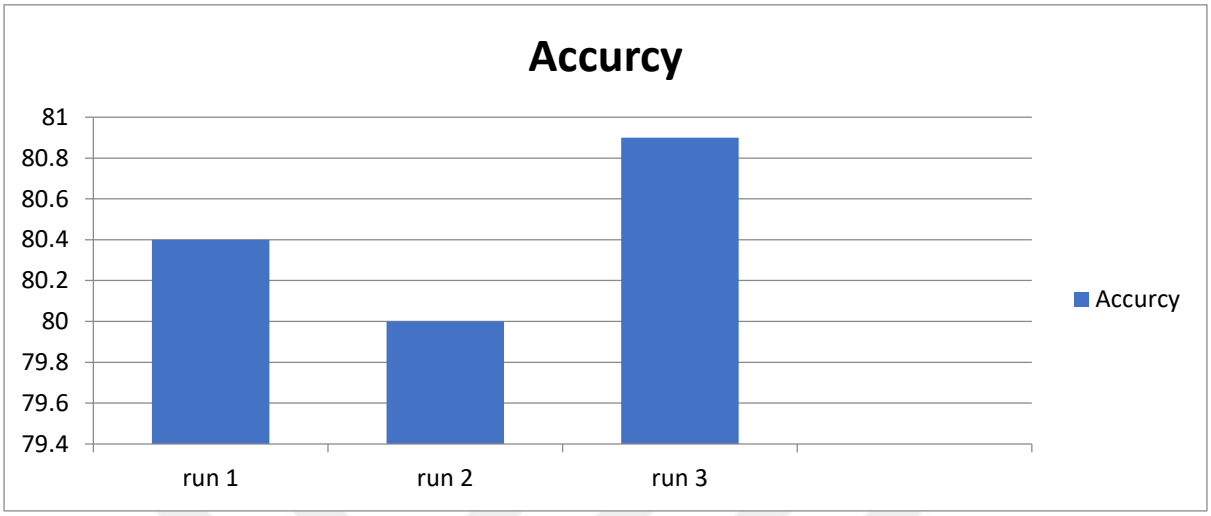

Figure 3.4: The colourization accuracy demonstration for all the runs.

Eventually, the time of processing more likely, the time required by the model to reach the final results that are been monitored and depicted in Figure 3.5. However, the results showed that minimum time had monitored in the second run of the model.

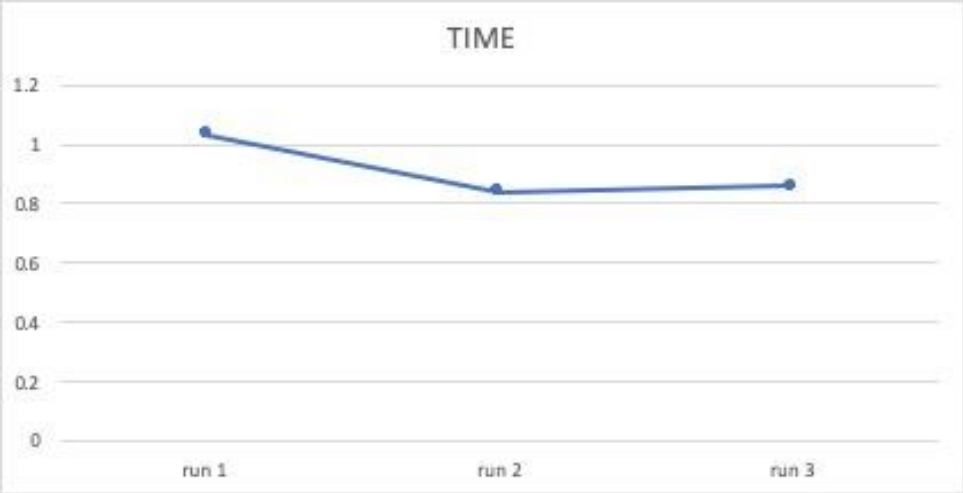

Figure 3.5: The time demonstration for all the runs.

3.2 OBSERVATIONS AND IMPRESSIONS

1> Results were taken after determining the performance metrics for the said model after it predicted the colorization in each image. That means seven hundred values of each performance metric are yielded. The procedure is involved determining the average value of each metric during each run of the model and the same is tabled in Table 3.1.

2> Results are realized with very little variance during all the runs of the model.

3> The difference between the best colorization accuracy and worse colorization accuracy is only 0.42 percent which occurred between run 2 and run 3.

4> Time is reported lesser in the run two while it the bigger slightly in the run three.

5> MSE, MAE, and RMSE are seen all in very small variance.

6> System reported with 80.91 percent of colorization accuracy in 0.86 seconds. That was the best result produced by the model.

4. CONCLUSION AND DISCUSSION

In our approach, image colorization is conducted using the smart machine learning approach with objectives to minimize the time and enhance the overall performance. The project made to draw a high performance of colorization by reducing the error to the minimum. Hence, average 81% accuracy is obtained. Time was notably low which highlights a big contribution over the current studies in image colorization. On the other hence, the system is less computational complex as CNN predictor is used with good training performance which dispensed the need of involving a complex computational method such as Pseudo coloring and histograms.

