International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 12 | Dec 2023 www.irjet.net p-ISSN: 2395-0072

A Comprehensive Survey of Identity Document Data Extraction Techniques for Efficient KYC Verification and Identity Management

Vridhi Sachdev[1], Chirag Sandil[2], Abhaysinh Landge[3], Ashish Kolhe[4], Prof. Bhagyashree Dhakulkar[5], Ritesh Sachdev[6]

[1,2,3,4]Student, Department of Artificial Intelligence and Data Science Engineering, Ajeenkya DY Patil School of Engineering, Lohegaon, Savitribai Phule Pune University, Pune, Maharashtra, India
[5]Head of Department, Department of Artificial Intelligence and Data Science Engineering, Ajeenkya DY Patil School of Engineering, Lohegaon, Savitribai Phule Pune University, Pune, Maharashtra, India
[6]CTO & Director, Mapalon Technology Solutions Pvt Ltd, Pune, Maharashtra, India

---***---

Abstract - *Know Your Customer (KYC) processes are crucial in various industries, ensuring compliance and establishing trust with clients. This study delves into a significant emphasis on the seamless integration of cutting-edge Optical Character Recognition (OCR) technology. The deployment of state-of-the-art OCR algorithms stands as the keystone in this investigation, showcasing how logistics providers can revolutionize their KYC workflows. This transformative shift not only leads to heightened operational efficiency, driven by enhanced OCR accuracy and data quality, but also ensures seamless compliance with industry standards. The combination of KYC procedures and OCR technology in the logistics industry creates a new standard for safe and effective customer onboarding procedures that is advantageous to both logistics service providers and their customers.*

Key Words: Logistics, Know Your Customer, Optical Character Recognition, Natural Language Processing, Convolutional Neural Network

1.INTRODUCTION

Know Your Customer (KYC) is a process that businesses and financial institutions use to verify the identity, credibility, and suitability of their customers or clients. It involves collecting and verifying specific information about customers, such as their full name, address, date of birth, identification documents, and sometimes additional details depending on the industry and regulatory requirements.

KYC practices are implemented across various industries and verticals to ensure compliance, mitigate risk, and maintain the integrity of business relationships. These industries include financial services, e-commerce and retail, telecommunications, healthcare, real estate, insurance, logistics and more. Different industries require specific sets of documents for their KYC processes based on their unique regulatory requirements and risk assessment considerations.

The documents required for KYC may include standard government issued documents such as Aadhar cards, PAN cards, Passports, Driving Licenses to verify customer identities. Additionally, organizations may request supplementary documents such as bank cheques, GST certificates, income certificates, health insurance certificates, and credit cards from diverse sources. This variety of non-standard documents adds complexity to the KYC process, presenting challenges in gathering and verifying the required information from different entities.

However, KYC requirements in most industries are simplified as they typically only require standard government-issued documents for customer identification and verification. Therefore, AI/ML techniques can be employed to support real time KYC process in logistics, enabling efficient data extraction from documents for further processing. Implementing AI/ML techniques improves efficiency, accuracy, scalability, consistency, adaptability, and cost-effectiveness, allowing logistics service providers to process documents more effectively, streamline operations, and focus on higher-value tasks while reducing the risk of errors and enhancing data quality. Furthermore, logistics service providers may require third-party services that offer a platform for efficient verification and validation of customer information. Potential implementation to use Setu APIs in seamless verification process for KYC.

Integrating AI into the Logistics KYC workflow further strengthens compliance measures, minimizes fraud risks, and safeguards the integrity of customer information and most importantly help businesses in deduplicating customer information.

This paper presents a comprehensive survey of Optical Character Recognition (OCR) techniques employed to extract data from various identity documents which will be the first step in developing an AI tool to assist the KYC procedure in most industries.

