International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Currency Detection for Blind using DL and AI

V.S. Tamboli[1], Bansode Tejas[2], Aher Rohan[3], Shaikh Juned[4], Labade Sangram[5]

[1]*Professor, Dept. of Computer Technology, P.Dr.V.V.P. Institute of Technology and Engineering, Loni, Maharashtra, India*

[2,3,4,5] *Final year Diploma Student, P.Dr.V.V.P. Institute of Technology and Engineering, Loni, Maharashtra, India*

---***---

Abstract - Visually impaired individuals face significant challenges in performing daily financial transactions due to their inability to visually recognize and differentiate between various currency denominations. This often leads to dependency on others, increased chances of fraud, and reduced independence. To address this issue, this project proposes an AI-based Currency Detection System designed specifically for blind and visually impaired users. The system utilizes Deep Learning (DL) and Computer Vision techniques to automatically identify and classify different denominations of currency in real time. The proposed system captures images of currency notes using a mobile or embedded camera. These images are then processed using advanced deep learning model MobileNet trained on a large dataset of various currency denominations under different lighting and background conditions. The model is capable of recognizing features such as color patterns, printed symbols, and texture variations unique to each denomination. Once the note is identified, the result is immediately communicated to the user, making the system completely accessible to the visually impaired. The implementation combines AI-based object detection with assistive technology, ensuring accuracy, speed, and user-friendliness. It also supports multiple currencies, making it adaptable for international use. Furthermore, the system can operate offline once trained, which enhances its practicality for real-world scenarios where internet connectivity may be limited. This project aims to promote financial independence, social inclusion, and accessibility for visually impaired individuals by harnessing the power of Artificial Intelligence and Deep Learning. The integration of image recognition, deep neural networks, and speech output provides an efficient, affordable, and real-time currency identification solution.

***Key Words*:** Currency Recognition, Deep Learning, Artificial Intelligence, Computer Vision, Convolutional Neural Networks, Assistive Technology, Text-to-Speech, Accessibility, Visually Impaired, Object Detection

1.INTRODUCTION

Money is one of the most essential aspects of human life, playing a crucial role in daily transactions, independence, and social participation. However, for visually impaired or blind individuals, identifying and handling currency notes can be a major challenge. Since most paper currencies have similar sizes and textures, it becomes nearly impossible for blind users to distinguish between denominations through touch alone. Although some countries have implemented tactile marks or differently sized notes to assist visually impaired citizens, these measures are often insufficient, inconsistent, or not universally adopted. As a result, visually impaired individuals are often forced to rely on others for financial transactions, which can lead to a loss of privacy, autonomy, and even vulnerability to fraud.

With the rapid advancement in Artificial Intelligence (AI) and Deep Learning (DL) technologies, new opportunities have emerged to create assistive systems that enhance accessibility and inclusivity for people with disabilities. AI-driven image recognition and deep neural networks model MobileNet have shown exceptional performance in visual tasks such as object detection, image classification, and pattern recognition. By leveraging these technologies, a reliable system can be developed to automatically detect and identify currency denominations through visual input, such as an image captured by a smartphone or camera. When integrated with audio feedback, such a system can act as a powerful assistive tool for visually impaired individuals, allowing them to recognize and confirm currency values in real time.

The proposed project, “Currency Detection for the Blind using Deep Learning and AI,” focuses on developing a smart, real-time currency recognition system capable of identifying currency notes accurately under various environmental conditions. The system captures an image of the note and processes it through a MobileNet trained on a diverse dataset containing images of currency notes from different angles, lighting conditions, and levels of wear and tear. The MobileNet extracts visual features such as printed numbers, unique patterns, embedded symbols, and color variations that help distinguish one denomination from another. Moreover, this project aligns with the goals of inclusive technology and social innovation, which aim to use emerging technologies to create solutions that empower differently-abled individuals. By integrating AI and Deep Learning with assistive applications, this system contributes to the broader mission of Digital Accessibility — ensuring that technology serves everyone equally, regardless of physical or sensory limitations. The adoption of such systems can significantly reduce the social and psychological barriers faced by blind individuals in their everyday lives.

