Table of Contents

Introduction

RPA and AI are progressively utilised in current enterprises to enhance operational efficiency and optimise procedures (Treacy et al., 2023). RPA automates tedious, rule-governed operations such as data input, invoice processing, and report generation, enabling organisations to concentrate on more valuable endeavours. The incorporation of AI augments its functionality by integrating cognitive processes to improve accuracy and decision-making. Invoice management is a domain where AI can mitigate challenges such as human data entry, absent or misplaced invoices, and conflicts between invoices and payments (Kanaparthi, 2023). However, businesses must understand their processes and requirements before implementing these technologies to fully leverage the benefits of RPA and AI.

To address the issue of processing large volumes of invoices, this project aims to build an attended RCA solution leveraging OCR from UiPath as the automation platform. UiPath's extensive RPA and AI capabilities (Potturu, 2023), including its intuitive interface, pre-built activities, and scalability, make it an optimal choice for creating a prototype suitable for large-scale deployment in retail settings. Based on the three OCR engines evaluated in the literature review, only the Tesseract OCR and Omnipage OCR approaches will be implemented and evaluated in this project. Although Google Vision OCR was considered as a potential OCR, it will not be included due to not being accessible within the UiPath Community Plan.

Not to mention, Omnipage OCR appears to outperform Tesseract OCR in document processing workloads based on prior studies. Unlike Tesseract, which requires significant pre-processing and struggles with intricate layouts, OmniPage's built-in language detection, advanced image correction, and adaptive learning capabilities ensure accurate and reliable data extraction with minimal pre-processing. This suggests that OmniPage OCR could be the superior choice for automating invoicing in this project. However, this observation remains preliminary, and the objective is to determine the effectiveness of each OCR engine in terms of data accuracy, processing efficiency, and overall reliability. The following section will cover the proposed model workflow, the implementation of RCA in UiPath, performance evaluation, and future improvements. By implementing this intelligent automation model, retail organisations can

achieve a streamlined, compliant, and cost-effective invoice management system, illustrating the profound impact of AI-enhanced RPA in optimising routine yet critical operations.

Proposed Architecture

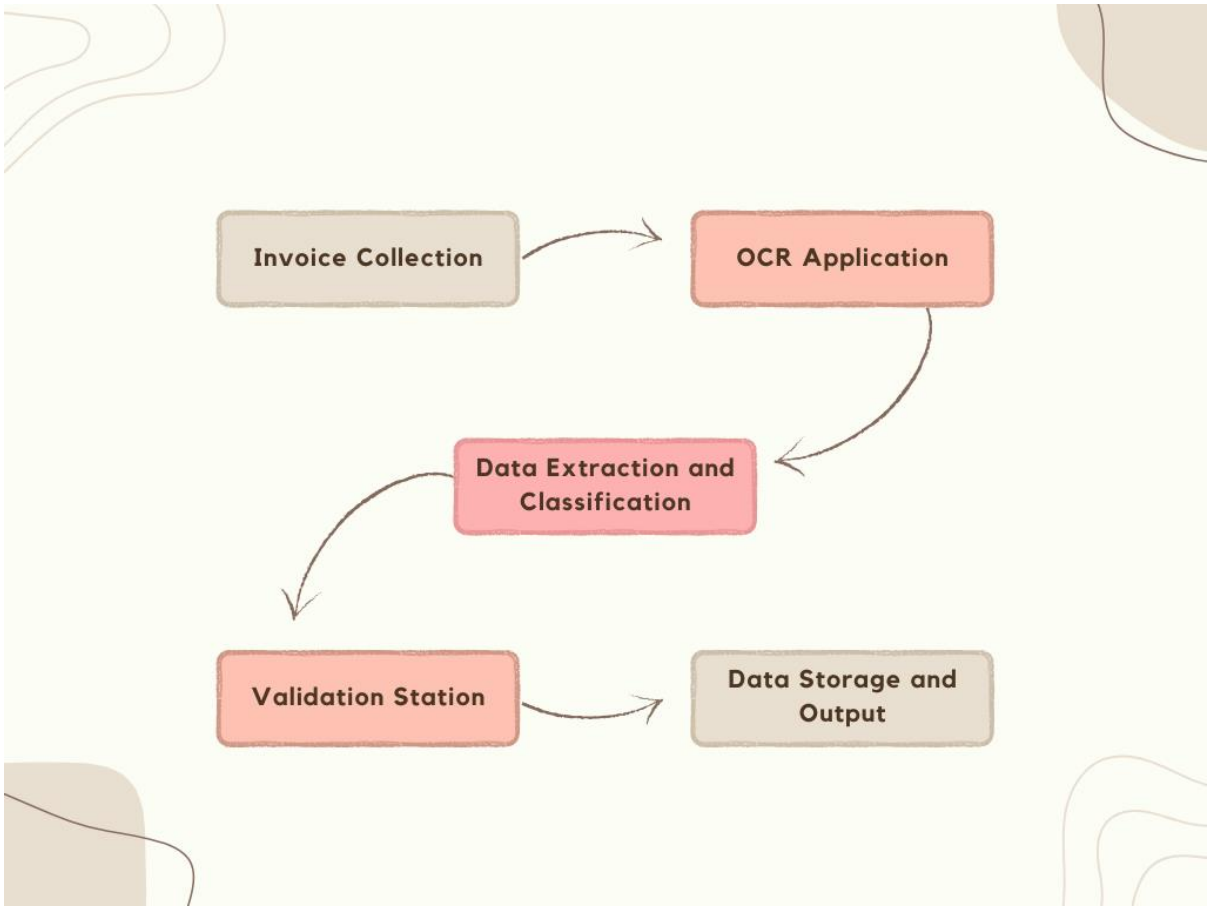


Figure 1: Proposed Architecture for Invoice Processing

System Architecture and Workflow

1. **Invoice Collection**

   A specific folder will be created to keep a pile of invoices in PDF format. PDFs will be the input for processing, as they are generally more compact to facilitate storage, transmission, and download. This will also allow users to systematically store and organise new invoices into the folder.

2. **OCR Application**

   Tesseract OCR and Omnipage OCR will be used and applied to transform invoices into machine-readable text. These OCR engines are capable of retaining key document elements such as graphics, form fields, and formatting features.

3. **Data Extraction and Classification**

   Through the identification of keywords and distinctive identities, text will be automatically extracted to determine the essential fields, like invoice number, due date, billed to, and total amount. This can be done by the OCR engines, which leverage machine learning algorithms for keyword-based classification.

4. **Validation Station**

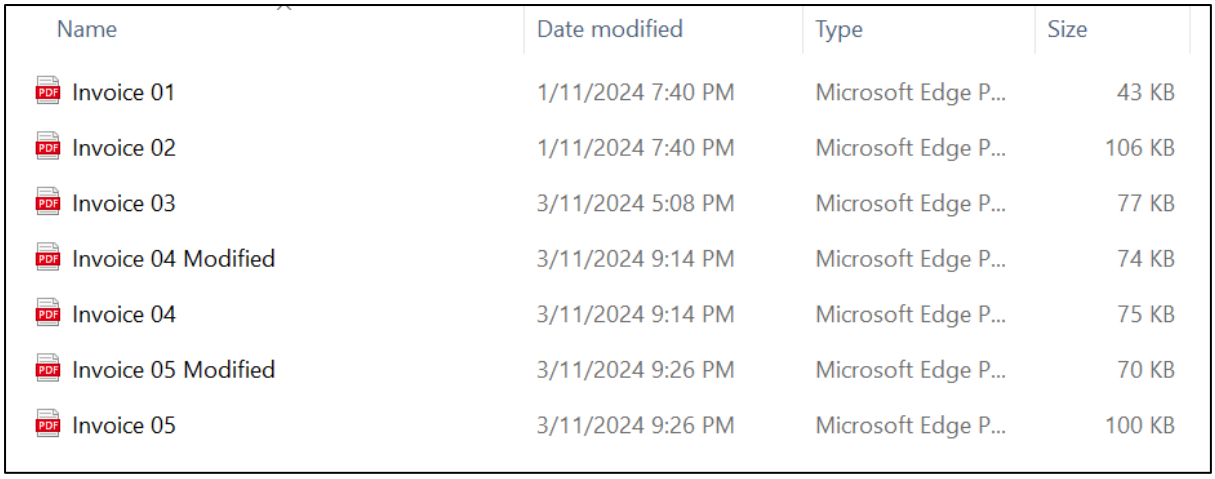
   A human review is initiated for extracted fields to check the confidence scores and data correctness. It is a built-in function from UiPath's Document Understanding package to allow users to rectify discrepancies by viewing the original document and extracted data side by side, improving data quality.

5. **Data Storage and Output**

   After extraction and verification, data is stored in an Excel file. This is done to maintain the project's simplicity and speed of access by standardizing the output format. In addition, Excel files are commonly used applications in every industry, especially reports on invoice data. If larger operations are required, then this development and proof of concept can be integrated in a SQL database or other business applications.

Prototype Development

In this phase, the end-to-end implementation will be described according to the conceptual architecture outlined earlier. All the steps are implemented in UiPath, with the goal to streamline invoice processing using Tesseract OCR and Omnipage OCR. Later on, the performance metrics for the automation solution will be further discussed in the subsequent section.



Name	Date modified	Type	Size
Invoice 01	1/11/2024 7:40 PM	Microsoft Edge P...	43 KB
Invoice 02	1/11/2024 7:40 PM	Microsoft Edge P...	106 KB
Invoice 03	3/11/2024 5:08 PM	Microsoft Edge P...	77 KB
Invoice 04 Modified	3/11/2024 9:14 PM	Microsoft Edge P...	74 KB
Invoice 04	3/11/2024 9:14 PM	Microsoft Edge P...	75 KB
Invoice 05 Modified	3/11/2024 9:26 PM	Microsoft Edge P...	70 KB
Invoice 05	3/11/2024 9:26 PM	Microsoft Edge P...	100 KB

Figure 2: Invoice Folder

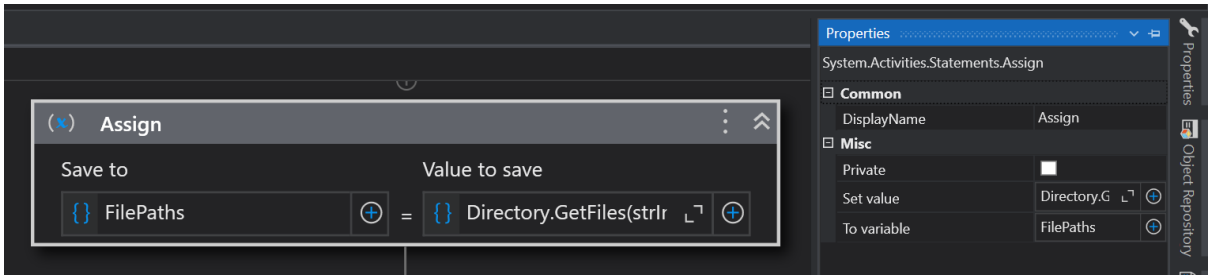


Figure 3: Activity – Assign

First of all, all invoices are stored in a local folder path to allow automation workflow to scan for any invoices. In this case, a folder named **Invoices** is created. Subsequently, the **Assign** activity is used to gather all PDFs from the respective folder. Next, a variable named **FilePaths** is generated for the **"Save to"** field, which records a list of invoice directories. **GetFiles(strInvoicePath)** is assigned to the **"Value to save"** field to find out all files from that

directory. This configuration facilitates batch processing of invoices and guarantees that each document is systematically managed inside the workflow.

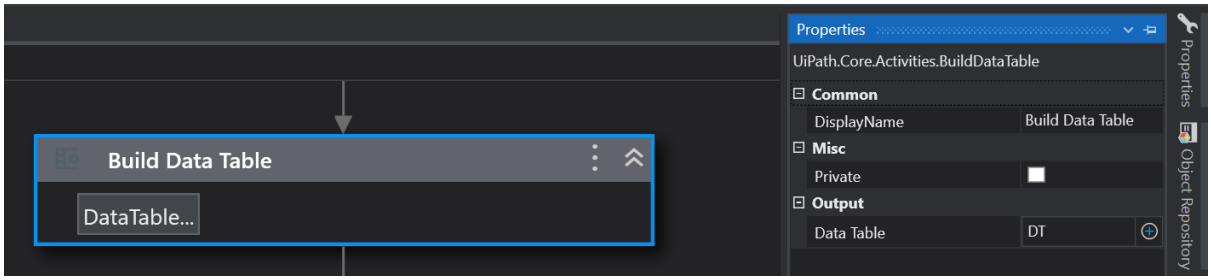


Figure 4: Activity – Build Data Table

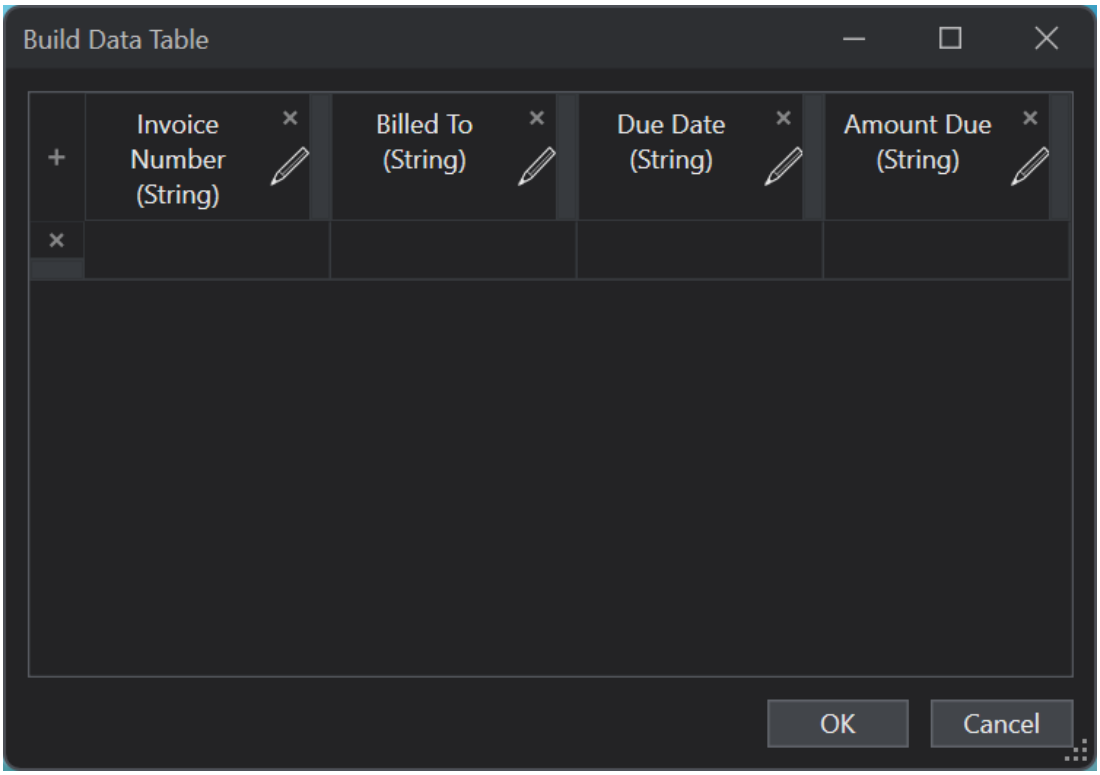


Figure 5: Activity – Build Data Table Configuration

Secondly, the **Build Data Table** activity is employed to build a customised table specifically to store important information on invoice data, such as invoice number, billed to, due date, and amount due. All columns will feature a string data type and contain empty rows. This framework standardises data extraction by predefining each element. Under the part of **Properties**, a variable called **DT** is also assigned to the **Data Table** as output.

