Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 08 Issue 01
2024

Research Article
https://doi.org/10.56979/801/2024

YOLOv9-Based YOLO-Enhanced Smart Glasses for Real-Time Recognition of Pakistani Currency: Empowering the Visually Impaired

Mamoona Khalid[1], Hafiz Burhan Ul Haq[1*], Zara Janjua[1], Hafiz Muhammad Muneeb Akhtar[1], and Hafiz Muhammad Yousaf[2]

[1]Department of Information and Technology, Faculty of Computer Sciences, Lahore Garrison University, Lahore, Pakistan.
[2]Department of Information Engineering and Computer Science, University of Trento, Italy.
*Corresponding Author: Hafiz Burhan Ul Haq. Email: burhanhashmi64@lgu.edu.pk

Received: March 24, 2024 **Accepted:** September 26, 2024

Abstract: People who are blind or visually impaired may find it difficult to manage daily tasks, particularly when it comes to knowing the different denominations of cash. We provide an innovative approach to tackle this problem, enabling real-time cash detection through the combination of wearable technology and deep learning. In this work, a deep learning model-integrated smart glasses system powered by a Raspberry Pi IoT device is introduced. The cutting-edge YOLOv9 algorithm is used by the system to accurately recognize cash notes. Training the model included using an extensive dataset of 3,611 photos with seven distinct rupees denominations: 10, 20, 50, 100, 500, 1000, and 5000. The smart glasses immediately alert the user of the denomination when they detect a bank note by means of voice feedback. Our technology boosts the transaction experience and gives visually impaired people more independence in their financial operations, all while maintaining an accuracy rate of up to 95%. As a practical and dependable instrument for daily transactions, the system's lightweight and portable design guarantees simplicity of use in a variety of environments. Providing a reliable and effective method for cash identification, this research marks a substantial breakthrough in assistive technology.

Keywords: Artificial Intelligence; Deep Learning; Currency Identification; Internet of Things; Text- to-Speech Conversion.

1. Introduction

Blindness or poor vision is a critical global problem that affects a significant portion of people on a worldwide scale. Even the World Health Organization (WHO) estimates that approximately two hundred million human beings have some difficulties with either near or distance vision. One hundred million of them had some form of treatment [1]. One of the findings carried out between 2019 and 2021 in the Third National Survey of Blindness showed that there are 9 028 073 people with vision difficulties within the population of the country, which varies from mild to totally blind students. Among them, 484,027 have lost their vision, which is not the same number recorded a year ago. In 2004, the population of the visually impaired was assessed to be 5 million. As per the study of Prof. Asad Aslam Khan, the Public Facilitator of Pakistan's Anticipation of Visual Impairment Program, the visual deficiency older females are 2.07%and the males 1.98%. The above figure, when made complete, is summarized to 2.02%. [2]. Blind individuals face many unique challenges in their daily lives. They often rely on alternative and flexible approaches to overcome these challenges. One of the difficulties, therefore, presents them with a major challenge whenever they are engaged in financial transactions. For the identification of currency notes, a blind person is fully dependent on another person. In this situation, there is a likelihood that another person can deceive the blind person. In any financial transaction, a visually impaired person must be able to determine the currency notes on his or her own without the help of another person. For the identification of currencies, various methods and techniques have been developed to enhance feasibility and reduce reliance on others. Some of these include tactile features such as raised dots or lines that are embedded on the banknotes to depict denominations of money, notch or corner-cut models, and embossed symbols. Other countries try to meet the needs of the blind as well as those who have low vision, whereby they design money so that the value of a banknote is expressed based on one or more features. Different dimensions on the amount of bills are one of the common techniques. A visually impaired person can determine the values of bills by differentiating them based on size using a template or by touching over time. Different colors for currency denominations are also used in other parts of the world besides the United States. It helps many people with poor sight. [3]. With the headway of innovation, electronic currency reader systems and frameworks, versatile applications considering picture handling strategies were created with perceptible and material criticism reporting the money group. Also, the combination of Internet of Things (IoT) and artificial intelligence (AI) advances is significantly developing currency recognition for people with visual impairments. Portable applications with computer-based intelligence help and IOT devices have been utilized to recognize the money denomination. An innovative approach to the problem of visually impaired people independently detecting different denominations of cash during financial transactions is presented in this research. Conventional approaches have drawbacks that don't offer total autonomy, such as tactile indications and differences in banknote size. With the integration of the YOLOv9 object identification model into smart glasses driven by a Raspberry Pi, our suggested solution takes advantage of state-of-the-art deep learning and IoT technologies. The model correctly recognizes the banknotes in real time using a dataset of Pakistani currency notes (10, 20, 50, 100, 500, 1000, and 5000 denominations). A vocal output is produced as soon as a note is recognized, enabling the user independently and quickly identify the denomination of money. By lowering their dependency on others and lowering the possibility of fraud or deceit during transactions, this creative method improves the financial independence and safety of people with visual impairments. Our approach is a useful tool for empowering the blind population as it is not only effective but also simple to utilize.

