

A Machine Learning Framework for Bills and Invoices Analytics

Sheeban E Tamanna¹ and Mohd Sakib² and Mohd Usama^{3*}

¹Department of Information Technology, National Institute of Engineering, Mysore, India

²Department of Computer Science, AMU, Aligarh, India

³Department of Computer Science and Engineering, KARE, Krishnankoil, India

*Corresponding author: Mohd Usama, Department of Computer Science and Engineering, KARE, Krishnankoil, India, Tel: 08604314845, E-mail: drmohdusama.cs@gmail.com

Received Date: November 05, 2025 **Accepted Date:** November 25, 2025 **Published Date:** November 29, 2025

Citation: Sheeban E Tamanna, Mohd Sakib, Mohd Usama (2025) A Machine Learning framework for Bills and Invoices Analytics. J Comput Sci Software Dev 4: 1-16

Abstract

Electronic invoices generated by various agencies exhibit diverse unstructured data that can be leveraged to address contemporary issues. For instance, grocery bills can serve as a source of information for predicting the demand of specific items within a given region, while travel invoices can provide insights into travel demand trends over a certain period. These data sets hold potential for data-driven decision-making and can be analyzed using modern techniques such as machine learning and data analytics to extract actionable insights. The challenge is to extract the data from different unstructured E-bills/invoices into some meaningful structured way to reuse it. Thus, in this paper, we present an end-to-end machine learning-based framework that extracts meaningful data from all kinds of bills/invoices and saves it into structures format (JSON). Our ML framework consists of classification, finetune YoloV5, and NER models. The classification model classifies the bill into specific categories; YoloV5 extracts the data from table-like objects, and the NER model extracts other metadata in bills/invoices. Our ML framework can generate structured data from any bills/invoices with table detection accuracy 92.00%, and NER accuracy 89.00%. This paper provides a comprehensive guide for implementing the proposed framework, which can be used to address current issues, such as predicting demand for items in a region using grocery bills or forecasting travel demand based on travel invoices.

Keywords: Bills/Invoices, OCR, Text Processing, YOLOv5, NER, Classification, Prediction, Deep Learning

Introduction

E-bills are important in our everyday lives since they serve as proof of payment and as a record of daily expenses. These bills can be used to address various financial issues. With the growth of information technology, a large number of bills/invoices are produced. Our model focuses on six types of bills: groceries, restaurant, travelling, electricity, water, and hospital bills. Extracting data from printed bills using computers would be a more efficient solution. In light of this, our research discusses the possibility of developing a system that can automatically extract information from bill images without human intervention using available extraction tools.

Extraction and identification of unstructured data has always been a challenging task. Processing such enormous amounts of data is difficult; however, recent developments in computer vision have made this task more manageable [1]. Our research aims to gather information from invoices and process them to determine current user preferences and how they change over time. This can also help identify the financial expenses of an organization or a family. Most invoice-issuing units still rely on manual identification techniques, which leads to issues such as low efficiency, errors, and difficulty in securing information. To address these challenges, this research proposes an end-to-end framework for bills and invoice recognition using machine learning.

The system first takes an invoice image as input and processes it through an optical image conversion tool, “Optical Character Recognition (OCR),” which extracts the information. OCR typically involves three steps: scanning the document, recognizing the text, and saving the output in a desired format, generally as text. This text contains two types of data: table data and metadata (header and footer). We used YOLOv5 to extract table data and SpaCy NER to extract metadata. The collected information is saved in multiple parts, including the organization and its address, bill ID, tax calculation, and total. These components are divided into sections so the data can be saved in JSON format, which can then be automatically stored in a MongoDB (NoSQL) database.

This service-oriented application maintains invoices for a user and makes them available for analysis and visualization. The machine learning approach for digitization is discussed in this work, including unsupervised models for entity recognition and bill processing. The work highlights how much unstructured information is created through everyday interactions. Unstructured data can be found in various forms, such as PDFs, SMS, and formal documents. It is an important aspect of data gathering for this project since it is extracted and used for decision-making, and the resulting information improves the system’s efficiency.

In this project, we utilized three models: NER, OCR, and YOLOv5. NER (Named Entity Recognition) identifies and classifies named entities in text, such as people, organizations, and locations. OCR (Optical Character Recognition) extracts text from images or scanned documents. YOLOv5 (You only look once, version 5) is a popular real-time object detection model that can detect and localize objects in images or videos. Using these models together enables the extraction of valuable information from multiple data sources, providing insights that support decision-making and enhance performance in various applications. Further sections present an overview of all these models.

Optical Character Recognition (OCR)

The field of Artificial Intelligence encompasses a number of subfields, one of which is computer vision. This subfield investigates how computers interpret digital images; unlike humans, computers don’t have eyes or a mind to interpret what they see. One use of computer vision is identifying digital images and converting them into strings through Optical Character Recognition (OCR) (Figure. 1). It involves extracting textual information from a picture and converting it into a format a machine can understand [2]. It also allows automated extraction of data from printed or handwritten text in scanned documents, which is then transformed into a machine readable format for tasks such as editing or searching [3]. OCR can be employed in a wide range of applications, from reading ZIP codes to reading number plates [4], [5], [6].

However, computerized OCR systems still face difficulties. Some letters and numbers look almost the same to computers, making them hard to differentiate. Additionally,

extracting text that is not clearly visible can be challenging. Tausheck, a German scientist, proposed the concept of OCR technology in 1929 during his work on pattern recognition [2]. After a decade of research, simple recognition systems such as postal code readers were developed, making it possible to automatically recognize codes on mail [7]. Similarly, China began analyzing OCR technology after 1970

and developed systems to recognize Chinese characters. In 1986, researchers at Tsinghua University, in collaboration with other universities, proposed an invoice recognition system based on OCR, leading to the development of the Chinese invoice OCR model [8]. Since then, various OCR models have been proposed for different purposes. In this study, we used the Easy OCR technique, although other OCR algorithms could also be applied as discussed above.