2. LITERATURE REVIEW

The authors of [1] introduce the ARPIC (Automatically Recognize the Personal Information in CIC) framework to address the challenge of accurately recognizing personal information on watermarked ID card copies—a task that existing OCR technologies, primarily designed for printed documents, encounter significant difficulties in resolving. The proposed solution aims to automate this recognition process by incorporating four key modules: ID card region detection (IRDM), watermark removal (WRM), key text locating and recognition (KLRM), and text correction (TCM). The comprehensive approach combines data-driven and rule-based models to offer a robust solution, addressing issues related to watermarked copies. IRDM employs template matching, WRM utilizes a Conditional Generative Adversarial Network (CGAN) based approach, KLRM relies on a Connectionist temporal classification (CTC) based deep learning model, and TCM refines results based on the characteristics of key text elements. The framework's effectiveness is demonstrated through training and evaluation on a synthetic dataset, emphasizing the performance on watermarked regions (WRM) and key text recognition and correction (KLRM). The synthetic dataset creation involves 500,000 names, 701,681 addresses, and 2,852 issuing authorities, along with key personal details such as gender, nationality, birthday, and ID card numbers. This dataset, comprising both class-balanced and CIC-similar samples, is crucial for training the different modules, addressing various challenges associated with watermarked ID card copies.

Paper [2] details the application of the OpenNMT architecture for text recognition, encompassing two key components: Text Detection and OCR Engine. For both camera captured and scanned document images, text detection is essential, and a differentiable binarization is used as a text detector. By using adaptive thresholding in the segmentation network, this method improves performance and creates a probability map that is transformed into a binary image. The binarization operation is crucial because post-processing methods, including pixel clustering, organize pixels into text instances. The non-differentiable standard binarization function is replaced with the Differentiable Binarization (DB) method, ensuring full differentiability when trained alongside the segmentation network. Strategies, such as learning rate warm-up, light backbone and FPGM (Filter Pruning via Geometric Median) further enhance efficiency. For feature extraction and localization, the OCR Engine leverages a Convolutional Neural Network (CNN) and Bidirectional Long Short Term Memory (BiLSTM) in the encoder, and an Long Short Term Memory (LSTM) in the decoder to generate sequences. For text detection and recognition, an extensive set of images made up of 187,000 training visuals is created from a real dataset as well as ICDAR 2015 and ICDAR 2019. This comprehensive strategy incorporates cutting-edge methods to enhance text detection and recognition operations in a variety of document pictures.

In the paper [3], the authors address the pervasive challenges within computer vision studies concerning identity document recognition on mobile devices. The proposed approach involves image detection and OCR. The research is grounded in the context of a mobile application interface, akin to those utilized by Kazakhstan's banks. Assumption maximization, thresholding, and K-Means cluster analysis are only a few of the examples of the algorithms implemented in the study to further explore the important topic of image segmentation, resulting in a comprehensive high-level image processing pattern. The model that was developed utilizes a sequential combo of CNN and BiLSTM and follows the Convolutional Recurrent Neural Network (CRNN) architecture. To maintain compatibility with the BiLSTM layer, the height dimension is removed from the output vector by adaptive average pooling after CNN processing.

A framework for implementing an OCR system for Vietnamese identity cards is provided in [4]. The study utilizes a private dataset containing 2500 Vietnamese ID card images, which are partitioned into separate train and test datasets. The proposed approach employs a two-stage process consisting of text detection and text recognition. A pre-trained Connectionist Text Proposal Network (CTPN) model is utilized for text region cropping in testing as well as training images. Through an OCR labeling tool, all the cropped text elements are labeled. The CRNN model is then trained using the cropped regions from the training images and their associated labels. The final results for the cropped sections in the testing images are subsequently predicted using the previously trained CRNN model. The research compares predicted outcomes with the testing labels to assess the OCR system's accuracy. Three very important fields—the ID number, the name, and the date of birth—are reported to have an average accuracy of 78.0%.

Using a combination of OCR and post-processing methods incorporating Natural Language Processing (NLP), Paper [5] describes the creation of an Indonesian identity card extractor. The dataset used in this study comprises 50 ID cards of two distinct types, including 25 scanned images and 25 camera images. The Python Tesseract OCR module is applied to extract text data from the ID card images. Subsequently, NLP techniques are employed to refine the OCR output. The Punctual Remover is utilized to eliminate unnecessary punctuations, while Regular Expressions (Regex) are employed to identify character sequences that define specific search patterns. Additionally, a Word to Number converter is implemented to address misspelled characters. Moreover, to capture face photos from the ID cards, the Haarcascade Frontalface Default, sourced from OpenCV, is utilized. The total F-score achieved in this study is 0.78. Based on the two categories of images, this F-score is divided as follows: 25 ID cards photographed with cameras yield an F-score of 0.67, whereas 25 scanned images have a total F-score of 0.89. These results demonstrate the effectiveness and accuracy of the proposed Indonesian ID card extractor, showcasing its potential practical applications.