2. Literature Review

M. Badawi et. al. [4] proposed the AI Amal Glasses, with functions including text-to-audio converter, currency difference, color recognition, obstacle detection, and face recognition. Raspberry Pi 4 B uses the functions incorporated by using OpenCV, Optical Character Recognition (OCR), and Python. The accuracy achieved is 95%. The complexity of currency images used to train the support vector machine model might be limited by the lack of additional datasets.

V. D. Mhaske et. al. [5] proposed the Android app for the identification of banknotes. Different ML techniques have been employed, which include brute force classification, the ORB algorithm, and the local binary pattern algorithm (LBP) and Principal Component Analysis (PCA) for dimensionality reduction and extraction of features. The OpenCV library has been used for image processing. Natural Language Processing (NLP) techniques have been used for voice integration. Flutter and Android Studio are used to build an app. The accuracy achieved is 99% for recognition and 90% for counterfeit.

Raghu et. al. [6] introduced the system to empower the visually impaired utilizing ML algorithms. COCO dataset has been used. Mobi Net's Single Shot Detector model, OpenCV, and Random Forest algorithm have been used. In addition, an emotional model based on the FER-2013 dataset has been employed to detect seven types of emotions. This technique can recognize all the faces in the webcam feed depending on the mood detected from the 4-layer CNN. The test accuracy achieved is 65% on 60 epochs. SG. Alamirew and GA.

Kebede [7] introduce an assistive system that can identify Ethiopian banknotes. Models for object detection include YOLOv5, Faster RCNN Inception v2 and SSD MobileNet v2. "Genzebe" mobile application has been combined with the model with voice output. The hardware components include the Raspberry Pi and Pi Camera. YOLOv5 performs well in terms of speed and accuracy, achieving a mAP of 97.9%. S.

Angsupanich and S. Matayong [8] proposed a mobile app for real time identification of Thai banknotes to aid visually impaired. Color and Haar-like features have been used for the detection and text-to-speech functionality. The mobile app architecture consists of a built-in camera, image detection event listener, image processing component, and vocal UI for communication. The accuracy achieved is 88%. Md.

Atikur et al. [9] designed a framework for the detection of indoor and outdoor objects and the Bangladeshi currency. The five modules in the methodology includes local finding, fall detection, object recognition, object detection and an IoT-enabled integrated smart system. SSD model, MobileNet, TensorFlow Lite, and obstacle sensors have been used. For currency detection, the SIFT algorithm is utilized. The accuracy achieved is 99.31% for detection and 98.43% for recognition. However, the accuracy of concrete currency detection is not stated.

RR. Mahmood et al. [10] proposed real-time Iraqi banknote identification for the visually impaired. The YOLOv3 algorithm is employed for detection. Google's Text to Speech (gTTS) is used for conversion. The accuracy achieved is 97.405%.

MA. Rahman et. al. [11] proposed an automatic navigation system to aid blind individuals. The system includes object detection, currency detection, distance determination, face identification, and speech devices. The system uses a camera, ultrasonic sensor, and complementary metal oxide semiconductor to run a distance-measuring algorithm. Only 350 images are included in the dataset. The systems achieve 98% accuracy. However, the percentage for currency identification is not stated. Images are collected by hand and do not contain all possible variations of objects in real-life cases.

K. Pandya and Dr. Bhargav C. Goradiya [12] proposed electronic glasses based on audio to aid visually impaired. Deep neural network algorithms coupled with optical character recognition (OCR) models and image processing techniques have been used in models that are able to identify human faces and currencies. The integration of the system includes the NVIDIA etson TK1 development kit, the SainSmart IMX219 AI camera, and Tesseract OCR software with the source IMAGE-NET database. The identification of currency was assessed under simulation over 500 pictures. The attained precision is 90%. Future improvements need to be made to the scale to make it universal.

S. A. Dande et al. [13] implemented a technique for the identification of Indian currency notes utilizing YOLOv5. A web app has been developed to ensure the identified structure detects currency notes. The proposed model successfully identifies 85 images out of 100 on test data with a 90% pass rate.

GDV Cortez et al. [14] designed a smart glass prototype capable of sensing objects and text signage with audio feedback. The model uses Raspberry Pi 2 Model B as a microprocessor. The YOLOv3 model has been employed for object detection and optical character recognition for text detection methods. The prototype is limited to detecting 15 objects only.

N.P. Hafiz Mohammed et. al. [15] proposed an AIoT-powered smart specs by using assistive technology. Arduino-based ultrasonic systems, SURF technology for object recognition, and portable navigation aids with synthetic speech have been used. But this device relies on the quality of the sensor and capability to differentiate between the designs of currency, especially in lightning conditions. AI-driven computer vision helps to increase the degree of accuracy. However, the specific detection rates cannot be pinpointed.

