



RuralConnect: Bridging digital divide between rural communities

Shraddha C¹, S R Suresh², P Likith³

Engineering Graduate, Information Science, Global Academy of Technology, Bangalore, India,

MS Ramaiah University of Applied Sciences^{1,2,3}

Abstract: This describes the development of an application designed to assist low-literacy rural communities to handle their bill and paper safely and efficiently. To make its functionality work, the application features a very robust login mechanism that avails itself of the fingerprint authentication mode for confirmed access. After the customers authenticate themselves, they can use the camera of their device to scan critical documents like utility bills, PAN cards, and Aadhar cards. The application uses NLP techniques, particularly that of SpaCy, to analyze the retrieved text and provide critical insights derived after using the Optical Character Recognition (OCR) technology to pull the text from these scanned images. This enables non-English-speaking users to access the application through transliteration options in Hindi, Kannada, Telugu, and Tamil together with the text to be read. The application will transform the words into speech by using Google Text-to-Speech (gTTS), depending on the needs of the user as it is unlikely that the audio would be in any other language. The app has a reminder alert about payments which are soon to be overdue with respect to bills, so one may not suffer from late fines and scams. This makes the application easy to use, accessible, and full of an inclusive approach in order to improve the management of documents and bills for the vulnerable communities of the rural fraternity.

Keywords: Rural development, fraud, extract text, insights, OCR, NLP, multi lingual support, text-to-speech, multi-language translation, document access, , camera, gTTS, notification reminder.

I. INTRODUCTION

Low literacy rates in many rural areas have led to huge problems in the management of official documents and even in the tracking of some bill payments, leaving the countryside citizen often at the mercy of fraudsters. We propose here a mobile app, which will help the people of the countryside manage their documents and track their bill payments in a secure and efficient manner. This app is equipped with a strong fingerprint-based authentication system that allows authorized users only to access the app. The users can scan documents like Aadhar cards, PAN cards, and utility bills, along with the application employs Optical Character Recognition (OCR) to image to text processing application will be in use of the NLP algorithms such as SpaCy to extract relevant information from the text but to overcome the language limitations, application can translate text into Hindi, Kannada, Telugu, and Tamil using Indic Transliteration. Audio output for blind readers or challenged readers could be achieved by using Google Text-to-Speech (gTTS) application will have the reminder system where it will remind the user of an upcoming bill. As these payments erode late charges and minimize fraud risk, the integrated solution using new technologies improves access, usability, and security for rural people and enhances their individual capacity in carrying out their personal affairs significantly.

II. LITERATURE SURVEY

[1] Named Entity Recognition is one of the main tasks for natural language processing, where structured information is retrieved from unstructured text. In this regard, this paper provides a comprehensive overview of NER methodologies and their applications as well as the challenges in the field. This research narrows its scope toward the different techniques regarding NER, from the classical rule-based approach to the sequence-based methods, together with the transformative approach brought about by the emergence of transformer models such as BERT. It talks about domain-specific NER models tailored for finance, legal, and healthcare domains, but also some reinforcement learning and deep learning initiatives in NER. The paper further indicates that there's a play between NER and Optical Character Recognition (OCR) and the practical applications of NER in finance and biomedicine.

[2] FastPitch is the fully-parallel text-to-speech model building on top of FastSpeech, improving voice synthesis quality via conditioning with fundamental frequency (F0) contours. Hence, by using pitch contour prediction and conditioning in inference time, FastPitch can produce more expressive speech in general, and also semantically better aligned. With



all this, it produces synthesized speech which is entirely more accurate than previous work and falls into the rank of the best autoregressive models available, all without requiring extra processing power. Real-time factor greater than 900x with the use of feed-forward Transformer architecture by the model enables fast and real-time mel spectrogram synthesis. Adding further linguistic information to the plain feed-forward Transformer improves convergence and pronunciation while not inducing quality issues, as found in FastPitch in its conditioning on F0.