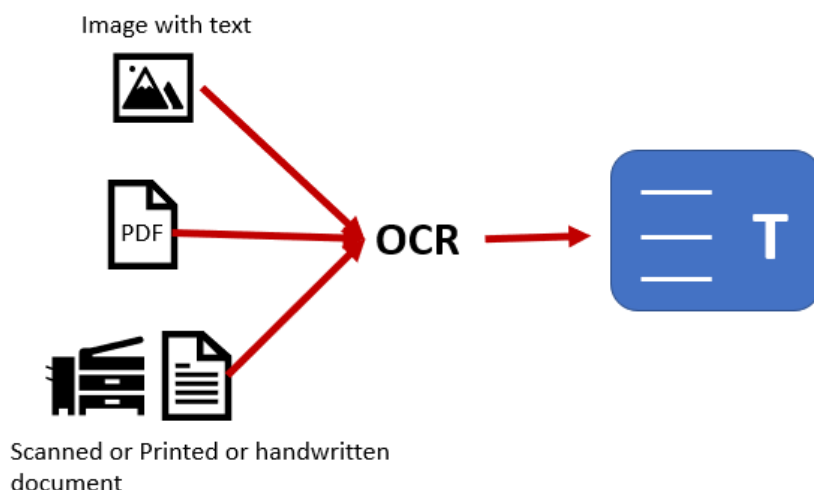


Figure 1: OCR Converts the Contents of Images, Pdfs And/or Scanned Document into Text Format.

Yolo

You Only Look Once (YOLO) is a single-shot object detection algorithm presented by Joseph Redmon et al. in 2015 [9]. It is extremely fast and suitable for real-time applications. Object detection algorithms are generally categorized into two types: region proposal-based methods such as R-CNN, Fast R-CNN, and Faster R-CNN [10]–[12], and regression-based methods such as YOLO. Both techniques generate Region Proposal Networks (RPNs) and classify these proposals.

Before YOLO, R-CNN models were widely used for object detection. Although they produced reliable results, they were slow due to the multiple steps involved in generating bounding box proposals. YOLO was developed to eliminate these multiple stages and perform detection in a single step, reducing processing time. In contrast, two-stage detectors emphasize accuracy, whereas one-stage detectors focus on speed. One-stage models like YOLO achieved high speed by simplifying the architecture and removing

the pipeline [9], though they initially had lower accuracy. Continuous improvements have been made, and a comprehensive review of YOLO variants (v1–v5) is provided in [13].

In this study, we used YOLOv5 to extract table content from bills and invoices. YOLOv5, created by Glenn Jocher (2020), incorporates mosaic data augmentation and automated anchor box learning.

Named Entity Recognition (NER)

Another subfield of Artificial Intelligence is Natural Language Processing (NLP), which focuses on enabling computers to understand and process human language. The concept of named entities was first introduced during the Sixth Message Understanding Conference (MUC-6), which aimed to extract structured information from unstructured text. Identifying references to names (person, organization, location) and numeric expressions (date, time, money) became known as Named Entity Recognition (NER) [14].

NER is a widely used NLP approach for recognizing and classifying key information in text into predefined categories [15].

SpaCy is a widely used Python NLP library that provides several functionalities, including NER. In this project, we used SpaCy to extract metadata such as headers and footers from bills and invoices. SpaCy's NER model uses statistical techniques like Conditional Random Fields (CRF) and neural network architectures to analyze text and predict named entities with corresponding labels. By leveraging these features, we accurately extracted metadata, enabling further analysis and efficient processing. Overall, SpaCy's versatility and advanced capabilities make it a powerful tool for tasks such as metadata extraction from bills and invoices.

Literature Survey

Before commencing this research, a large number of currently available technologies and previous work were examined and taken into account. This provided us with adequate knowledge and information that clarified the methodology that was used to get the most favorable outcome. Tesseract, an open source and well known OCR technique is still in demand. The Tesseract performs its fundamental processing using a pipeline technique that is comprised of a series of step-by-step procedures, during which a connected component analysis is carried out.

In this section, we are presenting an overview of previous studies concerning with the specific task of extracting structured information using machine learning. When it comes to the process of extracting text from invoices and bills, the input document is frequently more effectively represented as a dispersed arrangement of many text blocks rather than as a single contagious body of text [16, 17] started the very first review on image segmentation and recognition on a chine bank invoice which is a color picture and implemented using an encryption standard. The essential detailed information on the invoice is located at the right angle and is made up of 26 English cross-rows and 0-9 numerals. Each invoice has a distinct encoded version of this important information. The Chinese bank digitally controls computer-generated bills using an image processing technique. 92% of the letters are correctly identified.

Many researchers have utilized the text recognition or OCR technologies to translate the scanned document to a machine readable format. Tesseract's OCR engine produces low quality results if the images given to it are very noisy or contain any item that are not necessary. Therefore, images need to undergo preprocessing before being sent to the tesseract OCR system. In one of the studies, authors [18] analyzed the scanned document in two steps; skew detection and segmentation. Primary responsibilities of skew detection are visualized and coordinated the image accurately. Noise should not be there. The modules for text detection and localization are closely related with each other. Similarly, the name implies, text segmentation is the procedure of splitting up scanned text into useable parts like words, sentences, or ideas.

Furthermore, Open CV (Open Source Computer Vision Library), a machine learning and computer vision library, built to provide the common structure for computer vision applications, taken into consideration to remove the unnecessary and unwanted noise and objects from background [19]. The Open CV library allows for the processing of images with greater quality and having effective data extraction.