In recent years, several studies and prototypes have explored the application of image processing for currency recognition, but many of them have limitations such as low accuracy, poor adaptability to varying lighting conditions, or dependence on internet connectivity. The proposed system addresses these challenges through advanced neural network models, MobileNet, which are lightweight and optimized for mobile platforms. These models provide superior classification performance while maintaining computational efficiency, ensuring that the system can function smoothly even on low-end devices.

[1] Jindong Wang et al. proposed effective two-stage clustering algorithm based on a combination of hierarchical clustering and K-means clustering. A TF points energy method is employed to estimate the number of clusters before clustering. To get the used K-means clustering algorithm insensitive to the noise and overcome the drawback of randomly selecting initial clustering centers, the proposed two-stage clustering algorithm effectively eliminates the outliers to further improve the estimation accuracy of the mixing matrix. Extensive simulations and experiments are conducted to verify the effectiveness of the proposed method. These results show that the mixing matrix estimated by the proposed method is substantially more accurate than that of the classic K-means clustering alone. Rolling bearing fault experiments show that the proposed method can effectively separate multiple fault sources and can accurately find the location of the fault location with envelope analysis. The proposed method has some limitations. First, we focus on the linear instantaneous mixing system and it may not suitable for the convolution cases. Second, the time cost of the proposed two-stage clustering method is longer than that of single clustering. Using this paper's examples, the proposed method can be applied to bearing fault analysis, and it is anticipated that it will also be useful for diagnosing complex mechanical faults in future work.

[2] Sam Ansari et al. Proposed highlights the significance of BSS in enhancing signal processing capabilities in B5G and 6G. In the context of future wireless communication systems such as B5G and 6G, the process of transmitting a signal from the transmitter side and subsequently receiving it at the receiver side introduces the possibility of signal contamination due to undesired components in the transmission channel. This article presents a novel algorithm aimed at restoring the original signals from such contamination. This study introduces a novel three-way neural network architecture that combines transfer learning, a pre-trained DPRNN, and a transformer model. The proposed algorithm outperforms all the benchmarked techniques in terms of SI-SNR, SDR, and STOI metrics, showcasing the highest improvement percentages across the board compared to that of selected algorithms.The effectiveness of the proposed algorithm in tackling real-world challenges, such as complex acoustic environments characterized by noise and reverberation, is clearly evident. The robustness of the framework enables its applicability in practical domains such as speech enhancement, audio transcription, and audio-visual processing, particularly within the realm of B5G and 6G technologies. Future work in audio BSS could focus on exploring hybrid approaches that combine the strengths of different algorithms to further improve separation performance. Additionally, the development of novel evaluation metrics that capture additional aspects of audio quality and perceptual attributes could provide a more comprehensive assessment of separation algorithms.

[3] Alessandra Galli et al. Introduced method provides for the first time indications about the most suited electrode configuration for fHRmonitoring by multichannel electrophysiological recording, taking into account the characteristics of the source signals in combination with the algorithms employed for fHR estimation, and dynamically adapting to the fetal presentation. The proposed method, based on SVR, can predict in advance the accuracy of the fHR estimation based on features extracted from the raw data, prior to

performing complex fHR estimations. When more electrodes are available, the automatic selection of the best 5-electrode configuration proposed in this manuscript is the optimal solution to take advantage of the 16-channel recording while keeping the computational cost low. Indeed, the optimal configuration is made up of only four leads that ensure lightweight processing. The proposed solution, beyond being flexible, robust, and lightweight, can handle the change in fetal position and manage signal degradation, ensuring accurate and reliable estimation of the fHR also during wearable, long-term monitoring applications.

The second section of this study presents an examination of the prior studies deemed as the Literature Survey. Section 3 provides a comprehensive description of the proposed methodology. Part 4 examines the experimental evaluation, while Section 5 investigates prospective modifications before concluding the essay with a summary of the current plan.