C. Charleen and G. Putra Kusuma [16] compared different models to detect currency notes, which include ORB and SURF for the feature-based approach. For the deep learning approach, YOLOv3, YOLOv5, YOLOv7, YOLOv8, and EfficientNet have been utilized. ORB has been working effectively with high accuracy. YOLOv8 performed best in terms of accuracy.

S. Syed Ameer Abbas et. al. [17] proposed an intelligent device that utilizes ML and DL algorithms for the partially sighted. Flame sensors, ultrasonic sensors, and image processing techniques YOLOv8 and Haar-Cascade have been used. For the people, the accuracy achieved is 93%. For objects, the accuracy achieved is 83%.

D. Danai Brilli et. al. [18] proposed an Airis system including modules for face identification, scene understanding, object detection, text recognition, money counting, and note-taking using the Raspberry Pi. For currency detection, image processing, pattern recognition, and computer vision filters and OCR have been used. User feedback reveals certain problems, such as delays in response, uncomfortable ergonomics, and the requirement to improve the adjustments to a user interface.

Muhammad Imad et al. [19] proposed a CNN and SVM-based system for Pakistani currency recognition. In this system, real-time recognition has been attained using deep CNN pre-trained

AlexNet architecture with feature extraction using histogram of oriented gradients (HOG). Each Pakistani note category contains 100 images. The overall accuracy achieved is 96%. For Pakistani coins on SVM, the accuracy achieved is 76.19%. The baseline accuracy on SVM of the proposed system for Pakistani Note is 82.04%.

L. Dunai Dunai et al. [20] proposed a portable device for the visually impaired to detect Euro banknotes. The system uses a Raspberry Pi and a Pi NoIR camera glass, and the SURF approach has been utilized for value recognition. Viola and Jones algorithms have been used for detection. The success rate for detection has been 84%, and for recognition, the success rate achieved is 97.5%. From the experimental data, it can be stated that EUR50 has been identified with 100% accuracy in 9 to 13 seconds, while EUR5 banknotes occasionally contained errors in 69.

Essakiraja S. et al. [21] introduced the Blind People Currency Detector, which can recognize banknote images with audio signals. Real-time identification of currency has been assisted using OpenCV and Python Sx3. Image analysis techniques have been employed, which include object detection, extraction of features, and recognizing patterns. The system can easily recognize currency; however, it may fail to recognize deformed or wrinkled banknotes.

N. Foysal et. al. [22] proposed a system by modifying YOLOv7 and YOLOv8 algorithms for the detection of multi-class items and currency categories. The dataset was created, which includes 295 pictures divided into eight sets for currencies and 3,046 pictures with 20 categories for object detection. YOLOv7 obtained a mAP of 91.4% for multiclass object identification, and YOLOv8 obtained a 94.6% mAP for currency categorization.

U. Kadam et. al. [23] implemented a system using CNN model for cash detection embedded in a Raspberry Pi, integrated with a cane stick for the visually impaired. They added 300 more pictures to an existing collection of Indian National Rupees. The system captures live videos using a camera, analyzes the frames with the CNN model to segregate currency, and then voice output is generated. The testing accuracy achieved is 92%.

B. Nayal et. al. [24] developed a cash detection system to aid the visually impaired. The proposed algorithm for the identification of Indian banknotes utilizes TensorFlow and MobilenetV-224 architecture, which belong to the CNN family. Pre-image classification, segmentation, histogram equalization, and template matching techniques have been employed. The accuracy achieved for Rs. 10 is 84%; for Rs. 20, it achieves 92.66%; for Rs. 5, it's 85%. The refund rate by the amount has been 33% for Rs. 50, 78% for Rs. 100. Increasing the exposure level, the conversion rates result in 33% for Rs. 200 and 84.66% for Rs. 500.

Abilash CS et. al. [25] designed a mobile app that incorporates the machine learning (ML) model for currency identification. The differentiation of Indian currency has been performed by the SIFT and K-Nearest Neighbor algorithms. In the audible output module, Bag of Words (BoW) has been used to recognize the text, and the ORB-FREAK algorithm has been utilized for text-to-speech conversion. The system of accuracy has been enhanced by training additional datasets but does not specify the accuracy rates achieved.

S.S. Chakravarthi et. al. [26] developed a deep learning system to determine the identification of coins using YOLOv8 with ResNet50. The Indian Coin Image Dataset has been used, which includes 1796 images. The testing accuracy achieved is 98.2%. The model was not trained with different orientations and augmentations of coins, which could affect performance in real-world scenarios. It primarily focuses on visual characteristics that can be replicated by counterfeit coins.

3. Proposed Method

Roboflow is used for annotating Pakistani currency images. The model is trained by the YOLOv9 object detection algorithm. After annotating the images, they are fed into a model for training. The methodology for the proposed model, from preparing the dataset to the voice output, is shown in Figure 1.

