

# Currency Identification Device for Visually Impaired People Based on YOLO-v5

ABHIRAMI B <sup>1</sup>, GOPIKA I<sup>1</sup>, HASAN A R<sup>1</sup>, IJAZ MUHAMED N J<sup>1</sup>, AISWARYA S S<sup>2</sup>

<sup>1</sup>UG Scholar, Department of Computer Science and Engineering

<sup>2</sup>Asst.Prof, Department of Computer Science and Engineering  
UKF College of engineering and technology Kollam, Kerala, India

\*\*\*

**Abstract:** A normal human being can easily identify and recognize any banknote, but it is extremely difficult for a visually impaired or blind individual to do so. Currency plays a significant role in our daily lives and real-time detection and analysis are required for any commercial transaction. Detecting banknotes seems to have become a necessity for a person especially for those who are blind or have visual impairments. There is a large number of visually impaired persons in India. The currency notes are almost same in thickness and some are also in similar size, which make difficult to visually impaired person to identify currency notes. The YOLO-v5 CNN model-based banknote detection and recognition device is proposed for this purpose, and it is fast and accurate. Photos of various denominations and circumstances were initially collected, and then these images were enhanced with various geometric and visual alterations to make the system more resilient. These augmented photos are then manually labelled, and training and validation image sets are created from them. Later, the trained model's performance was assessed on a real-time scene as well as a test dataset. The suggested YOLO-v5 model-based technique exhibits detection and recognition accuracy of 96.5 percent according to the test results. The entire system is self-contained and operates in real time.

**Keywords:** Deep learning; Object detection; Currency Note; YOLO-v5, Real-time detection; Currency Recognition.

## 1. INTRODUCTION

Over 2.2 billion people worldwide suffer from visual impairment, including 1 billion people with severe or acute distance vision impairment or blindness, the majority of whom are over 50 years old [1]. Glaucoma, cataracts, untreated presbyopia, and refractive error are the most common causes of debilitation. According to the World Health Organization, the number of persons affected by visual impairment will more than double by 2020. Assistive devices, such as walking dogs or white canes, are commonly used by visually impaired people. The white cane is most usually chosen for reasons such as low cost, portability, and widespread acceptance within the blind population. However, when faced with a range of obstacles and conditions in their daily lives, these assistance devices have their own limitations. People frequently regard such individuals as a burden and leave them to fend for themselves. As a result, the visually impaired individual requires the assistance gadget on a regular basis, which can assist in their day-to-day responsibilities and rehabilitation. People in their eighties and nineties have a higher risk of vision loss for those with visual impairments, the assistive system plays an important role in social situations. It feels difficult without this assistive equipment they're reliant on others. In addition, the cost of rehabilitation is out of reach for low-income people.

Currently, India has around 12 million blind people which makes India home to one-third of the world's blind population [2]. So, a real-time Indian currency detection device will be very much beneficial for visually impaired persons in India. Several frameworks and strategies for healthcare services have been created in the last decade. The goal of these improvements is to lower the cost of medical diagnosis while also assisting the health sector with technology that allows people to self-manage their lives more readily than ever before without the need for direct supervision from an expert. People with impairments, on the other hand, were not the primary beneficiaries of these achievements. However, there is a pressing need for technology that may help and assist people in their daily lives, improve their living in a simple way, and lead to independence. Visual Impairment has to be the most serious of these disabilities.

Currencies are significant as a medium of buying and selling goods. Each country has its own currency, which comes in a variety of colors, sizes, shapes, and patterns. Visually challenged people find it difficult to detect and count different denominations of currency. Due to continual use, tactile marks on the banknote's surface evaporate or fade away, making it

difficult for visually impaired people to detect and identify banknotes accurately by touch. Digital image processing is a large field that provides solutions to problems like these, in which patterns and identification markings are searched for and extracted, and then compared to actual banknote images. The suggested banknote detection and recognition system's key contribution is to build a simple standalone system easy to use, which will assist individuals in identifying banknotes in a real-time scenario. After augmentation and human annotation, the demanding self-built dataset is constructed, and transfer learning is done on the YOLO-v5 model.