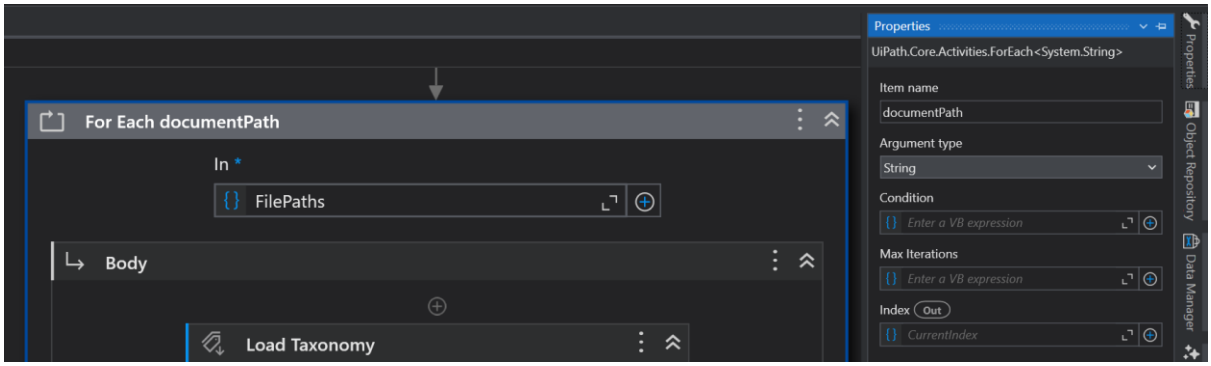


Figure 6: Activity – For Each

Thirdly, the **For Each** activity will iterate over each file from the variable **FilePaths**. This ensures that the workflow processes each invoice in a sequential manner. To clarify, a variable **documentPath** is created under the section of **Properties** for referencing each file. It is crucial to implement a loop mechanism, as it will only process a single file if this structure is not included.

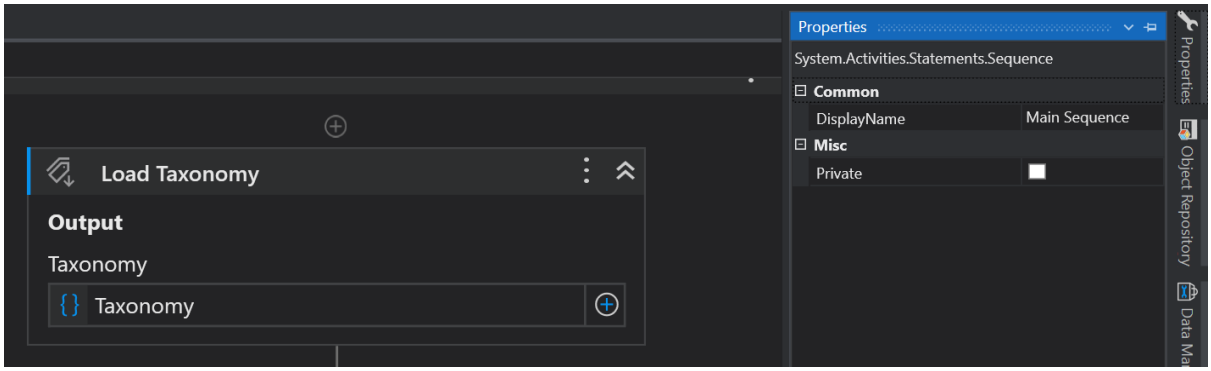


Figure 7: Activity – Load Taxonomy

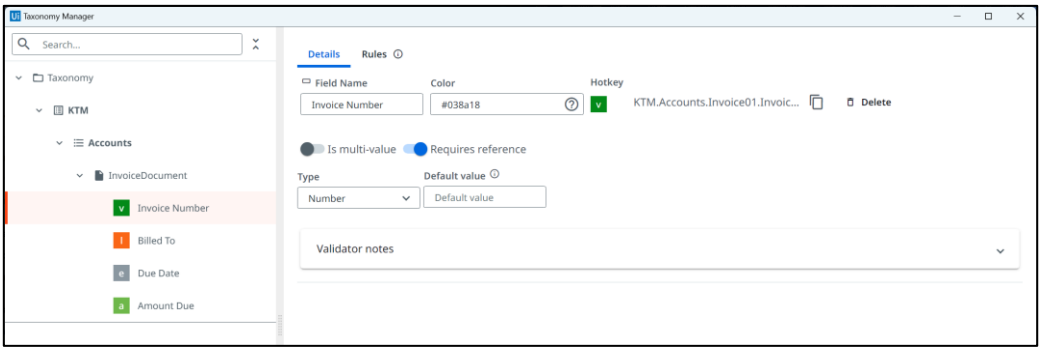


Figure 8: Taxonomy Manager

Furthermore, the **Load Taxonomy** activity is applied to extract document fields using a pre-defined taxonomy structure from **Taxonomy Manager**. To ensure all necessary invoice data is captured properly, the invoice processing taxonomy is loaded into a new variable called **Taxonomy**. The template must be generated in the **Taxonomy Manager** for the next few procedures, as it will allow the automation workflow to understand what needs to be extracted from each document.

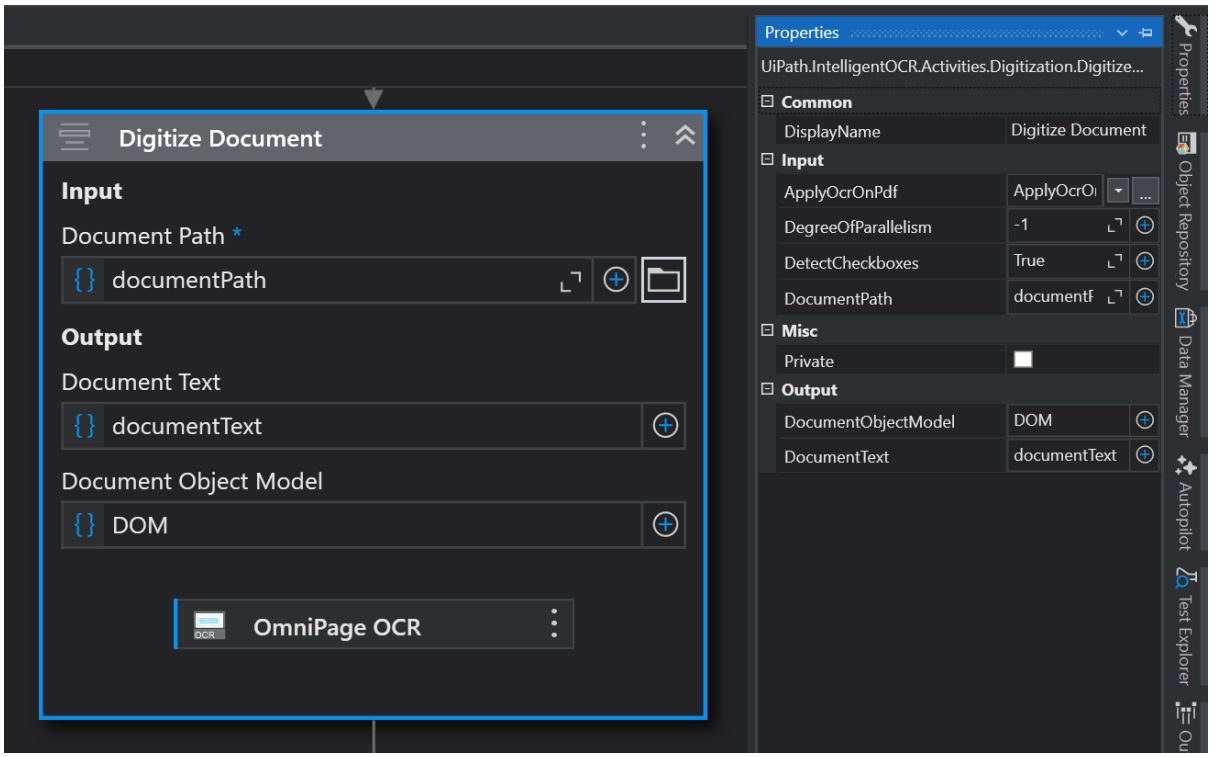


Figure 9: Activity – Digitize Document

The **Digitize Document** activity is the core component of this workflow, as it uses OCR to convert invoice documents into text. It transforms unstructured data into a comprehensible and usable format for automated workflows, which is crucial for data extraction. The previously created variable **documentPath** is used as input to load files for digitisation. By default, the input property of **ApplyOcrOnPdf** is set as **Auto**, which means the workflow will automatically assess whether the document should be applying the OCR algorithm based on the input document. Also, the **DegreeOfParallelism** under **Properties** is set to **-1** to let the workflow handle parallel processing automatically based on the number of cores on the machine. Two new variables are produced from the digitisation of the document, namely

documentText and **DOM**. The **documentText** variable will record the extracted text content, whereas the **DOM** variable will present a systematic depiction of the document. For this setup, Omnipage OCR and Tesseract OCR will be used interchangeably for the results and analysis in the following section.

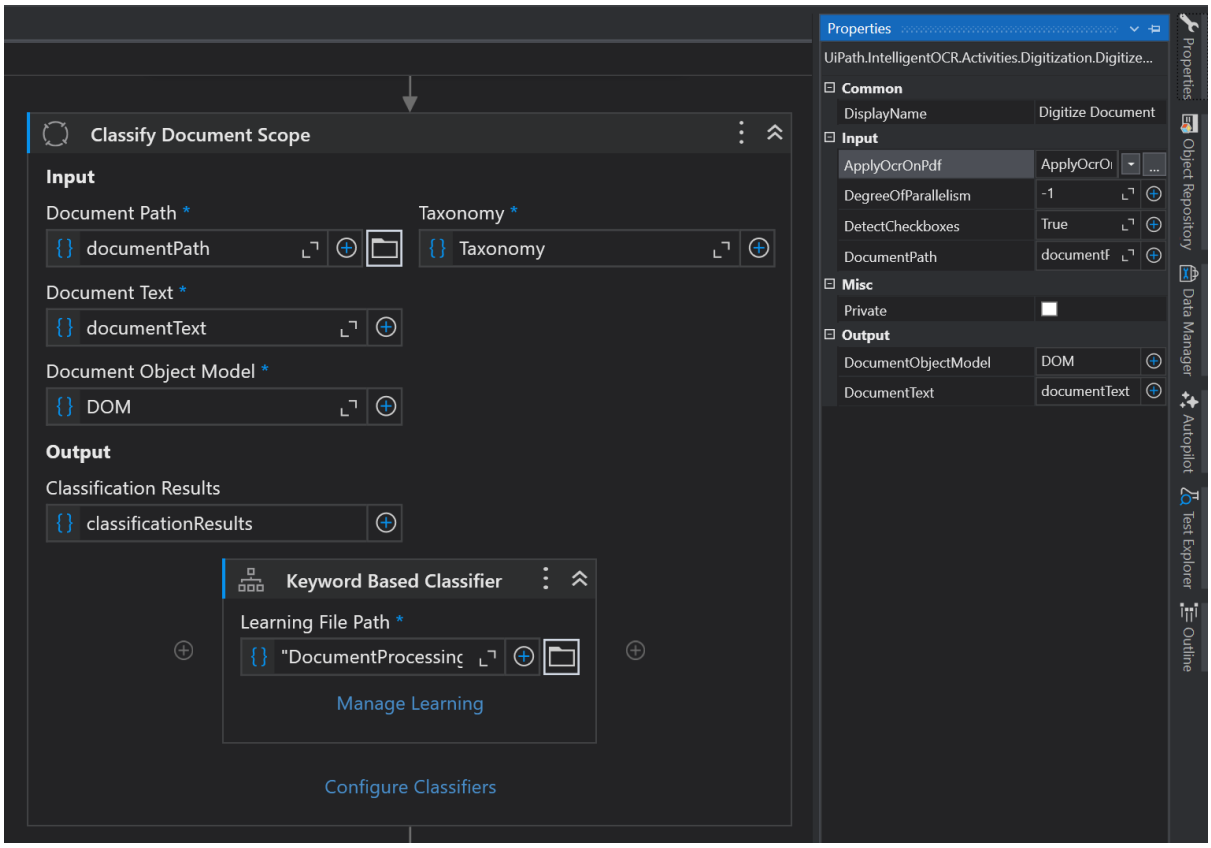


Figure 10: Activity – Classify Document Scope


```
[{"DocumentTypeId":"KTM.Accounts.Invoice01",
"Keywords":[{"Values":["invoice"],
"NumberOfTimesConfirmed":1,
"LastUpdated":"2024-11-03T15:10:00.9190794+08:00"}]}]
```


Figure 11: Keyword JSON File

Additionally, the **Classify Document Scope** activity is utilised for sorting documents into specified categories based on certain rules or templates to extract data. Most importantly, a **Keyword Based Classifier** is employed to compare document content based on the **Learning File Path**, which is a **JSON** file storing the patterns or keywords that belong to particular

document categories. This is a crucial step as the automation workflow necessitates an understanding of the type of document it is assessing. For example, an invoice and receipt may have different formats; therefore, it requires specific data extraction rules to classify the document type information. In other words, it can improve the automation intelligence to distinguish document types and apply tailored extraction logic. Ultimately, the **classificationResults** variable is generated to retain the classification process's results, including document types and classifier-detected unique features.