The techniques that noticed in the literature for performing similar projects was as following :

1. techniques mentioned in the first 3 approaches i.e. [22] [2] and [22] are simple and ensure no accuracy in the results.

2. other techniques such as probability theory and logical reasoning is quite complex and consuming a large amount of computational power and hence experiencing a large time delay.

3. a handful of approaches in the literature were using machine learning for image colorization.

4. our study contribution lies in the consistency of the method from different aspects and performance metrics as the model is tested for various performance metrics apart from the time and proven a good performance in all the tests.

5. The techniques that noticed in the literature for performing similar projects were as following:

Table (4.1): Comparison of the accuracies of previous studies.

Reference	Accuracy
[41]. Automatic colorization	48.7%
[42]. Learning representations for automatic colorization	59.41%
Our study	80.91%

4 REFERENCES

[1]. Lunke Fei , Member, IEEE, Guangming Lu, Wei Jia, Member, IEEE, Shaohua Teng, and David Zhang, Fellow, IEEE, “Feature Extraction Methods for Palmprint Recognition: A Survey and Evaluation”, IEEE, 2019.

[2]. C.W.Kok, Y.Hui, T.Q.Nguyen, “Medical Image Pseudo Coloring by Wavelet Fusion” 18th Annual International Conference of theIEEE Engineering in Medicine and Biology Society. Amsterdam,1996.

[3]. Ruifang Wang, Raymond Veldhuis, Daniel Ramos, Luuk Spreeuwers, Julian Fierrez1, Haiyun Xu, “ON THE USE OF SPECTRAL MINUTIAE IN HIGH-RESOLUTION PALMPRINT RECOGNITION”, IEEE, 2013.

[4]. Janitha R. Karunawallabha1, Thilini I. Wijesiriwardene2, Vajisha U. Wanniarachchi 3, Anuja T. Dharmaratne4, Damitha Sandaruwan, “A Novel Approach to Optimize Crime Investigation Process Using Palmprint Recognition”, 2013 International Conference on Advances in ICT for Emerging Regions ,IEEE

[5]. Pawan Dubey, Tirupathiraju Kanumuri and Ritesh Vyas, “Palmprint Recognition using Binary Wavelet Transform and LBP representation”, International conference on Recent Innovations is Signal Processing and Embedded Systems (RISE-2017) ,IEEE.

[6]. Yibin Yu, Yaofang Tang, Jinguo Cao, JunYing Gan, “Multispectral Palmprint Recognition Using Score-Level Fusion”, IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, 2013.

[7]. Essia Thamri, Kamel Aloui and Mohamed Saber Naceur, “New approach to extract palmprint lines”, IEEE, 2018.

[8]. Zhenhua Guo, Gang Wu, Qingwen Chen, Wenhuang Liu, “Palmprint Recognition by A Two-Phase Test Sample Sparse Representation”, China Postdoctoral Science Foundation under Grant (No. 20100480301 and No. 201104102) and Natural Science Foundation of China (NSFC) (No. 61101150), 2011.

[9]. S.Valarmathy, M.Arun Kumar, M.Sudha, “Improvement in Palmprint Recognition Rate Using Fusion of Multispectral Palmprint Images”, IEEE, 2012.

[10]. Antonio Iula, and Monica Micucci, “Palmprint recognition based on ultrasound imaging”, IEEE, 2019.

[11]. Sen Lin, Yuchen Bai, Yonghua Tang, “Design of Online Non-contact Palmprint Recognition Simulation System”, 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics(CISP-BMEI 2016).

[12]. Waziha Kabir, Student Member, M. Omair Ahmad, M.N.S. Swamy, “Palmprint Recognition Based on Histograms of Sparse Codes”, IEEE, 2017.

[13]. Jyotismita Chaki, “An efficient two-stage Palmprint recognition using Frangi-filter and 2 component partition method”, IEEE, 2018.

[14]. Xiyu Wang, Hengjian Li*, Changzhi Yu, Jian Qiu, “An Improved Palmprint Image Recognition Algorithm Via Image Restoration”, 2018 2nd IEEE Advanced Information Management,Communicates,Electronic and Automation Control Conference(IMCEC 2018).

[15]. Abidin ÇALIŞKAN1, Burhan ERGEN, “PALMPRINT RECOGNITION SYSTEM BASED ON GRAY LEVEL CO-OCCURRENCE MATRIX”, 2014 IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014).

[16]. Lunke Fei, Shaohua Teng, Jigang Wu, Yong Xu, Jie Wen, Chunwei Tian, “Combining Enhanced Competitive Code with Compacted ST for 3D Palmprint Recognition”, 2017 4th IAPR Asian Conference on Pattern Recognition.

[17]. Xin Wu*,Zhigang Zhao,Danfeng Hong ,Weizhong Zhang,Zhenkuan Pan, Jiaona Wan, “] A Palmprint Recognition Algorithm Based On Binary Horizontal Gradient Orientation and Local Information Intensity”, 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC) Dec 20-22, 2013, Shenyang, China.

