



Visionary Banknotes Denomination Recognition: Empowering the Visually Impaired with Deep SqueezeNet Model

Samia A. El-Moneim Kabel^{1*}, Walid El-Shafai², Fathi E. Abd El-Samie³, Rania M. Ghoniem³

¹ Department of Communications and Computer Engineering, Tanta High Institute of Engineering and Technology (THIET), Tanta 31511, Egypt

² Security Engineering Lab, Computer Science Department, Prince Sultan University, Riyadh 11586, Saudi Arabia

³ Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

Corresponding Author Email: engsamia35@gmail.com

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.430105>

ABSTRACT

Received: 16 July 2025

Revised: 22 August 2025

Accepted: 26 September 2025

Available online: 28 February 2026

Keywords:

currency denomination, classification, SqueezeNet, Support Vector Machines, Convolutional Neural Network, confidently

Vision impairment affects billions of people globally, severely limiting their ability to identify everyday objects including the crucial task of recognizing currency denominations. This paper introduces an innovative solution designed to assist visually-impaired individuals in accurately distinguishing between various Egyptian banknotes: 1 EGP, 5 EGP, 10 EGP, 20 EGP, 50 EGP, 100 EGP, and 200 EGP. The proposed approach focuses on the robust detection and classification of Egyptian currency notes, irrespective of their orientation or position. At the core of this system is the use of a pre-trained deep neural network, SqueezeNet, which excels in extracting distinguishing features from banknote images. These features are then processed by a Support Vector Machines (SVM) classifier to determine the correct denomination. Key attributes of the system include its simplicity, user-friendliness, low cost, and adaptability to users of different ages and social backgrounds. Observed outcomes highlight the effectiveness of the model, achieving an outstanding 99.5% accuracy in banknote recognition. For benchmarking, the proposed system was compared to a conventional five-layer Convolutional Neural Network (CNN), which attained a slightly lower accuracy of 96%. This comparison clearly demonstrates the superior performance and practical viability of our system in supporting the independence of visually-impaired individuals by enabling them to confidently identify Egyptian banknotes on their own.

1. INTRODUCTION

Identifying currency is a major challenge for individuals with visual impairments. With millions of people affected by blindness worldwide, there is a strong need for assistive technologies that can recognize currency [1]. Existing methods fall into two categories: scanner-based and camera-based systems. Scanner-based systems require the full image of a banknote for recognition, whereas camera-based systems can identify currency from just a portion of the image. Our approach uses a camera-based system that recognizes Egyptian banknotes even when only a part of the image is captured.

Deep learning plays a key role in image recognition. It automatically extracts important features from images using convolutional layers, which are effective in learning complex patterns. However, deep learning models often need extensive datasets for effective training [2]. To address this issue, we use transfer learning, which leverages knowledge from pre-trained models. This method reduces both the need for large datasets and training times [3].

Our method depends on SqueezeNet, a compact pre-trained

deep learning model, to derive visual features from images [4]. Once extracted, these features are classified using and Support Vector Machines (SVM) classifier. We created a dataset of Egyptian currency that includes notes of one, five, ten, twenty, fifty, one hundred, and two hundred pounds. The images include various orientations and partial views of banknotes, making the system adaptable and effective in real-world conditions.

The remainder of this paper is structured as follows. Section 2 provides an overview of the related research. Sections 3 and 4 cover the fundamentals of transfer learning and the details of the SqueezeNet architecture. Section 5 outlines the application of SVM for classification. Section 6 describes the proposed framework, and Section 7 focuses on its implementation. Finally, the last section summarizes the main conclusions and insights drawn from this study.

2. RELATED WORKS

The past few years have seen several approaches developed to support banknote recognition for visually-impaired

individuals. These approaches vary in technique, but share the goal of accurately identifying currency. Ali and Manzoor [5] proposed an intelligent system for recognizing Pakistani currency. They collected 100 samples for each of five denominations (10, 20, 50, 100, and 1000 rupees), creating a diverse dataset. Their system relied on extracting distinctive features from each note and achieved an impressive average accuracy of 98.57%, successfully identifying eight denominations, including 10, 20, 50, 100, 500, and 1000 rupees.

Dunai et al. [6] introduced a portable assistive system for blind individuals to detect and recognize Euro banknotes. They used a Raspberry Pi with a No Infrared (NOIR) camera integrated into sunglasses and employed the modified Viola-Jones algorithm for detection and the Speeded Up Robust Features (SURF) algorithm for recognition. Their system achieved 84% detection accuracy and 97.5% recognition accuracy. Smary et al. [7] used traditional image analysis approaches, including segmentation, histogram normalization, and defining regions of interest. Their system matched images using cross-correlation with a reference dataset. It successfully identified Egyptian banknotes with an accuracy of 89%, while also reducing processing time.

Orabi et al. [8] proposed a layered framework for recognizing Egyptian currency. The first layer focused on watermark detection to prevent counterfeiting. The second layer handled segmentation to extract values, textures, and patterns. The third layer incorporated local features, and the final layer performed classification. Their system demonstrated robust performance, achieving a recognition accuracy of 93% across all tested denominations.

3. DEEP LEARNING AND TRANSFER LEARNING

Convolutional Neural Networks (CNNs) have become a cornerstone of image-based research, as widely demonstrated in the literature [9, 10]. These models excel at extracting complex low-level image features. Yet, training deep CNNs typically requires large datasets. To overcome this limitation, researchers often use transfer learning, based on models pre-trained with extensive data collections like ImageNet, which are adapted for new tasks. These pre-trained models can recognize general features such as edges and contours that are common across various types of images, making them transferable to different domains [11-14] (see Figure 1).