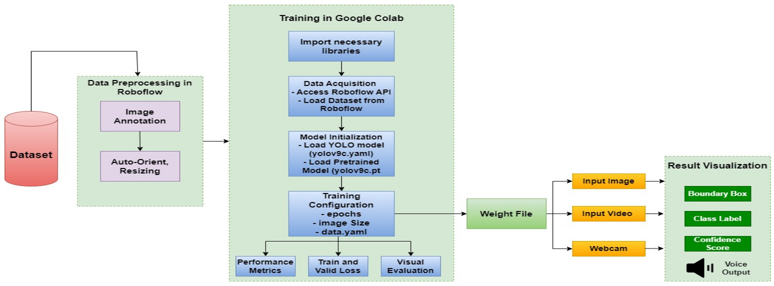

Figure 1. Methodology Diagram

3.1. Image Classification

Pakistani currency images are divided into seven classes. i.e., 10, 20, 50, 100, 500, 1000, and 5000. The classification of images is important to be labeled to predefined classes for accurately identifying currency notes and assigning them to the correct class labels. The classes assigned to Pakistani currency images are shown in figure 2.

Figure 2. Class Balance

3.2. Image Annotation

Image annotation is the process of marking and labeling specific objects within the image. Roboflow offers annotation tools for helping users to annotate images manually or by using automated techniques. (Figure 3) illustrates the annotation of the 50 rupee image in Roboflow.

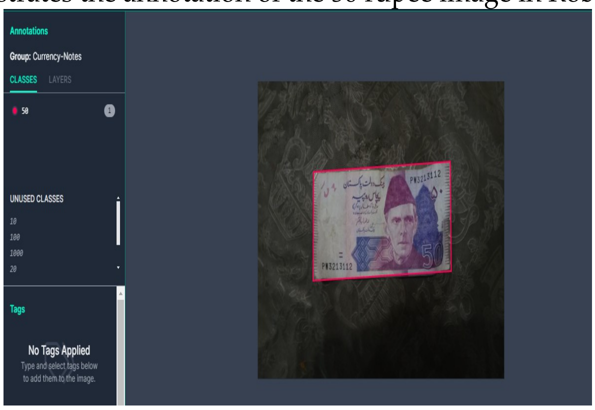

Figure 3. Annotation of Image

All 3611 images are assigned and annotated to the relevant class. A variety of modifications can be made to raw input. It entails cleaning, formatting, and improving the dataset to assure the quality of the training model. The performance is enhanced, and it helps to reduce training time. Two preprocessing techniques are employed, named Resizing and Auto Orient. The dataset is then exported in YOLOv9 format.

3.3. Model Training

3.3.1. Dataset Split and Training

The dataset was split into three subsets: training, testing, and validation, for appropriate assessment and optimization. 70% of the total data was in the training set, and 20% were in the testing and validation sets. The dataset was split and preprocessed using Roboflow, and then it was ready for training by being downloaded through the Roboflow API. The YOLOv9c model was trained using the annotated images, allowing it to recognize and categorize currencies according to the annotations that were supplied. The YOLO (You Only Look Once) v9 paradigm, which is well-known for its effectiveness and precision in real-time object recognition tasks, is employed in this procedure. A superior member of the YOLO family, YOLOv9c includes architectural improvements and improved training methods for increased performance.

3.3.2. Training Approaches

The YOLO model was trained using two primary techniques; the first is building from scratch. Yolov9c.yaml, is a YAML configuration file, was used to specify the model architecture. This file specifies the model's structure, including the layers, filter count, kernel sizes, and other parameters. This makes it possible to create a unique model that meets needs. The second technique involves loading a pre-trained model; an alternative approach is to use the Yolov9c.pt pre-trained model. This model can employ learnt features to improve training and performance because it already has pre-existing weights and biases from a previous training session.

3.3.3 Tools and Parameters for Training

Training was done using the Ultralytics library, a strong and valued tool for computer vision applications involving object recognition. 640x640 pixel photos from a dataset set up in data.yaml were used to train the model. There were 30 epochs of training. Hyperparameters like learning rate, batch size, and data augmentation methods were carefully adjusted to maximize outcomes.

3.4. Real Time Object Detection

The real-time object detection capabilities of the trained YOLOv9c model are the primary objective of this section. With this setup, the system can instantly identify items from a variety of input sources, such as static images, video files, and webcam feeds. Analyzing individual frames as we iterate over the selected input source is part of the real-time detection process. The model makes assumptions to determine the locations of objects for every frame. The detected items are then labeled with the confidence scores and class names that correspond to them.

3.5. Voice Output

The real-time object detection system was enhanced with a speech output option to improve accessibility for visually challenged users. The pyttsx3 text-to-speech library is used in this functionality to provide auditory feedback regarding objects that are recognized. The speech output system aims to detect class names and item counts in real time. This is accomplished by tracking detections and identifying those that have been detected. The system continuously examines the output of the YOLO model in each frame. Voice message generation: Based on objects recognized, the system generates spoken messages using the pyttsx3 package.