Authors [20] made an attempt to transform physical invoices to the digital environment. They have tesseract OCR with CNN to recognize raw materials from handwritten text. In this investigation, distinct CNN architectures were employed. LeNet-5 is a basic and tiny architecture that has been successfully used to read bank checks. The alternative architecture, VGG architecture, where instead of the vast filters like 9x9 and 11x1 utilized in prior architectures, tiny filters like 3x3 are employed to extract features. Because invoice photos are structurally identical, a more thorough inspection was necessary smaller size filters provide more detailed picture analysis. CNN architectures, which are popular in image classification, have also been applied in document classification research. In the Deep Document Classifier document classification study, a CNN architecture-based document classification system was demonstrated utilizing the Image Net data set, which contained millions of sample pictures. The system uses document pictures with dimensions of 3x227x227 and comprises of four convolutions and three completely linked layers to analyze and extract the fea-

tures in bill effectively.

Table 1: Relevant Studies on Text Extraction from Bills

Year, Ref	Benchmark	Techniques	Contribution
2018, [21]	Bank serial number of Renminbi (RMB)	Block-wise Prediction Network (BPN), OCR	Provided a full stage study of bank serial number recognition system with 99.92% character accuracy and 99.24% Renminbi accuracy.
2018, [22]	Edge detection system from images	Convolutional neural network (CNN), OCR	Preprocessing and post processing is applied with various phases. Testing accuracy is 95% if image taken at a distance of 10m. Model, failed when apply on blurred images.
2018, [23]	Total amount of the bill from daily uses	OpenCV, Tesseract, OCR	This research shown a good result when apply on digital images invoices but often for handwritten bills. Accuracy and efficiency relatively good.
2017, [24]	Printed hard copy documents	OCR	Extracted all kinds of blurred images and converted binary text object into ASCII
2009, [25]	Chinese bank bills, scanned input	Text extraction tools.	Applied preprocessing, image processing and then number and letter segmentation. They got 92% accuracy on letter recognition and 99 % accuracy on number recognition.
2017, [26]	Scanned and camera capture images of ERP system	Stroke Width Transform (SWT) text detection	Extracted the edge of images using the canny edge detector, define the character candidate region, and then clustering the region into text.
2019, [27]	Invoice information based on template matching	OCR along with Contour extraction, edge detection	Identify the information in four steps; preprocessing, matching of template, apply OCR, and exporting of information. They claimed that important information such as money, goods and purchases were accurately identified with 95% of accuracy.
2019, [28]	Visual text of information	CNN based feature extractor	EATEN (Entity-aware attention text extraction network) proposed which is a single shot method to extract information. Study provides a new perspective of text recognition.
2018, [29]	Arabic and Latin scanned invoice	OCR without layout	Entity extraction framework is proposed. Structure relationship graph built for entities. Failed to capture the large portion of Arabic invoices. Work carried out in around 150 real invoices.
2016, [30]	Chinese bank statement images	OCR engine for digits	Geometrical and sentimental knowledge is used to recognize the digit. Showing the page layout accuracy in 98% and recognition accuracy is 89%. Work carried out on 569 bank statement images

Other studies have focused on the use of natural language processing (NLP) techniques for bills and invoices analytics. For instance [17] developed an NLP-based approach for extracting structured information from invoices, such as line items and their corresponding prices. In addition,

some studies have proposed end-to-end machine learning frameworks for bills and invoices analytics. For example [31] developed a framework that combines optical character recognition (OCR) with machine learning techniques for extracting information from invoices. Their framework

achieved high accuracy in extracting invoice data, such as supplier name, invoice date, and total amount.

Overall, these studies demonstrate the potential of machine learning and NLP techniques for bills and invoices analytics, and highlight the importance of developing end-to-end frameworks for automating the process.

Materials and Methodology

A general invoice often contains three parts: the organizational details, item/commodity details, and total amount details. The organizational details can be further classified as organization name, address, phone number, and email. The item/commodity details part also contains item name, quantity, unit price, and total price of the item in a list. Similarly, the details in the total part include the total amount, tax amount, and total amount with tax.

The Azure Form Recognizer service [32] offers JSON-formatted output to facilitate data processing. For our project, we utilized the prebuilt Receipt model in Form Recognizer, which performs an analysis of the input image and returns a JSON-formatted output containing key-value

pairs for each extracted entity or line of text. The output includes information on the success or failure of the operation, as well as details on the Form Recognizer service, such as the version and prebuilt model used. The JSON output also includes a confidence level for each extracted field, indicating the degree of accuracy of the extracted information. Additionally, it provides information on the extracted entities, including the entity's type, value, text, and "Bounding Box" information, which specifies the position of the entity within the image using coordinates. Confidence values range from 0 to 1, with higher values indicating greater accuracy. Ultimately, we will leverage machine learning to automate this entire structured process, standardizing the extracted data for further analysis.

Easy OCR is a deep learning-based Optical Character Recognition (OCR) tool that utilizes Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to accurately recognize and extract characters from an image (Figure2). The text output generated by Easy OCR represents the extracted characters from the input image, and can be further processed and analyzed for various natural language processing tasks.