In this work Section 2 begins with a review of related work, followed by a discussion of proposed system in Section 3. Finally, in Sections 4 includes the conclusion and future scope

## 2. RELATED WORK

Deep learning has become one of the most common methods for solving computing and prediction problems where the dataset is trained over neural networks in past few years. The speed and precision of these models differ. A CNN-based study is conducted on the folded banknotes [3], and all of the banknotes are of the same denomination. A computer vision-based system for automatic banknote recognition for the benefit of visually impaired persons [4], where SURF properties were exploited for recognition. This research is conducted on US banknotes with distinct image of the person on the banknote, making it easier to distinguish between those with the identical front note features as those in India.

YOLO-v3 CNN model-based banknote detection work is performed in [5]. In this work, pictures of various denomination Indian currency notes and in various conditions were first collected, and then these images were enhanced with various geometric and visual alterations. These augmented images are then manually annotated, and training and validation image sets are created from them. Later, the trained model's performance was assessed on a real-time scene as well as a test dataset. An identity dataset is created, and transfer learning is performed. After augmentation and testing on the YOLO-v3 model. The digital image has been manually annotated. Processing is a broad term that refers to the process of finding and extracting information. After that, patterns and identification signs are performed. Compare to images of actual banknotes

Introduces a revolutionary banknote image processing method based on the Free-From-Deformation (FFD) model in [6], which can aid in the processing of low-quality banknotes and minimize the false rejection rate. In [7] describes a new architecture for a banknote recognition and verification system based on neural networks for categorization and verification. A method for recognizing paper currency that uses a sequential deep neural network and data augmentation to improve accuracy is proposed [8]. Due to excessive computation, the proposed job was designed to address small data problems and was unable to perform real-time processing. [9] discusses a money identification system for Ethiopian banknotes that uses a support vector machine and recognizes the front part of the coin well.

Deep convolutional neural networks are used to classify Turkish lira banknotes, which are developed and trained using the DenseNet-121 architecture [10]. Applies image processing techniques on the front and back sides of Myanmar currency (kyats) in three denominations in [11]. Zernike moments were employed for feature extraction, and the k-nearest neighbor method was applied for classification. Also, in [12] uses a neural network to overcome these types of problems for visually challenged people. Their findings suggest that further major research on cognition frameworks and neural activity could lead to more significant results in these types of challenges. In [13] describes a portable technology allowing blind persons to identify and recognize Euro currencies. The modified Viola-Jones algorithms are used to detect banknotes. The recognition of currencies centered on a modified Viola-Jones algorithm [14] and SURF (Speed Up Robust Features) algorithm [15].

According to article [16] the YOLO-v5 network [17] has higher accuracy and speed. The YOLO-v5 network [17] is a more sophisticated version of the YOLO network [18] and YOLO-v3 network [19]. YOLO-v5's network architecture which is divided into three sections: CSPDarknet is the backbone, PANet is the neck, and Yolo Layer is the head. The data is first supplied into CSP Darknet, which extracts features, and then into PANet, which fuses them. Finally, Yolo Layer gives you the results of your detection (class, score, location, size). YOLO-v5 is smaller, faster, and more accurate than other YOLO networks. As a result, a YOLO-v5 based CNN model can be used to create a rapid and accurate banknote detection and recognition system for visually impaired and blind persons to assist them in their daily life. As a solution, a YOLO-v5 based CNN model can be used to design a quick and precise currency identification and tracking device for visually impaired and blind individuals to support them in their everyday life.