[3] The extensive linguistic diversity of 22 major and 1618 spoken languages in India poses a radical roadblock to inclusive access to technology, especially for the most disadvantaged and visually impaired communities in the country. A consortium managed by the Indian Institute of Technology Madras, under funding from the Ministry of Communications and Information Technology, has developed a standard framework for text-to-speech systems of Indian languages. The framework provides non-English. It was supporting the smooth integration of TTS in a variety of applications for the speakers and the blind, as it was created with Festvox and Festival. Examples of applications are Android-based applications, browser plugins for reading highlighted text, OCR systems for turning scanned text into speech, and screen readers (Orca and NVDA). Pronunciation dictionaries and CART trimming solved some important issues about usability, including loading times and response optimization. Besides, only very minimum English support was available to handle the required English content and, hence, equally primitive English support was available as well. In reality, key benefits are achieved, such as enabling blind people to learn how to use office applications, e-mail, and the Internet, thus promoting additional integration in a digital environment.

[4] The work introduces a new dataset to be hopefully applied in order to move ahead the template-free invoice processing by means of AI techniques. The dataset is known as Multi-Layout Unstructured Invoice Documents Dataset and aims at dealing with the challenge of several layouts presented during real-world scenarios. The authors make it an important dataset in both training and evaluating AI models capable of handling various invoice structures without any predefined templates. The authors test several AI approaches to the dataset, showing good performance in extraction of the crucial information regarding an invoice, including billing details and quantities.

[5] It for the first time, offers a potential direction for training TTS models in a more expressive manner. The proposed framework combines frame and style reconstruction losses as a step towards improving expressive speech synthesis with respect to fine-grained temporal structures (frames) and stylistic nuances. Through this framework, the model can capture the natural variations of speech more effectively and output those variations better.

The proposed approach is evaluated experimentally, along with significant improvements in the quality and naturalness of synthesized speech across the different expressive styles.

III. IMPLEMENTATION

Our application has been designed with several integrated technological components to create a no-risky, user-friendly solution for the rural population. First, we created a robust system of authentication by integrating a fingerprint sensor to make sure that only registered users can access the application. For document handling, we integrate the Optical Character Recognition (OCR) technology for scanning and text extraction of various forms like Aadhar cards, PAN cards, and utility bills. Users would capture the image using the camera of the device. The OCR system would then proceed to extract relevant information from images, which would be processed on through Natural Language Processing algorithms such as SpaCy so that key insights are derived and presented. For the individuals who do not understand English, the implementation of Indic Transliteration was used in order to translate the text into regional languages like Hindi, Kannada, Telugu, and Tamil. Another benefit is the adoption of Google's Text-to-Speech, which translates written text into audio for individuals who have reading difficulties so that they can receive that information in their own preferred language. The mobile application includes a reminder tool that gives pop-up reminders to remind users when the payments of later deadlines on bills are due. The interface was planned to be easy and intuitive, meaning accessible for any viewer regardless of the level of digital literacy. Data security was ensured during all the stages of development and remained the highest priority; it was achieved through encrypted databases that stored the information entered by users. Extensive testing phases were part of the project to ensure because of its functionality and reliability, coupled with continuous updates and improvements premised on user feedback with the view of continually enhancing the user experience, this all-inclusive implementation ensures that the application successfully supports rural populations in safely managing their documents and bills.

Document image capture is started with camera input followed by authentication for safe entry. The images captured by the camera are processed using Optical Character Recognition and Natural Language Processing techniques and key insights are drawn from them. If the output text be in English, then the output is fed to thereby obtain an instant display of knowledge. If the output text be in a language other than English but for example the system makes it more convenient by adopting methods of language conversion in text such as Indic Transliteration. It also produces output both textually



and in audio via text-to-speech technology, recognizing the difference in the proficiencies of languages people may converse in or their need for accessibility. In general, the system combines authentication, image capture, OCR/NLP processing, language conversion, and accessibility features in building a more fluid document analysis and enhancement through user interface interaction.