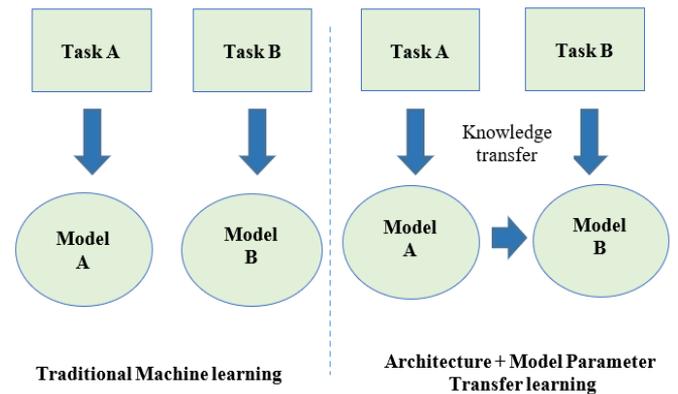
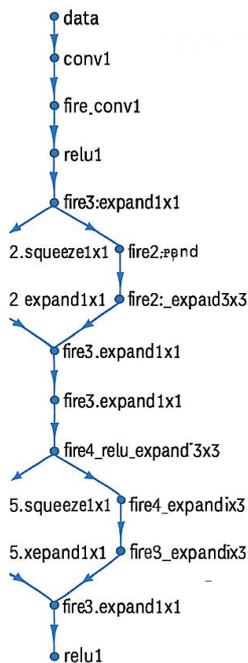


Figure 1. Traditional machine learning in comparison to transfer learning



ANALYSIS RESULT				
#	Name	Type	Activations	Learnables
1	data [227x23 images with zerror normalization]	Image input	227x227x3	Weights 1x3x3x64
2	(exams [758x3 convolutions with 1.2 x 32 and padding [0 0 0 0] [0 0 0]	Exams	703x227x3	Weights 1x1x6x64
3	pool1 relu1	Convolution	208x227x4	ReLU 1
4	fire2_expand1x1 [758x4 convolutions with 1.2 x 62 and padding [0 0 0 0] [0 0 0]	Convolution	56x56x64	Weights 1x1x6x64
5	pool2 relu1	Convolution	596x66x48	ReLU 1
6	fire3_expand2x1 [758x4 convolutions with 1.2 x 62 and padding [0 0 0 0] [1111]	Depth concatenation	56x56x128	Weights 1x3x8x64
7	pool3 relu1	Convolution	56x56x128	ReLU 1
8	fire5_expand3x1 [758x6 convolutions with 1.2 x 62 and padding [0 0 0 0] [1111] relu1	Pool1	56x56x48	Weights 1x3x6x64

Figure 2. SqueezeNet architecture

Transfer learning thus serves as a bridge, allowing knowledge from large-scale datasets to be applied to smaller, task-specific datasets. This not only reduces the need for extensive training data but also increases model efficiency and expands the applicability of deep learning across diverse use

cases [3].

In this study, we utilize the CNN architecture known as SqueezeNet [15] (see Figure 2). SqueezeNet is specifically designed to offer high performance, while significantly reducing model size, computational cost, and inference time.

Its compact architecture enables efficient feature extraction with minimal resource consumption, making it suitable for deployment on a wide range of devices and platforms. By balancing accuracy and efficiency, SqueezeNet supports practical deep learning applications, especially when computing resources are limited [16].

4. SQUEEZENET ARCHITECTURE

SqueezeNet is an 18-layer deep CNN with predetermined input dimensions of 224×224 [4]. It aims to extract detailed and meaningful visual features from images, which makes it ideal for transfer learning, especially in scenarios with limited training data. SqueezeNet core component is the Fire module, comprising two layers: a squeeze layer that precedes an expand layer. This structure, illustrated in Figure 3, plays a central role in the network efficiency. The squeeze layer minimizes the number of input channels by employing 1×1 convolutions, while the expand layer increases the network representational capacity through a mix of 1×1 and 3×3 convolutions.

This architectural design allows SqueezeNet to compress and expand information effectively, enabling rich feature extraction with fewer parameters. As a result, SqueezeNet achieves high performance with low computational cost, making it ideal for image analysis tasks on resource-constrained devices.

SqueezeNet architecture is composed of several key components designed for efficient feature extraction. It begins with a separate convolutional layer, conv1, followed by eight Fire modules (Fire2 through Fire9), and ends with a final convolutional layer, conv10. Together, these layers form a compact yet powerful structure for image analysis and classification as in Figures 2, and 3. Each Fire module consists of two parts: a squeeze layer with 1×1 convolution filters, and an expand layer that combines both 1×1 and 3×3 filters. As the network progresses from the early to the deeper layers, the set of filters in these modules gradually increases. This progressive growth enhances the model capacity to learn and process more complex visual patterns [15].

To further optimize computation and control spatial dimensions, **max-pooling** with a stride of 2 is applied at specific stages after **conv1**, **Fire4**, **Fire8**, and **conv10**. This down-sampling strategy reduces the size of feature maps, while preserving important information, improving both the efficiency and effectiveness of the network (see Figure 4 and Table 1).

The ReLU activation function plays a pivotal role within the architecture, being applied consistently between all the 'squeeze' and 'expand' layers integrated into the 'Fire module'. This activation function aids in boosting the model performance in capturing essential features, contributing to its efficiency.

To counteract the potential issue of overfitting, Dropout layers are thoughtfully introduced into the network structure after the 'Fire9' module. These Dropout layers act as a regularization mechanism, helping to avoid model overfitting and enhancing its ability to generalize to new data.

Notably, SqueezeNet distinguishes itself by opting not to employ any Fully-Connected (FC) layers in its architecture. This design strategy aligns with the network overall goal of reducing computational complexity and model size, while preserving performance.

During the training process, SqueezeNet is adjusted with a

learning rate initially set at 0.04, which is progressively decreased in a linear fashion. To optimize model parameters, the network leverages the Adam optimizer. Additionally, training batches are structured with a size of 32.

The utilization of SqueezeNet yields an impressive model size, with a remarkable $50 \times$ decrease relative to other networks [4]. This efficiency in model size, coupled with the network robust design and training strategies, underscores SqueezeNet as an impactful innovation in the area of deep learning and image analysis.