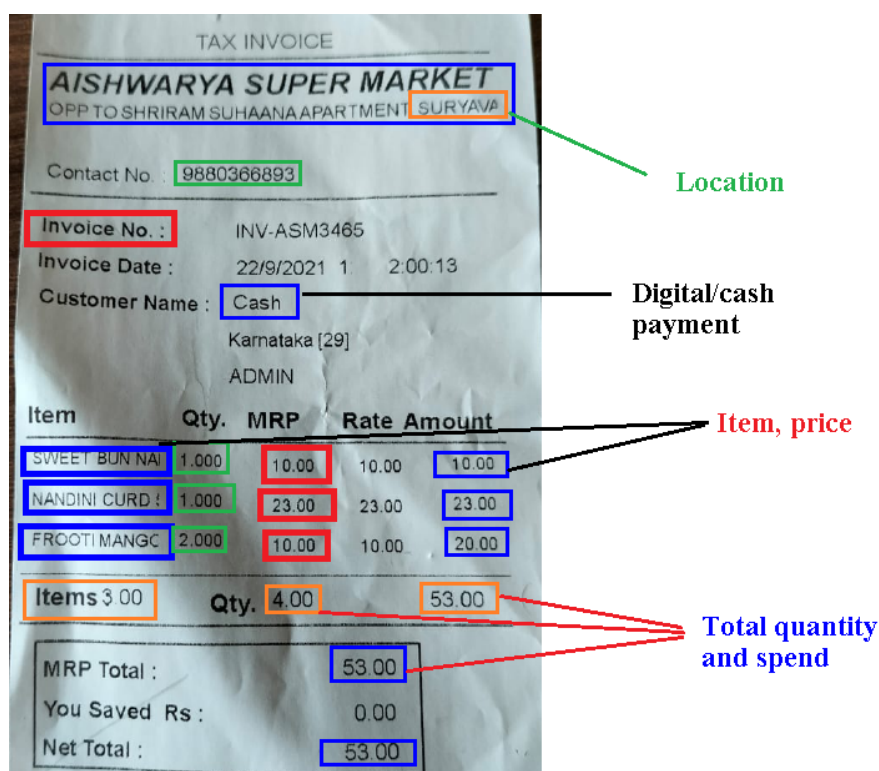


Figure 2: Text extracted from bill using Easy OCR

In the view of the above narration, our proposed framework consists of 4 components including image-to-text translation, invoice classification, meta-text recognition and table detection. For each part we used separate algorithms. Working of all components is described in sequences as follows

Image-To-Text Translation

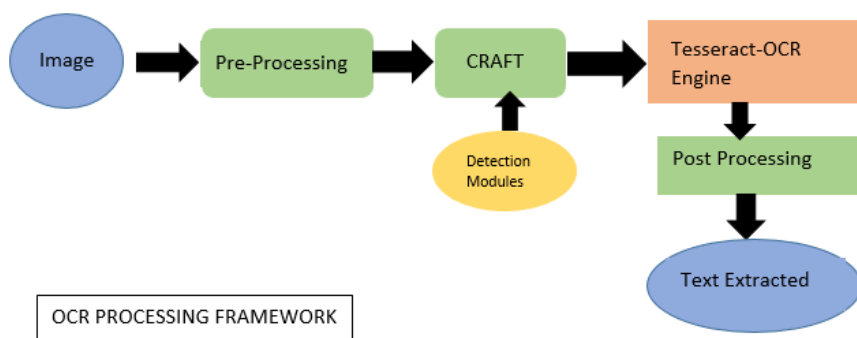


Figure 3: Framework of Easy-OCR

This process typically involves the following steps:

Pre-Processing

The pre-processing stage is the most crucial step in recognition of any hand written or printed invoices. It requires a number of processes, such as binarization, noise reduction, smoothing, base line detection, skew detection, filtering, skeletonization, or thinning, among other things.

Segmentation

In the segmentation process, the pre-processed input data is broken down into smaller pieces so that the specific features within the input picture may be identified. Once the lines have been segmented, bounding box methods may be used to extract the characters one by one. At last, the size of each character's image is standardised at 32 by 32 pixels.

Feature Extraction

The system extracts various features from the segments, such as shape, size, and texture, to identify the characters. For character recognition, we first employ a process

In the initial phase of our project, we employed Optical Character Recognition (OCR) technology to convert images into machine-readable text format. OCR technology utilizes advanced artificial intelligence algorithms, such as deep learning neural networks and pattern recognition, to accurately recognize and extract characters from the input image. The converted text data can then be further processed and analyzed for various natural language processing tasks. (Figure 3)

called "feature extraction," in which the most useful features of the text picture are extracted.

Character Recognition

The system uses machine learning algorithms to match the features of the segments to a database of known characters, such as the ASCII or Unicode character sets.

Post-Processing

The recognized text is post-processed to correct errors, fill in missing characters, and format the output to match the desired format.

NER Algorithm for Entity Recognition

In our project, we utilized a NER (Named Entity Recognition) algorithm to identify and extract entities from the meta data section of invoices/bills. The NER algorithm is a natural language processing technique that enables the identification of named entities, such as people, places, organizations, and other significant items from text data. The NER algorithm identifies and extracts entities, such as invoice/bill type, restaurant name, mobile number, invoice

number, and other relevant information from the meta data section of invoices/bills. By leveraging NER technology, we were able to automate the entity recognition process, which typically requires manual effort and time. The extracted entities can then be used for further processing and analysis, enabling organizations to efficiently manage and process their invoice/bill data.

We have performed this entity extraction task with Spacy v3. Spacy has a very fast statistical entity recognition engine that assigns tags to spans of consecutive tokens. With the use of its 'ner' pipeline component, Spacy can identify token spans that correspond to a given set of named entities. The 'ents' property of a Doc object contains information about them.

The purpose of using Spacy v3 for NER entity ex-

traction is to find various entities and their keys in a receipt or bill. Such as store_name, store_address, contact, gst_no, date, time, total etc. NER tagging solved the problem of rule-based entity extraction using regular expressions. As regular expressions would not work if an entity is not matching with the rules we created using regex. So, to solve this problem we use the NER method.

Data Extraction

Our bill storage approach is designed to be versatile and highly generalized, supporting all types of bills. We begin by storing store-related details, followed by item/-customary details, and finally total details. Our approach ensures that the storage of bills is comprehensive and adaptable to various formats, enabling our system to effectively handle a diverse range of billing scenarios.