2. LITERATURE SURVEY

[4] Fan Xiangyu et al. Observed is in accordance with the DCS theory proposed by predecessors to judge the feasibility of this method on the aspect of radar separation by conducting equation derivation, and to expand the existing algorithm from the two signals to multi routes signals by building Givens matrix. By simulation, the radar is separated and the practicability of this method is judged. Because the common noise is not equipped with cycle correlation, and the cyclisation frequencies of different radar signals are different. This paper aims to separate and identify the radar signal without priori information from cyclisation dimensionality. In comparison with the common stationary process, cyclisation theory can retain the phase information of system, which is extremely beneficial to the identification and parameter estimation of system.

[5] Chaofeng Lan et al.we optimize the traditional Wave-U-Net model structure to improve the speech separation performance. The SAM module is introduced in the skip connection of the Wave-U-Net model, which is used to selectively aggregate features to reduce the semantic gap between the encoding and decoding layers. The 1D convolution layer of the bottleneck layer in the Wave-U-Net model is replaced with the ASPP module, which can not only increase the receptive field of the network but also enhance the ability of the network to obtain multi-scale contextual information. The experimental results on the Musdb18 dataset show that the Wave-UNet+SAM+ASPP model proposed in this paper is superior to some existing baseline models. At the same time, the above changes are all helpful to improve the performance of themodel. However, the Wave-U-Net+SAM+ASPP model has a relatively long model training time, which requires further improvement and optimization.

[6] Mohammed E. Fouda et al. Narrate the applicability of blind source separation techniques for self-interference cancellation in IBFD systems. We experimentally compared the performance when applying real ICA versus complex ICA, demonstrating the better performance of complex ICA by 3dB at the expense of 4x runtime. ICA-based SIC can be used for FD-MIMO systems with a small number of Tx/Rx antennas without changing the communication standard. However, to accumulate more antennas, the frame length has to be much larger and pilots should be used. ICA algorithms are complex and require the computation of division, square-root, and logarithmic operations, as discussed in Section VI-B. Thus, the digital implementation of such algorithms is challenging, which could be a bottleneck in designing efficient VLSI circuits. Optimizing the ICA algorithms in favor of the hardware would of interest, in order to facilitate the deployment of such systems.

[7] Kyung-Ah Shim et al. discussed shown that the three lattice-based blind signature schemes and the blind ring signature scheme do not achieve blindness. Their vulnerabilities against our attacks are due to the exposure of the blinding factors used in the blinded messages and their related signatures. We have proposed a generic construction from a semantically secure homomorphic encryption scheme and a one-more unforgeable blind signature scheme which does not achieve blindness, whose resulting blind signature scheme achieves blindness as well as one-more unforgeability. However, the improved schemes are very inefficient since they require additional overhead for encryptions and decryptions to generate views. Thus, we can say that it remains an open problem to construct secure and efficient lattice-based blind signature and lattice-based blind ring signature schemes based on algebraic relations for hiding the blind factors without using the homomorphic encryption schemes.

[8] Mandhatya Singh et al. Observed focuses on the problem of Indian currency recognition for BVIP and presents an end-to-end automated solution. We

propose an extensive large-scale Indian currency dataset (approximately 10x larger images count than the existing ones). The dataset contains images from varied backgrounds conditions and different illumination and orientations. Apart from that, images with folded and partial views are included focusing on the BVIP scenario. The proposed lightweight network (IPCRNet) uses controlled multi-dilation and depthwise separable convolution schemes with dense connection, enabling local and global contextual information aggregation. IPCRNet offers the advantage of enlarging the receptive field without a larger resolution requirement. An extensive evaluation of the proposed framework on publically available datasets has been performed for assessing the generalization and prediction capabilities. The experimental analysis demonstrates the IPCRNet competence in capturing the currency-specific features. IPCRNet is simpler and efficacious in terms of parameters (3.6M) and accuracy. An android application Roshni is presented to recognize Indian currency denominations for BVIP. The proposed framework is suitable for a mobile-compatible environment offering a trade-off between memory, speed, and high accuracy. A preliminary user study and feature comparison has also been presented to showcase the effectiveness of App.

