

DEEP LEARNING BASED MALAYSIAN COINS RECOGNITION FOR VISUAL IMPAIRED PERSON

Lina Suhaili Rosidi, Nur Anis Jasmin Sufri, Muhammad Amir As'ari*

School of Biomedical Engineering and Health Sciences, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), 81310 UTM Johor Bahru, Johor, Malaysia

Article history

Received

08 June 2021

Received in revised form

06 August 2021

Accepted

21 November 2021

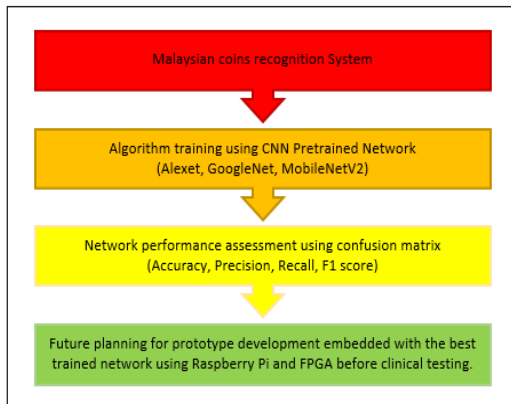
Published online

31 May 2022

*Corresponding author

amir-asari@biomedical.utm.my

Graphical abstract



Abstract

Currency recognition has been widely developed using various types of techniques and able to assist people who have a visual impairment. Machine learning is one of the methods implemented where deep learning architecture is one of them. The deep learning approach is reliable and can be used in detection and recognition of objects based on images. As currency recognition has been developed for other currencies, thus in this project, currency recognition using Malaysian coins has been developed by modeling Convolutional Neural Network (CNN) in recognizing coin images. Malaysian coins dataset was developed consist of 2400 images of four classes of coins, 5 sen, 10 sen, 20 sen, and 50 sen. In this study, pretrained CNN which are AlexNet, GoogleNet, and MobileNetV2 were formulated in recognizing such coins. Performance of each trained model was evaluated using confusion matrix and GoogleNet obtained the best performance with 99.2% testing accuracy, 99.2% precision, 99.18% recall, and 99.19% F1 score. From the trained model, it can be further developed and implemented in assisting visually impaired persons by producing a prototype using Raspberry Pi and FPGA before it can be clinically tested on the subject.

Keywords: Visual impaired, coins recognition, AlexNet, GoogleNet, MobileNetV2

© 2022 Penerbit UTM Press. All rights reserved

1.0 INTRODUCTION

As stated by World Health Organization (WHO) [1], it has been estimated that people who currently has visual impairment is about 1.3 billion of world population. Visual impairment can be related to visual acuity in which spatial resolution of human vision is measured [2]. Visual acuity is measured by having fractional number of which the numerator shows the distance of test chart in feet while the denominator is the distance of which human eyes can see the smallest letters in the chart [2].

Visual impairment can be categorized into few types which are decrement of light sensitivity, blurry vision, loss of vision and total blind [3]. The categorization of degree of impairment is necessary in providing assistive devices which can help them in daily life according to their condition and situation. Assistive devices have been developed by researchers over years with different function but same goals of aiding visually impaired person. Many types of devices have been developed including for navigation, detection

of obstacles, face recognition, text recognition and object recognition [3].

Specifically, for recognition of currency, researchers have developed vision-based currency recognition system which able to assist in performing daily activities. Different techniques and methods have been used to improve the system and obtained high accuracy of recognition output. Studies of currency recognition system mostly used machine learning and deep learning approach. Machine learning is a method where the machine itself as in computer learned by examples or data in which can be categorized into three ways of learning which are supervised learning, unsupervised learning, and reinforcement learning [4]. Machine learning approach is paired with conventional human crafted feature selection and extraction, image processing in which can manually select relatable information from the given data causing time constraint [5].

However, only a few focuses on deep learning intervention in recognizing the banknote. Deep learning approach of using

Convolutional Neural Network (CNN) architecture has surpassed machine learning in which have functionality to determine out their own most discriminative features relevant to the hassle and making it more promising, reliable, and highly accurate for computer vision task. Thus, in this research, CNN which is one of the deep learning was proposed and implemented for Malaysian coins recognition which will benefit in developing the assistive system for visual impaired person.

2.0 LITERATURE REVIEW

Assistive Technology for Visually Impaired Person

Throughout the year, due to advancement of technology and research made has resulted in many developments of assistive devices for the use of visually impaired person and blind people [3]. Thus, these different types of assistive devices can help them in their daily activities such as navigation and obstacle avoidance, information access, reading, and recognizing nonverbal communication cues and recognizing items in the surrounding. In a study which discussed on the assistive devices developed by researchers, it is mentioned that vision substitution with output of not based on visual sense, are categorized into 3 types including "Electronic Travel Aid (ETAs), Electronic Orientation Aid (EOAs), and Position Locator Devices (PLDs)" [6]. In addition to that, information from surroundings plays vital role in assistive devices as it is to be transferred to the user.