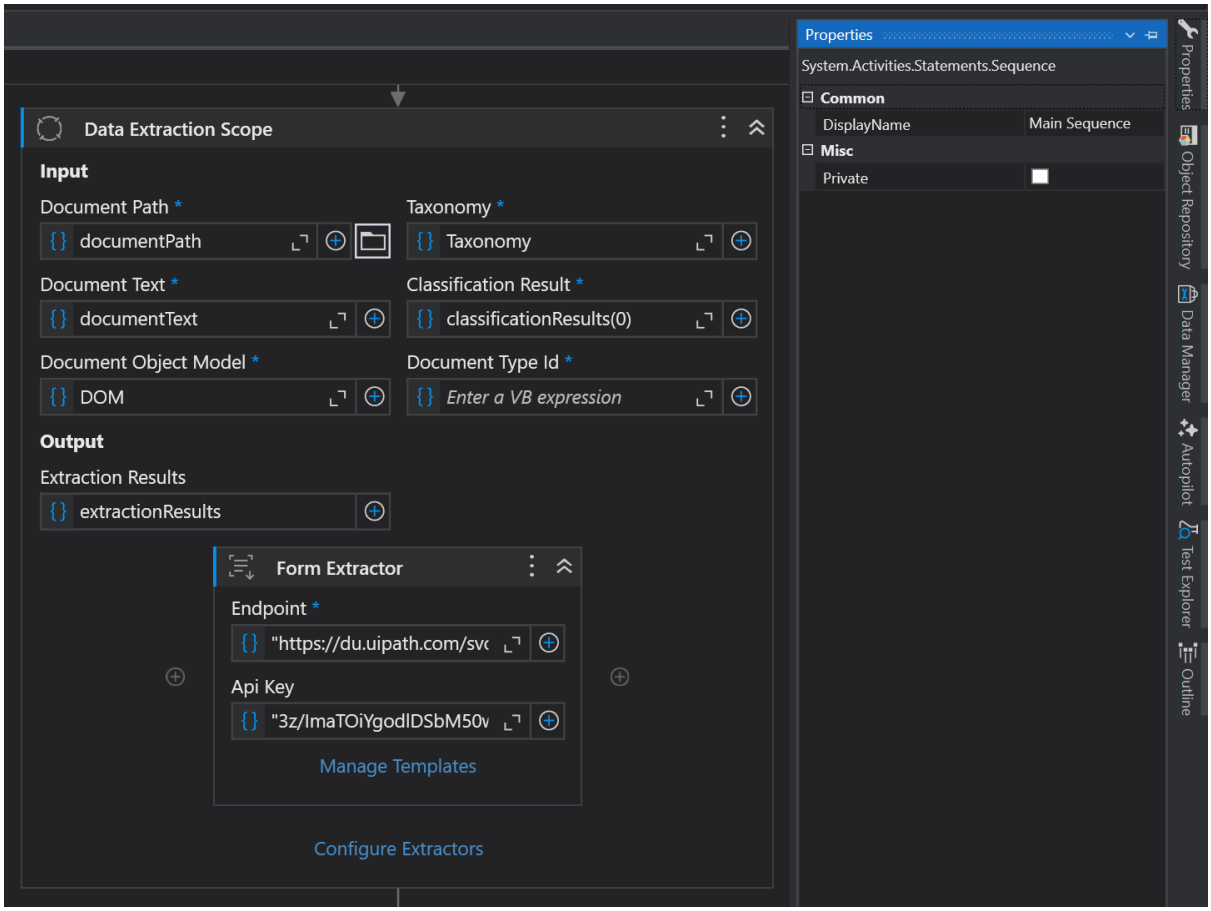


Figure 12: Activity – Data Extraction Scope

The next step involves the **Data Extraction Scope** activity, which gathers essential information from the document's taxonomy and classification. The integration with the **Form Extractor** component enables it to connect to the URL that points to the Document Understanding services from UiPath. Since this project is conducted using a community license, the maximum size of extractable documents is restricted to 2 pages and 4 MB (UiPath, 2024). The **Form Extractor** relies on the provided templates as they describe the data field document locations;

hence, managing templates is needed to accommodate various document formats. As can be seen, the extracted data and taxonomy field values are kept in the **extractionResults** variable, which can be supplied to other downstream processes.

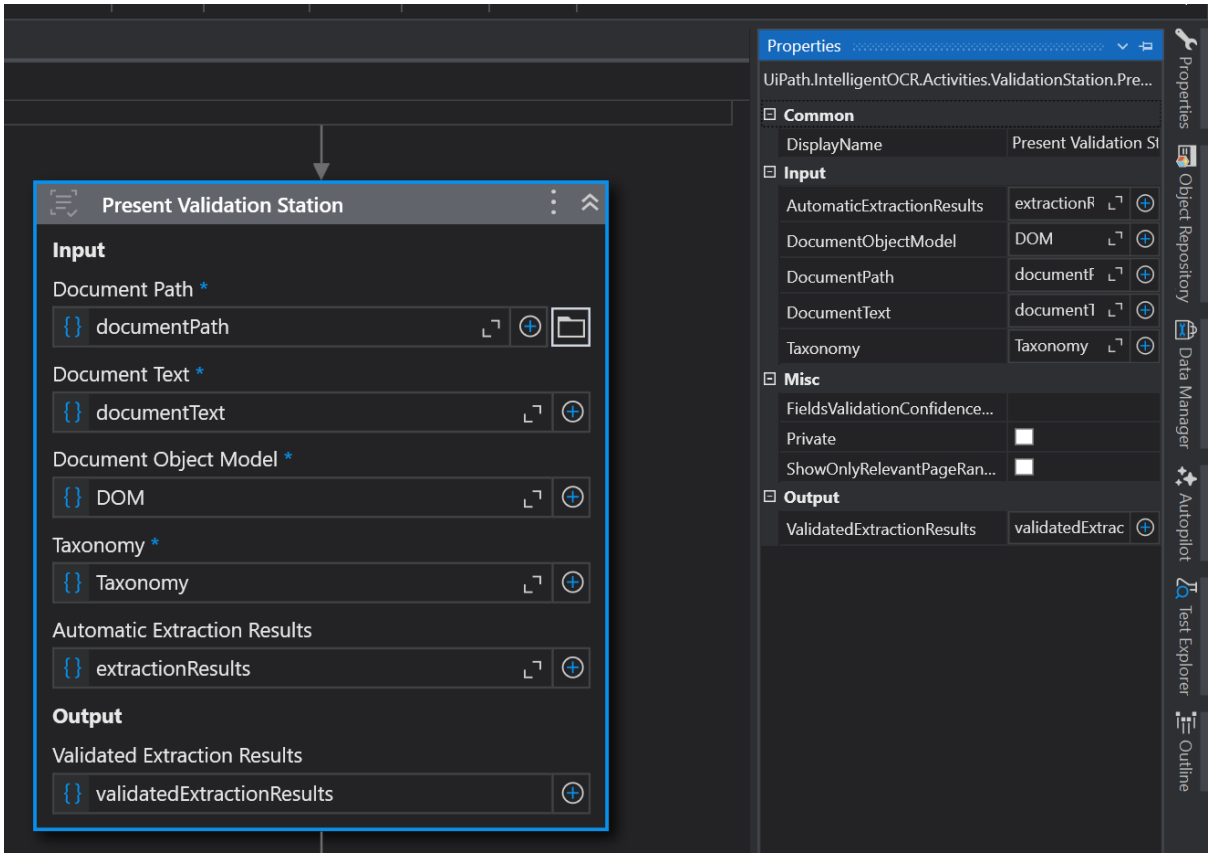


Figure 13: Activity – Present Validation Station

The activity of **Present Validation Station** is to allow users to verify invoice data that has been automatically retrieved by the automation. It marks fields with low confidence scores or potential mistakes for human intervention, which can ensure data quality and accuracy. Each document is loaded and visually compared with the extracted text from OCR, and then the results are saved in the **validatedExtractionResults** variable after human validation is completed. The verified values for each field are included in this output, so they are ready to be stored or processed further. The confidence threshold can be set from the **FieldsValidationConfidence %** under **Properties** to be marked automatically for human inspection.

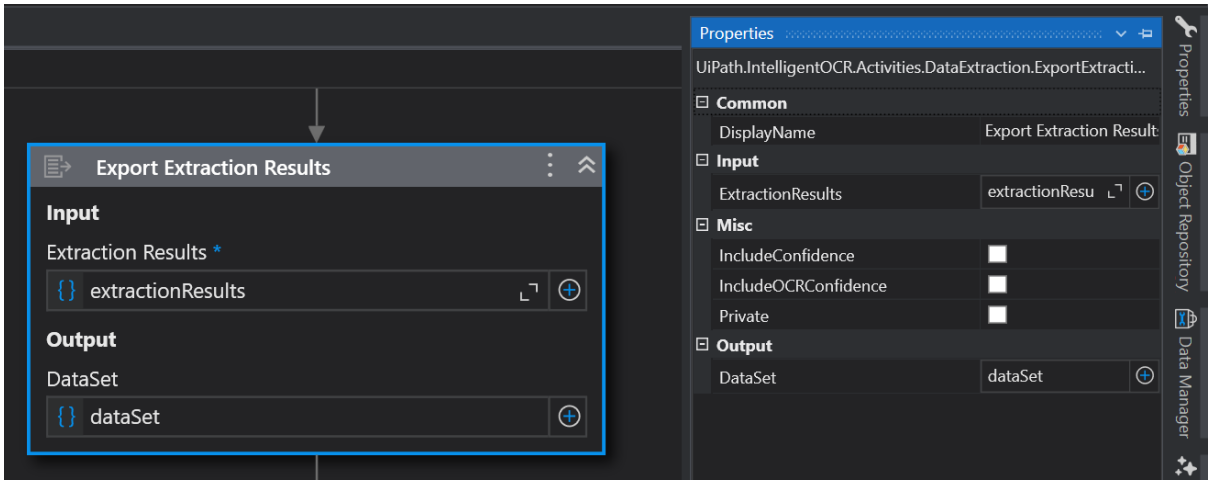


Figure 14: Activity – Export Extraction Results

In addition, the **Export Extraction Results** activity is applied to the **extractionResults** variable for organising and verifying process results for other automation purposes. For this case, it is expected that the exported data can be converted into a spreadsheet for inserting the results as a new record in an Excel file. Therefore, the **dataSet** output variable is created to store the extracted data in a tabular format. Since the **Present Validation Station** activity is already demonstrating the confidence level testing in the prior step, both settings for **IncludeConfidence** and **IncludeOCRConfidence** under **Properties** will not be checked for quality assessments.

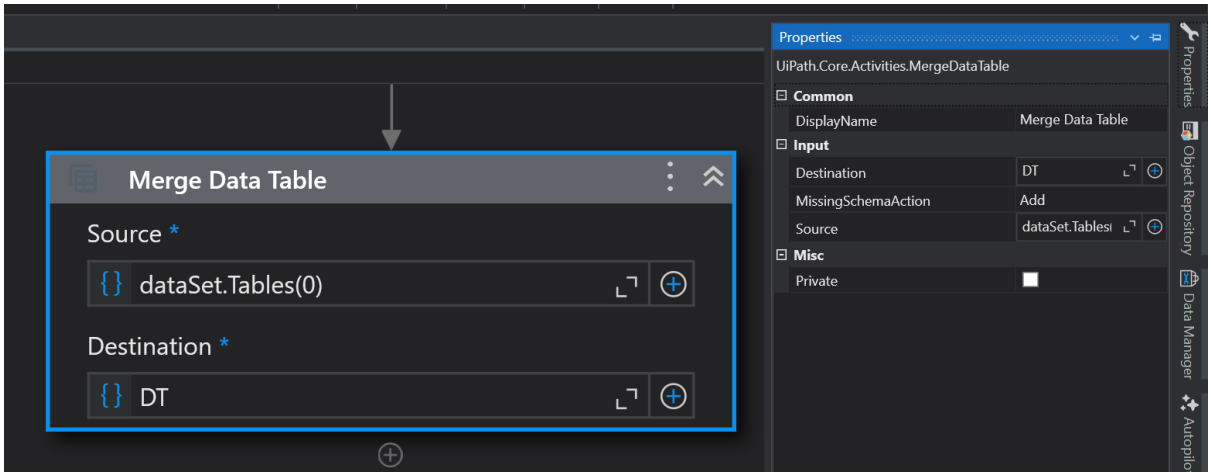


Figure 15: Activity – Merge Data Table

Before writing the verified output into an Excel file, the activity of **Merge Data Table** will consolidate data from several different tables and save it into the destination variable, **DT**. This is to simplify processes by instantly merging data, removing the requirement for tedious aggregation.

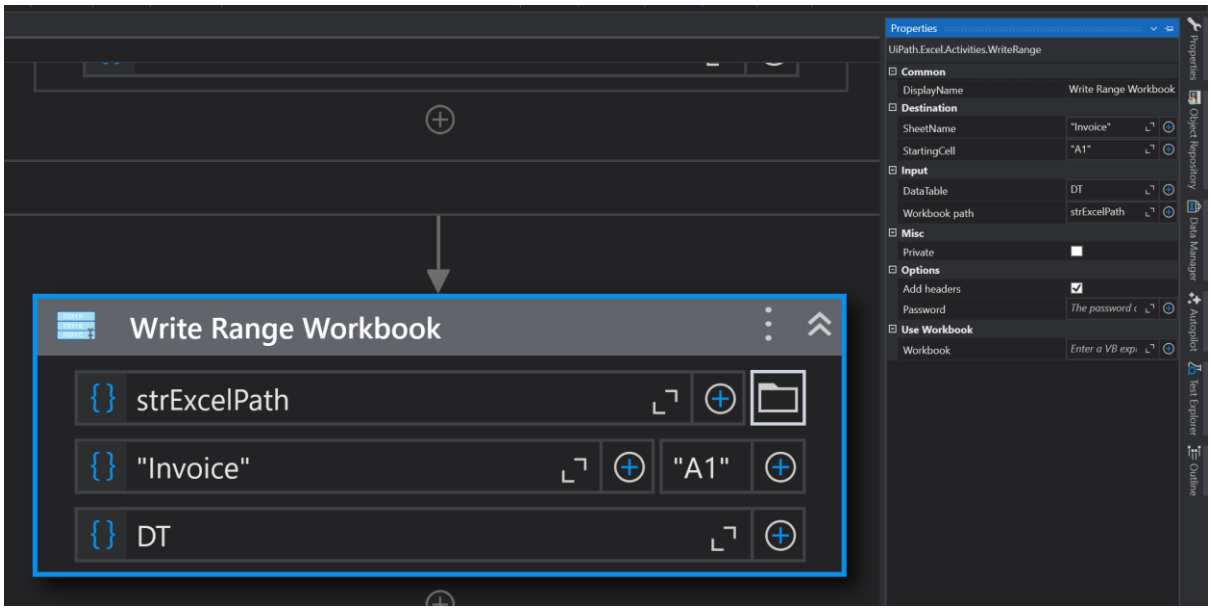


Figure 16: Activity – Write Range Workbook

	A	B	C	D
1	Invoice Number	Billed To	Due Date	Amount Due
2	847184	Melisa	5/8/2024	550.00
3	21893	KIM Kinra	8/15/2024	2,080.00
4	038229	Kris	10/23/24	52,000
5	6681688	Orange	18/9/2024	
6	6681688	Tre Xing	18/9/2024	$ 1,778.76
7		Satellite		
8	28298	Satellite E	3/20/24	$7,725.60
9	23117	Raymonc	19/9/2024	317.22
10	213119	Helen Mc	25/7/2024	RM1,400.00
11	3388666	Larisa Sin	15/05/202	RM1,700.00
12				

Figure 17: Extracted Output on Excel Workbook

The **Write Range Workbook** activity represents the final method for exporting **DT** data to an Excel file. Over here, all the extracted invoice data will be saved to a specified file location variable called **strExcelPath**. The **SheetName** is **Invoice**, and the starting cell will be **A1** as the checkbox for **Add headers** under **Properties** is checked.