4. Experimental Results

4.1. Dataset

The dataset for Pakistani currency was downloaded from Kaggle [27]. 3611 images are in total, and they stand for the following seven denominations: 10, 20, 50, 100, 500, 1000, and 5000 rupees. The collection includes the front and back sides of every currency note. Table 1 provides specifics on how the images were distributed among denominations.

Table 1. Pakistani Currency Dataset Information

Currency	Total Number of Front Side	Total Number of Back Side	Overall total number of images
10	268	279	547
20	248	258	506
50	240	253	493
100	280	281	561
500	262	261	523
1000	227	236	463
5000	257	261	518

4.2. Training

Several measures were employed to evaluate the model's performance, including precision, recall, and mean average precision (mAP) at various levels of intersection over union (IoU). To offer a thorough assessment, these metrics were computed for the training and validation datasets. A comparison of the precision, recall, and mAP values acquired during the YOLOv9 model training procedure is shown in Figure 4.

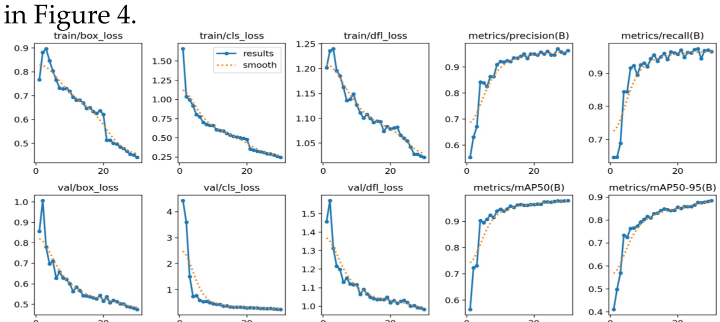

Figure 4. Training and Validation Graphs

The results are summarized as follows:

- Precision: Using the training dataset, the model achieved 95% accuracy. This demonstrates that 95% of the favorable events that the model had anticipated came to pass.

$$\text{Precision} = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (1)$$

- Recall: Using the training dataset, the model obtained a 95% recall rate, meaning that 95% of the real positive events were accurately detected.

$$\text{Recall} = \frac{True\ Positives(TP)}{True\ Positives(TP) + False\ Negatives(FN)} \quad (2)$$

- mAP50: At an Intersection over Union (IoU) criteria of 0.5, the model achieved a 90% mean Average Precision (mAP). This metric assesses how well the model can recognise items with a moderate level of overlap with the ground truth bounding boxes.

$$\text{mAP} = \frac{1}{C}\sum_{c=1}^{C} AP_c^{0.50} \quad (3)$$

- mAP50-95: The model's mean Average Precision (mAP) was 85% across a range of Intersection over Union (IoU) thresholds, from 0.5 to 0.95. This measure assesses the model's ability to detect objects with resilience by thoroughly analysing its output at different levels of overlap between the anticipated and ground truth bounding boxes.

$$\text{mAP} = \frac{1}{C}\sum_{C=1}^{C} \left(\frac{1}{10}\sum_{0.50}^{0.95} AP_c^t\right) \quad (4)$$

Where C is the number of classes, AP is the average precision of class C at IoU at threshold t.

4.3. Validation Performance

The precision, recall, mAP50 and mAP90-95 measures were also used to assess the performance of the model on the validation dataset. The following are the outcomes:

- Precision: On the validation dataset, the model obtained a precision of 92%, meaning that 92% of its positive predictions came true.
- Recall: A 92% recall rate was attained, meaning that 92% of the validation set's real positive events were accurately identified by the model.
- mAP50: The model's capacity to reliably identify objects with significant overlap on unseen data was demonstrated when it reached a mAP of 88% at an IoU threshold of 0.5.
- mAP50-95: Across IoU thresholds between 0.5 and 0.95, the model produced a mAP of 82% on the validation dataset. This illustrates how well the model works in identifying things on unseen data that overlap to varied degrees.

On both training and validation datasets, the model showed strong performance, attaining high precision, recall, and mAP scores. The consistent mAP values across different IoU thresholds demonstrated the model's capacity to generalize across varying object overlap levels. Loss curves, such as those for distributional focused loss, bounding box, and classification, shed light on the model's ability to concentrate on difficult samples and how it was learning. Overall, the accuracy and generalizability of the model showed encouraging results. The losses are computed using the following formulas: Mean Squared Error (MSE) can be used to compute the box loss, or Smooth L1 Loss (Huber Loss) can be used to determine it. The MSE, or mean squared error:

MSE = $\frac{1}{n}\sum_{i=1}^{n} (y_i - \widehat{y_i})^2$

(5) The Smooth L1 Loss,

Smooth L1 $(x) = \{0.5.x^2 \quad if\ |x| < 1\ |x| - 0.5\ \ otherwise$

Cross entropy loss is applied here to get the classification loss.