In another study, lower visual perception experienced by visually impaired person can be replaced by other senses which includes "vision enhancement, audition, somatosensory, visual prosthesis and olfactory and gestation" [3]. In the same study, many assistive devices developed have been summarized according to its type such as canes, glasses and others which compare its type of feedback, function as well as sensor used for the devices. One of the assistive devices mentioned is from Wicab, Inc which is available on the market called BrainPort Vision Pro [7]. This device is an "oral electronic vision aid" which can assist the user apart from the help of white cane or guide dog in direction, movement and to recognize object. It is designed as a headset comprised of video camera, user control and tongue array which consists of 394 electrodes. User who has experienced the device has described it as "see with your tongue".

DBG Crutch Based MSensors [8] is one of the examples of the device which can assist the user in their movement from one place to another as it is able to help user to detect and avoid anything which can hindrance their movement. This system used ultrasonic sensors which are attached to the cane in which function to obtain data on distance from surroundings. Another assistive device with same function but different implementation has been developed using Dynamic Vision Sensors (DVS) [9]. This study of developing assistive glasses converted surrounding information into 3D spatial sound as output to the user.

Apart from having assistive cane or glasses to aid visually impaired person in their daily life, other modality such as hat, belt, bracelet, robotic dog, jacket, wheelchair, hand-held cube and flashlight have been developed by the researchers. Froneman et al. [10] developed a wearable support system using ultrasonic sensor which was implemented as a belt for the user to put on their waist. This system would give output to the user through vibration which can help the user in moving around without

bumped into obstacles. In a study of developing an assistive device, a shape-changing interface has been implemented using 'The Animotus' which is a hand-held cube [11]. Its function is to facilitate indoor navigation by providing shape-changing tactus feedback to the user.

Introduction to Convolutional Neural Network

Deep learning is a subtype of machine learning. In machine learning, the relevant features of an image were manually extract but with deep learning, the raw images were feed directly into a deep neural network that learns the features automatically. The most well-known image processing structure of deep learning is Convolutional Neural Network (CNN). CNN has multiple layers that mainly consist of convolutional layer, non-linearity layer, pooling layer, and fully connected layer [12]. CNN has been used for image recognition and classification in many areas such as security and surveillance [13], biometric [14], medical imaging [15], agriculture [16] and many other applications. CNN can be trained in two common ways: 1) training from scratch, 2) transfer learning [17]. Training a deep CNN model from scratch would require a huge amount of data [18]. Meanwhile, transfer learning aims to reuse the existing pre-trained network like AlexNet, GoogLeNet and MobileNetV2.

AlexNet is a CNN architecture which was invented by Alex Krizhevsky, Geoffrey Hinton, and Ilya Sutskever [19]. In a study conducted by Minhas et al. [20], a method is proposed to classify shot effectively for cricket and soccer sports video based on AlexNet network where the accuracy performance is 94.07%. AlexNet also has been implemented in analysis of breast cancer histopathology images as conducted by Titoriya et al. [21] using 7909 images for training and obtained maximum accuracy of 95.7%. S. Liawatimena et al. [22] involving AlexNet to carry out classification on three types of fish images and gain a high accuracy of 99.63%.

GoogLeNet is a CNN architecture which has been selected as the winner in ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14) [23]. Hendrick et al. [24] used GoogLeNet to develop a model to classify chest x-ray images for diagnosis of tuberculosis and has achieved a reliable performance of 98.39%. GoogLeNet network proposed by Ma et al. [25] was applied in recognizing sprouting potato to prevent any harm to the consumer through image recognition. Al-Qizwini et al. [26] proposed a method for five accordance parameters in controlling the vehicle for autonomous driving by comparing performance of CNN and discovered that GoogLeNet architecture obtained the best result among all the pretrained networks.

MobileNetV2 is a CNN architecture in which a bottleneck depth-separable convolution with residuals is the basic building blocks [27]. In the study by Akiyama et al. [28], plant identification and classification using MobileNetV2 of 33 different types of plant species scored the best result among the architecture involved. MobileNetV2 architecture was involved in a study by Yuan et al. [29] regarding the detection of railway surface defect where it is a crucial aspect to measure to make sure the operation of the rail transit system is safe. In a study done by Ale et al. [30], MobileNetV2 was chosen in developing the model in which can recognize facial recognition.

Banknote Recognition System

Visually impaired person may have difficulties in recognizing object or anything in which require visual input including face, text, and currency [3]. Thus, previous studies involving currency recognition system have develop various techniques to ease visually impaired person in performing activities of daily living using both method of machine learning and deep learning. The recognition of banknotes can be divided into two categories: 1) vision-based system and 2) sensor-based system. The camera was mainly applied to aid the visually challenged individual in the vision-based system. In terms of sensor-based systems, most of the research suggested various sensing modalities. Existing sensor-based system studies involve many electrical components, and the intended outcome was not accurate due to sensor limitations. As a result, most studies on vision-based systems claimed to have a greater accuracy.