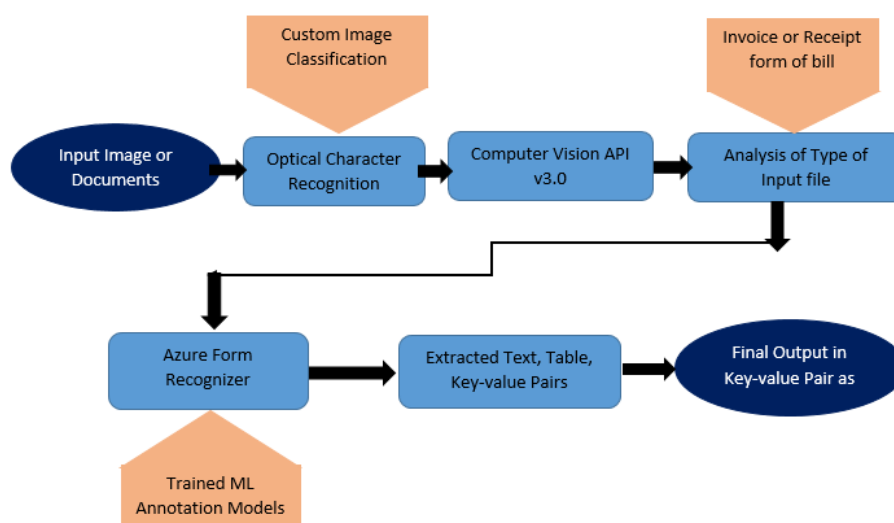


Figure 4: Data extraction

This flexibility extends to the tax details and total calculation, which may vary depending on whether an item's price is inclusive or exclusive of taxes. Precise and accurate calculation is essential, and we ensure this by performing the calculation concisely. Once the data is structured, we store it in a database using Mongo DB, a No SQL database that allows for structured data storage. Mongo DB's ability to store data in JSON record format aligns well with our chosen bill format, providing a flexible and easy-to-manage data storage solution.

Table detection (Yolo v5)

We have also utilized YOLOv5, a state-of-the-art object detection algorithm, to accurately identify the structure of tables within the given data. In conjunction with YOLOv5, we have employed Paddle OCR, a robust optical character recognition (OCR) framework, to extract text information from the detected table cells with high precision. By combining these powerful tools, we have achieved efficient and effective table analysis results, enabling us to extract valuable insights from complex datasets. YOLO v5 is the latest version of the YOLO series and is designed for improved accuracy and speed.

The model takes an input image and outputs bounding boxes around detected tables, along with a confidence score indicating how certain the model is about the detection. A novel technique, termed "dynamic anchor boxes," is used in this YOLO v5 to generate the anchor boxes. Clustering the ground truth bounding boxes and then employing the centroids as anchor boxes is the method employed here. Therefore, the anchor boxes may be adjusted to better fit the dimensions and contours of the objects that have been identified. Spatial pyramid pooling (SPP) is a new sort of pooling layer in YOLO v5 that is used to decrease the feature maps' granularity in space. Since SPP gives the model a broader perspective, it is able to better recognize even the smallest of things. Although YOLO v4 also makes use of SPP, the SPP architecture has been much im-

proved in YOLO v5, allowing for more impressive outcomes.

Overall Model Architecture

Lastly, by fusing the output of NER with YOLOv5, we potentially improve the accuracy of object detection by incorporating information about named entities in an invoice. For example, if the NER model identifies a person in the scene, the object detection model may be better able to identify and track that person. To perform this fusion, we developed a pipeline that takes in input data (such as text or images), passes it through both the NER and YOLOv5 models, and then combines the output in a meaningful way, as shown in figure 5.

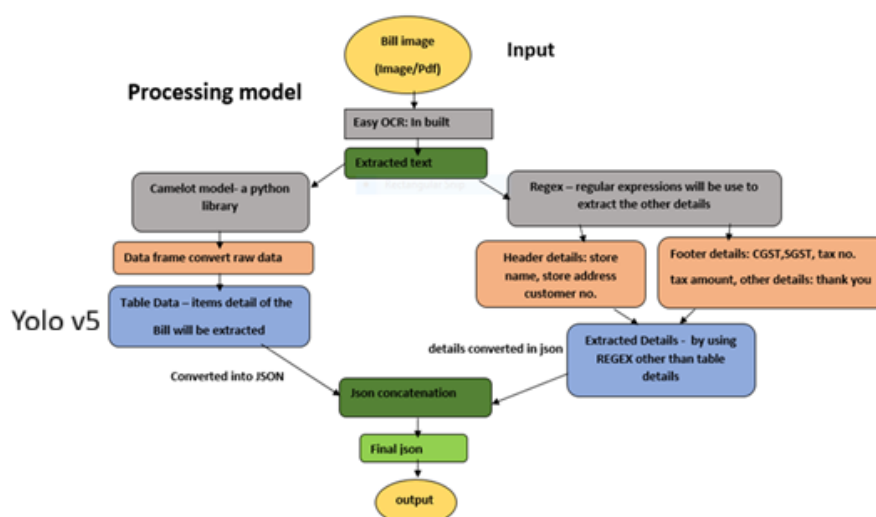


Figure 5: The Complete Processing System for the Deployment of the Proposed Approach.

The process begins by passing the input image or PDF document through a processing model that utilizes optical character recognition (OCR) technology to extract text from the input file. The data extraction stage is further divided into two parts, the first part extracts meta details from the bill or receipt, including store name, store address, and additional information, utilizing Regular Expression techniques. The second part deals with tabulated data, such as item names, prices, and quantities, which is extracted using Camelot model and other form recognizer techniques. These techniques generate output data in a key-value pair format, which is concatenated to obtain the complete bill data in JSON format. This JSON data can be saved as a separate

file or loaded into a database for further analysis.

Result & Discussion

Artificial intelligence (AI) can be applied to validate the accuracy of receipt information and provide clients with precise rate predictions. This can simplify the process of adjusting order calculations, and allow for testing of different algorithms. AI can also automate internal business processes, increasing efficiency. To improve optical character recognition (OCR) extraction accuracy, it is necessary to evaluate extraction algorithms and develop better OCR engines. Additionally, improving accuracy in the extraction of

text requires the extraction of different types of layouts and fields.