Cross-Entropy $(, q) = \sum p(i)\ log\ (q(i))$

Where p is the ground truth distribution and q is the predicted distribution. A specific loss function for bounding box regression in object identification models is called distribution focused loss.

Training Losses: Box Loss: From 0 it reduced to 0.9 to 0.5; Classification Loss: Copied from 1. 5 to 0.25; Distributional Focal Loss: It declined from 1.2 to 1.05. Validation Losses: Box Loss: From 1.0 it reduces to 0.5; Classification Loss: It has declined from 4.0 to 1.0; Distributional Focal Loss: It reduced from 1.5 to 1.0.

The model's performance in terms of classification is represented visually by the confusion matrix, which is shown in Figure 5. The percentage of correctly classified examples for a given currency denomination compared to the total number of actual instances of that class is represented by each cell in the matrix. The accuracy of the model for each class is shown by the diagonal cells, which underline accurate predictions. Misclassification rates are represented by off-diagonal elements. A numerical breakdown of the confusion matrix is provided in Table 2, which includes the accurate prediction percentages for every denomination of cash and misclassifications for some denominations.

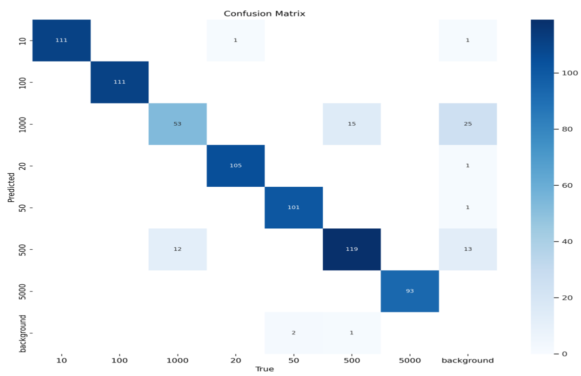

Figure 5. Confusion Matrix

Table 2. Detailed Analysis of Confusion Matrix

Currency Denomination	Diagonal Values (Correct Predictions)	Off-Diagonal Values (Misclassifications)
10	100%	No
10	100%	No
100	100%	No
1000	82%	Misclassified as 100 and 500
20	99%	Misclassified as 10
50	98%	Misclassified as 1000 and 500
500	88%	Misclassified as 1000 and 50
5000	100%	No

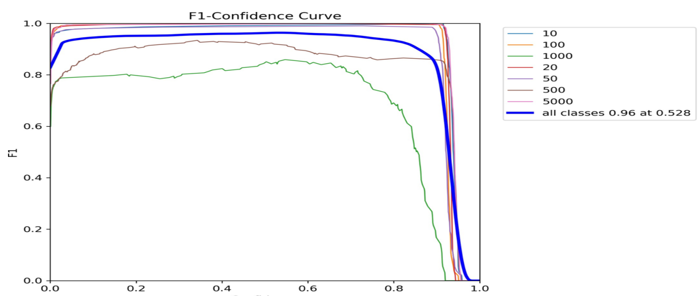

Figure 6. F1-Confidence Curve

Figure 6 illustrates the relationship between the F1-score and the confidence threshold for each class. All currency classes showed a significant trade-off between precision and recall, with the model producing an optimum F1-score of 0.96 at a confidence threshold of 0.528.

The link between precision and confidence threshold for each class is shown in Figure 7. When the model shows high confidence, it obtains perfect precision (1.00) for all classes at a confidence threshold of 94%, suggesting highly accurate predictions. The trade-off between recall (true positive rate over actual positive cases) and precision (true positive rate) is depicted in Figure 8 by means of the Precision-Recall curve. This graph shows how well the model performs at various classification levels.

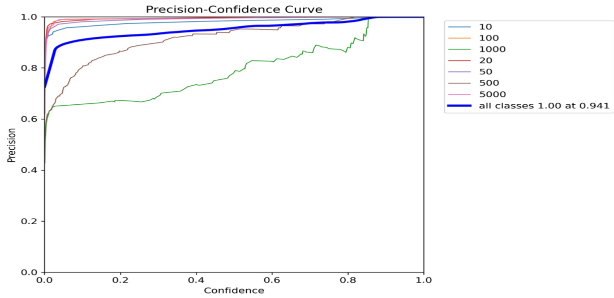

Figure 7. Precision Confidence Curve

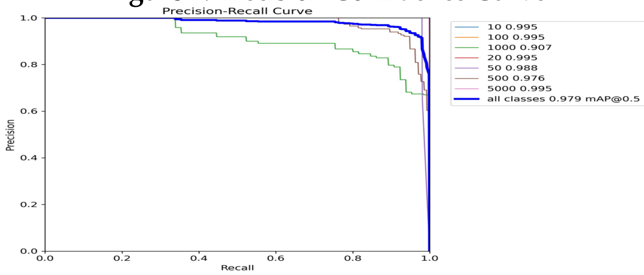

Figure 8. Precision Recall Curve

The object detection capabilities of the model are demonstrated in Figure 9, where banknotes are displayed inside colored bounding boxes. Every banknote is marked with the appropriate denomination and a confidence score that represents the degree of the model's predictive accuracy.