### 3. PROPOSED SYSTEM

#### A. IMAGE AUGMENTATION AND ANNOTATION

A camera with a resolution of 1280\*720 pixels is utilized in this method to capture images in various settings such as occlusion, illumination (lighting from the front, side, and back), and occlusion and so on. In total, 3720 banknote pictures were scanned. Images were also obtained from the camera and the web in many file types (.jpg,.jpeg,.png, etc.). This collection of data is called training, validation, and certification for various banknotes are separated into three categories. Set of trials for the competition, 65-70 percent of the photographs were chosen at random. Each denomination of banknotes has a training set and a resting set for testing and validation sets. It is done to enhance the image additionally, to create a huge image dataset that avoids Preventing overfitting in the training model and retaining the necessary details images from the data set After then, the resolution of these 3720 photos was enhanced to a total of 10,000 photos.

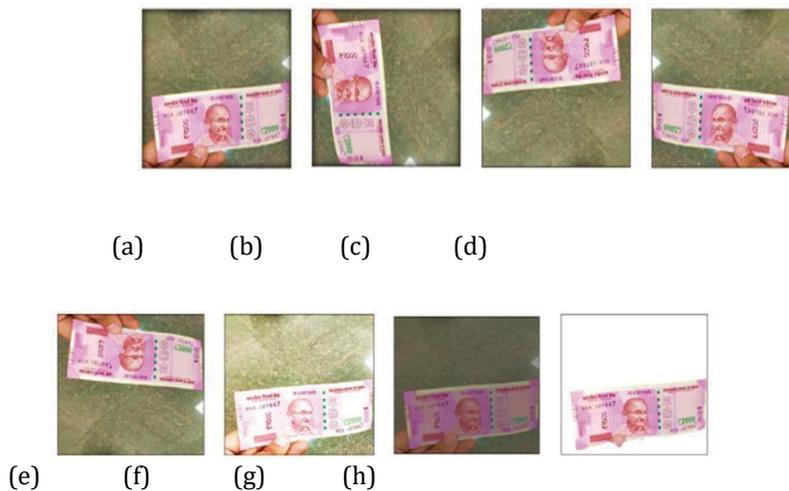


Fig 3.1. Various image augmentation methods on collected images, (a) Actual Image, (b) 90° Horizontal Tilting, (c) 180° Horizontal Tilting, (d) Horizontal Turn, (e) Vertical Turn, (f) Increased Horizontal Tilting(h) Background Elimination, (g) Noise Extension, (h) Brightness

Various image augmentation techniques are used to generate the dataset for all banknote categories. A number of image enhancement techniques are available such as resizing, shear, rotation, brightness, reflection, color elimination and translation of noise and background. To make it more interesting, make the changes shown in Fig. 1. The image data is augmented with transformation and augmentation techniques to increase the complexity. The images of several people after they've been augmented. The banknotes are sorted and numbered in a certain order, further the images were numbered and manually annotated. "Labellmg" is a utility that allows you to label images. For each dataset,.xml annotation files are saved. images. Bounding boxes are placed around the banknotes in order to do this. every image was drawn. Various denomination notes were manually categorized by different labels on each image. Positive examples to avoid over-fitting the neural network, images of banknotes that were insufficient or unclear were not accepted. Moreover, an area of the banknote at the corner was occluding more than 80% of the time, not used for a portion of the image with less than 20% area annotation. Annotation files and the training dataset. The YOLO-v5 detection model was trained using this data, while the rest of the photos were used to create a validation dataset for confirming the YOLOv5 model's detection and training performance. Figure 3.2 shows the denominations used in training.



(a)



(b)



(c)



(d)



(e)



(f)

Fig. 3.2. Samples of Banknote Images in Dataset (a) 10 (b) 20 (c) 50 (d) 100 (e) 200 (f) 500

## B. TRANSFER LEARNING AND MODEL TRAINING

YOLO is a real-time object recognition technique that uses neural networks. Because of its speed and efficiency, this algorithm is very popular. It has been used to identify traffic signals, pedestrians, parking metres, and animals in a variety of applications. This algorithm which detects and recognises different items in an image (in real-time). Object recognition in YOLO is performed as a regression problem, and the class probabilities of the discovered images are returned. Convolutional neural networks (CNN) are used in the YOLO method to recognise objects in real time. To detect objects, the approach just takes a single forward propagation through a neural network, as the name suggests. This indicates that a single algorithm run is used to forecast the entire image. At the same time, the CNN is utilised to predict various bounding boxes and class label. There are several variations of the YOLO algorithm. Tiny YOLO and YOLOv3 are two popular examples. For the following reasons, the YOLO algorithm is significant:

- Speed: Since it can predict objects in real time, this system improves detection speed.
- High precision: YOLO is a prediction technique that produces appropriate results with low background errors.
- Learning abilities: The algorithm has enriched learning abilities, allowing it to learn and apply object representations to object detection.

The YOLO version5 is most advanced object detection algorithm which is launched by Ultralytics, which was released in June 2020. It's a brand-new convolutional neural network (CNN) that identifies things in real time with great precision. This method uses a single neural network to evaluate the full image, then divides it into sections and predicts bounding boxes and probabilities for each component. These bounding boxes are weighted based on the estimated likelihood. In the sense that it produces predictions after only one forward propagation through the neural network, the approach "looks once" at the image. After non-max suppression, it provides discovered items, which assure the object recognition algorithm only identifies each object once. YOLOv5's initial launch, which is faster in performing, and simple to use. YOLOv5 is incredibly user-friendly and comes "out of the box" ready to use on bespoke objects. The majority of the performance gain in YOLOv5 is acquired from PyTorch training processes, while the model architecture remains similar to that of YOLOv4. The goal is to create an object detector model that is extremely fast (on the Y-axis) in terms of inference time (X-axis).

### Yolo v5's Benefits

- It's about a third of the size of YOLOv4 (27 MB vs 244 MB)
- It's around 180 percent faster than YOLOv4 (140 FPS vs 50 FPS)
- On the identical assignment, it's about as accurate as YOLOv4 (0.895 mAP vs 0.892 mAP)

### Yolo v5's drawbacks

The fundamental issue is that, unlike prior YOLO versions, there is no official paper for YOLOv5. Furthermore, because YOLO v5 is still in development and we receive frequent updates from ultralytics, various parameters may be updated in the future.

The YOLO network transforms the detection problem into a regression problem. By using regression, it creates a coordinated bounding box and probability for each class. Rather than employing local classifier-based systems, which have a high failure rate. After applying the scoring zones which are referred for detection. Model to an image is a variety of locations and scale. When compared to other methods, this improves detection speed. Faster R-CNN, RCNN by 100 and 1000 are examples of such approaches, time intervals, respectively the accuracy and detecting speed are significantly improving. YOLO is more common, rather than applying the concept to various situations, the detection model is applied to all image locations.