The following is a detailed technical summary flowchart that covers all the steps used for the prototype development.

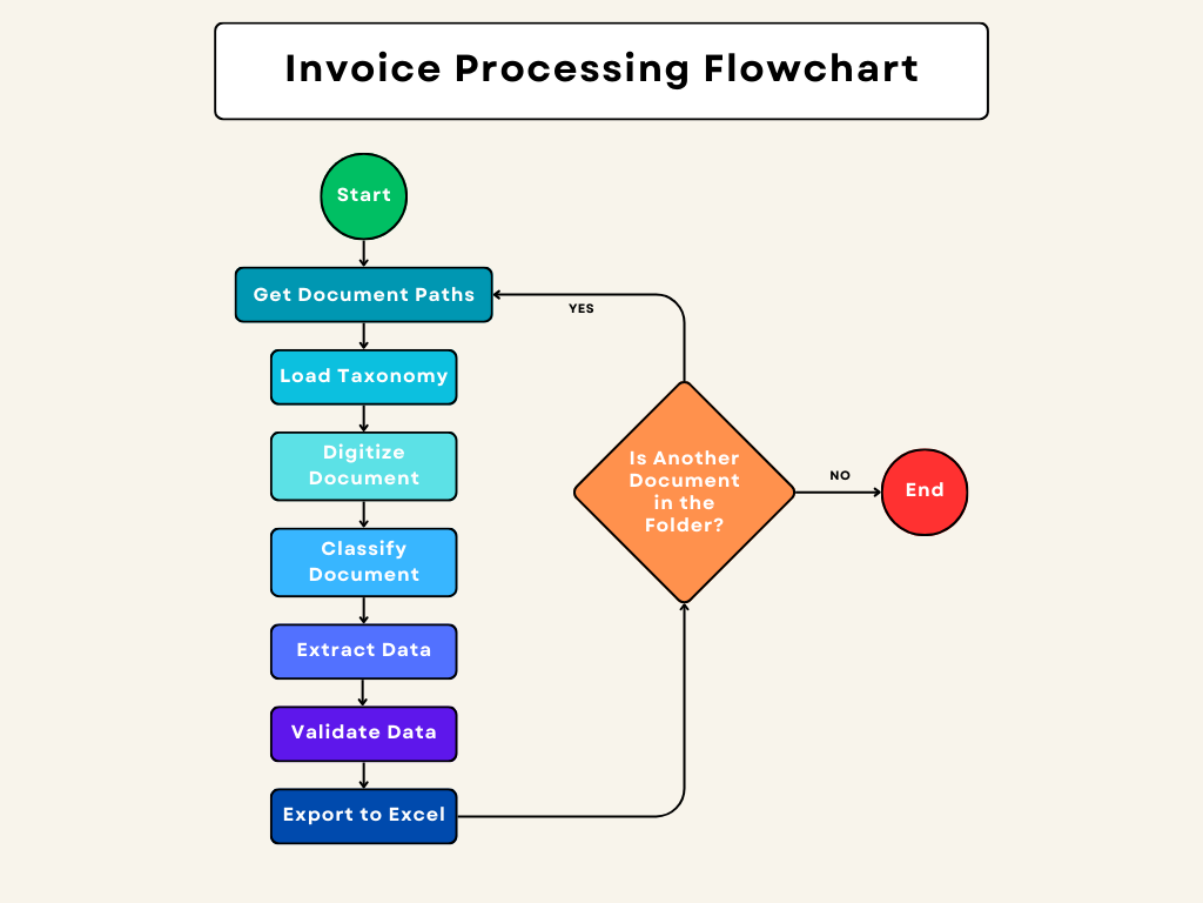


Figure 18: Invoice Processing Flowchart

Performance Evaluations

Performance evaluation is one of the principal steps used to measure and assess the efficiency and quality of the developed system. At this stage, the evaluation of automated invoice processing performance will be conducted on several metrics using UiPath's Validation Station, as stated below. After that, a few samples are provided in the next section to display the available information from the extracted results.

- **OCR Confidence Score:** This is to indicate the proportion of accurately identified characters to the total characters in the original document.
- **Extraction Confidence Score:** This measures the level of confidence after extracting and labelling information, indicating the performance of the extraction algorithm.
- **Validation Time:** This is to measure the time taken for the reviewer to validate and correct fields. If all the confidence scores from OCR and extraction are high, the validation time will be reduced, as it requires less validation to achieve excellent accuracy detection.
- **Corrections Made:** This serves to tally the number of manual adjustments carried out during validation.
- **Corrected Fields:** This section lists the fields that require manual correction, allowing us to observe which fields are commonly misclassified.
- **Comments:** This is to provide a statement to identify what particular corrections or results have been made.

To ensure insightful comparison, both Omnipage OCR and Tesseract OCR will be presented separately with their own metrics table summary of all the performance evaluations. The purpose is to provide a comprehensive view of each measurement associated with every processed document and summarise its overall performance. The testing will involve a total of 10 invoices, of which two are duplicated with modified structures, such as the adjusted invoice number's position. This will allow us to check whether the automated system can still correctly extract the required data in different formats. Any partial or fully incorrect extraction will be recorded with a sign of "(x)" next to the value on the table.