Dunai et al. [6] works on Euro banknote recognition system with dataset of 2000 images using Viola and Jones algorithm and Speed Up Robust Features (SURF) methods and able to have 97.5% recognition accuracy. A recognition system of Ethiopian banknote involved 500 images in dataset which include 80% training data and 20% testing data was proposed by Ayalew et al. [7] using Local Binary Pattern (LBP) techniques as feature extraction step before feed into Support Vector Machine (SVM) classifier and acquired 98% recognition accuracy. Another banknote recognition system was proposed by Abu Doush et al. [8] for 10 classes of Jordanian Dinar (JD) using Scale-Invariant Feature Transform (SIFT) algorithm. This study used 400 of training images and 100 of testing samples resulting color SIFT approach showed higher percentage of correct recognition performance compared to gray SIFT. Hlaing et al. [9] has developed Myanmar Kyat banknote recognition system applying Gray Level Co-occurrence Matrix (GLCM) approach during feature extraction before proceeding to k-Nearest Neighbor (k-NN) classifier using 500 images classified into 5 classes and resulted in recognition rate of 99.2%.

Banknote recognition system for Malaysian currency also has been proposed by Sufri et al. [10] applying machine learning architecture in which k-Nearest Neighbor (k-NN) and Decision Tree Classifier (DTC) have overall of 99.7%. rate of accuracy. In a study by Mittal et al. [15], MobileNet was used for 12160 images of Rupee banknote with 96.9% of recognition accuracy. Deep learning approach was also applied in study by Almisreb et al. [31] where performance of pretrained CNN model including GoogLeNet, AlexNet and Vgg16 were compared using Bosnian currency. The related works in banknote recognition system implemented conventional Machine Learning and Deep Learning are summarized in Table 1.

Table 1 Summary of study involving Banknote Recognition System using Machine Learning and Deep Learning

References	Currency	Dataset	Method	Accuracy
Dunai et al. [6]	Euro	2000	Viola and Jones algorithm, SURF	97.5%
Ayalew et al. [7]	Bir	500	LBP, SVM	98%
Abu Doush et al. [8]	Dinar	500	SIFT	71%
Hlaing et al. [9]	Kyat	500	GLCM, k-NN	99.2%

Jasmin et al. [10]	Ringgit	672	k-NN, DTC	99.7%
Mittal et al. [15]	Rupee	12160	MobileNet	96.9%
Almisreb et al. [29]	BAM	110	GoogLeNet, AlexNet, Vgg16	88.65%, 95.24%, 100%
Murad et al. [16]	Taka	8000	MobileNet	99.80%

Coin Detection and Classification System

Several studies have been conducted related to machine learning and deep learning in detecting and recognizing coin using different algorithm and features in which involved coins from different countries.

As presented in a paper by Kaur et al. [32], an Indian coin recognition system undergo pre-processing: image crop into circular shape then converted into gray scale and features were extracted using Polar Harmonic Transform (PHT) before feed into Artificial Neural Network (ANN). Image requisition, segmentation, edge detection, polar transform, and Fourier transform was applied to Indian coins before implementing as input image to Multi-layered back propagation Neural Network (MLBPNN) in the study by Roomi et al. [33] with 82% of accuracy.

In addition, Farooque et al. [34] have proposed a coin recognition system of Pakistani coins by using Scale Invariant Feature Transform (SIFT) algorithm and Principle Component Analysis (PCA) for feature extraction before output image is passed to ANN. The datasets consisting of 200 coins was converted to gray scale before extracting its features and the overall results showed 84% of accuracy. Capece et al. [35] have proposed a Euro coin recognition system by using AlexNet with dataset of 8320 images consist of 8 classes of. User can identify coins by using mobile device which are connected to client-server architecture.

Qiu et al. [36] have proposed a method to detect and recognize coin in uncontrolled environment in which 6 classes of coins including China and Hong Kong coins with 34000 coins image as input. Hough detection method was involved in detecting coins and multilayer CNN act as algorithm for recognizing coin value. Anwar et al. [37] also presented ancient Roman coins recognition where reverse motifs of the coins were used with dataset consisting of more than 18000 images. The CoinNet (Specialized CNN) has achieved 98% of accuracy.

Kim et al. [38] have proposed a method by using AlexNet to identify characteristic landmarks on 4256 images pairs of obverse and reverse side of ancient Roman imperial coins. Schlag et al. [39] have presented another work of involving ancient Roman coins where the obverse side of coins was used to identify emperor face profiles with three datasets of 29807 coins, 19164 coins and 600 images respectively using CNN.