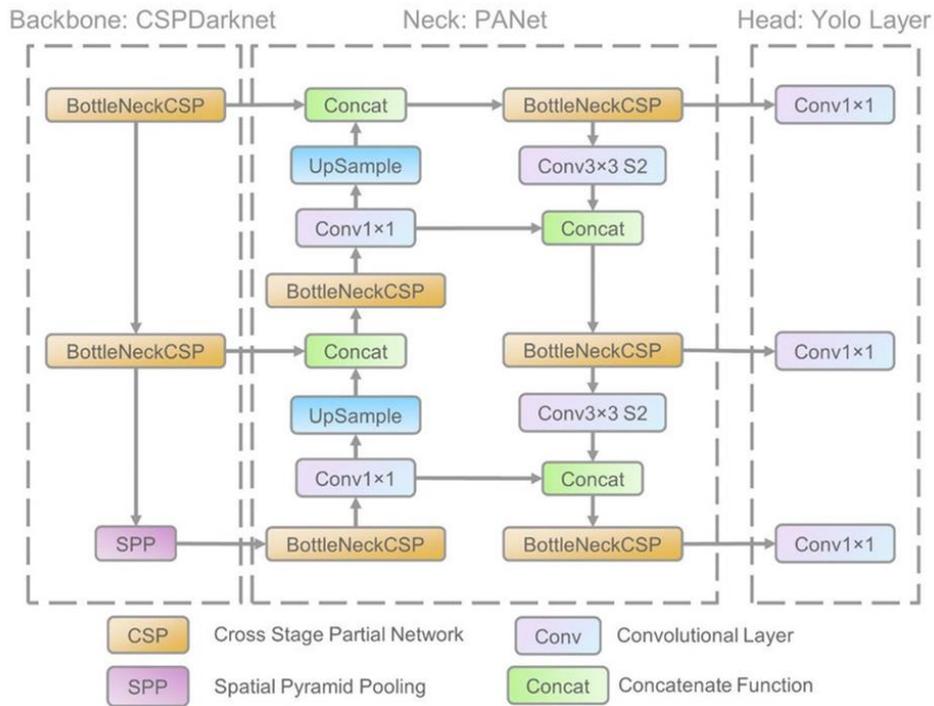


Fig 3.2.YOLO-v5 Network structure

Its architecture was basically made up of three components, namely-

1. **Backbone:** Backbone is typically used to extract essential features from an imagery input. In YOLO v5, the CSP (Cross Stage Partial Networks) backbone is utilised to extract rich in valuable attributes from an input image.
2. **Neck:** The Model Neck is primarily used to generate pyramids feature. When it comes to object scaling, feature pyramids help models generalise successfully. It makes it easier to recognise the same thing in different sizes and scales. Feature pyramids are quite useful for assisting models in performing well on previously unknown data. Other models, such as FPN, BiFPN, and PANet, utilise feature pyramid methods in various ways. In YOLO v5, PANet is utilised as a neck to generate feature pyramids.

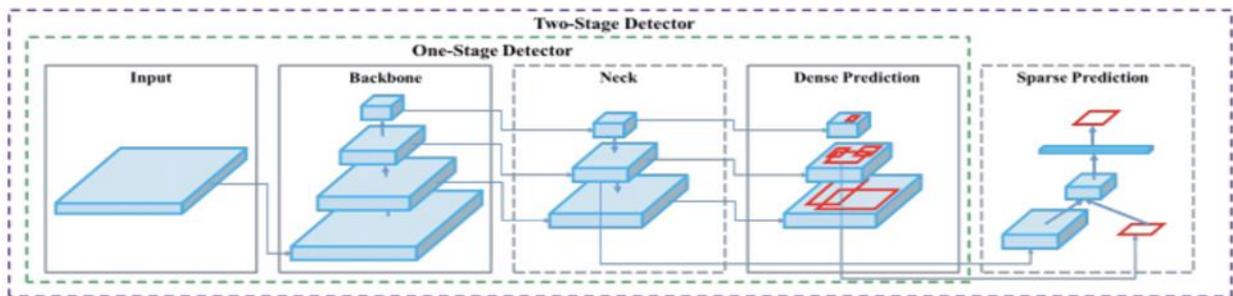


Fig.3.3. Object Detection

3. Head: The final detection step is largely handled by the model head. It creates final output vectors with class probabilities, objectness scores, and bounding boxes using anchor boxes.

The YOLO v5 model's head is identical to that of the preceding YOLO V3 and V4 variants. There are two kinds of architectural object sensing concepts. The YOLOv5 network is divided into three sections. a. Backbone - A CNN layer gathers picture features at various scales. b. Neck - Layers that combine image features and forward them to prediction. c. Head - Performs localization and classification using characteristics from the neck. The incoming picture features are compressed using a feature extractor (Backbone) and then forwarded to the object detector in all object detection systems (including Detection Neck and Detection Head). The Detection Neck (or Neck) functions as a feature aggregator, combining and mixing the features created in the Backbone in order to prepare for the next step in the Detection Head (or Head).