Sample Validation Station for OCR Confidence Level

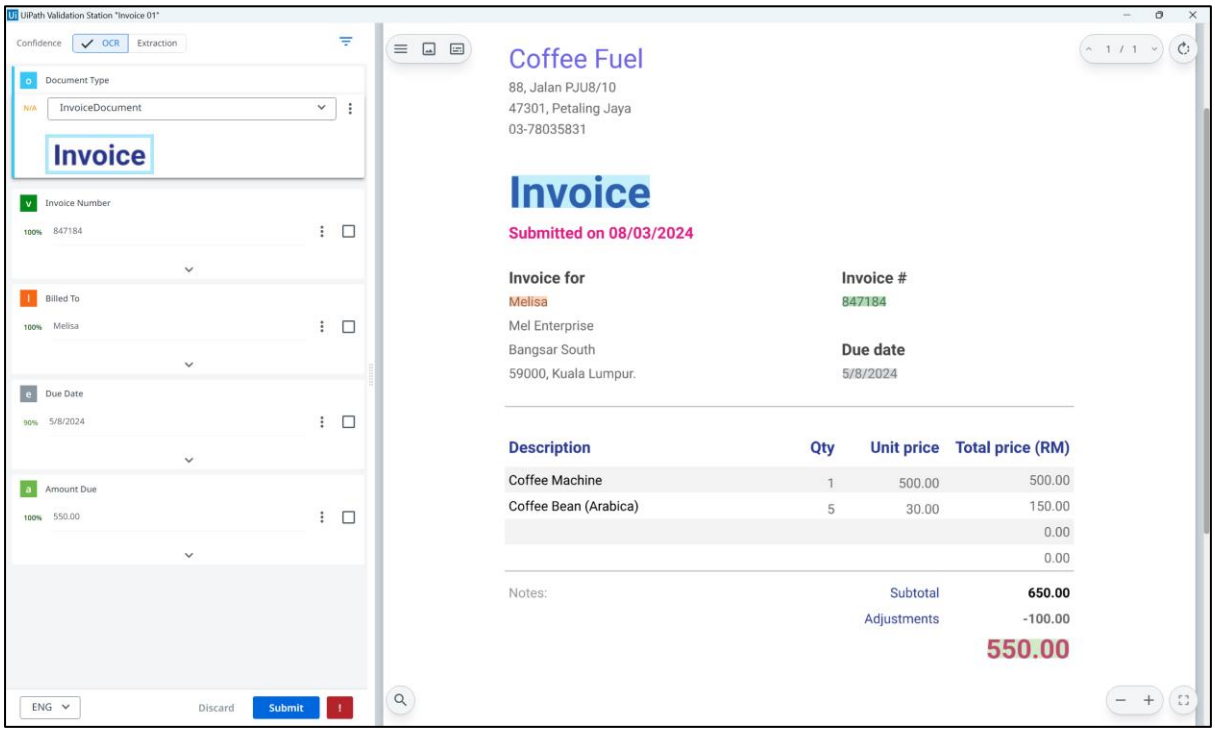


Figure 19: OCR Confidence Level for Invoice01

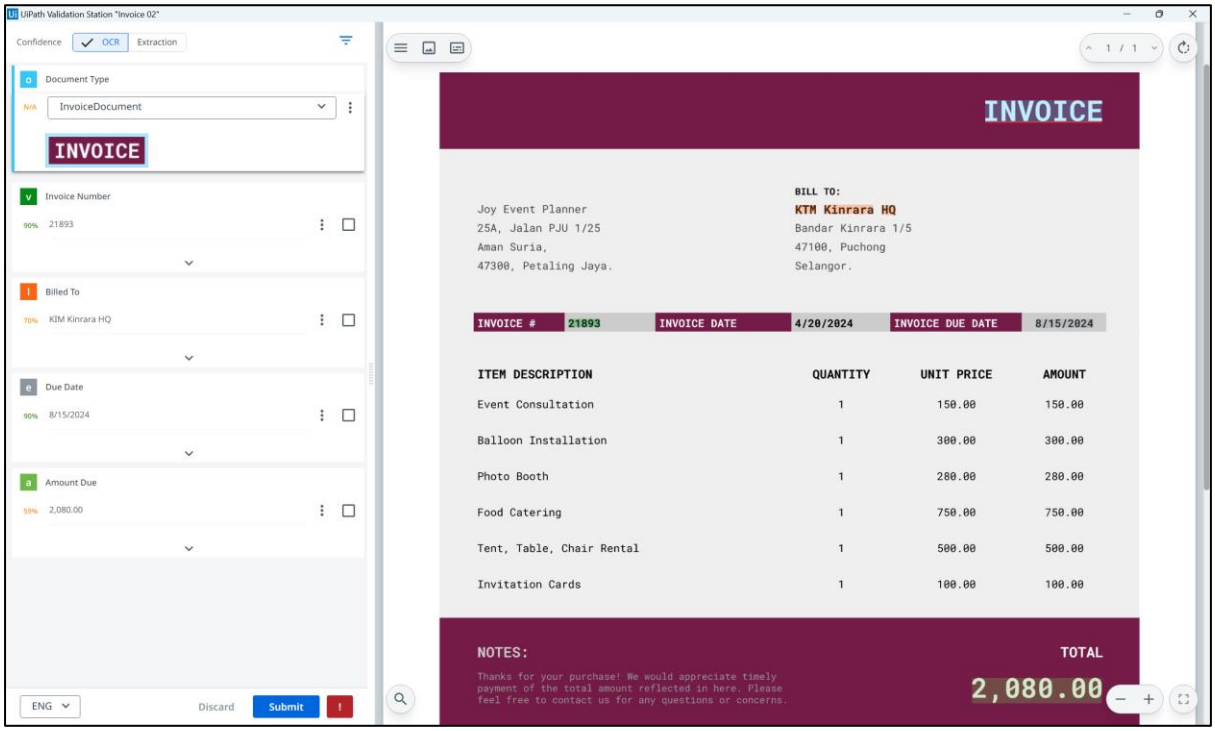


Figure 20: OCR Confidence Level for Invoice02

Sample Validation Station for Extraction Confidence Level

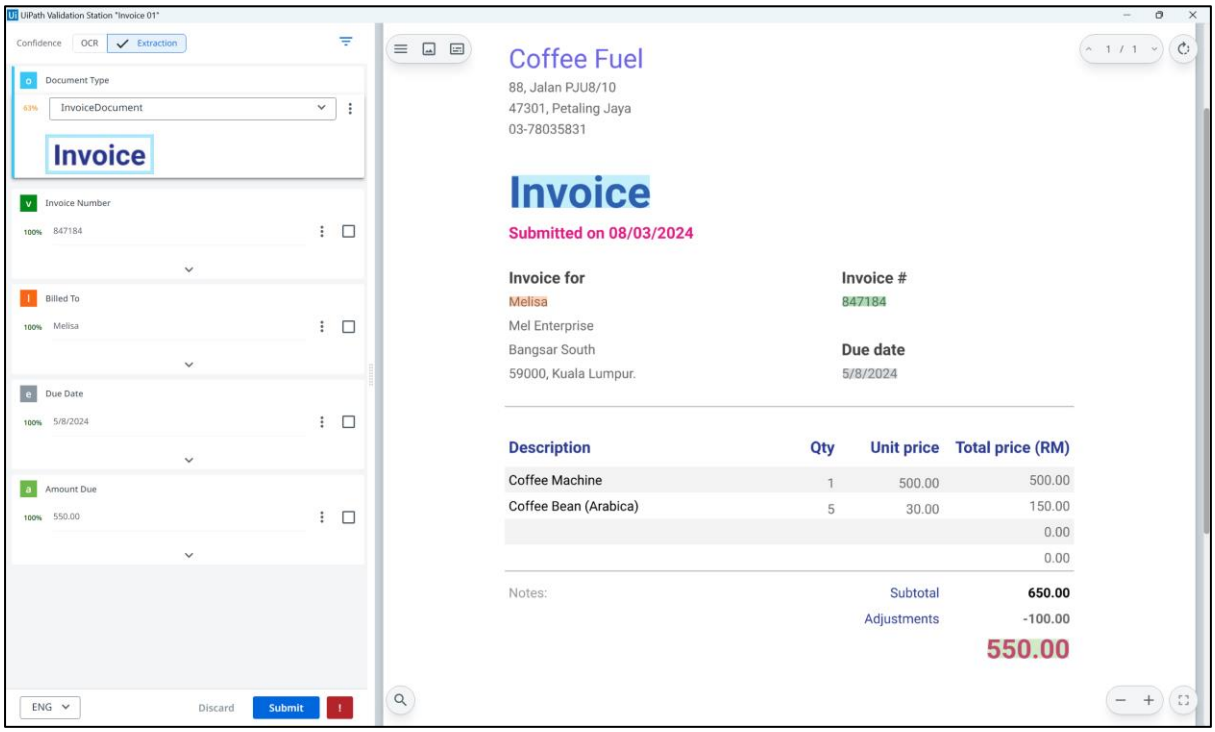


Figure 21: Extraction Confidence Level for Invoice01

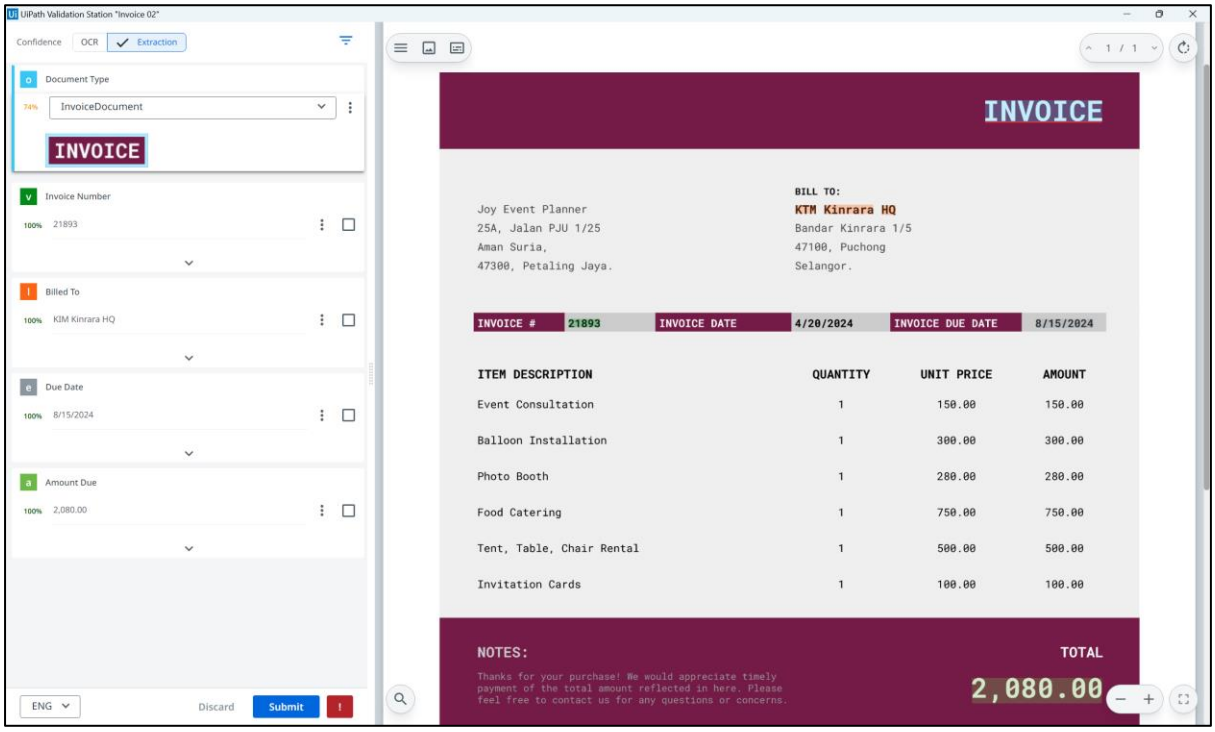


Figure 22: Extraction Confidence Level for Invoice02

Results Analysis and Discussion

OCR Engine – OmniPage

	OCR Confidence Score (%)				Extraction Confidence Score (%)			
Document ID	**Invoice Number**	**Billed To**	**Due Date**	**Amount Due**	**Invoice Number**	**Billed To**	**Due Date**	**Amount Due**
Invoice01	100	100	90	100	100	100	100	100
Invoice02	90	70	90	59	100	100 (x)	100	100
Invoice03	100	100	90	100	80	100	100	100
Invoice04 Modified	79	100 (x)	90	0 (x)	97	97 (x)	97	0 (x)
Invoice04	100	100	90	100	100	100	100	100
Invoice05 Modified	0 (x)	100 (x)	0 (x)	0 (x)	0 (x)	100 (x)	0 (x)	0 (x)
Invoice05	100	100	90	100	100	100	100	100
Invoice06	100	100	90	100	100	100	100	100
Invoice07	100	100	90	50	100	100	100	100
Invoice08	100	100	90	50	100	100	100	100

Table 1: Confidence Score of Omnipage for OCR and Extraction

Document ID	Overall OCR (%)	Overall Extraction (%)	Validation Time (seconds)	Corrections Made	Corrected Fields	Comments
Invoice01	97.5	100	13	0	None	None
Invoice02	77.25	100	30	1	Billed To	Billed To: “KIM” to “KTM”
Invoice03	97.5	95	12	0	None	None
Invoice04 Modified	64.25	72.75	60	2	Billed To, Amount Due	Billed To: “Orange” to “Tre Xing” Amount Due: Did not extract
Invoice04	97.5	100	15	0	None	None
Invoice05 Modified	25	25	130	4	Invoice Number, Billed To, Due Date, Amount Due	Invoice Number: Did not extract Billed To: “Satellite” to “Satellite Enterprise” Due Date: Did not extract Amount Due: Did not extract
Invoice05	97.5	100	15	0	None	None
Invoice06	97.5	100	20	0	None	None
Invoice07	85	100	18	0	None	None
Invoice08	85	100	17	0	None	None

Table 2: Overall Summary of Omnipage

Metric	Value
Average Document OCR Confidence	82.6%
Average Document Extraction Confidence	89.28%
Average Document Validation Time	33 seconds
Total Corrections Made	7
Most Corrected Field	Billed To

Table 3: Performance Metrics for Omnipage

OCR Engine – Tesseract

	OCR Confidence Score (%)				Extraction Confidence Score (%)			
Document ID	**Invoice Number**	**Billed To**	**Due Date**	**Amount Due**	**Invoice Number**	**Billed To**	**Due Date**	**Amount Due**
Invoice01	98	98	98	0 (x)	100	100	100	0 (x)
Invoice02	0 (x)	0 (x)	0 (x)	0 (x)	0 (x)	0 (x)	0 (x)	0 (x)
Invoice03	100	98	98	95 (x)	100	100	100	100 (x)
Invoice04 Modified	98	98 (x)	98	0 (x)	100	100 (x)	100	0 (x)
Invoice04	98	98	98	97	100	100	100	100
Invoice05 Modified	0 (x)	98 (x)	0 (x)	0 (x)	0 (x)	100 (x)	0 (x)	0 (x)
Invoice05	98	98	97	98	100	100	100	100
Invoice06	98	98	98	98	92	92	92	92
Invoice07	98	98	98	90	100	100	100	100
Invoice08	98	98	98	84 (x)	100	100	100	100 (x)

Table 4: Confidence Score of Tesseract for OCR and Extraction

Document ID	Overall OCR (%)	Overall Extraction (%)	Validation Time (seconds)	Corrections Made	Corrected Fields	Comments
Invoice01	73.5	75	60	1	Amount Due	Amount Due: Did not extract
Invoice02	0	0	410	0	None	Unable to proceed due to error
Invoice03	97.75	100	150	1	Amount Due	Amount Due: “52,000 |” to “52,000”
Invoice04 Modified	73.5	75	180	2	Billed To, Amount Due	Billed To: “Orange” to “Tre Xing” Amount Due: Did not extract
Invoice04	97.75	100	30	0	None	None
Invoice05 Modified	24.5	25	360	4	Invoice Number, Billed To, Due Date, Amount Due	Invoice Number: Did not extract Billed To: “Satellite” to “Satellite Enterprise” Due Date: Did not extract Amount Due: Did not extract
Invoice05	97.75	100	35	0	None	None
Invoice06	98	92	32	0	None	None
Invoice07	96	100	40	0	None	None
Invoice08	94.5	100	58	1	Amount Due	Amount Due: “RiMI1,700.00” to “RM1,700.00”

Table 5: Overall Summary of Tesseract

Metric	Value
Average Document OCR Confidence	75.18%
Average Document Extraction Confidence	76.7%
Average Document Validation Time	135.5 seconds
Total Corrections Made	9
Most Corrected Field	Amount Due

Table 6: Performance Metrics for Tesseract

Overall OCR Confidence

- According to the performance metrics of both OCR engines, it can be seen that the confidence score for Omnipage (82.6%) is higher than Tesseract (75.18%).
- Based on the OCR confidence level of Tables 1 and 4, Tesseract has captured significantly more incorrect data than Omnipage.
- In particular, the amount of invoices that were incorrectly identified for Tesseract is 6, and Omnipage is 2.
- This means that Omnipage extracts text from multiple fields better than Tesseract. It could possibly be due to its stronger optical character recognition and text structure management, which aligns with the initial findings from the literature review.

Overall Extraction Confidence

- Another insight is that OmniPage at 89.28% and Tesseract at 76.7% are confident in extraction.
- Although both performed reasonably well, the extraction confidence level of Tables 1 and 4 showed that Tesseract has extracted more incorrect data than Omnipage.
- To be specific, a total of 6 invoices were inaccurately captured by Tesseract, and Omnipage had 3.
- This discovery reveals OmniPage's text capture and structured information extraction capabilities, which can benefit invoices with varying layouts or fields.

Validation Time

- On average, document validation takes about 33 seconds in OmniPage and 135.5 seconds in Tesseract, based on Tables 3 and 6.
- The significant disparity demonstrates that OmniPage processes and verifies documents far more rapidly than Tesseract due to the fewer mistakes performed.
- A lower validation time is favourable because it is vital in high-volume processing settings where efficiency is top priority, especially in the invoice processing context.

Total Corrections Made

- Subsequently, the total count of corrections made for Omnipage is 7, whereas Tesseract is 9, according to Tables 2 and 5.
- It is evident that both OCR engines have made a very close number of corrections and have issues with invoices with modified structures according to their performance metrics.
- In this case, Tesseract is displaying a higher correction count compared to Omnipage due to its higher character error rate, which is also influenced by the OCR and extraction confidence levels.

Most Corrected Field

- The most corrected field for Omnipage is “Billed To”, which occurred 3 times on Table 2. This indicates that Omnipage may have problems correctly collecting names and company information because of font or spacing variations.
- On the other hand, the most corrected field for Tesseract is “Amount Due”, which appeared 5 times on Table 5. This suggests that Tesseract is facing difficulties in obtaining numerical values accurately.
- Therefore, each engine requires attention to specific scenarios, with Omnipage potentially improving its handling of text variations and Tesseract possibly enhancing its handling of numerical values.

Additional Insights

- According to Tables 1 and 4, both OCR engines have issues processing Invoice04 Modified and Invoice05 Modified.
- To point out, the fields for “Billed To” and “Amount Due” were incorrectly extracted for both OCR engines on Invoice04 Modified.
- Despite both OCR engines being able to extract the data from the “Billed To” field, the expected data was not retrieved.

- Even with the only extracted data for the “Billed To” field on both OCR engines, all fields from Invoice05 Modified were incorrect. For instance, the captured data was “Satellite”, when it should be “Satellite Enterprise” for Invoice05 Modified.
- After further investigation, it seems that reordering of the text and images will affect the taxonomy of the invoice, which became a potential source of confusion and added complexity for OCR.
- In addition, only Tesseract failed to extract data for Invoice02, which halted the entire program due to an “index out of bounds” error.
- This type of error is caused by the inability of Tesseract to distinguish text with a coloured background from the invoice.
- This proves that a complex background can impact data validation and lead to workflow exceptions.

Summary

This section has presented the performance assessment of OCR for Omnipage and Tesseract. Both OCR engines have demonstrated their own strengths and weaknesses. Based on the statistic, it is noticeable that the overall ratings of Omnipage are consistently higher than those of Tesseract, making it a much better selection for automated invoice processing. OmniPage exhibited a markedly reduced average validation time per document in contrast to Tesseract. This indicates that OmniPage excels in accuracy and enhances document processing speeds, which is essential for optimising operations in high-volume workflows. The performance difference between OmniPage and Tesseract emphasises how crucial it is to choose a top-notch OCR engine for jobs that are essential to business operations. Although Tesseract’s restrictions can pose a problem in processes where switching OCR engines is not practical, Aside from that, the automation workflow was stopped by Tesseract’s failure to extract any data from Invoice02, underscoring the importance of error handling procedures. Similar failures might interrupt processing and affect operational continuity if there are no backup plans.

Recommendations and Future Directions

Practical Recommendations

The implementation of AI-driven RPA has offered numerous advanced solutions to organisations for enhancing efficiency, accuracy, and scalability. In many industries, the integration of RCA invoice processing is a crucial function for business operations, but it still presents a formidable implementation challenge (Morshed et al., 2024). This section will discuss the practical recommendations for AI-driven RPA invoice processing with the supported findings in the earlier discussion.

Firstly, it is recommended to incorporate AI-driven RPA solutions with flexible templates for processing invoices. Traditional RPA systems, which rely on rigid templates to capture invoice data, may limit adaptability when processing various types of invoices. It is believed that this method can enable organisations to obtain key information from different invoice layouts without creating a new template for each layout. However, the limited scope of this assignment did not fully include flexible document extraction and classification templates. This is because UiPath uses community plans, which restrict the ability to create dynamic data extraction. Advanced document understanding from UiPath, such as Intelligent Keyword Classifiers and ML-based Extractors, is exclusive to the enterprise plan and enables more versatile data extraction across document formats. By creating semi-template-free solutions, the invoice processing systems will become more adaptable, which will reduce the time needed to configure templates from time to time. To get more accurate information, some hybrid models have been developed by combining OCR with CNN or BiLSTM models (Sharma et al., 2022). Not only can they capture complex invoice details from different layouts with better precision, they are also reducing the incorrect detection and maintaining the quality of the data.

The next recommendation for invoice processing automation would be investing in high-quality OCR technology. The performance metrics from the previous section demonstrate significant differences between the two OCR engines in terms of their confidence scores for OCR and extraction. Although not all fields have been identified correctly, Omnipage has shown consistently achieving higher accuracy and reliability in extracting key fields as

compared to Tesseract. This is particularly important for invoice processing because all financial records must be filled out precisely to prevent errors like incorrect invoice amounts or wrong customer billing data. Simultaneously, the previous analysis shows that Omnipage validates more quickly than Tesseract due to its superior ability to adapt to changing layouts and text densities. This will help businesses securely automate enormous document volumes, which can boost throughput and free up human resources for higher-value activities.

Furthermore, companies that are planning to employ sophisticated RCA systems should provide adequate training to their employees to maximise the benefits and maintain them. Since these systems are developed to reduce human intervention, skilled and knowledgeable employers are needed to understand the system's functioning and AI-driven automation. For instance, the person should have troubleshooting skills, be capable of utilising RPA applications, and manage workflows to fit the business requirements. In the context of OCR technology for invoice processing, understanding the end-to-end automation process will assist internal developers in learning the circumstances under which to perform manual intervention and manage error-handling procedures. This will ensure business continuity when there are issues with the automation workflow. With that, they will have more confidence in handling such unpredictable cases and keeping the process running smoothly.

Further Research Areas

Future research could integrate NLP to enhance the OCR solution for invoice processing, as it can analyse and retrieve information from intricate and diverse document layouts. Since invoices from different suppliers and transactions can have different formats, the OCR approach can be used to scan the required texts and NLP for understanding the document's context. This is certainly important, as multiple words in the invoice could have identical meanings. For example, the OCR system with integrated NLP would be able to differentiate related phrases better by detecting the surrounding language. This advancement will boost the accuracy of data extraction, enabling the extraction of data from a wider range of invoice documents, thereby enhancing the adaptability of RPA not only in the retail industry but also in other industries.