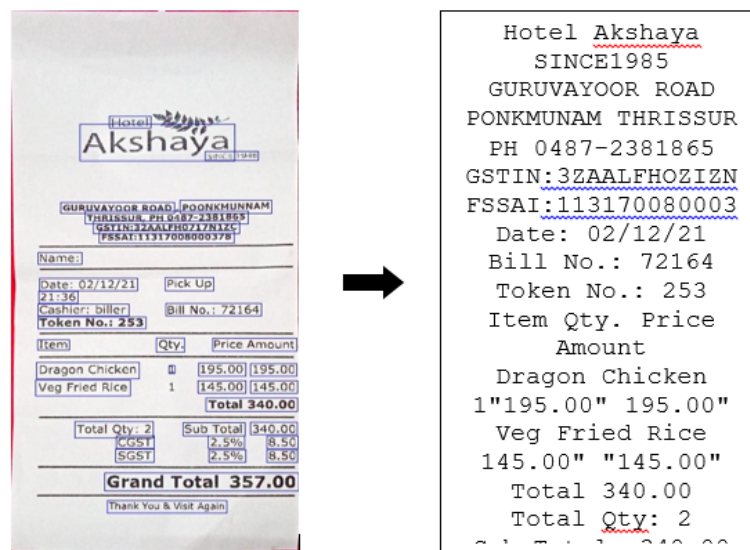


Figure 6: Text Extracted from PDF/ Image file

To increase system efficiency, larger datasets can be used to train algorithms, and more data fields can be extracted from all types of bills. By deploying the Azure API

process, individuals can use machine learning to build user-specific data models and classifications that are tailored to client needs.

```
{'Bill_type': 'TAX INVOICE',
'Total_money_spend': 53.0,
'contact_number': '9880366893',
'customer_name': 'Cash',
'invoice_date': '22/9/2021',
'invoice_number': 'INV-ASM3485',
'invoice_time': '2.00.13',
'items': {'item 1': {'amount': 10.0,
'name': 'SINEET BUN NA',
'quantity': 1.0,
'rate': 10.0},
'item 2': {'amount': 23.0,
'name': 'NANDINI CURD',
'quantity': 1.0,
'rate': 23.0},
'item 3': {'amount': 20.0,
'name': 'FROOTI MANGC',
'quantity': 2.0,
'rate': 10.0}},
'market_address': 'OPPTO SHRIRAM SUHAANAAPARTMENT; SURYAVA',
'market_name': 'AISHWARYA SUPER MARKET'}
```

Figure 7: Data Extracted in JSON Format Using Azure

Further Study

Once bills are stored in the database, users can retrieve and view the stored data. Filter options are available

to enable users to monitor and filter the records they wish to view. In cases where text prediction errors are identified, users may manually modify the data and incorporate the

corrected text into JSON format, tailored to the fields that

align with their requirements.

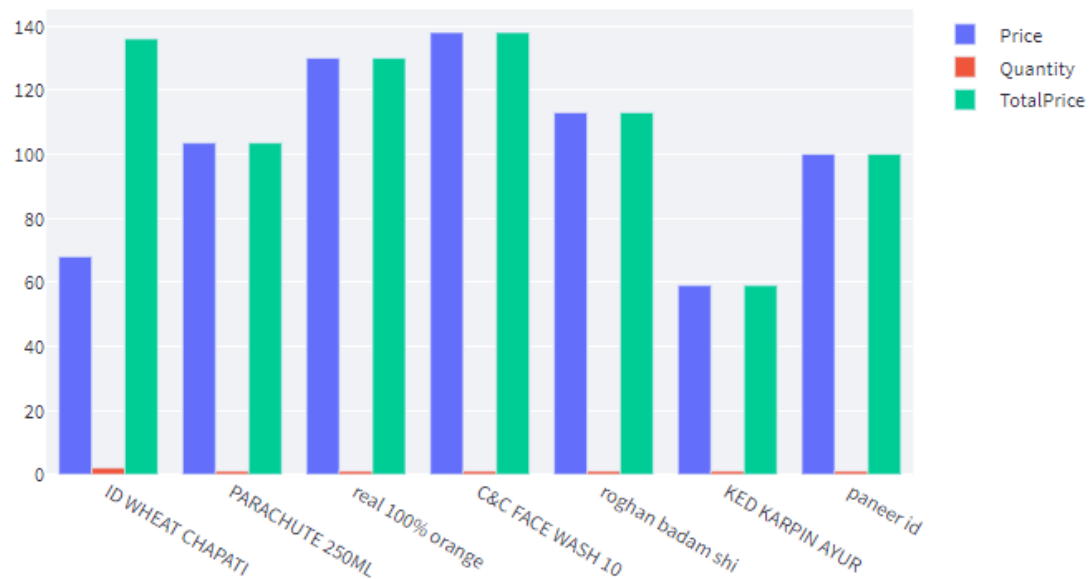


Figure 8a: Visualization of Items List Data – Bar Graph

As developers, we have the option to train or re-train the algorithm to detect specific features that clients desire by exposing the algorithm to images that include or exclude these features.

The visualized expenditure, based on itemized

bills, is presented through Figures 15a to 15d, within the Utility Analysis processing system, as per the user's requirement. This graphical representation facilitates a better understanding of the amount spent and where it was spent on. Furthermore, this visualization is instrumental in loading the data into the database.