[9] Firat Akba et al. Proposed the methods to detect manipulations in Bitcoin market. Accordingly, the most commonly used price estimation methods, involving time series forecasting, machine and deep learning techniques were implemented to determine the weekly/monthly rising and falling trends of Bitcoin pricing. The first priority was to make sure that price estimation methods could produce truly reliable results. To achieve this, we conducted our experiments by identifying a minimally manipulation-free zone. Afterwards, the leading sentiment analysis techniques were used in order to understand the effect of social media posts on prices. During these studies, we also enhanced the ISSFS method, which was previously suggested by us for sentiment analysis. As a result of the improvements applied, it was observed that the regressions encountered while working on big data were eliminated. It was clearly observed that experiments regarding pre-crisis data and virtual money stock market trend prediction achievements were really remarkable. According to the preliminary result, the comments made on the most popular social media platform we examined did have positive effects on the SVM and SARIMAX methods. LSTM had no positive or negative effects with sentiment results. Both weekly and monthly results of ARIMA did have negative effects while evaluating prices with sentiment results. In addition to the given conclusion, manipulators played a major role in changing trends during all periods when anomalies were detected.

[10] Jun Wei et al. Introduced Research and realized based on virtual currency trading MTCNN model video face detection algorithm, this method is a new cascade architecture to integrate multiple convolution neural network learning method, using three levels league architecture combined with carefully designed neural network algorithm, realization of virtual currency trading video face detection and rough localization of key points. At the same time, the traditional face detection technology is studied, and the classic face detection algorithm of Harrfeature and AdaBoost is realized to compare with the MTCNN algorithm. The experimental results show that the face detection method based on MTCNN model has more advantages and more robustness than the face detection method based on Harr feature and AdaBoost. Dots are commonly used in a variety of scenarios. Based on face characteristic such as diversity combining with the actual needs of the video, the traditional LeNet - 5 model adds a layer of convolution and sampling, extracts more comprehensive and more sophisticated technical link activation function, updates parameter rules, and redesigns the reverse derivation rules. Through the test of multiple public data sets, the advantages of virtual currency trading video face recognition of deep learning are very obvious, which proves the effectiveness of virtual currency trading video face recognition of deep learning for face recognition.

3. METHODOLOGY

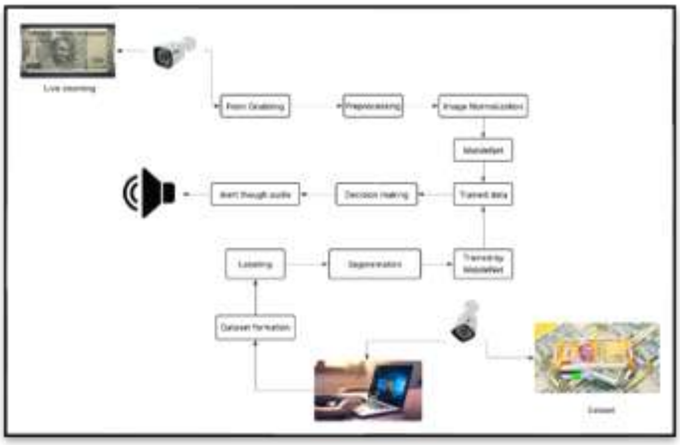

Fig 1: System Overview Design

The above system diagram represents the working process of the Currency Detection System for Blind People. The system uses computer vision and deep learning techniques to identify currency notes and provide an audio alert to visually impaired users.

Step 1: Live Streaming

The system begins with live video streaming using a camera. The camera continuously captures images of currency notes placed in front of it. These frames are sent to the processing module for further analysis.

Step 2: Frame Grabbing

In this stage, the system extracts individual frames from the live video stream. Using computer vision libraries such as OpenCV, each frame is captured and passed to the next stage for processing.

Step 3: Preprocessing

The captured frames are preprocessed to improve image quality and remove unnecessary noise. Preprocessing operations may include resizing, noise removal, and color conversion. This step helps improve the accuracy of the detection model.

Step 4: Image Normalization

After preprocessing, the image is normalized to ensure consistent pixel values and dimensions. Normalization adjusts brightness and contrast so that the neural network can process images effectively regardless of lighting conditions.

Step 5: Feature Extraction Using MobileNet

The normalized image is then passed to the MobileNet deep learning model. MobileNet extracts important visual features from the image, such as patterns, shapes, and textures that help distinguish between different currency denominations.