Another area of research would be to enhance the self-correction mechanism for automated invoice processing. As mentioned from the literature review, any errors in pricing, quantities, or other details can result in financial discrepancies. Therefore, the development of self-correction mechanisms within RCA workflows can minimise the dependency on human validation. Instead, the automation workflow will strengthen its ability to learn from past mistakes, thereby improving accuracy over time. In other words, self-learning AI models like NAS can be used together with OCR to extract accurate data, and it will likewise decrease the frequency of manual intervention.

Conclusion

In conclusion, this assignment has highlighted the current capabilities of RPA and AI-driven OCR technology using UiPath. A prototype AI-driven RPA solution has also been developed to investigate the performance of Omnipage OCR and Tesseract OCR for invoice processing. The final analysis revealed that Omnipage OCR is significantly better than Tesseract OCR in terms of data extraction accuracy, validation time, and reliability. However, both OCR engines also have drawbacks for extracting data when different layouts are introduced. This signifies that further development should be focused on improving the data extraction for diverse invoice formats. In the same fashion, these findings underscore the significance of selecting resilient OCR solutions capable of accommodating document diversity, an essential factor for enterprises managing various invoice forms. Besides that, some key AI-driven RPA challenges were also discussed, including reliance on inflexible template-based invoices and limited functions using community licenses. The restricted access to advanced document understanding features has brought about future research on other advanced AI methods, such as NLP-based contextual understanding and self-learning mechanisms. Moreover, more training should be provided to internal employees or developers to make sure they will be able to support any post-deployment monitoring efforts. Although the proposed RCA solution offers certain automation advantages, the potential for future enhancements suggests that enterprises can achieve superior accuracy, efficiency, and flexibility by embracing emerging AI technology. The ongoing progress in AI-driven RPA can facilitate the development of more intelligent and robust automated systems that can address the changing requirements of contemporary business landscapes (Baviskar et al., 2021).

References

Baviskar, D., Ahirrao, S., Potdar, V., & Kotecha, K. (2021). Efficient Automated Processing of the Unstructured Documents Using Artificial Intelligence: A Systematic Literature Review and Future Directions. *IEEE Access*, *9*, 72894–72936. https://doi.org/10.1109/access.2021.3072900

Kanaparthi, V. K. (2023). Examining the Plausible Applications of Artificial Intelligence & Machine Learning in Accounts Payable Improvement. *FinTech*, *2*(3), 461–474. https://doi.org/10.3390/fintech2030026

Morshed, A., Ramadan, A., Maali, B., Khrais, L. T., & Baker, A. a. R. (2024). Transforming accounting practices: The impact and challenges of business intelligence integration in invoice processing. *Journal of Infrastructure Policy and Development*, *8*(6), 4241. https://doi.org/10.24294/jipd.v8i6.4241

Potturu, S. M. (2023). UIPATH BOT FRAMEWORK: ACCELERATING RPA DEVELOPMENT AND INNOVATION. *IJRDO -Journal of Computer Science Engineering*, *9*(4), 1–15. https://doi.org/10.53555/cse.v9i4.5853

Sharma, C., Bharadwaj, S. S., Gupta, N., & Jain, H. (2022). Robotic process automation adoption: contextual factors from service sectors in an emerging economy. *Journal of Enterprise Information Management*, *36*(1), 252–274. https://doi.org/10.1108/jeim-06-2021-0276

Treacy, S., Adyanthaya, A., Kearny, C., Anand, J., O'Sullivan, K., & Xu, Y. (2023). From Hype to Reality: Navigating the Challenges of RPA Implementation. *European Conference on Innovation and Entrepreneurship*, *18*(2), 875–882. https://doi.org/10.34190/ecie.18.2.1721

UiPath. (2024, October 31). *Document Understanding Modern Projects User Guide*. UiPath Documentation. https://docs.uipath.com/document-understanding/automation-cloud/latest/user-guide/public-endpoints

Appendices

List of Abbreviations

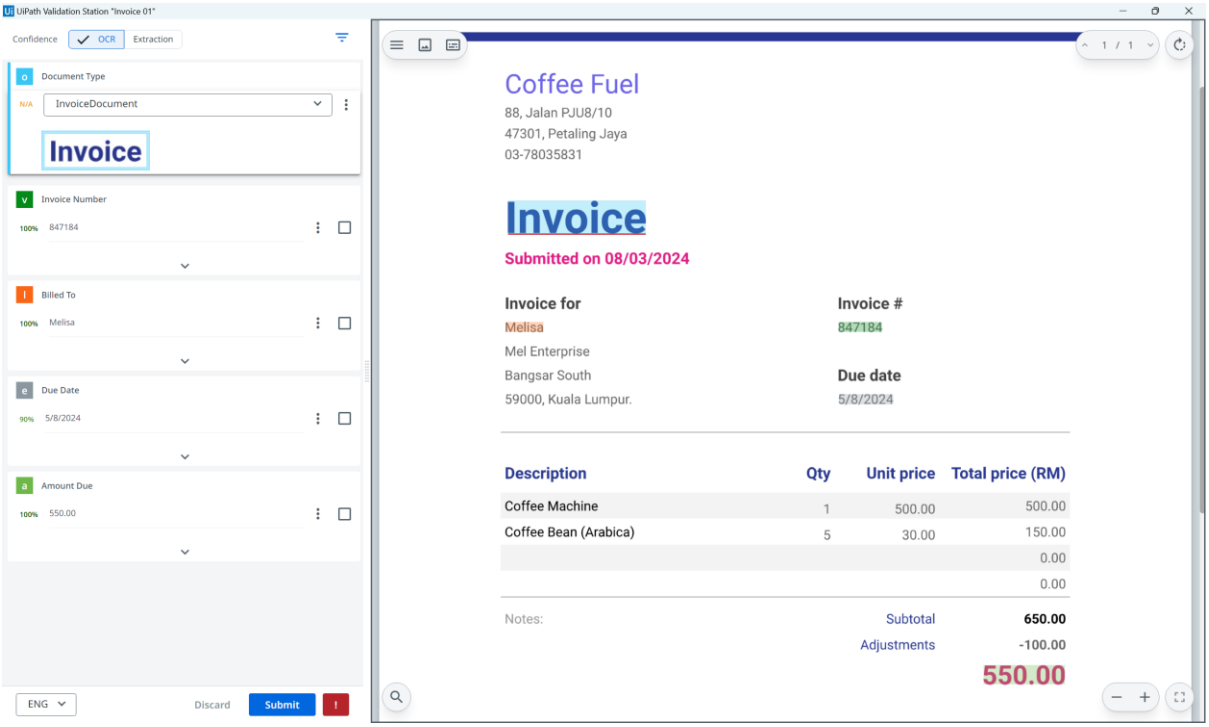
AI	Artificial Intelligence	ML	Machine Learning
RPA	Robotic Process Automation	NLP	Natural Language Processing
CRNN	Convolutional Recurrent Neural Network	BiLSTM	Bidirectional Long Short-Term Memory
RCA	Robotic Cognitive Automation	NAS	Neural Architecture Search

