IFERP

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 3 March 2024

Machine Learning-based Currency Detection for Visually Impaired Assistance using ESP Camera Module

[1] Naseer Hussain Gajula, [2] Ayyappa Kandagatla, [3] Siva Rama Raju Penumathsa, [4] Koteswara Rao Naik R

[1] [2] [3] [4] Gitam School of Technology

Corresponding Author Email: [1] naseerghussainndl@gmail.com, [2] kandagatlaayyappa@gmail.com, [3] sivaramaraju4912@gmail.com, [4] kramavat@gitam.edu

Abstract— *The "Machine Learning-based Currency Detection for Visual Impairment Assistance" project is a great example of how technology can be used to improve the lives of people with disabilities. The system uses a camera module and pre-trained machine learning models to classify banknote images, providing real-time audio feedback and a user-friendly interface. This means that visually impaired individuals can accurately identify currency denominations without the need for sighted assistance.*

The system works by first capturing an image of the banknote using the camera module. The image is then processed by the machine learning models, which identify the banknote denomination. The classification results are then spoken aloud by the system, allowing the visually impaired individual to identify the banknote easily. The system is designed to be user-friendly and affordable. The use of ESP camera modules makes the system relatively inexpensive to build, making it accessible to a wider range of people. The "Machine Learning-based Currency Detection for Visual Impairment Assistance" project is a promising new technology that has the potential to make a real difference in the lives of visually impaired individuals. The system is accurate, affordable, and user-friendly, making it a valuable tool for people who are blind or have low vision.

Index Terms— *ESP32 camera module, FTDI, Machine Learning, Currency Detection, Speech to Audio*.

I. INTRODUCTION

At least 2.2 billion people worldwide suffer from near- or far-sightedness impairments, with uncorrected refractive errors, cataracts, diabetic retinopathy, glaucoma, and age-related vision deterioration being the main causes. A 2017 National Centre for Biotechnology Information (NCBI) report revealed that 217 million of the 253 million persons with visual impairments worldwide were also blind. These conditions not only hinder vision but also significantly impact daily life, creating a profound need for innovative solutions.

The "Machine Learning-based Currency Detection for Visual Impairment Assistance" project aims to address the fundamental issue faced by the visually impaired community: the ability to identify and differentiate banknote denominations. By harnessing the power of machine learning, this innovative system transforms the lives of those with visual impairments, offering them newfound independence and inclusivity in financial transactions and daily activities. By exploring its components, operational processes, and tangible benefits, we can understand how technology can be harnessed to improve the lives of those affected by visual impairment, ultimately fostering a more inclusive and equitable world for all. This paper explores the transformative potential of this innovative system, highlighting the profound impact of technology in enhancing the lives of individuals affected by visual impairment.

II. PROJECT OVERVIEW

The project aims to assist visually impaired individuals in identifying Indian currency notes, that lack braille markings. It uses deep learning-based object identification and speech synthesis technologies to identify the notes based on size, color, and other features. The project will create a dataset of Indian currency note images, which will be used to train a deep-learning model to identify the notes. This technology could revolutionize the way visually impaired people interact with currency notes and gain financial independence. The system will recognize distinct features and patterns on the notes, such as colors, sizes, and unique symbols, often the only means for visually impaired individuals to distinguish them. The integration of speech synthesis technology will also enable the system to communicate the note's value audibly, empowering visually challenged individuals to understand the value of the notes independently Submit your manuscript electronically for review.

III. KEY FEATURES

A. Camera Module and Image Capture

The system incorporates a camera module, such as the ESP camera module, which captures images of banknotes. This hardware component is crucial for gathering visual data to be processed by the machine learning models.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 3 March 2024

B. Machine Learning Model

Pre-trained machine learning models are at the heart of the system. These models are trained to recognize and classify different currency denominations based on the images captured by the camera module. Machine learning algorithms enable the system to make accurate predictions swiftly.

C. Real-time Audio Feedback

One of the standout features of this system is its ability to provide real-time audio feedback. After processing the banknote image, the system verbally announces the denomination of the currency aloud. This auditory feedback is critical for visually impaired users to immediately and confidently identify the banknotes they are handling. Authors and Affiliations

D. User-Friendly Interface

The project prioritizes user-friendliness, ensuring visually impaired individuals can easily interact with and operate the system. This user-friendly interface is designed to be intuitive, making it accessible even to those who may not have extensive technical expertise.