Other deep learning-based system for recognizing coins was mentioned in a research by Tajane et al. [19] where CNN were used by adapting AlexNet model with dataset of more than 1600 images. The proposed method was able to give fast response to user and its performance was measured by recognition accuracy and time of response. The existing works in coin recognition implemented conventional machine learning and deep learning are summarized in Table 2.

As conclusion, existing works for Malaysian currency mostly only involving Ringgit banknote recognition [10] and there were no existing studies specifically using human crafted feature and machine learning approach or deep learning approach for Malaysian coin.

Table 2 Summary of study involving Coin Recognition using Machine Learning and Deep Learning

References	Currency	Dataset	Method	Accuracy
Kaur et al. [32]	Rupee	4 classes	Polar Transform; ANN	High
Roomi et al. [43]	Rupee	48	Hough Transform; Polar Transform; Fourier Transform	82%
Farooque et al [34]	Pakistani Rupee	200	SIFT; PCA; Confusion Matrix	84%
Capece et al. [35]	Euro	8320	AlexNet	72.21%
Qiu et al. [36]	Yuan, HongKong Dollar	34000	Hough detection; CNN	87.15%
Anwar et al. [37]	Ancient Roman Republican	18000	CoinNet (Specialized CNN)	98%
Kim et al. [38]	Ancient Roman Imperial	4526	AlexNet	Outperform SVM
Schlag et al. [59]	Roman Imperial	49571	CNN	Outperform state-of-art by magnitude
Tajane et al. [19]	Rupee	1600	AlexNet	Outperform conventional system

3.0 METHODOLOGY

The aim is to develop automated Malaysian coin recognition system using deep pretrained CNN model like AlexNet, GoogLeNet and MobileNetV2. The performance of each model in recognizing Malaysian coins is evaluated on MATLAB software. The general method involved were illustrated according to the block diagram in Figure 1 below.

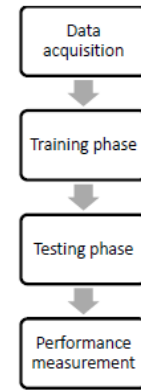


Figure 1. Block diagram of methodology

Data Acquisition

The own datasets of Malaysian coins image consist of 2400 labelled images captured using smartphone camera (a single coin in each image) were divided into 4 value of classes: 5 sen, 10 sen, 20 sen and 50 sen. Each class equally have 600 images which consist of both obverse (head) and reverse (tail) side of the coin. Table 3 below shows the total number of images for each class of coins for different sides.

Table 3 Number of Images for Each Class for Different Classes

Classes	Obverse (head)	Reverse (tail)	Total
5 sen	438	162	600
10 sen	405	195	600
20 sen	448	152	600
50 sen	436	164	600

The images for dataset were taken at a constant distance between the subject and the camera lens in same illumination level. In addition, the same setting of white plain background for each image was used to avoid variations of background. All images in the dataset were taken using SM-A105G camera model with output resolutions of 4128x3096 pixels.

Before proceeding to the next stage, the dataset was split into training, validation, and testing Training and testing dataset were divided in the proportion of 0.9:0.1 and then the training dataset was further divided into training and validation dataset by 0.7:0.3. Data augmentation was performed to the dataset (see Table 4) before training phase and the input image size was resized to fit with the input size requirement for each pretrained network. Figure 2 illustrates the sample images of Malaysian coins in the established dataset while Figure 3 displays the close-up view of the taken images samples at both obverse and reverse sides.

Table 4 Total Images for Each Augmented Dataset

	Training	Validation	Testing
Number of images	1512	648	240



Figure 2 Sample coin image of 4 different classes



Figure 3 Close up sample coin image of 4 different classes

Training Phase

Pretrained CNN model like AlexNet, GoogleNet and MobileNetV2 were used in training and testing phase of the coin images. In the training phase, the purpose is to train the network to identify value of different coins from the input images. In this phase, the significant step is to introduce the transfer learning technique where pretrained CNN model is used as the basis to learn new classification task. Steps taken in training the network are as illustrated in Figure 4.

Images from the dataset is load before proceeding to data augmentation where the images were resized according to its respective image input layer size. Then, the pre-trained network is load and the architecture of the network was observed. After that, CNN layers were replaced for the classification of the new task in which to classify coin images into four classes. Before the network is trained, the training option was specified where the size of mini batch, maximum epoch, initial learn rate and others were set. Lastly, the network is trained on the dataset for the new classification task.