In the context of ID document analysis, Paper [6] juxtaposes the steps of DIB (Document Image Binarization) and OCR. The experiments were conducted on the MIDV-500 and MIDV-2020 ID document image datasets. For ID document binarization, a trained model was attained. The study indicated that, with the exception of situations where the input was the image's binary ground truth, algorithms for recognition performed best on non-binarized images. In terms of recognition rate, PPOCR2 algorithm was particularly better than all other techniques. When using image templates, the Otsu method performed admirably, but the domain-specific retrained U-Net network outperformed Otsu in terms of error rates.

A method for obtaining textual characters from personal identity documents is put forward in [7]. The algorithm is evaluated using various random sample datasets. The initial phase involves image pre-processing to eliminate noise and document boundaries enclosing the textual regions. After pre-processing, edge detection is employed to remove photographic images, leaving only textual characters in the binarized image. The next phase involves Segmentation, where line and character segmentation techniques are applied to segment the textual characters. The experiments were conducted on three datasets: Dataset1 (50 random images) achieved an accuracy of 90.81% in extracting textual characters, Dataset2 (70 random images) obtained 93.46% accuracy, and Dataset3 (100 images with poor scan quality) achieved 89.63% accuracy. The overall methodology resulted in a promising 91.21% accuracy rate for extracting textual characters from images.

Paper [8] focuses on improving user authentication for online services including e-commerce and personal finance by introducing a complete system for automatically obtaining information from Vietnamese ID cards acquired by camera devices. The current issue is the variation in identification card formats brought on by the various text fonts, sizes, and placements, as well as the potential for image quality problems owing to lighting, perspective, and blurring. The proposed system has numerous phases, including: 1. Using corner detection and national emblem detection, the ID card is oriented and aligned to a standard aspect ratio. 2. Finding and trimming the exact foreground target region of the orientated ID card in different settings. 3. Text localization, which identifies the necessary texts using a deep convolutional neural network model, is followed by classification based on the texts' locations. 4. Text recognition utilizing the Attention OCR architecture to extract pertinent information fields like an ID number, name, and date of birth. Experiments show that the proposed method yields a mean accuracy of over 91% and cuts processing time. The paper also covers data augmentation, training methods, testing procedures, and evaluations in addition to discussing the particular difficulties associated with using this technology in the Vietnamese context.

A powerful deep YOLOv3 detection network is used in [9] to provide an automated system for the real-time visual detection and recognition of numerous fields on a driver's license. The authors present an approach that enables simultaneous identification of several field elements in a single shot in order to preserve high detection accuracy and adaptability. This is accomplished by implementing and optimizing the recently created YOLOv3-608 detector, which demonstrates an outstanding 97.5% accuracy in recognizing 11 fields from the new Taiwan driver's license. The authors conducted examinations in top, bottom, left, and right tilting view configurations in order to further examine the system's capabilities. These configurations' respective obtained accuracies were 93.3%, 90.2%, 97.5%, and 94.3%.

An efficient traffic management system has been developed, incorporating a driver's license recognition system based on OCR technology in [10]. The identification number on the license consists of either 17 digits and one capital letter X or 18 digits, and extracting the necessary features from it proves to be relatively straightforward. To achieve accurate character recognition, a convolutional neural network classifier is trained using segmented characters from the identification number. Through a series of experiments, the recognition rate of the identification number area has been successfully improved to reach an impressive 94.81%.

The development of a method to identify Indonesian electronic identification cards using a mix of image processing and OCR is covered in [11]. The purpose is to facilitate the automatic processing of customer data for online and offline purchases. The researchers' image processing methods and OCR allowed them to recognize ID cards with 98% accuracy. The system was integrated into the website interface of an Indonesian automotive company. The paper highlights the rapid development of Information Technology and how businesses can benefit from automating data input processes. Information from the ID card images is retrieved and pre-processed via image processing to obtain the elements that are needed. OCR is then utilized to instantly identify text characters, even handwritten writing, inside the photos. The performance of the OCR depends on how well the document is provided. The study contrasts the outcomes of name and identity number recognition using two different Tesseract models. One model employs manually generated training data from five ID cards using Tesseract 3.05, and the second model employs pre-existing training data in Tesseract 4.0, which incorporates a neural network model (LSTM) and includes Indonesian language text in several fonts.