Figure 9. Currency Notes with Class and Confidence Score

Table 3. Comparison with previous studies

Serial No.	Authors' Name	Paper	Accuracy
1	Raghu et al. [6]	AI & ML Integration	65%
2	Angsupanich et.al [8]	Thai Banknote Recognition	88%
3	Pandya et.al [12]	Audio-Based Electronic Glasses	90%
4	Dande et al. [13]	Indian Currency Detection	90%
5	Ameer et al. [17]	Smart Device for Visually Impaired	90%
6	Dunai et. al. [20]	Euro Banknote Detection	84%

7	Proposed System	Pakistani Currency identification -YOLOv9 Model	88% to 95%.

A comparison of the suggested approach with current Pakistani currency classification methods found in the literature is shown in Table 3. The findings show that the suggested model outperforms previously documented techniques in terms of accuracy and functionality.

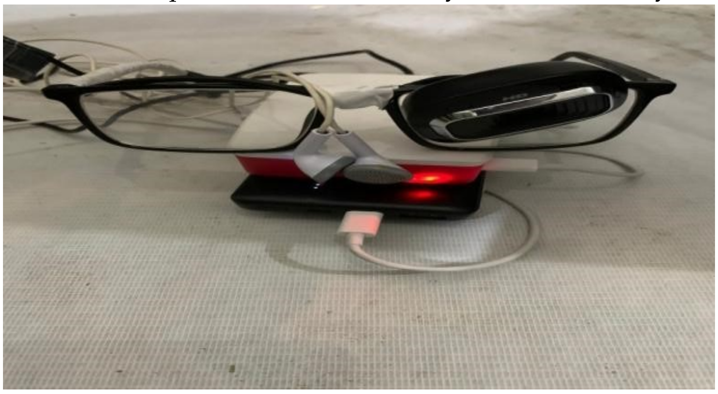

Figure 10. Integration of Components

5. Deployment on Raspberry Pi

After the completion of model training, the model is deployed to Raspberry Pi. It is an important step in this project as it enables currency identification on the edge device. For configuring the Raspberry Pi, connect it to the LED display using the HDMI port. The power bank serves as the power source for the Raspberry Pi. The camera and earphones are also connected to it for the detection of currency and voice output. Figure 10 illustrates the integration of components. Transfer the files to the Raspberry Pi by using a flash drive. Next, install the Python packages listed in the requirements.txt file. Some libraries used include Matplotlib, pytorch, opencv, pyYAML, and torchvision. After all the libraries and dependencies are successfully installed, it is ready to check for currency detection. By connecting the camera, the currency can be captured, and next it will be processed. After successful processing, the user will get voice output through earphones.

6. Conclusion

By creating a smart glasses solution that uses currency identification using the YOLOv9 model and speech output to enable visually impaired people to independently recognize Pakistani cash, this project offers a significant development in assistive technology. The smooth interaction of the device is made possible by the integration of IoT. Our goal is to improve the quality of life for the community of blind people. This technology can greatly improve the lives of people with vision impairments across the world by fostering inclusion and self-sufficiency. Future recommendations can be that other functionalities can be integrated into smart glasses, such as obstacle detection, and designing the system to enhance its maximum usability by visually impaired people.