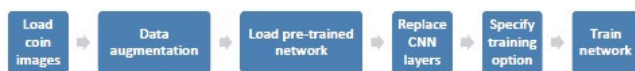


Figure 4 Training workflow of the model

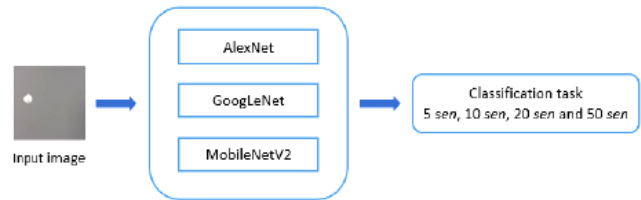


Figure 5 Implementation of CNN transfer learning

Figure 5 shows how the pre-trained CNN was implemented in this process. In general, the final layers were replaced with 4 fully connected layers and new classification layer to classify 4 different classes of coin images: 5 sen, 10 sen, 20 sen or 50 sen. The replaced final layers which were the new layers will be responsible to learn the new distinct features of the coin dataset. During training process, stochastic gradient descent optimization was chosen with mini batch size of 32, maximum training epoch of 10 and the initial learn rate was set to 0.0001.

At the end of the training phase, the training accuracy was obtained together with validation accuracy, training loss, and validation loss. The next step was to proceed for testing the network using the testing dataset.

Testing Phase and Performance Measurement

The total number of testing dataset images is 240 images where the number of images of each of the class is equivalent. The input image size for testing dataset was also resized according to each network input image size specification: 227x227 pixels for AlexNet and 224x224 pixels for both GoogleNet and MobileNetV2.

The result of the testing phase was obtained by referring at the testing accuracy of the trained network. For performance measurement, confusion matrix was used to assess the reliability of the trained network. Confusion matrix is the common method used in evaluating the performance of the classification model. From the matrix, the value of precision, recall (sensitivity), and F1 score can be acquired.

Equations (1-4) below show the calculation to obtain each of the parameter which includes accuracy, precision, recall and F1 score where TP is 'true positive', FP is 'false positive', TN is 'true negative' and FN is 'false negative'. TP denotes correctly classified coins of each class (i.e., 10 sen if the actual label is 10 sen) while TN represents the correctly classified of other classes compared to TP class. Meanwhile, FP represents the negative class misclassified as positive class and FN signifies the positive class misclassified as negative class. The formula to obtain precision, recall and F1 score is as shown as below:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1\ Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4)$$

4.0 RESULT AND DISCUSSION

From the training phase, a training progress graph was obtained for every network trained. Table 5 summarizes the performance

during training process for each of the network and Table 6 listed the training time acquired for each pretrained model

Table 5 Training Progress Result for Each Network

Pretrained model	Training accuracy (%)	Validation accuracy (%)	Training loss (%)	Validation loss (%)
AlexNet	93.75	98.15	11.10	5.95
GoogLeNet	100.00	97.22	2.44	8.38
MobileNetV2	96.88	97.53	11.00	8.17

Table 6 Training Time for Each Network

Pretrained model	Training time (minutes, seconds)
AlexNet	134 minutes and 16 seconds
GoogLeNet	155 minutes and 31 seconds
MobileNetV2	173 minutes and 12 seconds

It can be observed that the network with the highest training accuracy is GoogLeNet with significant value of final training accuracy of 100% compared to others with accuracy of 93.75% and 96.88% for AlexNet and MobileNetV2, respectively. AlexNet managed to obtain validation accuracy of 98.15% while GoogLeNet and MobileNetV2 gained 97.22% and 97.53%, respectively. By referring to the validation accuracy of each network, no signs of overfitting happened during the training progress of the network. GoogLeNet surpassed other trained network by recording the lowest loss training value of 0.0244 compared to others with both having almost the same value which is 0.1110 and 0.1100. As for validation loss, all the trained network achieved the value below the highest training loss recorded where AlexNet, GoogLeNet and MobileNetV2 achieved 0.0595, 0.0838 and 0.0817 value of validation loss, respectively. From the testing phase, the performance of trained model was evaluated by measuring the testing set accuracy and other parameters such as precision, recall and F1 score by plotting confusion matrix.

Figure 6 displays the confusion matrix of AlexNet on testing dataset in which the overall testing accuracy is 98.3%. For this model, all 60 images of 20 sen and 50 sen in the testing dataset were accurately classified. On the other hand, 59 images of 10 sen were correctly classified where 1 image was wrongly predicted as 5 sen. Also, 3 images of 5 sen wrongly predicted as 10 sen and 57 images were accurately predicted. Thus, according to Equations (2-4), the precision is 98.38.2%. Percentage of recall obtained for this class is 98.33% and F1 score is 98.35%.

Figure 7 presents the confusion matrix of GoogLeNet with the accuracy of testing surpassed testing accuracy of AlexNet, which is 99.2%. It can be observed that all 60 images of 10 sen, 20 sen and 50 sen were correctly classified. However, only two models managed to achieve percentage of precision and recall of 100% which are 20 sen and 50 sen. As for 5 sen, there were 58 images correctly predicted while 2 images were wrongly classified as 10 sen. Thus, according to previous Equations (2-4), the precision is 99.20%. Percentage of recall obtained for this class is 99.18% and F1 score is 99.19%.