E. Affordability

By utilizing affordable hardware components like Arduino or ESP camera modules, the project aims to keep costs low. This affordability is significant because it increases the accessibility of the technology to a wider range of individuals who may have limited financial resources.

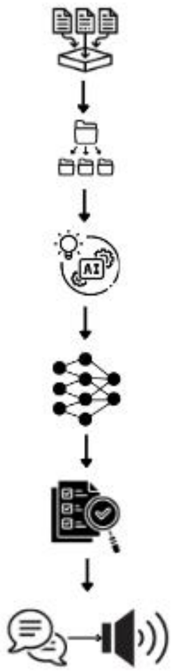
F. Independence and Inclusivity

Perhaps the most impactful aspect of this project is its potential to enhance the independence and inclusivity of visually impaired individuals. Allowing them to confidently identify banknotes on their own, reduces their reliance on others and promotes their active participation in financial transactions and daily life.

G. Scalability

The system is designed to be scalable so that it can be easily adapted to different currencies and denominations. This makes the system a valuable tool for people who travel to different countries or who need to handle multiple currencies.

IV. METHODOLOGY



Methodology

A. Data Collection:

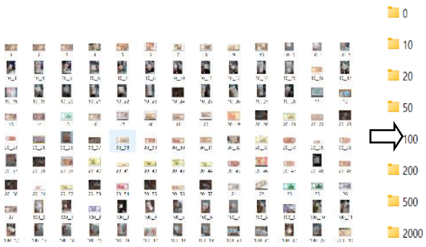
A comprehensive dataset of Indian currency notes is essential for model training. Select denominations from ₹10 to ₹2000, and capture images in different lighting conditions, backgrounds, angles, and focus points. Include distinct colors and patterns for a diverse dataset, ensuring sharpness and versatility.

S.No	Denomination	Image count
1	0	263
2	10	479
3	20	485
4	50	482
5	100	493
6	200	529
7	500	470
8	2000	468

Total Data Collected

B. Data Preprocessing:

Data preprocessing is a crucial process for machine learning model training and evaluation. It involves data cleaning, resizing images, and dividing the dataset into training, validation, and testing subsets. Accurate data labeling is crucial for the accurate classification of images. The preprocessed dataset is then organized and saved in a structured format, ensuring a robust foundation for effective model training and assessment. This meticulous process ensures randomness, diversity, and accuracy in the training and evaluation of machine learning models.



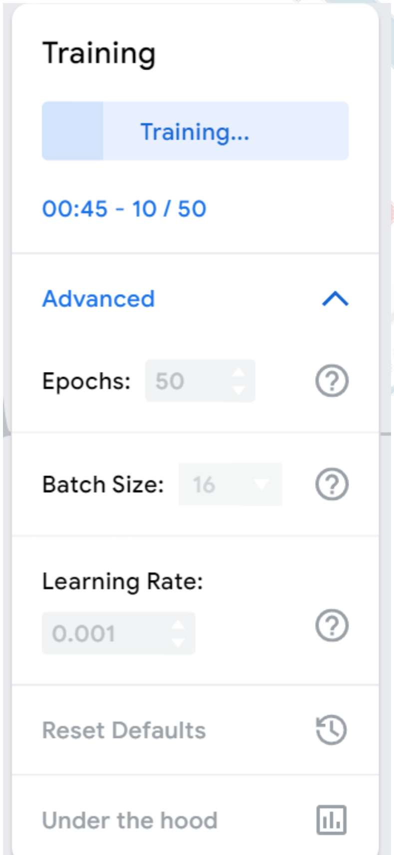
Data Preprocessing

C. Model Creation:

Choosing a deep learning architecture for object detection depends on project needs and accuracy goals. Popular options include SSD, YOLO, and Faster R- Google's Teachable Machine simplifies model creation without extensive programming knowledge. Custom models with TensorFlow or PyTorch offer greater control but require a deeper understanding of machine learning. Each approach offers unique advantages and suits different proficiency levels.