List of Figures

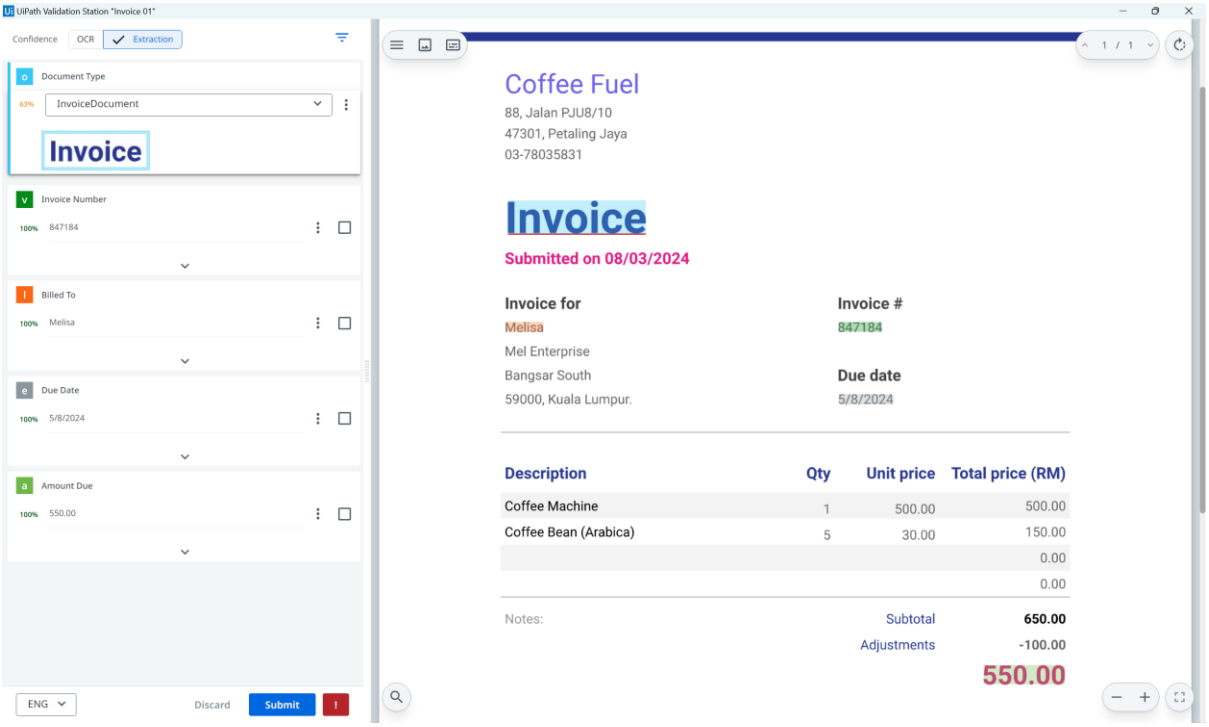
List of Tables

Test Results of Invoice Processing with OmniPage OCR and Extraction

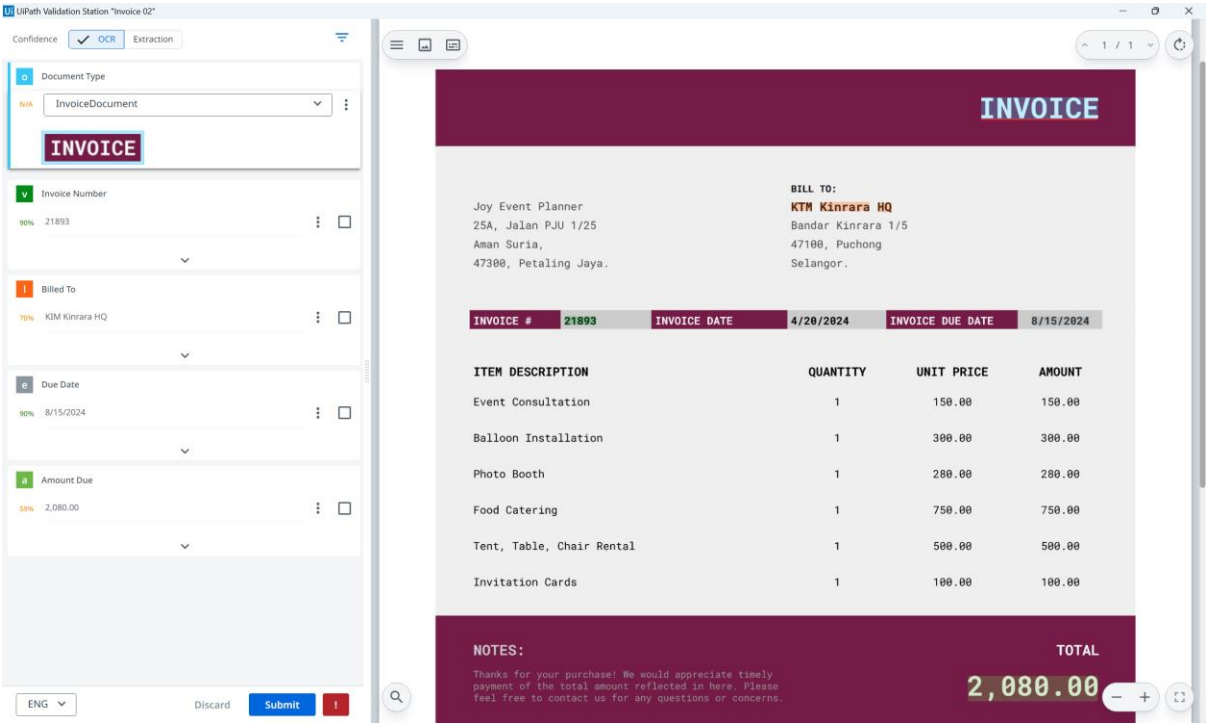
Invoice01 – OCR



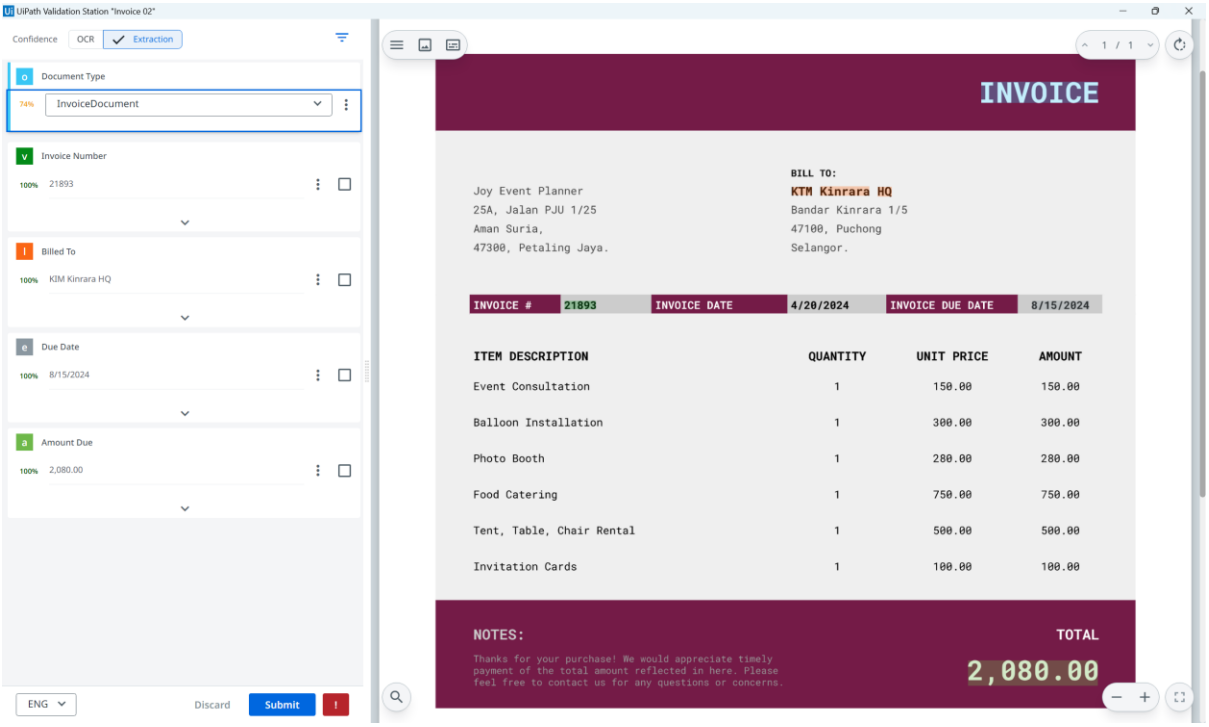
Invoice01 – Extraction



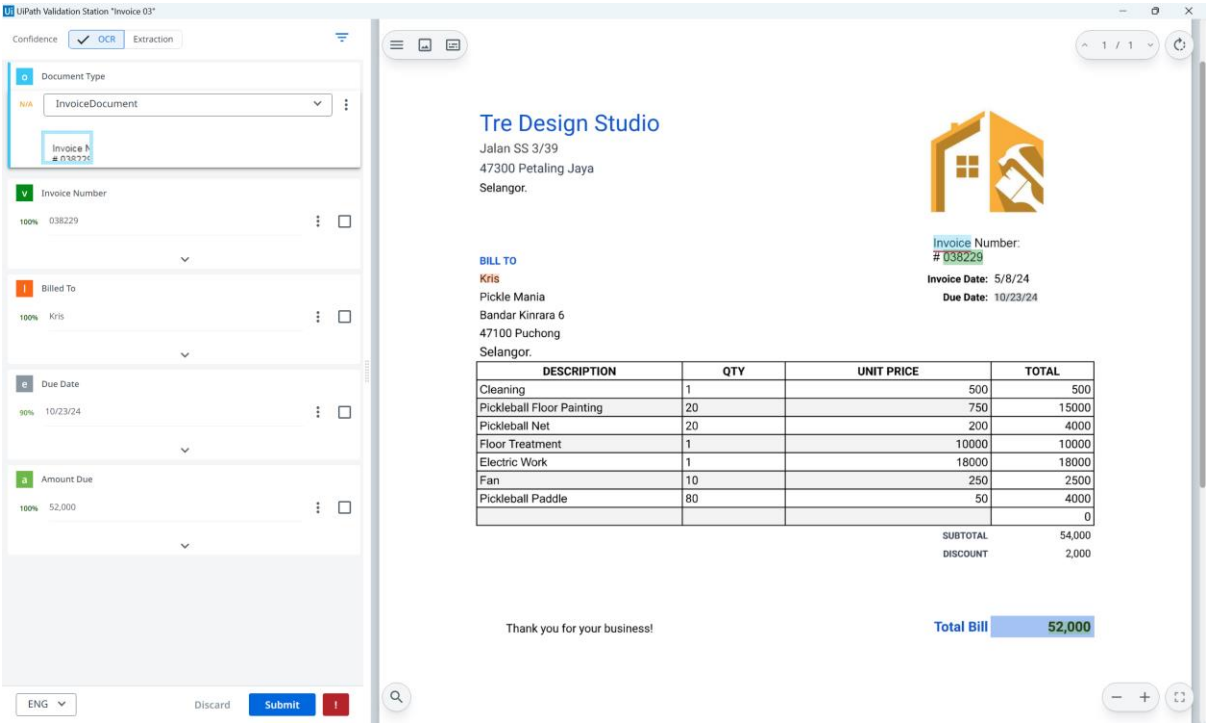
Invoice02 – OCR



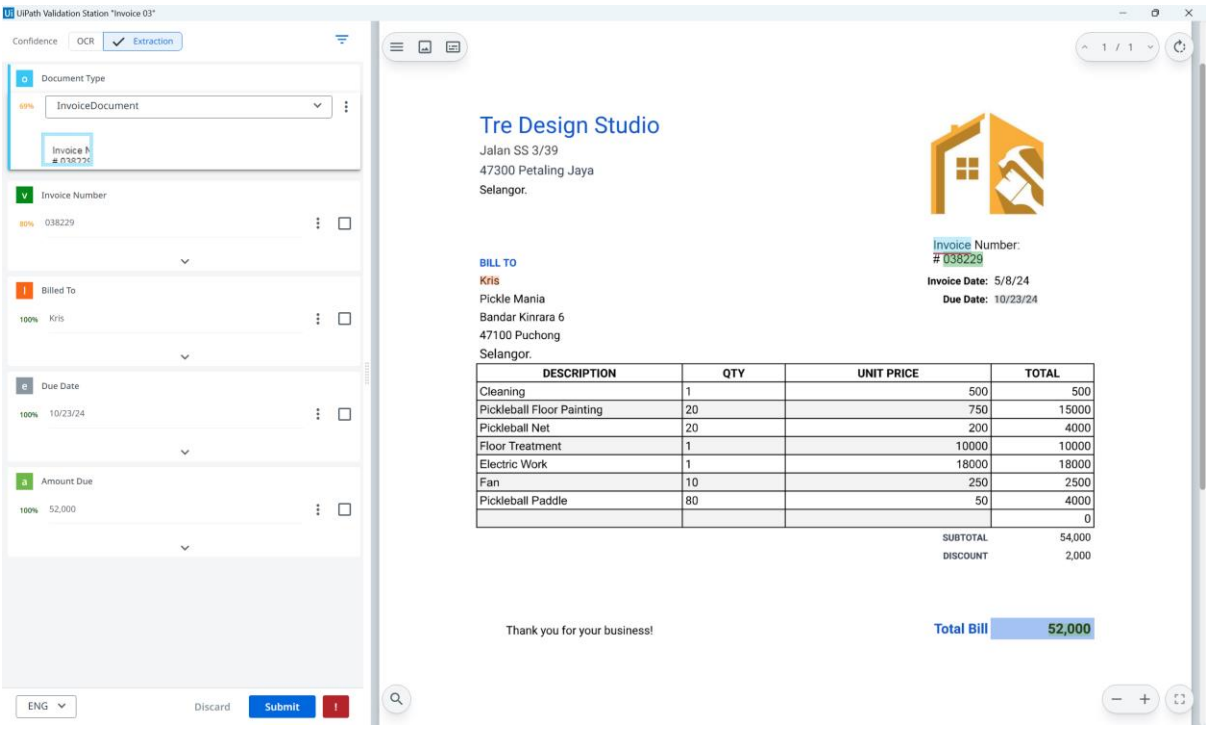
Invoice02 – Extraction



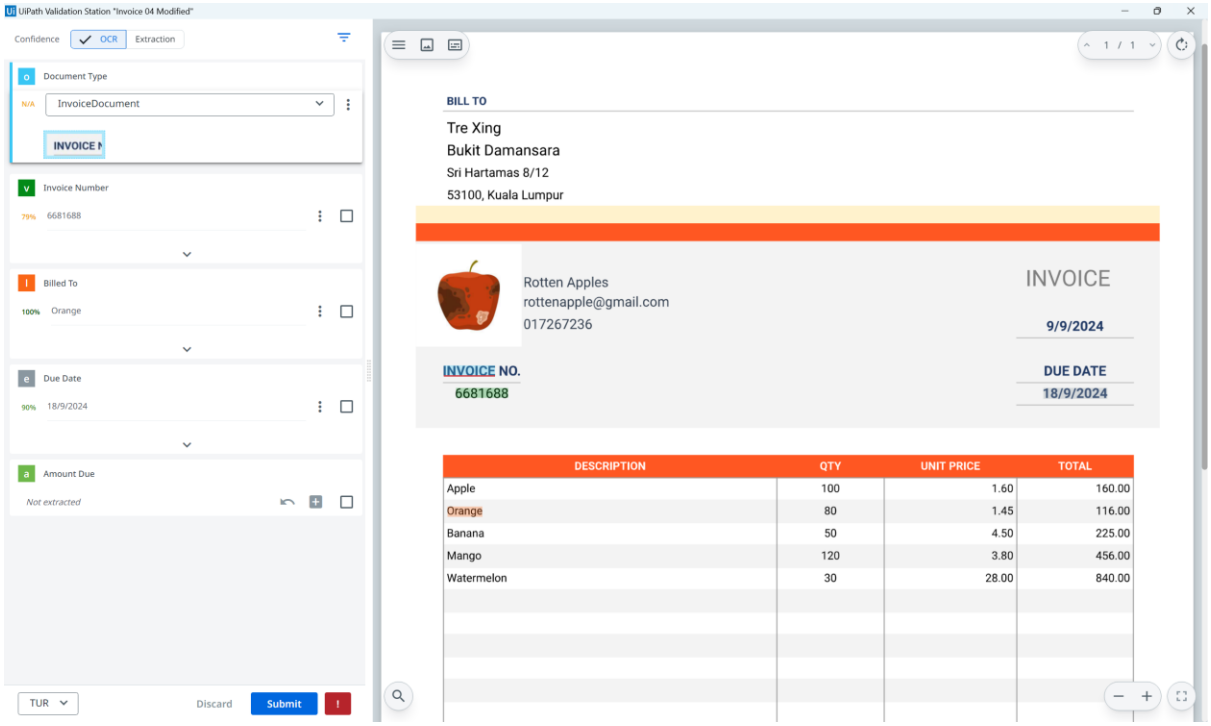
Invoice03 – OCR