From Figure 8, the overall accuracy of the testing data for MobileNetV2 is 94.6% which is lower than that of AlexNet and

GoogLeNet. This model successfully classified all 60 images for both 20 sen and 50 sen. For 10 sen, 51 images were accurately classified, and 9 images were wrongly classified as 5 sen. Compared to AlexNet and GoogLeNet, it has the least number of images correctly classified for 10 sen. Also, 4 images of 5 sen were mistakenly predicted as 10 sen. For 5 sen, there were 56 images which has been classified accordingly to its class. 4 images of 5 sen were not classified according to its class. Therefore, the percentage of precision, recall and F1 score are 94.73%, 94.58% and 94.65%, respectively.

		Confusion Matrix					
		10 sen	20 sen	5 sen	50 sen	Accuracy	Precision
Output Class	10 sen	59 24.6%	0 0.0%	3 1.2%	0 0.0%	95.2%	4.8%
	20 sen	0 0.0%	60 25.0%	0 0.0%	0 0.0%	100%	0.0%
	5 sen	1 0.4%	0 0.0%	57 23.8%	0 0.0%	98.3%	1.7%
	50 sen	0 0.0%	0 0.0%	0 0.0%	60 25.0%	100%	0.0%
		98.3%	100%	95.0%	100%	98.3%	1.7%
		1.7%	0.0%	5.0%	0.0%	0.0%	0.0%
		Target Class					

Figure 6 Confusion matrix of AlexNet on testing dataset

		Confusion Matrix					
Output Class	10 sen	60 25.0%	0 0.0%	2 0.8%	0 0.0%	96.8%	3.2%
	20 sen	0 0.0%	60 25.0%	0 0.0%	0 0.0%	100%	0.0%
	5 sen	0 0.0%	0 0.0%	58 24.2%	0 0.0%	100%	0.0%
	50 sen	0 0.0%	0 0.0%	0 0.0%	60 25.0%	100%	0.0%
		100% 0.0%	100% 0.0%	96.7% 3.3%	100% 0.0%	99.2%	0.8%
		Target Class					
		10 sen	20 sen	5 sen	50 sen		

Figure 7 Confusion matrix of GoogleNet on testing dataset

		Confusion Matrix					
Output Class	10 sen	51 21.2%	0 0.0%	4 1.7%	0 0.0%	92.7%	7.3%
	20 sen	0 0.0%	60 25.0%	0 0.0%	0 0.0%	100%	0.0%
	5 sen	9 3.8%	0 0.0%	56 23.3%	0 0.0%	86.2%	13.8%
	50 sen	0 0.0%	0 0.0%	0 0.0%	60 25.0%	100%	0.0%
		85.0% 15.0%	100% 0.0%	93.3% 6.7%	100% 0.0%	94.6%	5.4%
		Target Class					
		10 sen	20 sen	5 sen	50 sen		

Figure 8 Confusion matrix of MobileNetV2 on testing dataset

For AlexNet, precision of the trained model is 98.38% while GoogleNet and MobileNetV2 obtained 99.20% and 94.73%, respectively. As precision value is one of the parameters to measure the performance of the network, comparing the value achieved by the networks, percentage gained by GoogleNet is the highest among all the networks trained. This implies that the network is dependable when the model predicted coin images belong to their class.

Referring to confusion matrix, recall for each of trained model can be obtained by averaging the value from the last row of the matrix. Recall value for AlexNet, GoogleNet and MobileNetV2 is 98.33%, 99.18% and 94.58%, correspondingly. Similarly, among the trained networks, GoogleNet performed better than the others by achieving the highest recall value. It shows the efficiency of the model in detecting the class.

F1 score calculated for AlexNet achieved 98.35% while GoogleNet and MobileNetV2 having F1 score of 99.19% and 94.65%, respectively. Likewise, GoogleNet has the highest F1 score compared to GoogleNet and MobileNetV2. Moreover, as for testing accuracy, GoogleNet has outperformed the other trained network, AlexNet and MobileNetV2, by having the highest testing accuracy. GoogleNet obtained 99.20% of testing accuracy while AlexNet and MobileNetV2 achieved testing accuracy of 98.3% and 94.60%, accordingly. As a summary, Table 7 shows the complete performance of each of the different trained models which are AlexNet, GoogleNet and MobileNetV2.

Table 7 Summary of Trained Network Performance

Pretrained Model	AlexNet	GoogleNet	MobileNetV2
Training accuracy (%)	93.75	100	96.88
Validation accuracy (%)	98.15	97.22	97.53
Training loss	0.1110	0.0244	0.1100
Validation loss	0.0595	0.0838	0.0817
Precision (%)	98.38	99.20	94.73
Recall (%)	98.33	99.18	94.58
F1 score (%)	98.35	99.19	94.65
Testing accuracy (%)	98.30	99.20	94.60

5.0 CONCLUSION

The main goal is to have an automated coin recognition system specifically for Malaysia coins to assist visually impaired person in their daily activities especially in performing groceries activity. Currency recognition is one of the systems in assistive technology that have major interest of researchers from all over the world. Thus, currency recognition for different country banknote and coins has been studied using different techniques and approach. However, for Malaysia currency, only banknote recognition has been presented. Therefore, a coin recognition system is proposed by using deep learning approach involving CNN for Malaysia coins.