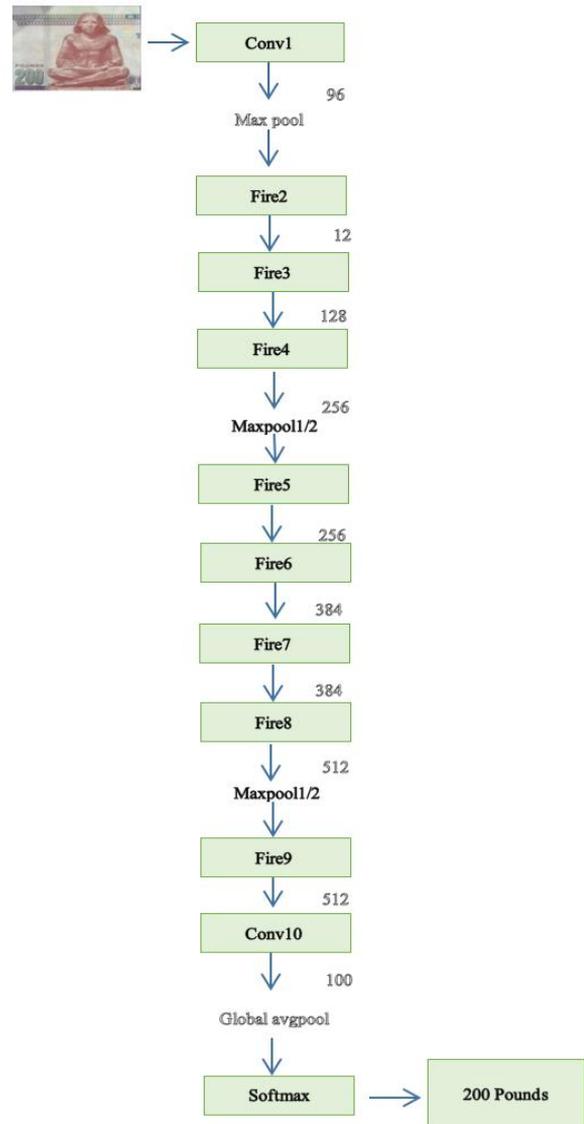


Figure 3. SqueezeNet for classification

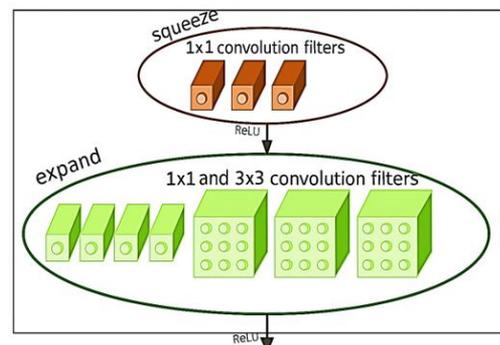


Figure 4. Squeeze and expand layers in SqueezeNet

Table 1. SqueezeNet detailed architecture

Layer Name/Type	Output Size	Filter Size / Stride (If not A Fire Layer)	Depth	s1×1 (1×1 Squeeze)	e1×1 (1×1 Expand)	e3×3 (3×3 Expand)	s1×1 Sparsity	e1×1 Sparsity	e3×3 Sparsity	# Bits	# Parameter Before Pruning	# Parameter After Pruning
input image	224×224×3											
conv1	111×111×96	7×7 / 2 (×96)	1				100% (7×7)			6	14,208	14,208
maxpool1	55×55×96	3×3 / 2	0									
fire2	55×55×128		2	16	64	64	100%	100%	33%	6	11,920	5,746
fire3	55×55×128		2	16	64	64	100%	100%	33%	6	12,652	6,352
fire4	55×55×256		2	32	128	128	100%	100%	33%	6bit	45,344	20,645
maxpool4	27×27×256	3×3 / 2										
fire5	27×27×384		2	32	128	128	100%	100%	33%	6	49,440	24,742
fire6	27×27×384		2	48	192	192	100%	100%	33%	6	104,880	44,800
fire7	27×27×384		2	48	192	192	100%	100%	33%	6	106,496	42,036
fire8	27×27×512		2	64	256	256	100%	50%	33%	6	188,992	77,581
maxpool8	13×13×512	3×3 / 2	0									
fire9	13×13×512		2	64	256	256	50%	100%	30%	6	197,184	77,581
conv10	13×13×1000	1×1 / 1 (×1000)	1				20% (3×3)			6	513,000	103,400
avgpool10	1×1×1000	13×13 / 1	0									
	Activations			Parameters			Compression info				1,248,424 Total	491,098 Total

5. SUPPORT VECTOR MACHINES

The SVM classifier is a prominent machine learning technique employed for the classification of sample data, leveraging input features acquired during the learning process [17]. The SVM is a highly valuable tool in discerning the optimal differentiating hyperplane among distinct classes within the target feature space.

One of SVM classifier fundamental strengths lies in its ability to discern intricate patterns within data, making it particularly well-suited for tasks involving binary classification. By meticulously identifying the most suitable hyperplane, SVM classifier optimally separates data points representing diverse classes. This powerful capability is at the core of the SVM classifier effectiveness in solving classification problems across various domains (see Figure 5).

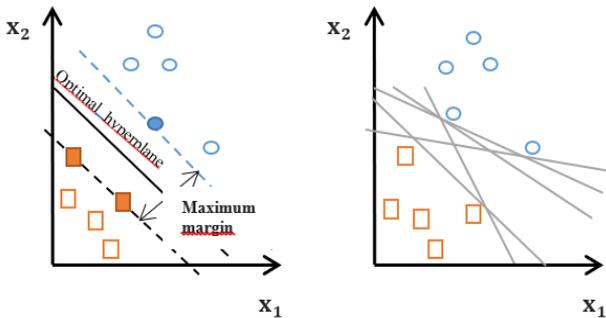


Figure 5. Support Vector Machines classifier concept

The primary goal here is to determine a hyperplane that possesses the maximum margin, which refers to the greatest separation distance between data points representing different classes. Maximizing this margin distance holds significant importance as it imparts robustness to the classification model, enhancing its ability to classify future data points with greater confidence and precision [18].

By striving to create a broader margin, we essentially carve out a more substantial and well-defined boundary between classes. This not only aids in the accurate classification of existing data but also fortifies the model resilience in handling

novel data points, ensuring that the classification decisions made are more reliable and effective [19].

6. PROPOSED FRAMEWORK

This study presents a new framework for the recognition of Egyptian banknotes. It has been meticulously developed to effectively identify seven distinct categories of Egyptian banknotes, encompassing denominations of one, five, ten, twenty, fifty, one hundred and two hundred Egyptian pounds. A visual representation of these denominations is provided in Figure 6 for reference.



Figure 6. Samples of the Egyptian currency

The framework primarily aims to enhance the precision and effectiveness of banknote recognition, eventually facilitating the seamless differentiation of Egyptian currency denominations. This achievement holds significant promise for various applications, where precise banknote recognition is of paramount importance.

These currency notes exhibit varying orientations, and it is noteworthy that only a portion of the banknote paper image

can be effectively captured, as depicted in the above figure. To assemble a comprehensive dataset for the study, images of different Egyptian banknotes, featuring diverse orientations and possible folding conditions, were meticulously collected. The dataset encompasses a total of 150 images for each denomination of Egyptian banknotes, with each image being standardized to a size of 224×224 pixels. To enhance model generalization and reduce overfitting, we applied a series of data augmentation techniques during training [20]. The following transformations were randomly applied to each input image:

- Rotation: Random rotations within the range of -10° to $+10^\circ$.
- Translation: Random horizontal and vertical shifts within the range of ± 3 pixels.
- Scaling: Random scaling of the image along both axes independently, with scale factors ranging from 0.9 to 1.1.