Table -1: Comparative Study of Literature Review

Paper	Dataset	Techniques Employed	Accuracy/Score
[1]	Synthetic dataset of 6,000,000 class-balanced and 500,000 CIC-similar samples	ID card region detection (IRDM), watermark removal (WRM), key text locating and recognition (KLRM), and text correction (TCM)	99.71 %
[2]	36,000 ID numbers with and without colored backgrounds	Differentiable Binarization (DB), CNN and BiLSTM	100%
[3]	2000 scanned and camera-captured document images	CRNN architecture, employing a sequential combination of CNN and BiLSTM	Train set – 99.26% Validation set – 98.33% Test set – 98.71%
[4]	2,500 Vietnamese ID card images	2 stage approach: text detection - pretrained multi-language CTPN and text recognition - custom CRNN model	78.0%
[5]	50 ID cards of two types, 25 scanned images, and 25 camera images	Pytesseract OCR followed by post processing using NLP	Total F-score – 0.78 F-score of 25 scanned images – 0.89 F- score of 25 camera images – 0.67
[6]	MIDV-500 and MIDV-2020	2 stages: Document Image Binarization (DIB) and Optical Character Recognition (PPOCR2)	97%
[7]	3 experimental datasets: Dataset1 - 50 images Dataset2 - 70 images Dataset3 - 100 images	Image preprocessing and line and character segmentation techniques	Dataset1 - 90.81% Dataset2 - 93.46% Dataset3 - 89.63% Overall - 91.21%
[8]	2500 samples	Image Preprocessing and deep CNN	91%
[9]	617 Taiwan driver licences	YOLOv3-608 network	97.5%
[10]	Chinese Driving Licence	Image preprocessing, character segmentation, character recognition using CNN	94.81%
[11]	Indonesian Electronic ID cards	Tesseract OCR	98%
[12]	10,000 collection of alphanumeric characters taken from an Indonesian citizen ID card	CNN	91%
[13]	3256 ID Card images (1628 front-side and 1628 back-side) with different resolutions of 200 dpi, 300 dpi and 400 dpi.	CNN with multidimensional LSTM	precision, recall and f-score between 95 and 99% for different fields
[14]	1000 Images	RetinaNet for text detection and Inception-v3 CNN network for text recognition	68% - 87%
[15]	1000 Document Images	morphological, edge, texture, and region-based approaches	Clear images – 98.1%, Burt images – 62.5% Blurred images – 96%
[16]	3,256 Vietnamese identity cards, 400k manual annotations, and more than 500k artificially generated texts for verification	U-Net, VGG16, contour detection, and Hough transformation, followed by Optical Character Recognition using CRAFT and Rebia neural networks	Segmentation accuracy is 94%, classification accuracy 99.4% and recognition accuracy 98.3%.
[17]	8,562 government human resources documents	Foxit, PDF2GO, Tesseract	PDF2GO F1-score – 86.27%. Foxit has the highest F1-Score for tabular structures – 84.01% Tesseract has the highest F1-Score for non-tabular structures – 92.46%

Convolutional Neural Networks, a cutting-edge Deep Learning model, is used by the authors of [12] to propose a model for citizen ID card detection. A dataset consisting of 10,000 alphanumeric characters was created from Indonesian citizen ID cards and encoded in HDF5 format. To enhance the dataset's quality, various image processing techniques were applied, including Grayscale, binary, and morphological transformations. The training, validation, and testing sets were then created from the original dataset. The training data was used to develop the model, the validation data was used to evaluate each training epoch, and the testing data was utilized for assessing model performance. The classification results were combined to form ID card fields, such as names, numbers, or addresses. After conducting 100 epochs of training, the CNN model with pre-processing achieved training accuracy of 91% and validation accuracy of 90%. On the other hand, the SVM model with pre-processing attained training accuracy of 64% and validation accuracy of 62%.

Through the combination of CNN and LSTM, [13] proposes an innovative and effective approach for identifying information fields in Vietnamese ID cards. The work entails doing experiments on a dataset of 3256 ID card photos, including 1628 photographs of the front and a matching quantity of images of the back. These identification cards were gathered from various provinces and displayed differences in quality, font sizes, and printing techniques. They were scanned at resolutions of 200dpi, 300dpi, and 400dpi. With precision, recall, and f-measure scores ranging from 95% to 99% across all identified information categories, the suggested implementation outperforms the findings of earlier studies.