Figure 3.4 shows the flow diagram for the training model based on the collection of banknote pictures and augmentation, as well as the flow chart for the suggested system's operation. The image input from the camera and the frame is captured and then pre-processed before being delivered. When the image is presented to the client, model which is a text output, is then transformed into voice for different labels or banknote recordings, which may then be played. As a result, each detected and recognised banknote's audio output is relayed to the visually impaired person via speaker or earphones. Which was trained on the dataset, produces output for recognised objects. Generate the label and boundary box for each banknote of the corresponding notes. It is capable of detecting multiple numbers. There is no such form of banknote that contains all of the banknotes in a single image.

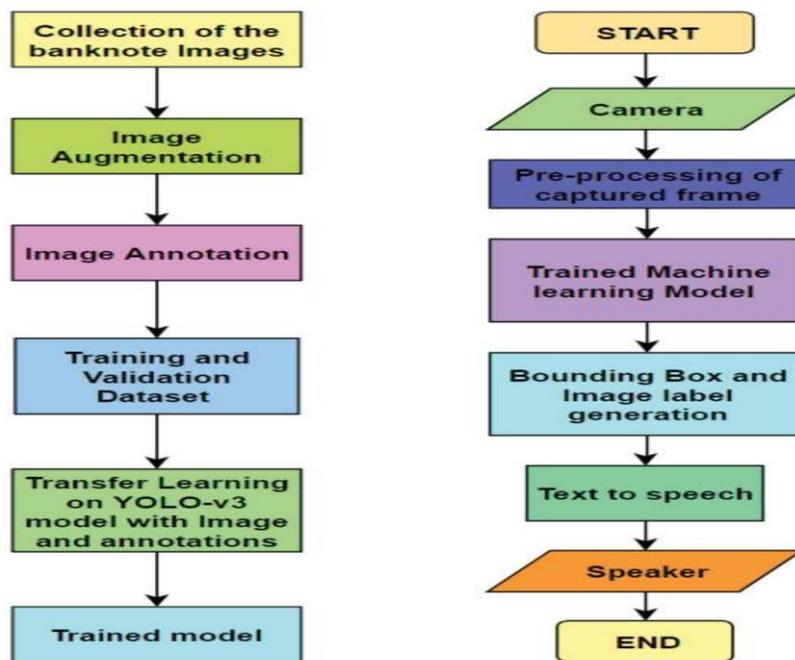


Fig 3.4. Model training and Proposed system

### C.FAKE NOTE DETECTION

The development of color printing technology has accelerated the manufacture and duplication of counterfeit money notes on a huge scale. As a result, the issue of fake notes instead of the genuine ones has skyrocketed. Now a days a lot of visually impaired people are cheated by simple fake currencies. In order to resolve the problem, we proposed a UV based system which is integrated into the currency detection device. The opacity of the original currency fake one is different, so the intensity of the UV light penetrated through the original currency note is less as it is compared with fake currencies. So the

difference in output is taken for evaluation which indicates whether the inserted one is fake or not. Alarm system implemented in the device produces a beep sound when it catches a fake one.

## ACKNOWLEDGEMENT

We would wish to thank everyone who had contributed to the successful completion of the project. We would like to express our gratitude to our project guide Ms.Aiswarya.SS Assistant Professor in Computer Science and Engineering Department who gave valuable suggestion and guidance for our project. We express our deep felt gratitude to beloved HOD Dr.Ramani ,Head of the Department for providing necessary information regarding the project and also her support in completing it. We also thank our project coordinator, Mr.Jithin Jacob, Assistant Professor who gave expert supervision, encouragement and constructive criticism amidst his busy schedule throughout the project. We also grateful to all the authors of books and papers which have been referred to publish this paper.