Step 6: Training with Dataset

Before the system can detect currency notes, the MobileNet model must be trained using a dataset. The dataset is prepared using the following steps:

Dataset Formation

Images of different currency notes are collected and stored in a dataset. These images represent various denominations captured under different lighting and orientations.

Labeling

Each image in the dataset is labeled using annotation tools. Bounding boxes are drawn around the currency notes to mark their positions.

Segmentation

Segmentation separates the currency note from the background so that the model focuses only on relevant features during training.

Training by MobileNet

The labeled and segmented dataset is used to train the MobileNet model. During training, the neural network learns patterns and features that distinguish one currency denomination from another.The trained model generates trained data (model weights) which are used during detection.

Step 7: Decision Making

Once the model detects a currency note in the frame, the system analyzes the prediction results. The decision-making module determines the denomination of the detected currency based on the model's output.

Step 8: Audio Alert Output

Finally, the detected currency denomination is converted into speech output using a text-to-speech module. The system announces the detected note value (for example, "100 Rupees detected"), helping visually impaired users identify the currency independently.

4. RESULTS AND DISCUSSIONS

The Windows-based system, which has 16 GB of main memory and an Intel Core i7 processor, is used to implement the proposed model. The Anaconda IDE repository for Spyder and Jupyter IDEs is used by the model for the experiment. The created model is rigorously evaluated using the parameters of the confusion matrix. You may understand the parameters of the confusion matrix using the following equations: accuracy, precision, recall, and macro F1.

$$Accuracy=(TP+TN)/(TP+FP+TN+FN) \quad -(1)$$

$$Precision(P)=TP/(TP+FN) \quad -(2)$$

$$Recall(R)=TP/(TP+FP) \quad -(3)$$

$$Macro\text{-}F1=(2*P*R)/(P+R) \quad -(4)$$

At this point, we have TP for true positives, TN for true negatives, FP for false positives, and FN for false negatives.

The obtained results are shown below,

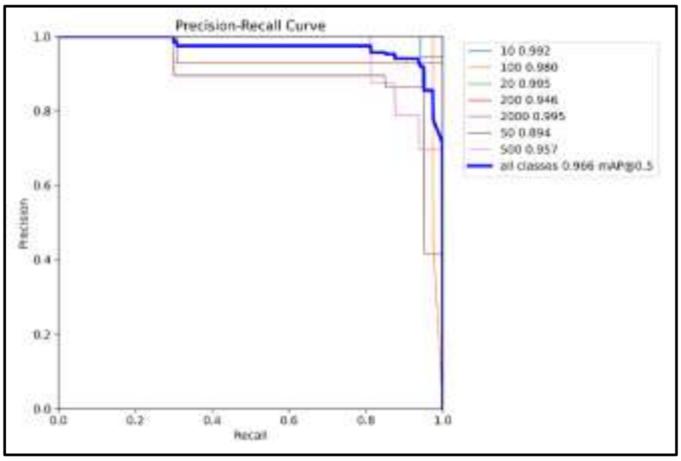

Figure 2: Precision–Recall (PR) Curve

The precision–recall curve evaluates the detection performance of the currency recognition model. The x-axis represents recall while the y-axis represents precision for different currency classes. The curves for most denominations remain close to the upper-right region, indicating high precision and recall. The model achieves strong Average Precision values such as 0.992 for ₹10, 0.995 for ₹20 and ₹2000, and 0.980 for ₹100. The overall mean Average Precision (mAP@0.5) is 0.966, demonstrating excellent detection performance. This indicates that the model accurately identifies currency notes with minimal false detections.