Paper [14] presents a comprehensive OCR system that leverages a two-step approach: text detection and recognition. For text detection, the proposed method employs RetinaNet, while Inception-v3 CNN network is utilized for feature extraction in the recognition model. Following the extraction of the features, the data is input into a BiLSTM network that has been improved with Spatial Attention for precise mask prediction based on the current RNN state. The experiments were conducted on a dataset of 1000 images containing ID cards with arbitrary positions. The results indicate an overall accuracy of 86.7% for ID number recognition, 85.7% for Name, 79.7% for Date of Birth, and 70.24% for Address recognition. Combined accuracy for ID number, Name, and DoB, is 68.5%.

Paper [15] discusses the challenges and techniques for extracting text from images, particularly in the context of document image analysis. The research examines various text extraction approaches, including morphological, edge, area, and texture-based ones. It highlights the significance of text extraction for various applications, such as document processing and video content summary. The implementation of the text extraction techniques is demonstrated, and the paper presents a performance evaluation using precision and recall rates. The evaluation reveals that morphological and edge-based techniques show better results compared to region and texture-based methods. The authors implement an interface for text extraction and achieve efficient results using built-in functions of pytesseract and OpenCV libraries. Accuracies achieved for clear images, burt images and blurred images are 98.1%, 62.5% and 96%.

A complete method for information extraction from pictures of Vietnamese identification cards is laid out in Paper [16]. Manual extraction takes a lot of time and is prone to mistakes in developing nations like Vietnam. The suggested solution consists of 3 stages and makes use of 2 image processing methods and 4 neural networks. Preprocessing using U-Net, VGG16, contour detection, and Hough transformation are among the different stages. Optical character recognition using CRAFT and Rebia neural networks is the subsequent step. Levenshtein distance and regular expressions are used for post-processing. 3,256 Vietnamese ID cards, 400k manual annotations, and more than 500k synthetic texts were used to build the dataset. The accuracy of segmentation is 94%, classification is 99.4%, and recognition is 98.3%.

In [17], Named Entity Recognition (NER) is discussed in relation to Information Extraction (IE) from text documents, especially those produced by OCR. To identify which OCR engine performs best for IE using NER, the study analyses three - Foxit, PDF2GO, Tesseract. Through the use of six categories, two structures, and four metrics, they examined 8,562 government human resources documents. Names, employee IDs, document numbers, publishing dates, employee positions, and names of family members were among the significant details that were automatically extracted. Preprocessing operations were carried out independently for each OCR engine. The results show that every OCR engine has benefits and limitations; Tesseract performs best in terms of NER extraction and conversion time with superior accuracy but falls behind in terms of the quantity of entities acquired. The F1-score leader on average is PDF2GO (86.27%). The greatest F1-Score (84.01%) in terms of tabulated document arrangement belongs to Foxit. For non-table document structure, Tesseract has the greatest F1-Score (92.46%). According to those results, Tesseract is more suited for non-table documents, while Foxit is better suited for tabular documents.

3. CONCLUSION

This study emphasizes the pivotal role of OCR technology in identity verification within logistics per se. The integration of advanced OCR models, especially tailored for identity documents, holds immense potential for streamlining processes. By investing in cutting-edge OCR, logistics companies can optimize data extraction, ensuring a seamless and secure verification process. This advancement not only enhances operational efficiency but also sets new industry standards, benefiting both providers and customers.

Additionally, the implementation of new enhanced OCR models for identity documents presents a significant leap forward in accuracy and efficiency. These models, leveraging the latest advancements in machine learning, offer a refined approach to extracting critical information from various document formats. By harnessing the power of these state-of-the-art OCR solutions, logistics service providers can further elevate the integrity of their KYC processes. This not only reduces the margin of error but also reinforces compliance measures, ultimately fostering a more secure and reliable environment for all stakeholders involved.

REFERENCES

[1] W. Miao, C. Zhang, Z. Hu, and C. Zhang, "ARPIC: Personal Information Recognition of ID Card with Interference of Watermark," in 2021 7th International Conference on Computer and Communications, ICCC 2021, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1505–1509. doi: 10.1109/ICCC54389.2021.9674471.