## CONCLUSION AND FUTURE SCOPE

In this paper, a standalone and real-time bank currency detection and identification model based on the YOLO-v5 model is proposed. The dataset is a collection of banknotes created under various conditions such as rotation, obstruction, brightness level, scale, and so on. The wide range of image dataset which helps to make detection and identification easier, Robust and precise. The entire system is self-contained and does not need access to the internet to conduct its recognition function. This device utilized for real-time banknote detection and identification with a live video stream recognition system, With the assistance of this device any visually challenged or blind individual can easily identifies the note. On the dataset, there was detection and a 100% average recognition rate and is capable of recognizing banknotes in a variety of situations like wrinkled or ripped currency notes, as well as partial occlusions can also detected by the device. Besides the currency recognition, the device uses UV beam approach which detect the counterfeit notes successfully. The future scope on this system will be focused on improving the system's performance as well as expanding the training dataset by including more banknotes from various countries. Additionally, work will be done to design an interactive system which including more functions like e wallet, which automatically count and fund the notes.

## REFERENCE

- [1] Blindness and vision impairment, Available Online: Blindness and vision impairment (who.int) (Accessed on 25 April 2022)
- [2] Estimation of blindness in India from 2000 through 2020: implications for the blindness \_control policy - PubMed (nih.gov) (Accessed on 25 April 2022)
- [3] M. JIAO, J. HE and B. ZHANG, "Folding Paper Currency Recognition and Research Based on Convolution Neural Network," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 18- 23.
- [4] F. M. Hasanuzzaman, X. Yang and Y. Tian, "Robust and Effective Component-Based Banknote Recognition for the Blind, " in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no.6, pp. 1021-1030, Nov.2012.
- [5] Rakesh Chandra Joshi, Saumya Yadav and Malay Kishore Dutta, "YOLO-v3 Based Currency Detection and Recognition System for Visually Impaired Persons" 2020 International Conference on Contemporary Computing and Applications (IC3A)
- [6] L. Tang, Y. Jin and M. Du, "A Hierarchical Approach for Banknote Image Processing Using Homogeneity and FFD Model," in IEEE Signal Processing Letters, vol. 15, pp. 425- 428, 2008.
- [7] P. Priami, M. Gori and Frosini, "A neural network-based model for paper currency recognition and verification," in IEEE Transactions on Neural Networks, vol. 7, no. 6, pp. 1482-1490, Nov. 1996

- [8] H. Vo and Hoang, "Hybrid discriminative models for banknote recognition and anti-counterfeit," 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, 2018, pp. 394-399.
- [9] M. T. Kebede Bahiru M, B. Ramani and M. E. Ayalew Tessfaw , "Ethiopian Banknote Recognition and Fake Detection Using Support Vector Machine," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2018, pp. 1354-1359.
- [10] G. Baykal, U. Demir, I. Shyti and G. Ünal, "Turkish lira banknotes classification using deep convolutional neural networks," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.
- [11] A. K. Gopalkrishnan and K. N. Oo, "Zernike Moment Based Feature Extraction for Classification of Myanmar Paper Currencies," 2018 18th International Symposium on Communications and Information Technologies (ISCIT), Bangkok, 2018, pp. 1-6.
- [12] A. Khashman and O. K. Oyedotun, "Banknote recognition: investigating processing and cognition framework using competitive neural network", Cognitive Neurodynamics, Springer, vol. 11, issue 1, pp. 67-79, 2017
- [13] Mónica & Peris-Fajarnés, Guillermo & Lengua, Ismael, Dunai, Larisa & Chillarón Pérez, (2017). Euro Banknote Recognition System for Blind People. Sensors. 17. 184. 10.3390/s17010184.
- [14] Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 8–14 December 2001.
- [15] Bay, H.; Ess, A.; Tuytelaass, T.; van Gool, L. Speed-Up Robust Features (SURF). Comput. Vis. Image Underst. J. 2008, 110, 346– 359.
- [16]: Nepal, U.; Eslamiat, H. Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. Sensors 2022, 22, 464.
- [17] M. Karthi, V. Muthulakshmi, R. Priscilla, P. Praveen and K. Vanisri, "Evolution of YOLO-V5 Algorithm for Object Detection: Automated Detection of Library Books and Performace validation of Dataset," 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICES), 2021, pp. 1-6
- [18] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 779-788.
- [19] J. Redmon and A. Farhadi. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.