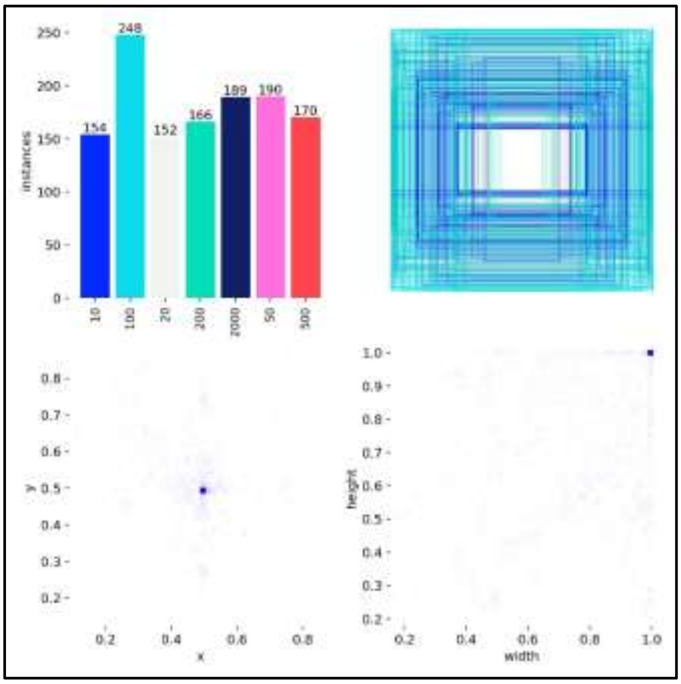

Fig 3: Dataset Distribution and Annotation Analysis

The above figure 3 shows the dataset statistics and annotation distribution used for training the currency detection model. The bar chart indicates the number of instances available for each currency class such as ₹10, ₹20, ₹50, ₹100, ₹200, ₹500, and ₹2000. From the graph, ₹100 notes have the highest number of samples, while the remaining classes are fairly balanced. The bounding box distribution visualizes the location and size variation of annotated currency notes in the dataset. The x–y center heatmap shows that most objects are located near the center of the images. The width–height plot indicates that the dataset contains notes at different scales, improving the robustness of the detection model.

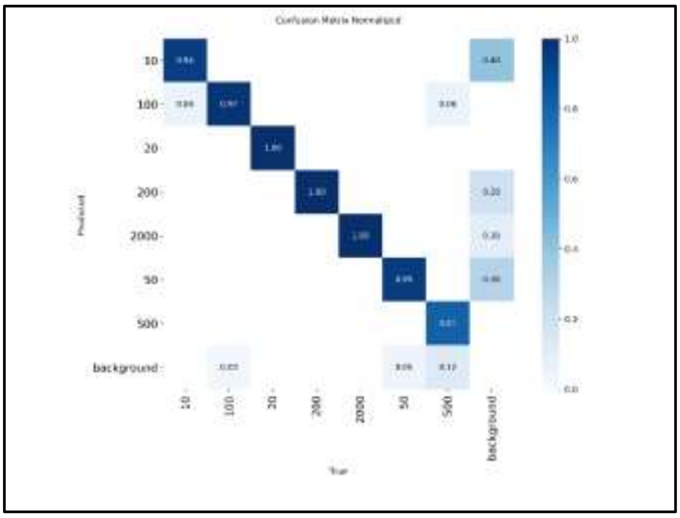

Fig 4: Confusion Matrix

The confusion matrix represents the classification performance of the currency detection model across different denominations. Each row corresponds to the actual currency class while each column represents the predicted class. The diagonal elements indicate correctly classified currency notes, whereas off-diagonal values represent misclassifications. A strong concentration of values along the diagonal indicates that the model correctly distinguishes between most currency denominations. Minor misclassifications may occur between visually similar notes due to lighting or orientation variations. Overall, the confusion matrix confirms the high accuracy and reliability of the proposed currency detection system.

Fig 5: Currency Detection Results

The above figure 5 presents sample outputs of the trained currency detection system. Each image shows a detected currency note with a bounding box and its predicted class label along with a confidence score. The model successfully detects multiple denominations such as ₹10, ₹20, ₹50, ₹200, and ₹2000 under different lighting conditions and orientations. The bounding boxes accurately surround the notes, indicating correct localization by the model. The high confidence scores demonstrate the reliability of the trained detection model. These results confirm that the proposed system can effectively recognize currency notes in real-time scenarios.