[2] M. K. Gupta, R. Shah, J. Rathod, and A. Kumar, "SmartIdOCR: Automatic Detection and Recognition of Identity card number using Deep Networks," in Proceedings of the IEEE International Conference Image Information Processing, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 267–272. doi: 10.1109/ICIIP53038.2021.9702703.

[3] B. Gulnara and A. Yerassyl, "Using Image Processing and Optical Character Recognition to Recognise ID cards in the Online Process of Onboarding," in SIST 2022 - 2022 International Conference on Smart Information Systems and Technologies, Proceedings, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/SIST54437.2022.9945823

[4] D. P. Van Hoai, H. T. Duong, and V. T. Hoang, "Text recognition for Vietnamese identity card based on deep features network," International Journal on Document Analysis and Recognition, vol. 24, no. 1–2, pp. 123–131, Jun. 2021, doi: 10.1007/s10032-021-00363-7.

[5] F. M. Rusli, K. A. Adhiguna, and H. Irawan, "Indonesian ID Card Extractor Using Optical Character Recognition and Natural Language Post-Processing," Dec. 2020, [Online]. Available: http://arxiv.org/abs/2101.05214

[6] R. Sánchez-Rivero, P. V. Bezmaternykh, A. V. Gayer, A. Morales-González, F. J. Silva-Mata, and K. B. Bulatov, "A joint study of deep learning-based methods for identity document image binarization and its influence on attribute recognition," Computer Optics, vol. 47, no. 4, pp. 627–636, 2023, doi: 10.18287/2412-6179-CO-1207.

[7] P. B. R and V. K. R, "Robust text extraction for automated processing of multi-lingual personal identity documents." [Online]. Available: https://www.researchgate.net/publication/304888413

[8] H. T. Viet, Q. Hieu Dang, and T. A. Vu, "A Robust End-To-End Information Extraction System for Vietnamese Identity Cards," in 2019 6th NAFOSTED Conference on Information and Computer Science (NICS), IEEE, Dec. 2019, pp. 483–488. doi: 10.1109/NICS48868.2019.9023853.

[9] C.-M. Tsai, J.-W. Hsieh, M.-C. Chang, and Y.-C. Lin, "Driver License Field Detection Using Real-Time Deep Networks," 2020, pp. 603–613. doi: 10.1007/978-3-030-55789-8_52.

[10] K. Kang and H. Xie, "Design and Implementation of Driver's License Recognition System," in 2018 13th International Conference on Computer Science & Education (ICCSE), IEEE, Aug. 2018, pp. 1–5. doi: 10.1109/ICCSE.2018.8468822.

[11] W. Satyawan et al., "Citizen Id Card Detection using Image Processing and Optical Character Recognition," in Journal of Physics: Conference Series, Institute of Physics Publishing, Jul. 2019. doi: 10.1088/1742-6596/1235/1/012049.

[12] M. O. Pratama, W. Satyawan, B. Fajar, R. Fikri, and H. Hamzah, "Indonesian ID Card Recognition using Convolutional Neural Networks," in 2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), IEEE, Oct. 2018, pp. 178–181. doi: 10.1109/EECSI.2018.8752769.

[13] N. T. T. Tan and N. Ha Nam, "An Efficient Method for Automatic Recognizing Text Fields on Identification Card," VNU Journal of Science: Mathematics - Physics, vol. 36, no. 1, Mar. 2020, doi: 10.25073/2588-1124/vnumap.4456.

[14] H. D. Liem et al., "FVI: An End-to-end Vietnamese Identification Card Detection and Recognition in Images," in 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), IEEE, Nov. 2018, pp. 338–340. doi: 10.1109/NICS.2018.8606831.

[15] O. K. Barawal and D. Y. Arora, "Text Extraction from Image," International Journal of Innovative Research in Engineering & Management, pp. 89–92, Jun. 2022, doi: 10.55524/ijirem.2022.9.3.12.

[16] K. Nguyen-Trong, "An End-to-End Method to Extract Information from Vietnamese ID Card Images," International Journal of Advanced Computer Science and Applications, vol. 13, no. 3, 2022, doi: 10.14569/IJACSA.2022.0130371.

[17] T. W. Ramdhani, I. Budi, and B. Purwandari, "Optical Character Recognition Engines Performance Comparison in Information Extraction," International Journal of Advanced Computer Science and Applications, vol. 12, no. 8, 2021, doi: 10.14569/IJACSA.2021.0120814.