These augmentations simulate variations in viewpoint, position, and object size, helping the model to gain greater resilience to such distortions in real-world scenarios. We created an imageDataAugmenter object to define preprocessing parameters for image augmentation, including resizing, rotation, translation, and reflection. We applied random translations of up to three pixels both horizontally and vertically, and allowed rotations of the images by angles up to 10 degrees.

This randomization introduced variability in the amount of blur across the dataset. This augmentation technique helps the model generalize better by exposing it to images with varying degrees of blur, which commonly occur in real-world settings due to motion or defocus.

Moreover, in order to improve the robustness and generalization capability of the proposed model, color jitter augmentation was applied to the original banknote images. Specifically, the original image was first resized to 25% of its original size to reduce computational load. Then, color jittering was performed using the jitterColorHSV function, which randomly adjusts the brightness of the image within the range of $[-0.3, -0.1]$. This process simulates variations in lighting conditions that may occur during real-world image acquisition. The augmented image was subsequently saved for use in model training. This process helps the network learn invariant features under different brightness conditions, thereby enhancing the classification accuracy on images captured in diverse environments.

Prior to the utilization of these images in the proposed model, a preprocessing step is performed, involving image augmentation. This process ensures that the images are resized to conform to the desired output dimensions, aligning with the technical requirements of the recognition system. A visual depiction of the proposed system is thoughtfully provided in Figure 7 for better clarity and understanding.

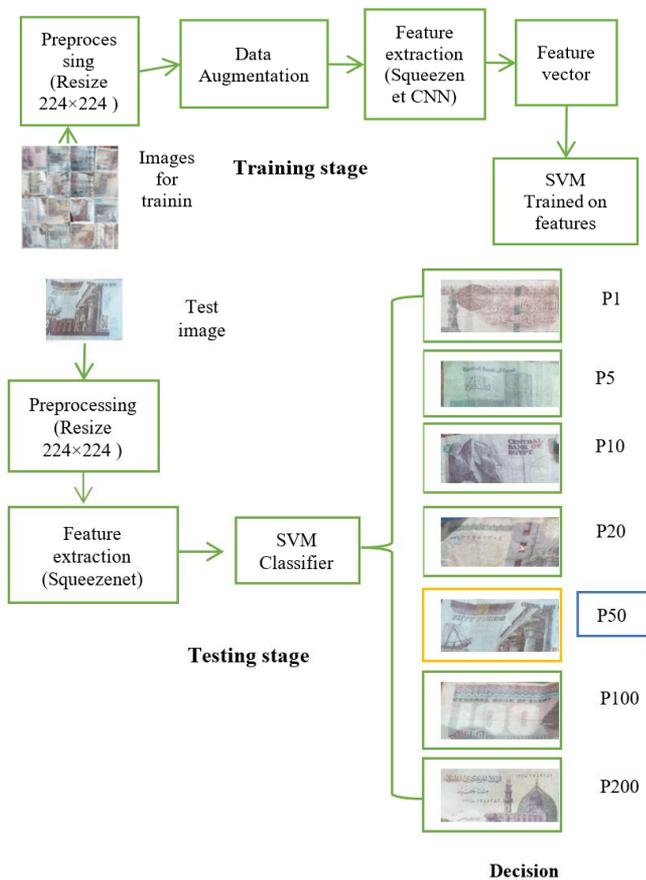


Figure 7. Proposed system for Egyptian currency recognition

To simulate realistic image degradations and increase the robustness of the model to blurred inputs, we applied Gaussian blur. Each image was first resized to 25% of its original dimensions to reduce computational load. The resized image was then convolved with a Gaussian and randomly sampled.

7. SIMULATION RESULTS

One of the significant advantages derived from the utilization of SqueezeNet lies in its remarkable efficiency in handling smaller datasets. Leveraging a pre-trained network, the system adeptly extracts features from the training images, with the dataset thoughtfully partitioned into 70% of the images for training, and the last 30% dedicated for testing.

In the training phase, the SVM classifier is effectively trained using the extracted features. During the testing phase, SqueezeNet continues to play a pivotal role, extracting features from the test images and subsequently forwarding them to the SVM classifier for accurate labeling, a process vividly represented in Figure 7 for clarity. To measure the efficacy pertaining to the proposed framework, several metrics have been used, namely accuracy, recall, and F1-score [21].

Accuracy reflects the overall accuracy by comparing the sum of true positives and true negatives against the total predictions [21].

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P}$$

where,

T_P : True Positives (correctly predicted positive class)

T_N : True Negatives (correctly predicted negative class)

F_P : False Positives (incorrectly predicted positive class)

F_N : False Negatives (incorrectly predicted negative class)

Recall measures how many actual positive instances were correctly predicted [21].

$$Recall = \frac{T_P}{T_P + F_N}$$

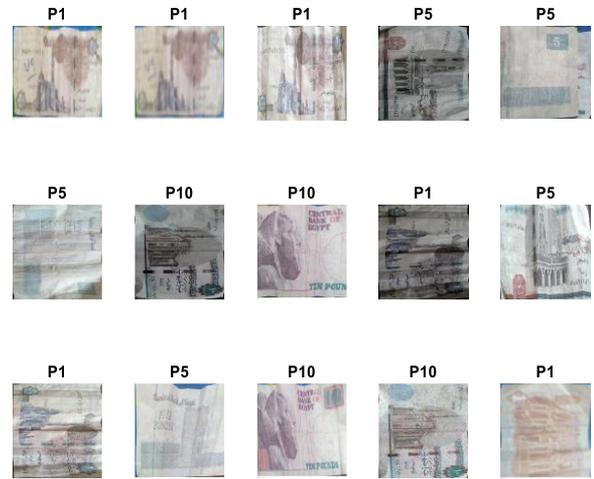
F1-Score is a metric that balances precision and recall through their harmonic mean, and is particularly useful in

imbalanced datasets [21].

$$F1 - Score = \frac{Precision \cdot Recall}{Precision + Recall}$$

Confusion matrix is a performance measurement tool used to evaluate the results of a classification model. It shows how often actual classes are correctly and incorrectly predicted.

The culmination of this meticulous process is revealed through simulation results, which unequivocally affirm the effectiveness of the proposed system. The system demonstrates an impressive recognition accuracy of 99% without augmentation and 99.5% with augmentation, as prominently depicted in Figures 8, and 9, respectively. These results underscore the system capacity to accurately identify Egyptian banknotes, reinforcing its practicality and utility in real-world applications.