The outcome of this study has proven that deep learning approach able to achieve higher performance compared to previous works regarding currency recognition in which GoogleNet obtained the best performance among the pretrained networks. GoogleNet outperform AlexNet and MobileNetV2 and has achieved testing accuracy of 99.2%. As a conclusion, the developed model is reliable and has achieved its objective in recognizing the coins by modelling deep CNN, evaluated the performance of CNN in recognizing coins and analyzed the performance of different CNN model. Some suggested future works are to extend the developed system by integrating with embedded system such as Raspberry Pi or Field Programmable Gate Arrays (FPGA) for the development of prototype before it can be tested clinically on visual impaired subject.

Acknowledgement

The authors would like to express their gratitude to Universiti Teknologi Malaysia (UTM) for supporting this research and Ministry of Higher Education under Fundamental Research Grant Scheme (FRGS/1/2018/ICT02/UTM/02/9).

References

- [1] W. H. Organization, 2019. "Blindness and Vision Impairment," [Online]. Available: <https://www.who.int/news-room/factsheets/detail/blindness-and-visual-impairment>. [Accessed: 19-Oct-2019].
- [2] E. E. Freeman and E. W. Gower, 2013 "Visual Impairment," in *Women and Health (Second Edition)*, Second Edi., R. T. and K. M. R. Marlene B. Goldman, Ed. Elsevier. 1463–1472.
- [3] M. Hu, Y. Chen, G. Zhai, Z. Gao, and L. Fan. 2019., "An Overview of Assistive Devices for Blind and Visually Impaired People," *International Journal of Robotics and Automation*, 34(5) :
- [4] S. Sumit. 2018., "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way," [Online]. Available: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>. [Accessed: 19-Oct-2019].
- [5] Y. Lecun, Y. Bengio, and G. Hinton. 2015., "Deep Learning," *Nature*, 521(7553) : 436–444,
- [6] L. D. Dunai, M. C. Pérez, G. Peris-Fajarnés, and I. L. Lengua, 2017. "Euro Banknote Recognition System for Blind People," *Sensors (Switzerland)*. 17(1) : 184
- [7] "BrainPort Vision Pro. 2019 [Online]. Available: <https://www.wicab.com/brainport-vision-pro>. [Accessed: 19-Oct-2019].
- [8] I. Abu Doush and S. AL-Btoush, 2017. "Currency Recognition using A Smartphone : Comparison between Color SIFT and Gray Scale SIFT Algorithms," *Journal of King Saud University - Computer and Information Sciences.*, 29(4) : 484–492.
- [9] K. N. N. Hlaing and A. K. Gopalakrishnan, 2016. "Myanmar Paper Currency Recognition using GLCM and k-NN," *2016 2nd Asian Conf. Def. Technol. ACDT 2016*, 67–72,
- [10] N. A. Jasmin Sufri et al., 2017. "Image Based Ringgit Banknote Recognition for Visually Impaired," *Journal of Telecommunication, Electronic and Computer Engineering.*, 9(3–9) : 103–111,
- [11] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, 2016. "Road Crack Detection using Deep Convolutional Neural Network," *IEEE International Conference on Image Processing (ICIP)*. ICIP, 3708–3712
- [12] S. Albawi, T. A. Mohammed, and S. Al-Zawi, 2018. "Understanding of A Convolutional Neural Network," *International Conference on Engineering and Technology (ICET) ICET 2017*; 1–6,
- [13] P. Harsha Vardhan and P. S. Uma Priyadarsini, 2016. "Transfer Learning using Convolutional Neural Networks for Object Classification within X-ray Baggage Security Imagery," *Res. J. Pharm. Biol. Chem. Sci.*, 7 : 222–229
- [14] J. Zhu, S. Liao, D. Yi, Z. Lei, and S. Z. Li, 2015. "Multi-label CNN Based Pedestrian Attribute Learning for Soft Biometrics," *International Conference on Biometrics (ICB) ICB 2015*, 535–540
- [15] S. Miao, Z. J. Wang, and R. Liao, 2015. "A Convolutional Neural Network Approach for 2D/3D Medical Image Registration,"
- [16] S. Kumbhar, S. Patil, A. Nilawar, B. Mahalakshmi, and M. Nipane, 2019. "Farmer Buddy-Web Based Cotton Leaf Disease Detection Using CNN," 14(11) : 2662–2666,
- [17] "Convolutional Neural Network." [Online]. Available: <https://www.mathworks.com/solutions/deeplearning/convolutional-neural-network.html>. [Accessed: 09-Jul-2020].