D. Model Training:

To train an object detection model, use machine learning frameworks like TensorFlow and PyTorch to achieve the desired accuracy on the validation dataset. The model should be fine-tuned on a specific dataset, with 50 epochs chosen for optimal performance. The choice of epochs depends on the specific dataset and problem, and experimentation is typically used.



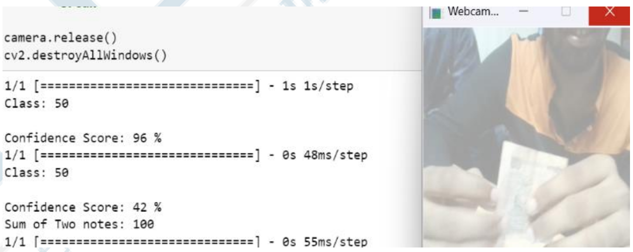
Training Model

Batch size refers to the number of examples processed in each iteration before updating the model's parameters. Smaller batch sizes may lead to noisy updates but may help the model generalize better. Larger batch sizes provide more stable updates but may require more memory. The choice of batch size affects convergence speed and memory requirements and is often selected based on available hardware and empirical testing.

The learning rate determines the step size at which the model's weights are updated during training, controlling how quickly or slowly the model converges to a solution. A learning rate of 0.001 is relatively small, allowing the model to make small adjustments during each update. Setting it too high can result in overshooting the optimal solution while setting it too low can lead to slow convergence or getting stuck in local minima. Experimentation is often required to find an appropriate learning rate.

E. Model Evaluation:

Evaluate the performance of the trained model on the testing dataset. This will help ensure that the model does not overfit the training data. The model's performance can be evaluated using a variety of metrics, such as Mean Average Precision (mAP).



The result of the 50 rupee note

Result of No Input

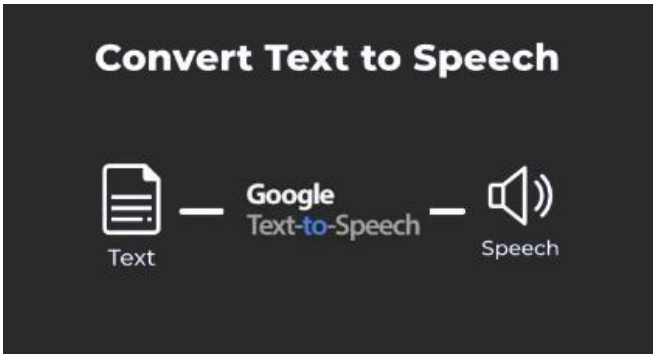
The result of the 500 rupee note

F. Speech Integration:

The object detection model has been integrated with speech synthesis libraries like pyttsx3 or gTTS to enhance

ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Volume 11 Issue 3 March 2024

user interaction and accessibility. This allows the model to generate audio feedback, such as the denomination of a detected currency note. The model also includes multi-language audio speech support, catering to a diverse user base and making the experience more inclusive. This integration not only makes the model visually informative but also audibly informative, making it a valuable tool for assistive technologies and automation.

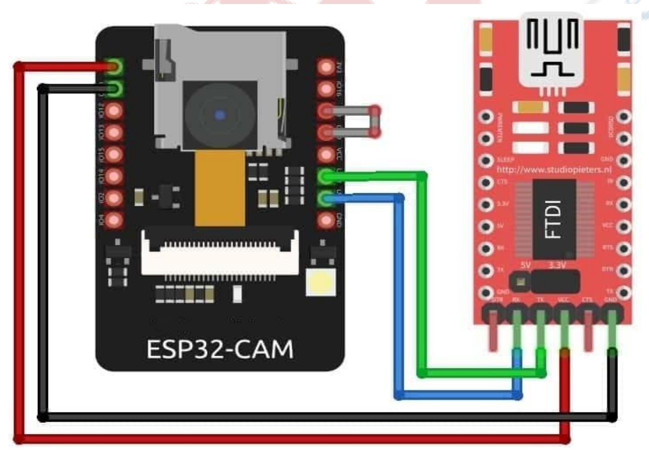


Text to Speech

Integrating multi-language audio speech synthesis fosters cross-cultural communication and understanding, allowing users from different linguistic backgrounds to access vital information easily. The real-time, spoken feedback is beneficial in scenarios where the user's visual attention may be occupied elsewhere. This broadens the applicability of the object detection model, making it a versatile solution for diverse sectors, including visual impairments and automated processes.