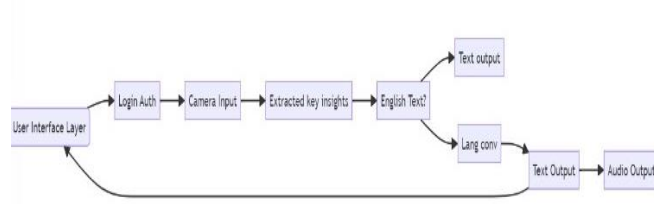


Fig 1.1 Block diagram

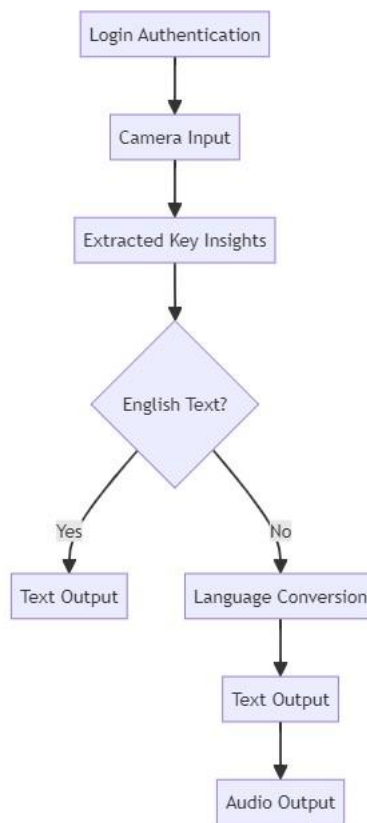


Fig 1.2 Work Flow



V. EXPERIMENTAL RESULTS

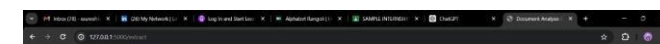


Document Analysis

Output



Fig 5.1 GUI



Document Analysis Output

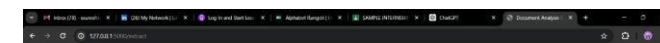
Electric Bill Information

Net Amount Due: Nilair

Due Date: 15/09/2017



Fig 5.2 Extracted text of Electricity Bill



Document Analysis Output

GAS Bill Information

Name : Subudhi Anil

Net Payable Amount: 8000

Due Date: 20/10/2020



Fig 5.3 Extracted text of Gas Bill with audio output

IV. CONCLUSION

Hence, our software is an important tool for digital inclusion and literacy in rural areas. A fingerprint sensor-based login authentication guarantees personalized access and thus user confidence. Advanced NLP algorithms of SpaCy can help the scanning and text extraction capability from important documents like voter IDs, passports, energy bills, PAN cards, water bills, and gas bills extract important information. It can also assist users to understand.

This will allow vital information and documents to be dealt with efficiently, hence improving the daily activities of the user. The application is also very accessible and user-friendly as it can translate into Hindi, Kannada, Telugu, and Tamil through Indic transliteration for even lesser-abled users in using English. Additionally, the gTTS text-to-speech function enables readers who might not read well by allowing them.

The content can be listened to in a preferred language. Another big tool is the bill reminder system, with an early reminder on approaching due dates to help their customers avoid missed payments and service interruptions.

**REFERENCES**

- [1] Kalyani Pakhale, "COMPREHENSIVE OVERVIEW OF NAMED ENTITY RECOGNITION: MODELS, DOMAIN-SPECIFIC APPLICATIONS AND CHALLENGES".
- [2] Adrian Lancucki, "Fastpitch: Parallel Text-to-Speech with Pitch Prediction".
- [3] Pranaw Kumar¹ , Gandhi Annamalai² , Sajini T³ , Anand Konjengbam⁴ , Praveen M² , G R Kasthuri² , Anil Prabhakar² , Shuo Qiao² , Sushanta K Pani¹ , Vishal Maral¹ , Binil Kumar S L³ , Arun Gopi³ , Neethu E A³ , Bira Chandra Singh¹ , Ranbir Singh⁴ , Hema A Murthy² , "Seamless Integration of Common Framework Indian Language TTSes in Various Applications".
- [4] Dipali Baviskar¹, Swati Ahirrao¹, Ketan Kotecha², "Multi-Layout Unstructured Invoice Documents Dataset: A Dataset for Template-Free Invoice Processing and Its Evaluation Using AI Approaches", IEEE 2021, DOI: 10.1109/ACCESS.2021.3096739
- [5] Rui Liu, Member, IEEE, Berrak Sisman, Member, IEEE, Guanglai Gao, Haizhou Li, Fellow, IEEE, "Expressive TTS Training with Frame and Style Reconstruction Loss", IEEE 2020, DOI: 10.1109/TASLP.2021.307636.