- [18] J. E. Luján-García, C. Yáñez-Márquez, Y. Villuendas-Rey, and O. Camacho-Nieto, 2020 "A Transfer Learning Method for Pneumonia Classification and Visualization," *Applied Sciences.*, 10(8) : 2908
- [19] A. U. Tajane, J. M. Patil, A. S. Shahane, P. A. Dhulekar, S. T. Gandhe, and G. M. Phade, 2018 "Deep Learning Based Indian Currency Coin Recognition," *International Conference On Advances in Communication and Computing Technology (ICACCT)*, 130–134,
- [20] R. A. Minhas, A. Javed, A. Irtaza, M. T. Mahmood, and Y. B. Joo, 2019. "Shot Classification of Field Sports Videos Using AlexNet Convolutional Neural Network," *Applied Sciences.*, 9(3): 483
- [21] A. Titoriya and S. Sachdeva, 2019. "Breast Cancer Histopathology Image Classification using AlexNet," *2019 4th International Conference on Information Systems and Computer Networks (ISCON)*, 708–712,
- [22] S. Liawatimena et al., 2019. "A Fish Classification on Images using Transfer Learning and Matlab," 1st 2018 Indones. *Indonesian Association for Pattern Recognition International Conference (INAPR)*. 2018 - Proc. 108–112,
- [23] C. Szegedy, S. Reed, P. Sermanet, V. Vanhoucke, and A. Rabinovich, 2014. "Going deeper with convolutions," 1–12,
- [24] H. Hendrick, W. Zhi-Hao, C. Hsien-I, C. Pei-Lun, and J. Gwo-Jia, 2019. "IOS mobile APP for tuberculosis detection based on chest X-ray image," *Proc. ICAITI 2019 - 2nd International Conference on Applied Information Technology and Innovation (ICAITI)*, 122–125,
- [25] J. Ma, J. Rao, Y. Qiao, and W. Liu, 2018. "Sprouting Potato Recognition Based on Deep Neural Network GoogLeNet," *IEEE 3rd International Conference on Cloud Computing and Internet of Things (CCIoT)*, 502–505
- [26] M. Al-Qizwini et al., 2017. "Deep Learning Algorithm for Autonomous Driving using GoogLeNet," no. Iv,
- [27] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, 2018, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4510–4520.
- [28] T. Akiyama, Y. Kobayashi, Y. Sasaki, K. Sasaki, T. Kawaguchi, and J. Kishigami, 2019. "Mobile Leaf Identification System using CNN Applied to Plants in Hokkaido," *IEEE 8th Global Conference on Consumer Electronics (GCCE) 324–325*,
- [29] H. Yuan, H. Chen, S. Liu, J. Lin, and X. Luo, 2019. "A Deep Convolutional Neural Network for Detection of Rail Surface Defect," *2 IEEE Vehicle Power and Propulsion Conference (VPPC) - Proc.*, 6–9,
- [30] L. Ale, X. Fang, D. Chen, Y. Wang, and N. Zhang, 2020. "Lightweight Deep Learning Model for Facial Expression Recognition," 707–712,
- [31] A. A. Almisreb and M. A. Saleh, 2019. "Transfer Learning Utilization for Banknote Recognition: A Comparative Study Based on Bosnian Currency," *Southeast Europe Journal of Soft Computing* 8(1): 1–5
- [32] S. Kaur and M. Kaur, 2015. "Coin Recognition System with Rotation Invariant using Artificial Neural Network,"
- [33] S. M. M. Roomi and R. B. J. Rajee, 2015. "Coin Detection and Recognition using Neural Networks," in *IEEE International Conference on Circuit, Power and Computing Technologies*, ICCPCT 2015
- [34] G. Farooque, A. B. Sargano, I. Shafi, and W. Ali, 2017 "Coin Recognition with Reduced Feature Set SIFT Algorithm Using Neural Network," in *Proceedings - 14th International Conference on Frontiers of Information Technology*, FIT 2016, :93–98.
- [35] N. Capece, U. Erra, and A. V. Ciliberto, 2017. "Implementation of a Coin Recognition System for Mobile Devices with Deep Learning," *Proc. - 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. 186–192,
- [36] Z. Qiu, P. Shi, D. Pan, and D. Zhong, 2017. "Coin Detection and Recognition in the Natural Scene," *Proc. 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* .653–657.
- [37] H. Anwar, S. Anwar, S. Zambanini, and F. Porikli, 2019. "CoinNet: Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention," 1–34,
- [38] J. Kim and V. Pavlovic, 2016. "Discovering Characteristic Landmarks on Ancient Coins using Convolutional Networks," *Proc. - 23rd International Conference on Pattern Recognition (ICPR)*. 1595–1600,
- [39] I. Schlag and O. Arandjelovic, 2017. "Ancient Roman Coin Recognition in the Wild Using Deep Learning Based Recognition of Artistically Depicted Face Profiles." 2898–2906