G. Deployment:

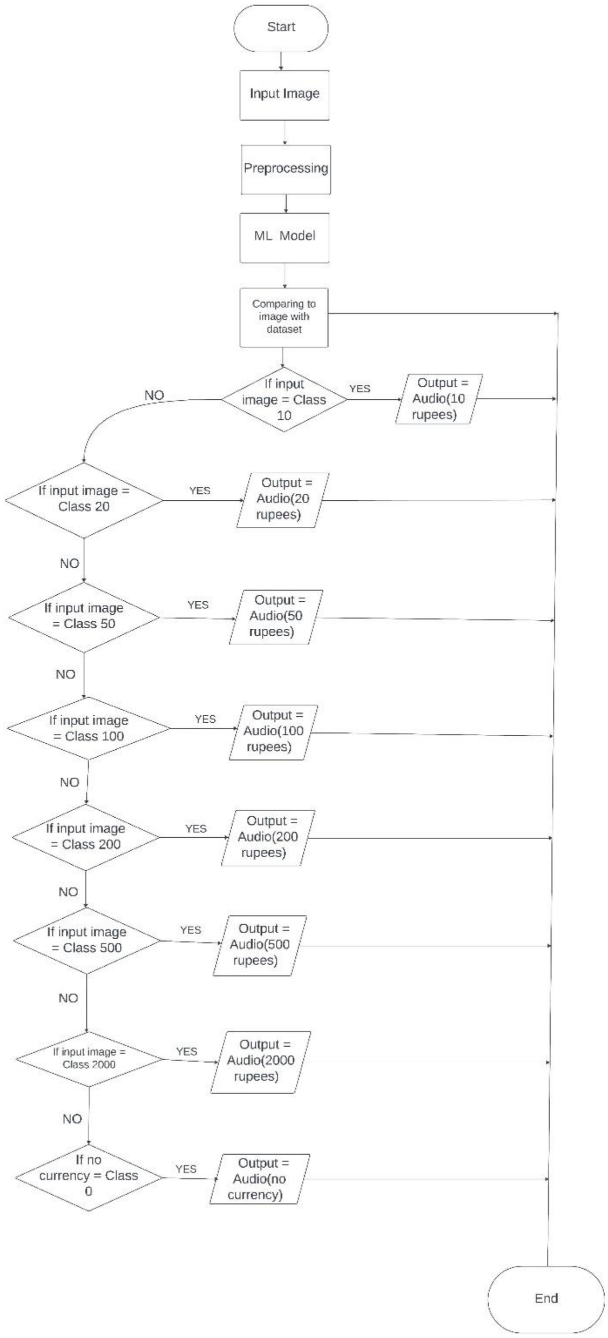
Below is the circuit diagram illustrating the hardware connections for deploying the code, incorporating the ESP32 Cam Module, FTDI, Speaker, and Jumper wires.



Hardware Connections

The ESP32 Cam is programmed using an FTDI programmer. By connecting pins GND, U0T, and U0R of the ESP32 Cam module to pins GND, RX, and TX of the FTDI module, respectively, the ESP32 Cam module and the FTDI module are linked in this circuit. A 5V battery powers the entire system. Note: The esp32 cam module's GPIO 0 pin must be connected to GND to upload the code.

V. FLOW CHART



Methodology Flowchart

The flow chart outlines a project that uses an ML model to process input images, predict a specific class, and generate corresponding audio output.

- The process starts with the input images, which are visual data that the ML model will analyze.
- The images undergo preprocessing, such as resizing or normalization, to prepare them for the ML model.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 3 March 2024

- The preprocessed images are then fed into the ML model, which has been trained on a dataset of images with seven different classes.
- The model compares the input image with the trained images, evaluating similarity or patterns to determine its class.
- If the input image is classified as "Class 0," the system generates an audio output indicating that Class 0 corresponds to a specific audio response.
- If not, the system iterates through the classes, repeating the image comparison and class prediction steps until it finds a matching class or exhausts all possibilities.