References

1. World Health Organization. Vision impairment and blindness. WHO 2024. Available online: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (accessed on 17 July 2024).
2. Junaidi, I. Blindness, vision problems on the decline, survey shows. Dawn 2024. Available online: https://www.dawn.com/news/1708864 (accessed on 17 July 2024).
3. National Academies of Sciences, Engineering, and Medicine. Currency features for visually impaired people. Available online: https://nap.nationalacademies.org/read/4828/chapter/3#10 (accessed on 17 March 2024).
4. Badawi, M.; Al Nagar, E.; Mansour, R.; Ibrahim, Kh.; Hegazy, N.; Elaskary, S. Smart bionic vision: An assistive device system for the visually impaired using artificial intelligence. International Journal of Telecommunications 2024, 4(1), 1-12. irjmets.com.Availableonline: https://www.irjmets.com/uploadedfiles/paper/issue_2_february_ 2024/49433/final/fin_irjmets1708412335.pdf. (accessed on 17 July 2024).
5. Raghu, C.N.; Raghavendra, G.; Malashree, G.; Gopinath, K.; Pavithra, G.; Manjunath, T.C. Enhancing mobility for the visually impaired using artificial intelligence and machine learning in IKSHANA. Grenze International Journal of Engineering & Technology (GIJET) 2024, 10.
6. Alamirew, S.G.; Kebede, G.A. Developing an assistive technology for visually impaired persons: Ethiopian currency identification. Available at SSRN 4697484.
7. Angsupanich, S.; Matayong, S. A prototype real-time mobile application for Thai banknote recognition to support visually impaired persons. Universal Access in the Information Society 2023, 1–16.
8. Rahman, M.A.; Sadi, M.S. IoT-enabled automated object recognition for the visually impaired. Computer Methods and Programs in Biomedicine Update 2021, 1, 100015.
9. Mahmood, R.R.; Rasheed, M.K. Currency detection for visually impaired: Iraqi banknote as a study case. Turkish Journal of Computer and Mathematics Education (TURCOMAT) 2021, 12(6), 2940-2948.
10. Rahman, M.A.; Siddika, S.; Al-Baky, M.A.; Mia, M.J. An automated navigation system for blind people. Bulletin of Electrical Engineering and Informatics 2022, 11(1), 201–212.
11. Pandya, K.; Goradiya, B.C. Audio assisted electronic glasses for blind & visually impaired people using deep learning. Reliability: Theory & Applications 2021, 16(SI 1[60]), 143–151.
12. Dande, S.A.; Uppunuri, G.R.; Raghuvanshi, A.S. YOLOv5-based web application for Indian currency note detection. International Research Journal of Engineering and Technology (IRJET) 2022, 2801–2807.
13. Cortez, G.D.V.; Valenton, J.C.D.; Ibarra, J.B.G. Low-cost smart glasses for blind individuals using Raspberry Pi 2. Journal of Positive School Psychology 2022, 6(3), 6081–6088.
14. Nihal, H.M.P.; Sasi, A.P.; Saj, G.; Sabu, M.; PS, S.K. AIoT smart specs: An assistive technology for the visually impaired—a review. Journal of Network & Information Security 2024, 12(1).
15. Charleen, C.; Putra Kusuma, G. Evaluation of deep learning models for detection of Indonesian Rupiah. International Journal of Computing and Digital Systems 2024, 16(1), 315–327.
16. Hareharan, M.; Sivakumar, I. Integrated intelligent assistant for visually impaired. Research Square 2024.
17. Brilli, D.D.; Georgaras, E.; Tsilivaki, S.; Melanitis, N.; Nikita, K. Airis: An AI-powered wearable assistive device for the visually impaired. arXiv preprint 2024, arXiv:2405.07606.
18. Imad, M.; Ullah, F.; Abul Hassan, M.; Naimullah. Pakistani currency recognition to assist blind persons based on convolutional neural network. Journal of Computer Science and Technology Studies 2020, 2(2), 12–19. Available online: https://al-kindipublisher.com/index.php/jcsts/article/view/529 (accessed on 17 July 2024).
19. Dunai, L.; Chillaron Pérez, M.; Peris-Fajarnés, G.; Lengua, I. Euro banknote recognition system for blind people. Sensors 2017, 17(1), 184.
20. Essakiraja, S.; Kasi Vishwanathan, M.; Anton Gino, J.A.; Subbulakshmi, T.C. An application for currency detection for visually impaired. International Journal of Novel Research and Development 2024, 9(4).
21. Foysal, N.U.; Thamid, T.J.; Uddin, M.S.; Islam, M.F.; Islam, M.N.; Al Faisal, F. Advancing AI-based assistive systems for visually impaired people: Multi-class object detection and currency classification. In Proceedings of the 2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), 438–442, IEEE, 2023.
22. Kadam, U.; Meena, A.; Abuzar, C.; Singhal, G.; Srivastava, D. Currency detection for visually impaired people using edge devices. In Proceedings of the 2023 International Conference on Modeling, Simulation & Intelligent Computing (MoSICom), 498–503, IEEE, December 2023.

23. Nayak, B.; Thota, N.P.; Kumar, J.D.; Dhanalakshmi, R.; Bairavel, S. Currency detector for visually impaired using machine learning. In Proceedings of the 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), 1–5, IEEE, 2021.
24. Abilash, C.S.; Dinakaran, M.; Hari Vignesh, R.; Anju, A. Currency recognition for the visually impaired people. In Proceedings of the 2022 IEEE Delhi Section Conference (DELCON), 1–3, IEEE, February 2022.
25. Chakravarthi, S.S.; Jaiadityan, S.; Yeshwanth, A.; Sreeman, R.; Sountharrajan, S. Deep intelligent system for precise coin identification. In Proceedings of the 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), 1–6, IEEE, April 2024.
26. Kaggle.com. Pakistani currency notes image dataset. Available online: https://www.kaggle.com/datasets/muhammadammarabid/pakistanicurrencydataset. (accessed on 17 March 2024).