		Confusion Matrix							
		P1	P10	P100	P20	P200	P5	P50	
Output Class	P1	41 14.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	P10	0 0.0%	41 14.3%	1 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	97.6% 2.4%
	P100	0 0.0%	0 0.0%	40 13.9%	0 0.0%	1 0.3%	0 0.0%	0 0.0%	97.6% 2.4%
	P20	0 0.0%	0 0.0%	0 0.0%	40 13.9%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	P200	0 0.0%	0 0.0%	0 0.0%	0 0.0%	40 13.9%	0 0.0%	0 0.0%	100% 0.0%
	P5	0 0.0%	0 0.0%	0 0.0%	1 0.3%	0 0.0%	41 14.3%	0 0.0%	97.6% 2.4%
	P50	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	41 14.3%	100% 0.0%
		100% 0.0%	100% 0.0%	97.6% 2.4%	97.6% 2.4%	97.6% 2.4%	100% 0.0%	100% 0.0%	99.0% 1.0%
		Target Class							

Figure 8. Confusion matrix of the proposed system (SqueezeNet+SVM classifier) without augmentation

		Confusion Matrix							
		P1	P10	P100	P20	P200	P5	P50	
Output Class	P1	28 14.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	P10	0 0.0%	27 14.0%	1 0.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	96.4% 3.6%
	P100	0 0.0%	0 0.0%	26 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	P20	0 0.0%	0 0.0%	0 0.0%	28 14.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	P200	0 0.0%	0 0.0%	0 0.0%	0 0.0%	27 14.0%	0 0.0%	0 0.0%	100% 0.0%
	P5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	28 14.5%	0 0.0%	100% 0.0%
	P50	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	28 14.5%	100% 0.0%
		100% 0.0%	100% 0.0%	96.3% 3.7%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	99.5% 0.5%
		Target Class							

Figure 9. Confusion matrix of the proposed system (SqueezeNet+SVM classifier) with augmentation

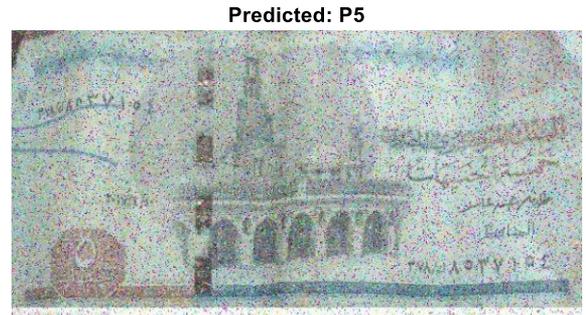


Figure 10. Samples of tested images with folding, blurring, reduced resolution and salt & pepper noise

Furthermore, a selection of the sample images that underwent testing is showcased in Figure 10. This visual representation offers a practical glimpse into the real-world applicability of the system and how it accurately performs in the context of recognizing Egyptian banknotes. These sample images serve to illustrate the system precision and its ability to effectively classify and identify banknote denominations.

The experimental findings confirm the efficiency of the proposed system in accurately recognizing banknotes, even

under challenging conditions such as partial occlusion, folding, blurring, or reduced image resolution and in the case of images with salt & pepper noise as in Figure 10. The system consistently achieves high recognition accuracy, demonstrating notable robustness and adaptability in practical, real-world environments.

The results in Table 2 show that data augmentation has a positive impact pertaining to the results of the proposed system. Without augmentation, the SqueezeNet+SVM model achieved an already high accuracy of 99%. However, with data augmentation, the accuracy is further improved to 99.5%, indicating enhanced generalization and robustness of the model when exposed to variations in the training data. This demonstrates that even small improvements in data diversity can contribute to more reliable classification, especially in real-world scenarios where input images may vary in quality or conditions.

Output Class	P1	P10	P100	P20	P200	P5	P50	Accuracy
P1	28 14.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.5%	96.6% 3.4%
P10	0 0.0%	26 13.5%	1 0.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	96.3% 3.7%
P100	0 0.0%	0 0.0%	25 13.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
P20	0 0.0%	0 0.0%	0 0.0%	28 14.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
P200	0 0.0%	1 0.5%	1 0.5%	0 0.0%	26 13.5%	0 0.0%	0 0.0%	92.9% 7.1%
P5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.5%	28 14.5%	0 0.0%	96.6% 3.4%
P50	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	27 14.0%	100% 0.0%
	100% 0.0%	96.3% 3.7%	92.6% 7.4%	100% 0.0%	96.3% 3.7%	100% 0.0%	96.4% 3.6%	97.4% 2.6%

Figure 11. Confusion matrix of the SqueezeNet+Random Forest method

Output Class	P1	P10	P100	P20	P200	P5	P50	Accuracy
P1	29 14.0%	2 1.0%	1 0.5%	0 0.0%	2 1.0%	0 0.0%	0 0.0%	85.3% 14.7%
P10	0 0.0%	21 10.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
P100	0 0.0%	1 0.5%	27 13.0%	0 0.0%	1 0.5%	0 0.0%	0 0.0%	93.1% 6.9%
P20	1 0.5%	0 0.0%	0 0.0%	30 14.5%	0 0.0%	0 0.0%	0 0.0%	96.8% 3.2%
P200	0 0.0%	1 0.5%	1 0.5%	0 0.0%	26 12.6%	0 0.0%	1 0.5%	89.7% 10.3%
P5	0 0.0%	4 1.9%	0 0.0%	0 0.0%	0 0.0%	30 14.5%	0 0.0%	88.2% 11.8%
P50	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	29 14.0%	100% 0.0%
	96.7% 3.3%	72.4% 27.6%	93.1% 6.9%	100% 0.0%	89.7% 10.3%	100% 0.0%	96.7% 3.3%	92.8% 7.2%

Figure 12. Confusion matrix of the SqueezeNet+Multi-Layer Perceptron (MLP) method

Output Class	P1	P10	P100	P20	P200	P5	P50	Accuracy
P1	28 13.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
P10	2 1.0%	28 13.5%	0 0.0%	1 0.5%	0 0.0%	0 0.0%	0 0.0%	90.3% 9.7%
P100	0 0.0%	0 0.0%	28 13.5%	1 0.5%	0 0.0%	0 0.0%	0 0.0%	96.6% 3.4%
P20	0 0.0%	0 0.0%	0 0.0%	28 13.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
P200	0 0.0%	0 0.0%	0 0.0%	0 0.0%	29 14.0%	0 0.0%	0 0.0%	100% 0.0%
P5	0 0.0%	1 0.5%	0 0.0%	0 0.0%	0 0.0%	30 14.5%	0 0.0%	96.8% 3.2%
P50	0 0.0%	0 0.0%	1 0.5%	0 0.0%	0 0.0%	0 0.0%	30 14.5%	96.8% 3.2%
	93.3% 6.7%	96.6% 3.4%	96.6% 3.4%	93.3% 6.7%	100% 0.0%	100% 0.0%	100% 0.0%	97.1% 2.9%

