AI-Powered Financial Management for Visually Impaired Individuals (Sightpay)

Ahmed Makboul	Youssef Abdelmajeed	Mariam Zaki	Jihan Khalil
Helwan University	Helwan University	Modern Academy	Benha University

Supervised by
Eng. Mahmoud Talaat
AI Engineer @ MCIT
(Ministry of Communication and information technology)

Abstract

This project presents an AI-powered solution aimed at empowering visually impaired individuals to independently manage their invoices and finances. By integrating advanced machine learning, image recognition, and voice assistance technologies, the solution allows users to effortlessly capture invoices, verify details, calculate totals, and compare expenses with available   funds through intuitive voice commands. This paper explores the   project's motivation, the financial management challenges faced by visually impaired individuals, and the technical implement-ation of the proposed solution.

Keywords

Machine Learning, Image Recognition, Voice Assistance,
Financial Management, Visually Impaired, Invoice Management, Accessibility Technology

1- Introduction

Background: Discuss the increasing reliance on technology for personal finance management and the challenges faced by visually impaired individuals in this area

Motivation: Highlight the importance of financial independence for visually impaired individuals and the societal impact of creating accessible financial tools.

Objectives: the goal of project to develop an AI solution that enhances financial management for visually impaired users.
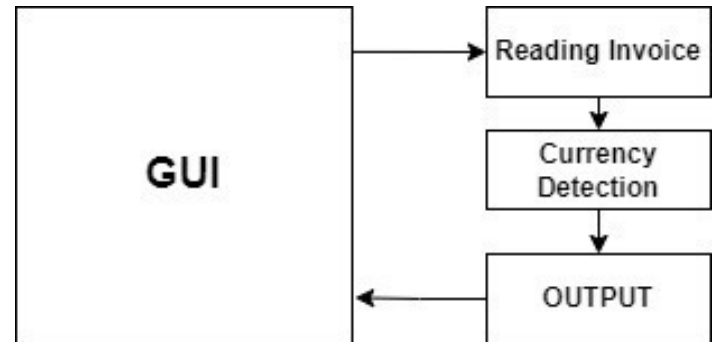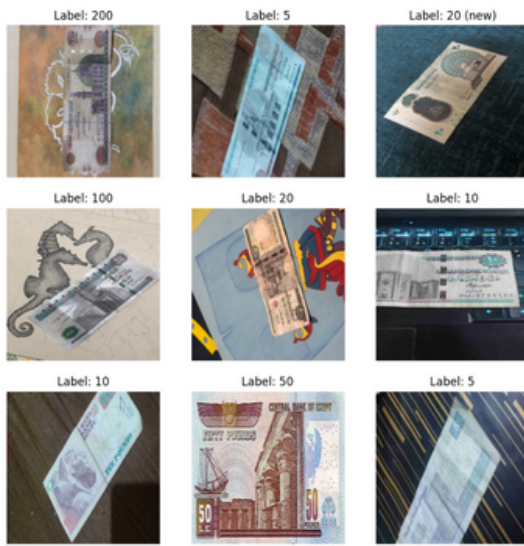

Fig 1.1

Problem Statement

Challenges:

·Inability to Read Bills: Visually impaired individuals often struggle to read and understand printed invoices or bills without assistance, leading to reliance on others for financial management.

·Difficulty in Identifying Amounts: Users may find it challenging to determine the total amounts owed and categorize expenses accurately.

·Impact: Discuss the implications of these challenges on financial independence and overall quality of life.

Experiments

Data

The dataset comprises manually gathered images of Egyptian banknotes in denominations of 10, 20, 50, 100, and 200 pounds. Images were captured under various conditions to ensure robustness, with data augmentation techniques like rotation and brightness adjustments applied to simulate real-world scenarios. The dataset was labeled by denomination and divided into training, validation, and test sets to ensure diverse model training and reliable evaluation.

Label: 200   Label: 5   Label: 20 (new)
Label: 100   Label: 20   Label: 10
Label: 10   Label: 50   Label: 5

EfficientNet Architecture

2- Related Work

2. 1 Optical Character Recognition (OCR):

·For extracting textual information from invoices, we utilize Gemini Flash for OCR, which provides accurate recognition of printed text across different fonts and formats. This allows the system to convert physical invoices into digital format for further processing.
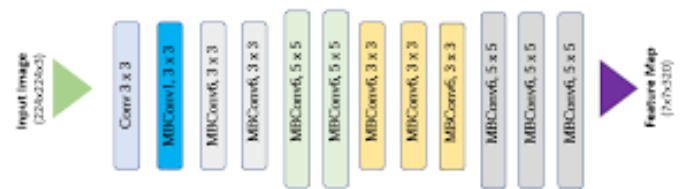
.

2.2 Voice Interaction:

·To enable hands-free interaction, we integrate Google Cloud services for text-to-speech and speech-to-text functionalities. The text-to-voice service provides audio feedback for the details extracted from the invoice, while the voice-to-text recognizer allows users to issue commands or queries through voice input
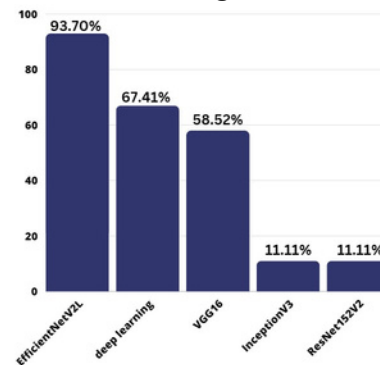
3- Model

For this project, we utilize EfficientNet as the primary image recognition model due to its superior ability to accurately recognize and process photos. EfficientNet is a state-of-the-art convolutional neural network architecture that balances model accuracy and computational efficiency through compound scaling. This architecture achieves better performance by scaling up depth, width, and resolution in a systematic way, leading to significant improvements in accuracy and efficiency compared to traditional models

4- Results

Fig 1.2

as results EfficientNet is best model achieving high accuracy 0.93% compared with others

5-Limitations

Limited Functionality for Complex Transactions: SightPay is effective for basic tasks but may struggle with complex financial scenarios requiring deeper integration.

Dependence on Voice Commands: The voice interface enhances accessibility but may face challenges in noisy environments or for users with speech impairments.

Invoice and Transaction Variability: Differences in invoice formats or payment systems across vendors could result in interpretation or processing errors.

6-CONCLUSION AND FUTUREWORK

Future improvements to SightPay will target enhanced accuracy through a larger dataset and advanced techniques, alongside refining text extraction and voice recognition. The user interface will be optimized with better voice commands and new input options like touch or gesture controls.

In conclusion, SightPay shows promise in helping visually impaired individuals manage finances independently, with basic invoice reading and payment capabilities. Continued enhancements will boost reliability and accessibility, aiming to fully meet users' needs.

References

BANKNOTE RECOGNITION SYSTEM FOR VISUALLY IMPAIRED PEOPLE
EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks
OCRBench: On the Hidden Mystery of OCR in Large Multimodal Models
An Application for Currency Detection for Visually Impaired
Currency Recognition System for Blind people using ORB Algorithm
End-to-End Speech Recognition: A Survey

Google Generative AI API