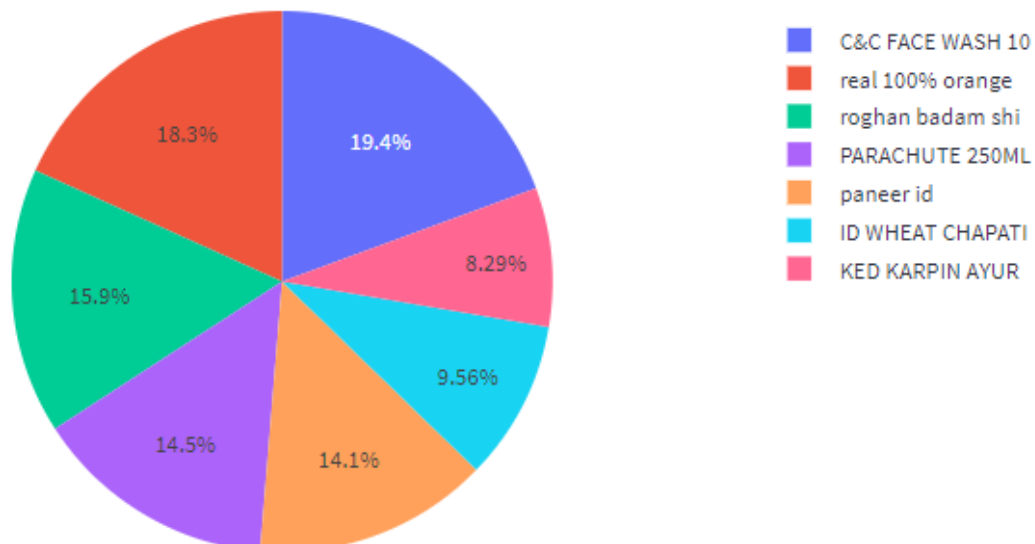


Figure 8b: The Visualization Cost Price of Each Item

TotalPrice

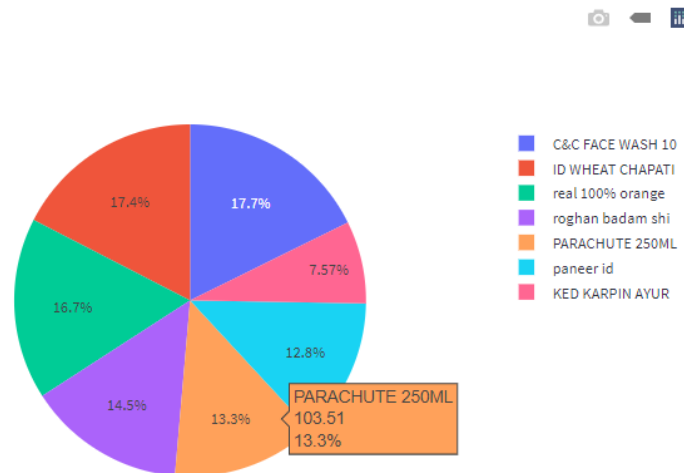


Figure 8c: Visualization of Total Prices Spent on Each Item

This approach enables us to continually refine and improve the accuracy of our text prediction, ensuring that our system consistently meets the evolving needs of our clients. Finally, the filtered data can be sorted based on a range of criteria, such as date, store name, item name, and more, enabling users to identify and access the relevant records with ease and convenience.

The culmination of our bill processing and storage approach is the visualization of data. The dashboard page of the application presents key fields such as the count of bills data, the overall number of bills, and the total amount spent within a given timeline. These essential metrics provide users with a comprehensive overview of their spending habits and enable them to make informed decisions based on this data.

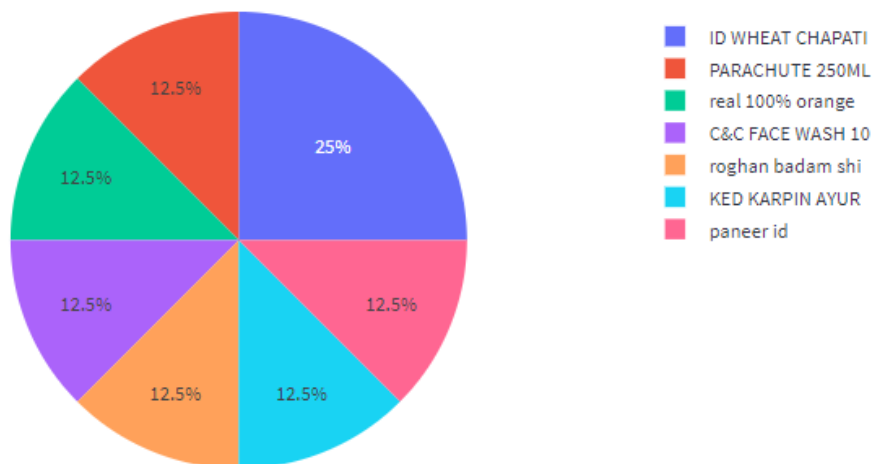


Figure 8d: Visualization of the Number of Products Purchased

The final part of our system involves visualizing the data. This is achieved by displaying store details, bill total amounts, tax amounts for all uploaded bills, and data-based visualizations. The third part of the system focuses on item-level details, and utilizes tools such as Pandas, Plotly, and Streamlit for user interface. The image in Figure 5

provides an overview of the system's processing, beginning with the conversion of PDFs or images into a standardized format, followed by text extraction from the bill and passing it through Form Recognizer to retrieve the required data in JSON format. This data can then be added to a database for further analysis, based on customer requirements. The anal-

ysis can be used for financial calculations on a daily basis.

Conclusion

The Utility Bill Analysis is a comprehensive solution that empowers users to perform bill analysis through visualizations created using Plotly. The components of the bill can be conveniently filtered to suit user preferences. The solution extracts bill text and stores it in a structured JSON format in MongoDB. It utilizes EasyOCR and Azure form recognizer for the purpose of text extraction and analysis. EasyOCR facilitates the extraction of text from JPEG/PNG bills, while Azure form recognizer enables the identifica-

tion, extraction, and analysis of text, tables, and form fields from various types of scanned documents and mobile phone images without manual intervention. The solution features a user-friendly interface that enables easy bill upload and analysis without any security concerns. Users have the flexibility to decline the save option if they prefer not to store their bill data in the database. Additionally, they can manually modify the items list from the bill if required, either to correct errors detected by OCR or to reflect any modifications. This end-to-end solution provides a user-friendly way to analyze and visualize bills of various kinds, which can help users keep track of their expenses.