Figure 13. Confusion matrix of the MobileNet2 + Support Vector Machines (SVM) method

Table 2. Recognition accuracy of the proposed method with and without augmentation

Method	Accuracy %	
	Images with Salt & Pepper Noise	
	Without Augmentation	With Augmentation
Proposed method SqueezeNet+Support Vector Machines (SVM)	99	99.5

Table 3. Assessment of the proposed system relative to other different methods

Method	Accuracy %	Recall	F1 Score	Processing Speed (Sec)
2-layer-Convolutional Neural Network (CNN)	86.97	--	--	--
3-Layer-Convolutional Neural Network (CNN)	90.48	--	--	--
4-Layer-Convolutional Neural Network (CNN)	96.19	--	--	--
SqueezeNet +Random Forest	97.4%	97.37	97.38	7.4
SqueezeNet +Multi-Layer Perceptron (MLP)	92.8%	92.64	92.54	0.0195753
Proposed method SqueezeNet +Support Vector Machines (SVM)	99.5%	99.47%	99.47%	0.8696

To evaluate the performance of the proposed system, it was compared with several other classification algorithms, namely the Multi-Layer Perceptron (MLP) [22] and Random Forest classifier [23]. The same SqueezeNet model was used for feature extraction in all cases to ensure a fair comparison. The proposed system (SqueezeNet+SVM) achieves high scores in accuracy, recall, and precision, in addition to maintaining an acceptable inference time (see Figures 11-13 and Table 3).

The comparative results in Table 2 clearly demonstrate that the SqueezeNet+SVM combination (the proposed system) outperforms both the Random Forest and MLP classifiers in

terms of accuracy, precision, and recall, achieving an impressive 99.5% accuracy with high consistency across metrics. While Random Forest classifier performs well with a strong accuracy of 97.4%, it requires a longer inference time (7.4 seconds), indicating potential inefficiencies in real-time applications. On the other hand, the MLP classifier offers the fastest inference (~0.02 seconds), but lags in recognition performance, especially with a recall of 92.8%, showing that it may struggle with generalization. The MobileNetv2 gives an accuracy of 97.4%, and it requires a time of 0.91 seconds as in Figure 13.

The proposed SqueezeNet+SVM model provides the best balance, combining exceptional classification performance with a reasonable inference time (~0.87 seconds). This demonstrates its robustness and suitability for real-time or near-real-time banknote recognition tasks, especially under non-ideal conditions such as partial visibility or image degradation and images with Salt & Pepper noise.

Also, a comparative analysis was conducted, pitting the results obtained against those of CNN models featuring two,

three, and four layers. A summary of these comparative results is tabulated in Table 3. This comparative analysis underscores the superior performance of the proposed system in terms of banknote recognition accuracy, especially under challenging conditions, showcasing its practicality and versatility (see Figures 14-16).

While the proposed currency recognition system achieves high accuracy and robustness in classification, its practical deployment requires careful consideration of computational efficiency and hardware constraints. In particular, implementing the system on embedded platforms or mobile devices necessitates optimization of model size, memory footprint, and inference speed. The use of lightweight architectures such as SqueezeNet already contributes to lower computational demands (model size ~5 MB), making the model suitable for resource-constrained environments. Preliminary experiments indicate an average inference time of approximately 0.87 seconds on a standard desktop CPU, showing potential for real-time or near-real-time processing with further optimization.