Once the correct class is identified, the system generates an audio output corresponding to that class

VI. FUTURE SCOPE

The future scope includes possible areas for development such as increased accuracy, worldwide currency support, interaction with financial systems, adherence to accessibility standards, and the implementation of strong privacy and security measures. Continuous refining of algorithms and machine learning models might enhance accuracy, while global currency support could expand utility for visually impaired people worldwide. Collaboration with banks and financial organizations may help improve accessibility during transactions.

VII. CONCLUSION

Machine learning is an artificial intelligence technology that enables computers to learn without explicit programming. It can be used for currency detection by learning features of currency notes, such as size, shape, color, and patterns. Convolutional neural networks (CNNs) are particularly well-suited for image recognition tasks, as they extract features from images in a hierarchical manner. Machine learning algorithms can identify currency notes even in different lighting conditions, as they don't rely on specific features only visible in certain conditions. They can be trained on small datasets, making it easier to collect large datasets. Additionally, they can be implemented on portable platforms like ESP modules, making them ideal for visually impaired individuals. This technology has the potential to significantly improve the quality of life for visually impaired individuals, making it easier to identify currency notes and enabling them to be more independent and self-sufficient.

REFERENCES

[1] K. Gautam, "Indian Currency Detection using Image Recognition Technique," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2020, pp. 1-5, doi: 10.1109/ICCSEA49143.2020.9132955.

[2] M. Pasumarthy, R. Padhy, R. Yadav, G. Subramaniam and M. Rao, "An Indian Currency Recognition Model for Assisting Visually Impaired Individuals," 2022 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE), Tainan, Taiwan, 2022, pp. 1-5, doi: 10.1109/RASSE54974.2022.9989624.

[3] P. Garkoti, P. Mishra, N. Rakesh, Payal, M. Kaur and P. Nand, "Indian Currency Recognition System Using Image Processing Techniques," 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2022, pp. 628-631, doi: 10.23919/INDIACom54597.2022.9763134.

[4] R. Vishwasi, S. K. Holla, Y. D, and K. Kulkarni, "A Raspberry Pi based CNN model for Indian Currency Detection for Visually Impaired People," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 1170-1175, doi: 10.1109/ICESC54411.2022.9885317.

[5] Joshi, M., Shukla, A., Srivastava, J., & Rastogi, M. (2023). DRISHTI: Visual Navigation Assistant for Visually Impaired. ArXiv. /abs/2303.07451

[6] Deshpande, Samruddhi & Shriram, Revati. (2016). Real time text detection and recognition on hand held objects to assist blind people. 1020-1024. 10.1109/ICACDOT.2016.7877741.

[7] V. Abburu, S. Gupta, S. R. Rimitha, M. Mulimani and S. G. Koolagudi, "Currency recognition system using image processing," 2017 Tenth International Conference on Contemporary Computing (IC3), Noida, India, 2017, pp. 1-6, doi: 10.1109/IC3.2017.8284300. keywords: {Currencies; Image color analysis; Feature extraction; Image recognition; Classification algorithms; Impedance matching; Image Processing; Feature extraction; Aspect ratio; Denomination; Binary image; Currency recognition},

[8] S. V. T. Sangeetha, T. Porselvi, A. Venkateshwaran, M. Gokul, I. Sathmikan and P. Tharun Kumar, "Currency Detection App for Blind People Using MIT App Inventor," 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-4, doi: 10.1109/ICPECTS56089.2022.10047784. keywords: {Presses; Blindness; Detectors; Control systems; Artificial intelligence; Currencies; Image classifier; MIT app inventor; Currency Detector},

[9] A. Koti, A. Khare and J. P. Kumar, "Netrahitha: Currency Detector Jacket for visually impaired people," 2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), Phuket, Thailand, 2023, pp. 364-368, doi: 10.1109/ECTIDAMTNCON57770.2023.10139655. keywords: {Face recognition; Visual impairment; Blindness; Detectors; Feature extraction; Character recognition; Currencies; Technology for Blind; Currency Detector/ recognition; Visual Impaired Assistance; Denomination; Image Processing Multi-factor Computation; Multi-sensor Integration; Image comparison},