References

1. X Yao, H Sun, S Li, W Lu (2022) "Invoice Detection and Recognition System Based on Deep Learning," *Secur. Commun. Networks*, vol. 2022.
2. R Smith (2007) "An overview of the tesseract OCR engine," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 2: 629–33.
3. AA Shinde, DG Chougule (2012) "Text Pre-processing and Text Segmentation for OCR," *Int. J. Comput. Sci. Eng. Technol.* vol. 2: 810–12.
4. Y Boulid, A Souhar, M. Ouagague (2017) "Spatial and Textural Aspects for Arabic Handwritten Characters Recognition," *Int. J. Interact. Multimed. Artif. Intell.*, vol. InPress, no. InPress: 1.
5. Y Wen, Y Lu, J Yan, Z Zhou, KM Von Deneen (2011) "An algorithm for license plate recognition applied to intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 12: 830–45.
6. X Fan, G Fan (2009) "Graphical models for joint segmentation and recognition of license plate characters," *IEEE Signal Process. Lett.* vol. 16: 10–13.
7. A Gonzalez, LM Bergasa, JJ Yebes (2014) "Text detection and recognition on traffic panels from street-level imagery using visual appearance," *IEEE Trans. Intell. Transp. Syst.* vol. 15: 228–38.
8. ELC Lai, X Yu (2015) "Invoicing currency in international trade: An empirical investigation and some implications for the renminbi," *World Econ.* vol. 38: 193–229.
9. J Redmon, S Divvala, R Girshick, A. Farhadi (2016) "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* vol. 2016: 779–88.
10. R Girshick (2015) "Fast R-CNN," *Proc. IEEE Int. Conf. Comput. Vis.* vol. 2015: 1440–48, 2015.
11. R Girshick, J Donahue, T Darrell, J. Malik (2014) "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit:* 580–87.
12. YP. Chen, Y Li, G Wang (2018) "An enhanced region proposal network for object detection using deep learning method," *PLoS One*, vol. 13: 1–26.
13. P Jiang, D Ergu, F Liu, Y Cai, B. Ma (2021) "A Review of Yolo Algorithm Developments," *Procedia Comput. Sci.* vol. 199: 1066–73.
14. Y Wen, C Fan, G Chen, X Chen, M. Chen (2020) "A Survey on Named Entity Recognition," *Lect. Notes Electr. Eng.* vol. 571 LNEE: 1803–10.
15. LF Rau (1990) "Extracting company names from text," *Proc. Conf. Artif. Intell. Appl:* 29–32.
16. B. K. Chang et al., (2015) "1995 Index IEEE Transactions on Systems, Man, and Cybernetics Vol. 25," *IEEE Trans. Syst. Man. Cybern.* vol. 25: 1–17.
17. X Holt, A Chisholm (2018) "Extracting structured data from invoices," *Proc. Australas. Lang. Technol. Assoc. Work.* 2018: 53–9.
18. S Mizan, Chowdhury Md, Chakraborty, Tridib, Karmakar (2017) "Text Recognition using Image Processing," *Int. J. Adv. Res. Comput. Sci.*, vol. 8.
19. Ö Arslan, SA Uymaz (2021) "Classification of Invoice Images by Using Convolutional Neural Networks," *J. Adv. Res. Nat. Appl. Sci.*
20. A Umam, JH. Chuang, DL Li (2018) "A Light Deep Learning Based Method for Bank Serial Number Recognition," *VCIP 2018 - IEEE Int. Conf. Vis. Commun. Image Process.*
21. RF Rahmat, D Gunawan, S Faza, N Haloho, EB Nababan (2018) "Android-based text recognition on receipt bill for tax sampling system," *Proc. 3rd Int. Conf. Informatics Comput. ICIC.*
22. H Sidhwa, S Kulshrestha, S Malhotra, S Virmani (2018) "Text Extraction from Bills and Invoices," *Proc. - IEEE 2018 Int. Conf. Adv. Comput. Commun. Control Networking, ICACCCN 2018:* 564–8.

23. S Mizan, Chowdhury Md, Chakraborty, Tridib, Kar-makar, "Text Recognition using Image Processing," *Int. J. Adv. Res. Comput. Sci.*, vol. 8.
24. W. Song, S. Deng (2009) "Bank Bill Recognition Based on an Image Processing," in 2009 Third International Conference on Genetic and Evolutionary Computing: 569–73.
25. F Jiang, LJ Zhang, H Chen (2017) "Automated Image Quality Assessment for Certificates and Bills," in 2017 IEEE International Conference on Cognitive Computing (ICC-C):1–8.
26. Y Sun, X Mao, S Hong, W Xu, G Gui (2019) "Template Matching-Based Method for Intelligent Invoice Information Identification," *IEEE Access*, vol. 7: 28392–401.
27. H Guo, X Qin, J Liu, J Han, J Liu, E Ding (2019) "EATEN: Entity-Aware Attention for Single Shot Visual Text Extraction," in 2019 International Conference on Document Analysis and Recognition (ICDAR): 254–59.
28. N Rahal, M Tounsi, M Benjlaiel, AM Alimi (2018) "Information Extraction from Arabic and Latin scanned invoices," in 2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR): 145–50.
29. L Xu, W Fan, J Sun, X Li, S. Naoi (2016) "A knowledge-based table recognition method for Chinese bank statement images," in 2016 IEEE International Conference on Image Processing (ICIP): 3279–83.
30. J Hom, J Nikowitz, R Ottesen, J C (2022) Niland, "Facilitating clinical research through automation: Combining optical character recognition with natural language processing." *Clin. Trials*, vol. 19: 504–11.
31. D. Chappell, "Introducing the windows azure platform," David Chappell Assoc. White Pap., 2010.

Submit your manuscript to a JScholar journal and benefit from:

- ¶ Convenient online submission
- ¶ Rigorous peer review
- ¶ Immediate publication on acceptance
- ¶ Open access: articles freely available online
- ¶ High visibility within the field
- ¶ Better discount for your subsequent articles

Submit your manuscript at
<http://www.jscholaronline.org/submit-manuscript.php>