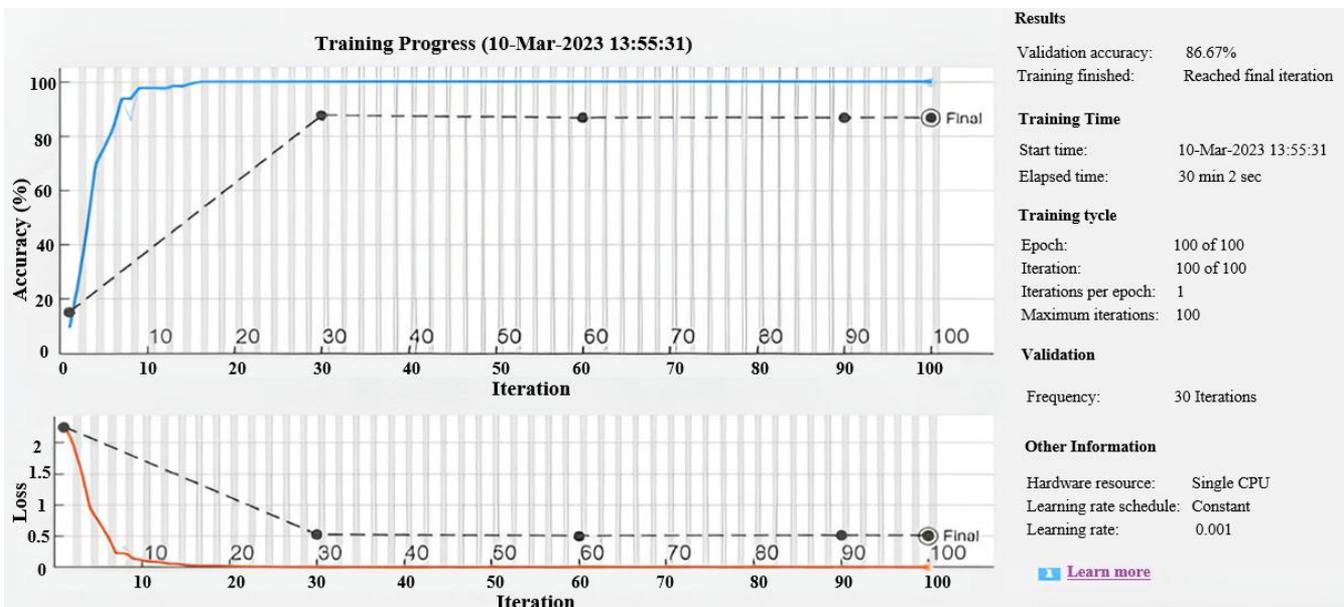


Figure 14. Two-layer Convolutional Neural Network results

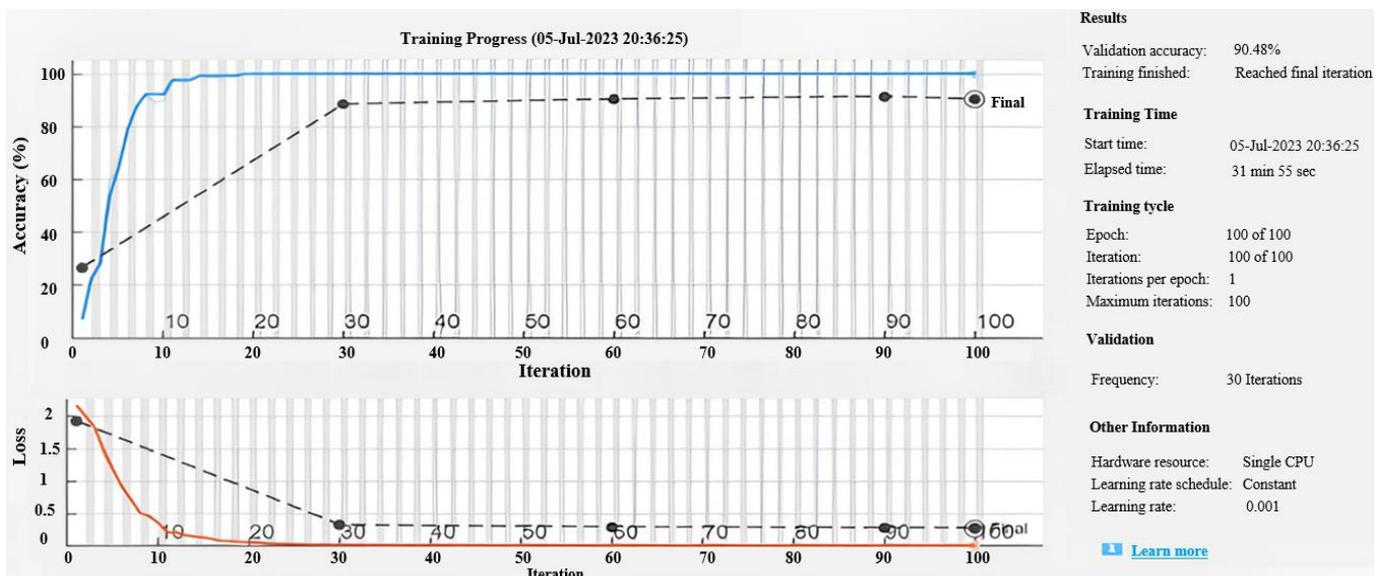


Figure 15. Three-layer Convolutional Neural Network results

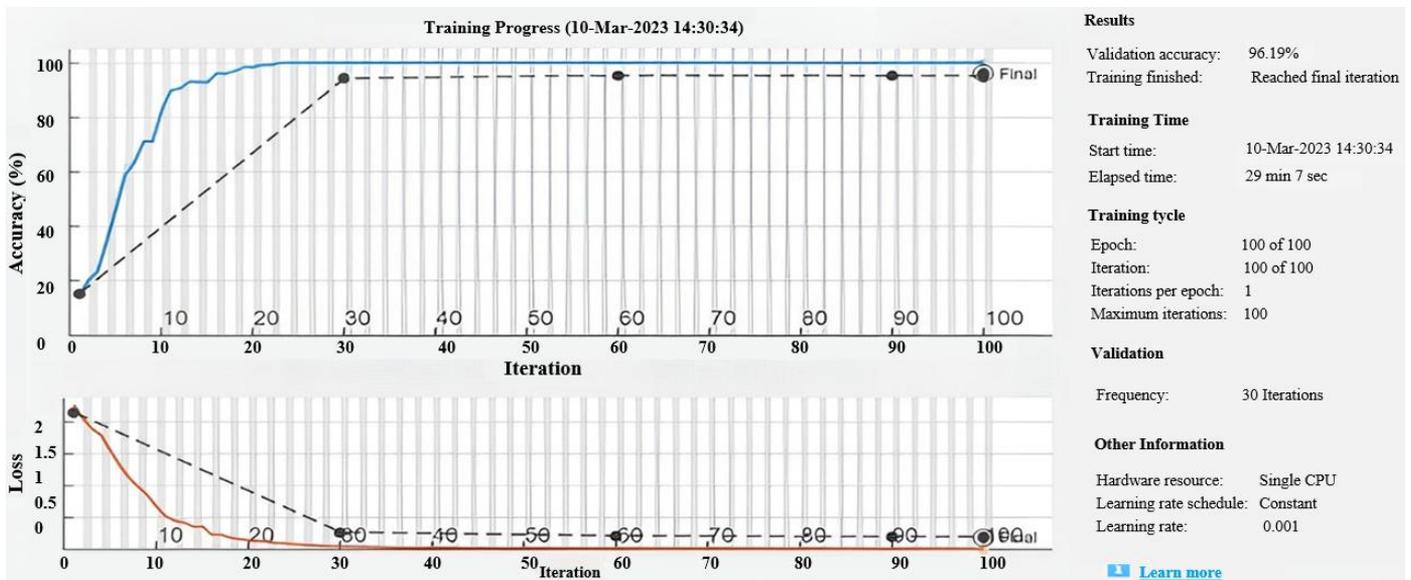


Figure 16. Four-layer Convolutional Neural Network results

To enhance compatibility with mobile and embedded devices, model compression techniques such as quantization, pruning, and conversion to formats like TensorFlow Lite or ONNX can be applied. These techniques allow for reduced model size and faster execution without significant loss in accuracy. Potential deployment targets include Android smartphones, which offer high processing capabilities, integrated cameras, and support for mobile Artificial Intelligence (AI) frameworks (e.g., TensorFlow Lite, ONNX Mobile). Alternatively, platforms such as the Raspberry Pi or NVIDIA Jetson Nano could serve as cost-effective embedded options, though they may require GPU acceleration or additional optimization to meet real-time constraints.

Future work will explore these deployments, ensuring the system usability in real-world settings, where fast and reliable currency recognition is essential.

8. CONCLUSION AND FUTURE WORK

This research paper presented a novel approach to facilitate the identification of Egyptian banknotes, particularly aimed at assisting individuals with visual impairments. The methodology hinges on the utilization of a pre-trained neural network, SqueezeNet, to extract salient features from currency images. Subsequently, these extracted features are directed into an SVM classifier. The proposed system demonstrates a remarkable level of accuracy, surpassing the performance of conventional CNN models. In fact, the accuracy achieved by our proposed system stands at an impressive 99.5% level, underscoring its efficacy and proficiency in the domain of banknote recognition. This remarkable level of accuracy is a testament to the system potential to make a significant positive impact on the lives of visually-impaired individuals, enhancing their ability to independently distinguish and identify Egyptian banknotes.

The future work can be in the following points:

- Extending the system to support additional currencies, which would involve collecting and annotating diverse datasets for different denominations and designs, and adapting the model to handle multi-currency recognition and classification.