[18]. Vijay N. Gangapurea, Rahul Sarkarb, Ananda S. Chowdhuryb, “2.5D Palmprint Recognition using Signal level Fusion and Graph based Matching”, IEEE,2017.

[19]. Dian, Liu, Dongmei, Sun, “Contactless Palmprint Recognition Based On Convolutional Neural Network”, IEEE, 2016.

[20]. Yufei Han, Zhenan Sun,Tieniu Tan, “Combine Hierarchical Appearance Statistics for Accurate Palmprint Recognition”, IEEE, 2008.

[21]. Masaki Watanabe, Toshio Endoh, Morito Shiohara and Shigeru Sasaki, "Palm vein authentication technology and its applications", *Proceeding of The Biometric Consortium Conference*, September 10th-21st 2005.

[22]. Ye Ji, Yan Chen, “ Rendering Grayscale Image using Color Feature” IEEE Int. Conf. on Machine Learning and Cybernetics, Vol.5, pp. 3017-3021, Kunming, China, 2008

[23]. W. Niblack, An introduction to digital image processing, Cliffs NJ:Prentice (Hall) Englewood, pp. 115-116, 1986.

[24]. A.K. Jain, L. Hong and R. Bolle, "On-line Fingerprint Verification", *IEEE Transactions on Patter Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 302-313, 1997.

[25]. E. Zhu, J. P. Yin and G. M. Zhang, "Fingerprint matching based on global alignment of multiple reference minutiae", *Pattern Recognition*, vol. 38, no. 10, pp. 1685-1694, 2005.

[26]. S. Prabhakar, S. Pankanti and A. K. Jain, "Biometric Recognition: Security and Privacy Concerns", *IEEE Security & Privacy*, pp. 33-42, March/April 2003.

[27]. GovindHaldankar, AtulTikare and JayprabhaPatil, “Converting Gray-Scale Image to Color Image” Proceedings of SPIT-IEEEColloquium and International Conference, Vol. 1, pp. 189-192,Mumbai, India, June 2011.

[28]. Y.-B. Zhang, Q. Li, J. You and P. Bhattacharya, "Palm vein extraction and matching for personal authentication", *Advances in Visual Information Systems International Conference Visual 2007*, pp. 154-164, June, 2007.

[29]. Jiangtao Xu, Kaige Lu, Xingping Shi, Shuzhen Qin, Han Wang, Jianguo Ma, A DenseUnet Generative Adversarial Network for Near-Infrared Face Image Colorization, Signal Processing, 2021.

[30]. Minghua Wang, Qiang Wang, Hypergraph-regularized sparse representation for single color image super resolution, Journal of Visual Communication and Image Representation, Volume 74, 2021.

[31]. Li Li, Menghan Xia, Chi Liu, Liang Li, Hanyun Wang, Jian Yao, Jointly optimizing global and local color consistency for multiple image mosaicking, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 170, 2020.

[32]. Leyuan Wu, Xiaogang Zhang, Hua Chen, Yicong Zhou, Unsupervised quaternion model for blind colour image quality assessment, Signal Processing, Volume 176, 2020.

[33]. Junxing Yang, Lulu Liu, Jiabin Xu, Yi Wang, Fei Deng, Efficient global color correction for large-scale multiple-view images in three-dimensional reconstruction, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 173, 2021.

[34]. Yonggang Su, Wenjun Xu, Tianlun Li, Jie Zhao, Shuaiqi Liu, Optical color image encryption based on fingerprint key and phase-shifting digital holography, Optics and Lasers in Engineering, Volume 140, 2021.

[35]. Decheng Liu, Qingtang Su, Zihan Yuan, Xueting Zhang, A fusion-domain color image watermarking based on Haar transform and image correction, Expert Systems with Applications, Volume 170, 2021.

[36]. Gang Qu, Xiangfeng Meng, Yongkai Yin, Huazheng Wu, Xiulun Yang, Xiang Peng, Wenqi He, Optical color image encryption based on Hadamard single-pixel imaging and Arnold transformation, Optics and Lasers in Engineering, Volume 137, 2021.

[37]. Jifei Miao, Kit Ian Kou, Wankai Liu, Low-rank quaternion tensor completion for recovering color videos and images, Pattern Recognition, Volume 107, 2020.

[38]. Xueting Zhang, Qingtang Su, Zihan Yuan, Decheng Liu, An efficient blind color image watermarking algorithm in spatial domain combining discrete Fourier transform, Optik, Volume 219, 2020.

[39]. XIAOYAN LEI, HUIBIN WANG, JIE SHEN, ZHE CHEN, WEIDONG ZHANG, A novel intelligent underwater image enhancement method via color correction and contrast stretching, Microprocessors and Microsystems, 2021.

[40]. M. Yamni, H. Karmouni, M. Sayyouri, H. Qjidaa, J. Flusser, Novel Octonion Moments for color stereo image analysis, Digital Signal Processing, Volume 108, 2021.

[41]. Dahl, R.: Automatic colorization. In: http://tinyclouds.org/colorize/. (2016).

[42]. Larsson, G., Maire, M., Shakhnarovich, G.: Learning representations for automatic colorization. European Conference on Computer Vision (2016)