5. CONCLUSIONS

By combining computer vision with deep learning, the suggested currency identification system offers a practical means by which people with visual impairments can recognize paper currency. The system is capable of real-time precise denomination recognition thanks to image preprocessing, feature extraction, and a trained object detection model based on MobileNet. The experimental results show that the model produces a strong mAP score and delivers high detection accuracy. The outcomes include dataset analysis, detection outputs, and precision-recall evaluation. Users who are visually impaired can nevertheless grasp the recognized value of the currency thanks to the system's aural feedback. For the visually challenged, this means more accessibility and the ability to sustain themselves financially. In terms of efficiency and reliability, the proposed method works well for real-world currency recognition tasks. To further increase accuracy, future work can involve increasing the dataset with more currency samples and experimenting with different lighting situations. One way to make the system more portable and user-friendly is to incorporate the model into a mobile application. It is possible to enhance the system's usability by adding support for detecting numerous currencies from different nations and by adding recognition for damaged or folded notes.

REFERENCES

[1] J. Wang, X. Chen, H. Zhao, Y. Li and D. Yu, "An Effective Two-Stage Clustering Method for Mixing Matrix Estimation in Instantaneous Underdetermined Blind Source Separation and Its Application in Fault Diagnosis," in IEEE Access, vol. 9, pp. 115256-115269, 2021, doi: 10.1109/ACCESS.2021.3105538.

[2] S. Ansari, K. A. Alnajjar, T. Khater, S. Mahmoud and A. Hussain, "A Robust Hybrid Neural Network Architecture for Blind Source Separation of Speech Signals Exploiting Deep Learning," in IEEE Access, vol. 11, pp. 100414-100437, 2023, doi: 10.1109/ACCESS.2023.3313972.

[3] A. Galli, E. Peri, C. Rabotti, S. Ouzounov and M. Mischi, "Automatic Optimization of Multichannel Electrode Configurations for Robust Fetal Heart Rate Detection by Blind Source Separation," in IEEE Transactions on Biomedical Engineering, vol. 70, no. 4, pp. 1196-1207, April 2023, doi: 10.1109/TBME.2022.3212587.

[4] F. Xiangyu, L. Bin, D. Danna, C. You and W. Yuancheng, "Blind Radar Signal Separation Algorithm Based on Third-Order Degree of Cyclostationarity Criteria," in Journal of Systems Engineering and Electronics, vol. 35, no. 6, pp. 1441-1453, December 2024, doi: 10.23919/JSEE.2024.000117.

[5] C. Lan, J. Jiang, L. Zhang and Z. Zeng, "Blind Source Separation Based on Improved Wave-U-Net

Network," in IEEE Access, vol. 11, pp. 125951-125958, 2023, doi: 10.1109/ACCESS.2023.3330160.

[6] M. E. Fouda, C. -A. Shen and A. E. Eltawil, "Blind Source Separation For Full-Duplex Systems: Potential and Challenges," in IEEE Open Journal of the Communications Society, vol. 2, pp. 1379-1389, 2021, doi: 10.1109/OJCOMS.2021.3086105.

[7] K. -A. Shim and Y. An, "Cryptanalysis of Lattice-Based Blind Signature and Blind Ring Signature Schemes," in IEEE Access, vol. 9, pp. 134427-134434, 2021, doi: 10.1109/ACCESS.2021.3113938.

[8] M. Singh, J. Chauhan, M. S. Kanroo, S. Verma and P. Goyal, "IPCRF: An End-to-End Indian Paper Currency Recognition Framework for Blind and Visually Impaired People," in IEEE Access, vol. 10, pp. 90726-90744, 2022, doi: 10.1109/ACCESS.2022.3202007.

[9] F. Akba, I. T. Medeni, M. S. Guzel and I. Askerzade, "Manipulator Detection in Cryptocurrency Markets Based on Forecasting Anomalies," in IEEE Access, vol. 9, pp. 108819-108831, 2021, doi: 10.1109/ACCESS.2021.3101528.

[10] J. Wei, "Video Face Recognition of Virtual Currency Trading System Based on Deep Learning Algorithms," in IEEE Access, vol. 9, pp. 32760-32773, 2021, doi: 10.1109/ACCESS.2021.3060458.