- Developing a mobile application to bring the system into practical use, especially for visually-impaired users. This will require optimizing the model for resource-constrained devices, ensuring real-time performance, and integrating user-friendly interfaces and accessibility features.
- Robustness to real-world conditions: It is required to improve performance under extreme conditions such as crumpled or torn notes, partial visibility and handwritten markings or stamps.
- Support for Coins: It is recommended to extend the recognition system to handle coins, which present different challenges (e.g., reflective surfaces, engravings, wear).

FUNDING

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R138), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

REFERENCES

- [1] Nasir, H.M., Brahin, N.M.A., Aminuddin, M.M.M., Mispan, M.S., Zulkifli, M.F. (2021). Android based application for visually impaired using deep learning approach. IAES International Journal of Artificial Intelligence, 10(4): 879-888. <https://doi.org/10.11591/ijai.v10.i4.pp879-888>
- [2] Pak, M., Kim, S. (2017). A review of deep learning in image recognition. In 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT), Kuta Bali, Indonesia, pp. 1-3. <https://doi.org/10.1109/CAIPT.2017.8320684>
- [3] Pleded, J., Gedeon, T. (2022). Deep transfer learning for image classification: A survey. arXiv Preprint arXiv: 2205.09904. <https://doi.org/10.48550/arXiv.2205.09904>
- [4] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv Preprint arXiv: 1602.07360. <https://doi.org/10.48550/arXiv.1602.07360>

- [5] Ali, A., Manzoor, M. (2013). Recognition system for Pakistani paper currency. *World Applied Sciences Journal*, 28(12): 2069-2075, <https://doi.org/10.5829/idosi.wasj.2013.28.12.300>
- [6] Dunai D.L., Chillarón Pérez, M., Peris-Fajarnés, G., Lengua Lengua, I. (2017). Euro banknote recognition system for blind people. *Sensors*, 17(1): 184. <https://doi.org/10.3390/s17010184>
- [7] Semary, N.A., Fadl, S.M., Essa, M.S., Gad, A.F. (2015). Currency recognition system for visually impaired: Egyptian banknote as a study case. In 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA), Marrakech, Morocco, pp. 1-6. <https://doi.org/10.1109/ICTA.2015.7426896>
- [8] Orabi, A.A.A., Nasr, A.A., Al-Marakeby, A. (2020). Banknote recognition system for visually impaired people. *Journal of Al-Azhar University Engineering Sector*, 15(57): 1169-1177. <https://doi.org/10.21608/aej.2020.120377>
- [9] Browne, M., Ghidary, S.S., Mayer, N.M. (2008). Convolutional neural networks for image processing with applications in mobile robotics. In *Speech, Audio, Image and Biomedical Signal Processing using Neural Networks. Studies in Computational Intelligence*, pp. 327-349. https://doi.org/10.1007/978-3-540-75398-8_15
- [10] Zhang, Q., Yan, W.Q. (2018). Currency detection and recognition based on deep learning. In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, pp. 1-6. <https://doi.org/10.1109/AVSS.2018.8639124>
- [11] Tan, C.J., Wong, W.K., Min, T.S. (2020). Malaysian banknote reader for visually impaired person. In 2020 IEEE Student Conference on Research and Development (SCOREd), Batu Pahat, Malaysia, pp. 74-79. <https://doi.org/10.1109/SCOREd50371.2020.9250979>
- [12] Jangir, H., Raghav, N., Kashyap, N., Tanwar, P., Kumar, B. (2020). Homer: Cryptography based currency detection system for visually impaired people. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, pp. 762-768. <https://doi.org/10.1109/ICSSIT48917.2020.9214204>
- [13] Tong, C., Lian, Y., Qi, J., Xie, Z., Zhang, A., Feng, J., Zhu, F. (2017). A novel classification algorithm for new and used banknotes. *Mobile Networks and Applications*, 22(3): 395-404. <https://doi.org/10.1007/s11036-016-0771-z>
- [14] Awad, S.R., Sharef, B.T., Salih, A.M., Malallah, F.L. (2022). Deep learning-based Iraqi banknotes classification system for blind people. *Eastern-European Journal of Enterprise Technologies*, 1(2(115)): 31-38. <https://doi.org/10.15587/1729-4061.2022.248642>
- [15] Nettur, S.B., Karpurapu, S., Nettur, U., Gajja, L.S., Myneni, S., Dusi, A., Posham, L. (2025). UltraLightSqueezeNet: A deep learning architecture for malaria classification with up to 54x fewer trainable parameters for resource constrained devices. *IEEE Access*, 13: 89428-89440. <https://doi.org/10.1109/ACCESS.2025.3571696>
- [16] Hassanpour, M., Malek, H. (2019). Document image classification using SqueezeNet convolutional neural network. In 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Shahrood, Iran, pp. 1-4. <https://doi.org/10.1109/ICSPIS48872.2019.9066032>
- [17] Chen, Y., Councill, I.G. (2003). An introduction to support vector machines: A review. *AI Magazine*, 24(2): 105. <https://doi.org/10.1609/aimag.v24i2.1710>
- [18] Bridgelall, R. (2022). Tutorial on support vector machines. Preprints. <https://doi.org/10.20944/preprints202201.0232.v1>
- [19] Burges, C.J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2): 121-167. <https://doi.org/10.1023/A:1009715923555>
- [20] Kumar, T., Brennan, R., Mileo, A., Bendeche, M. (2024). Image data augmentation approaches: A comprehensive survey and future directions. *IEEE Access*, 12: 187536-187571. <https://doi.org/10.1109/ACCESS.2024.3470122>
- [21] Öztürk, T., Turgut, Z., Akgün, G., Köse, C. (2022). Machine learning-based intrusion detection for SCADA systems in healthcare. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 11(1): 47. <https://doi.org/10.1007/s13721-022-00390-2>
- [22] Al Bataineh, A., Kaur, D., Jalali, S.M.J. (2022). Multi-layer perceptron training optimization using nature inspired computing. *IEEE Access*, 10: 36963-36977. <https://doi.org/10.1109/ACCESS.2022.3164669>
- [23] Li, X. (2013). Using "random forest" for classification and regression. *Chinese Journal of Applied Entomology*, 50(4): 1190-1197. <https://doi.org/10.7679/j.issn.2095-1353.2013.163>

NOMENCLATURE

CNN	Convolutional Neural Network
SVM	Support Vector Machine
T_p	True Positives (correctly predicted positive class)
T_N	True Negatives (correctly predicted negative class)
F_p	False Positives (incorrectly predicted positive class)