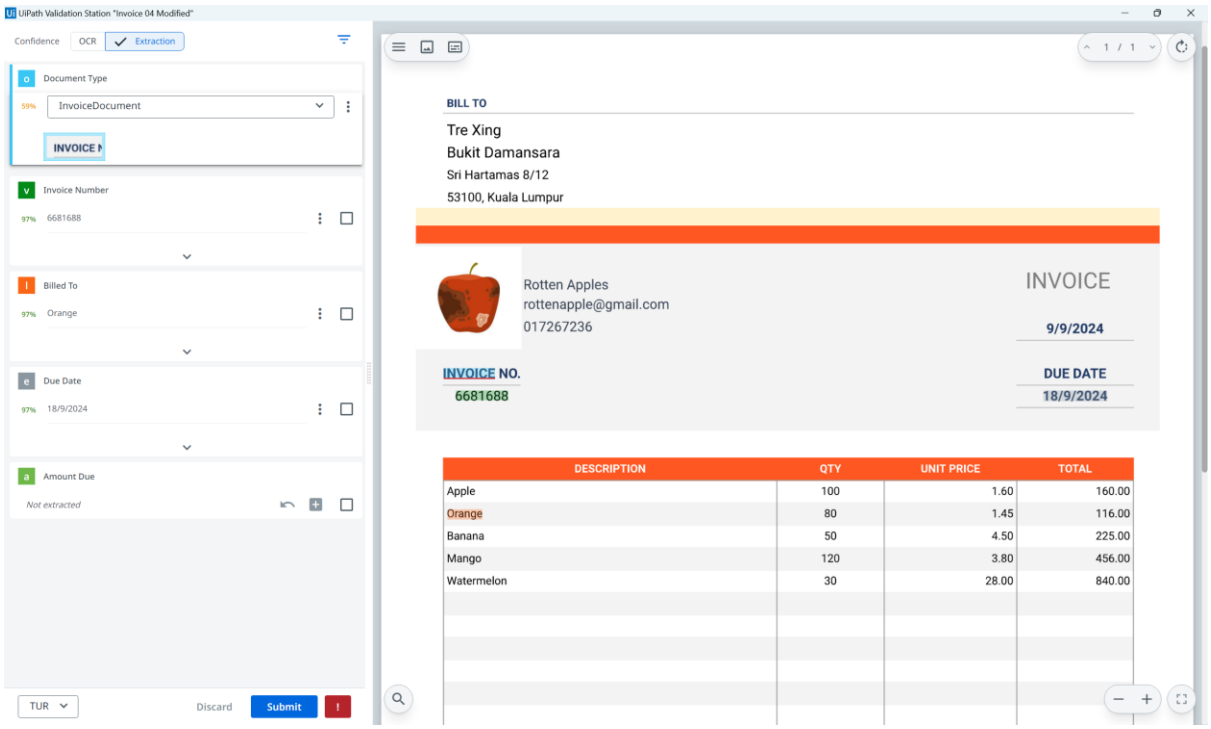
Invoice03 – Extraction



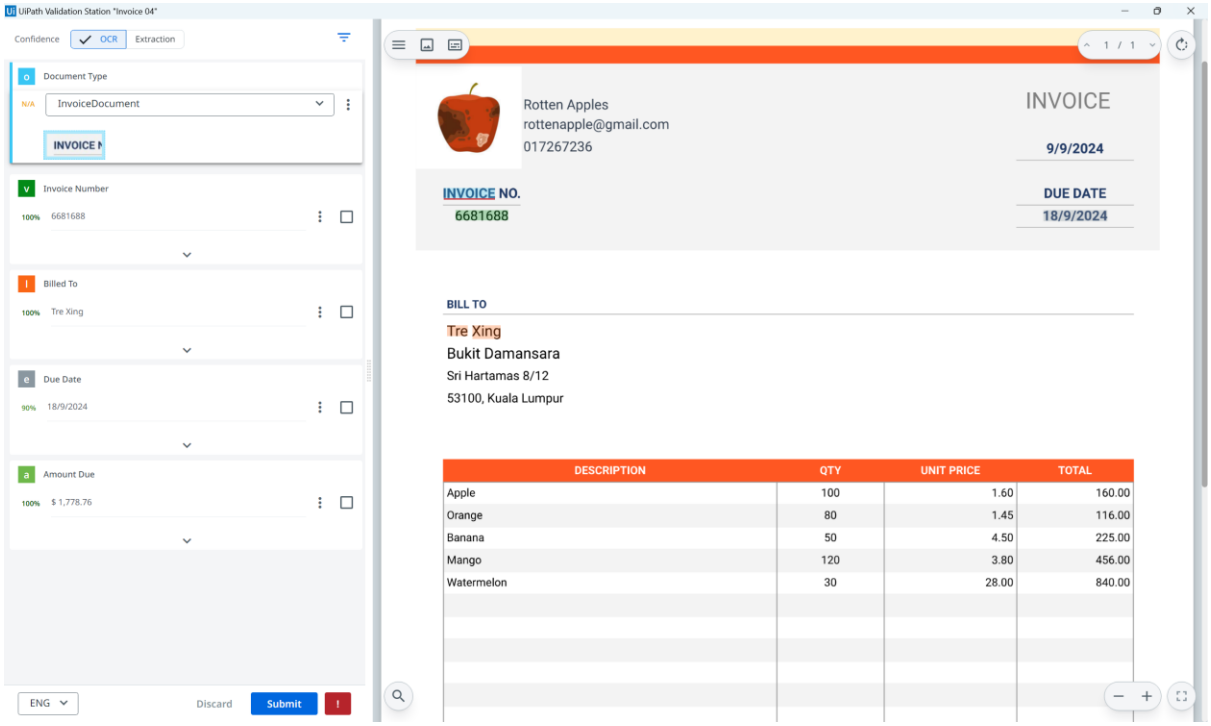
Invoice04 Modified – OCR



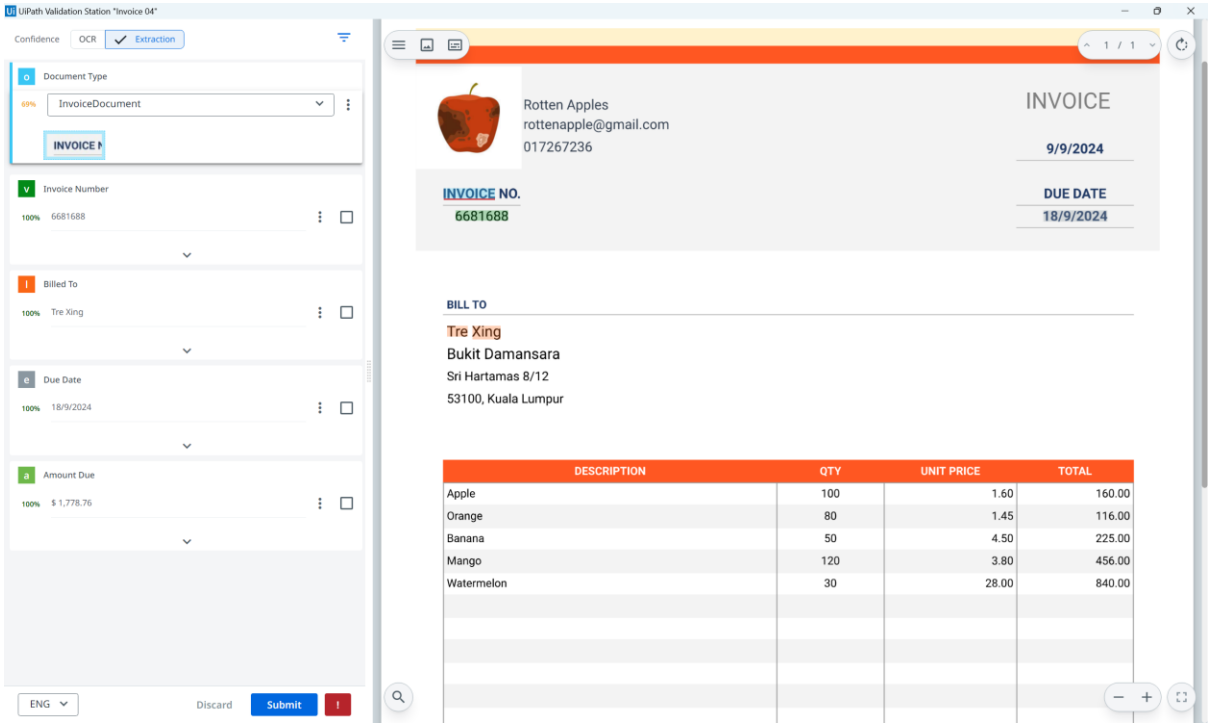
Invoice04 Modified – Extraction



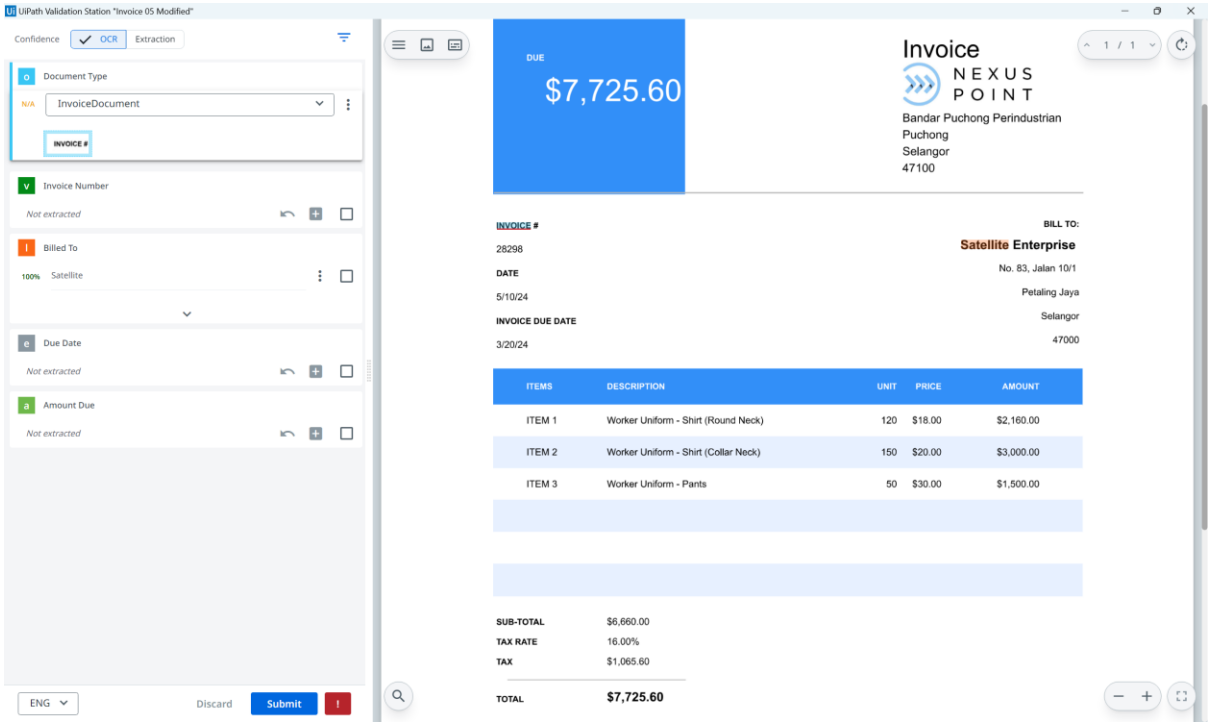
Invoice04 – OCR



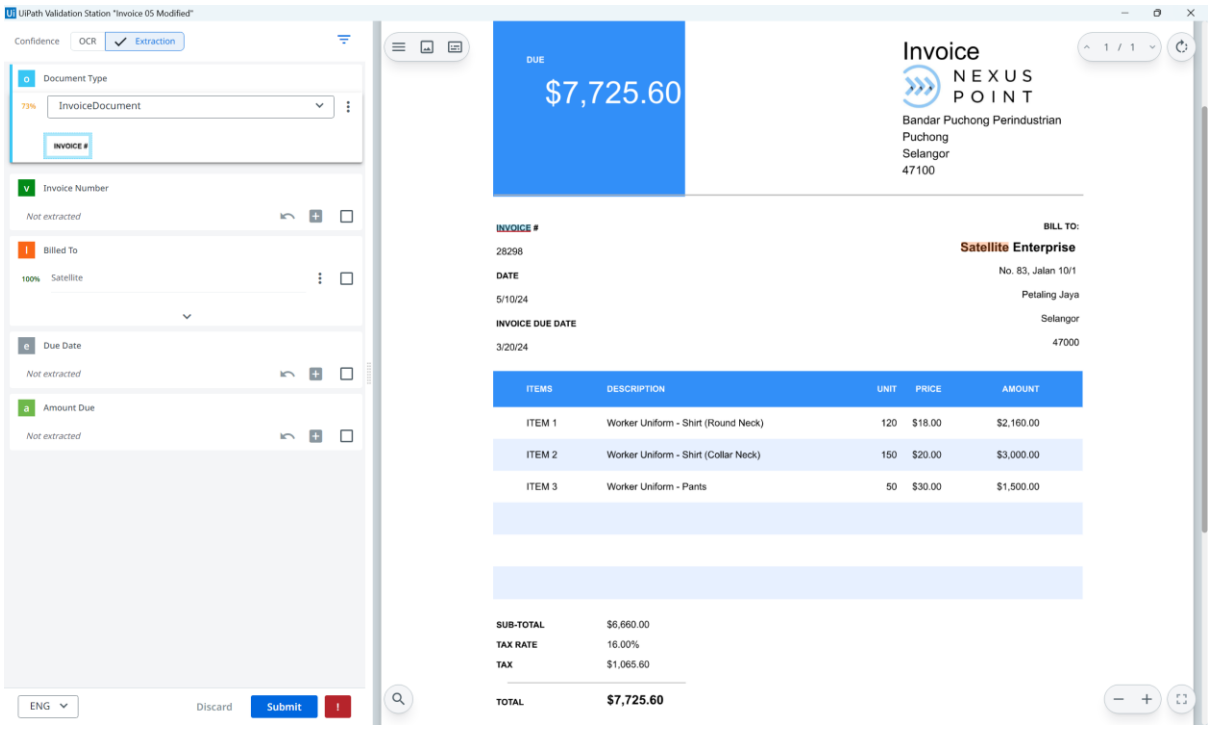
Invoice04 – Extraction



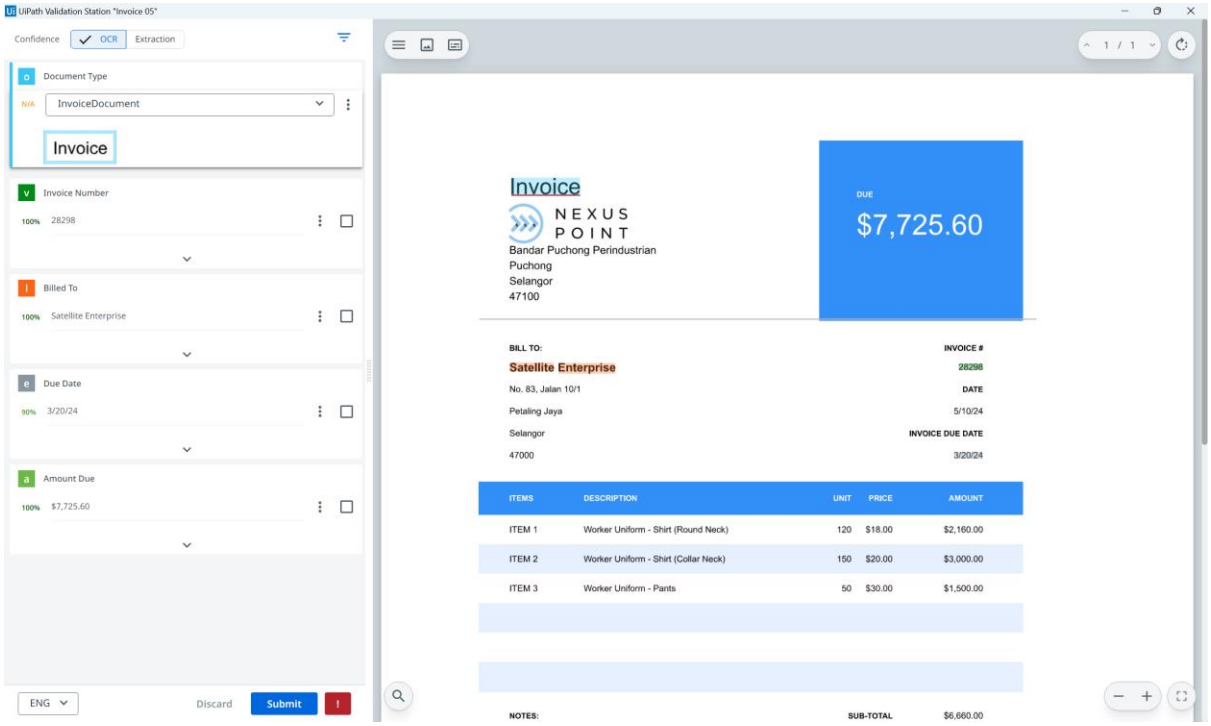
Invoice05 Modified – OCR



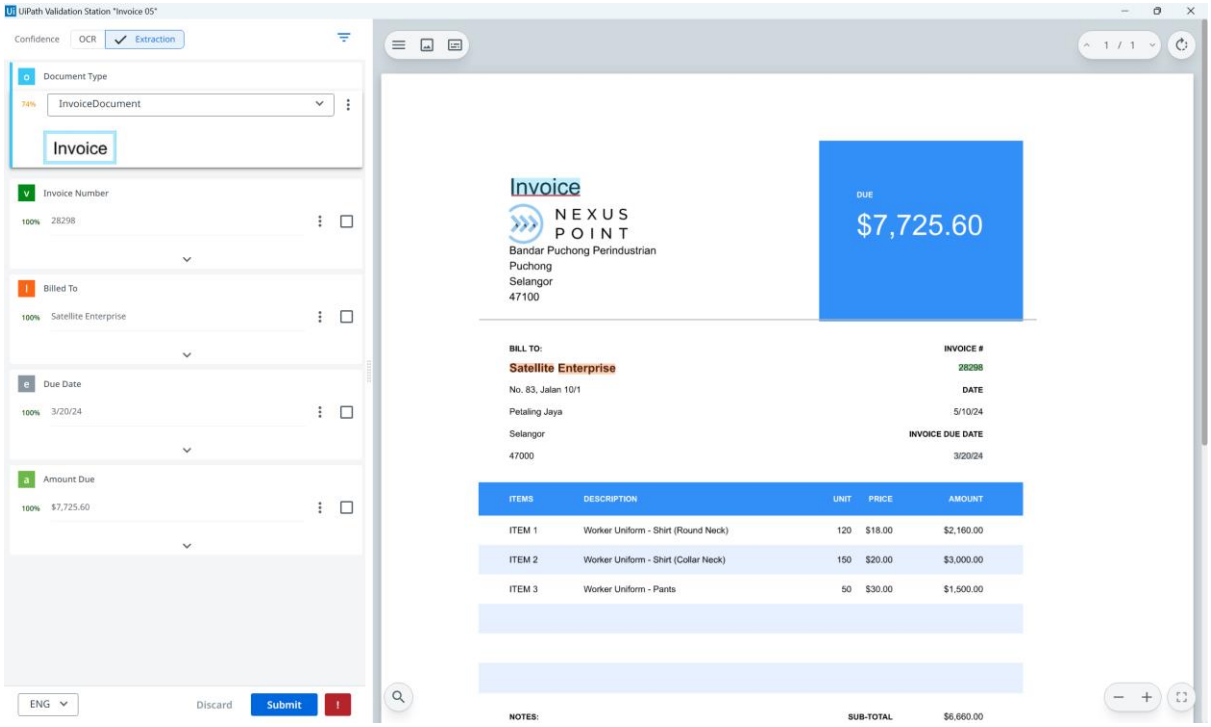
Invoice05 Modified – Extraction



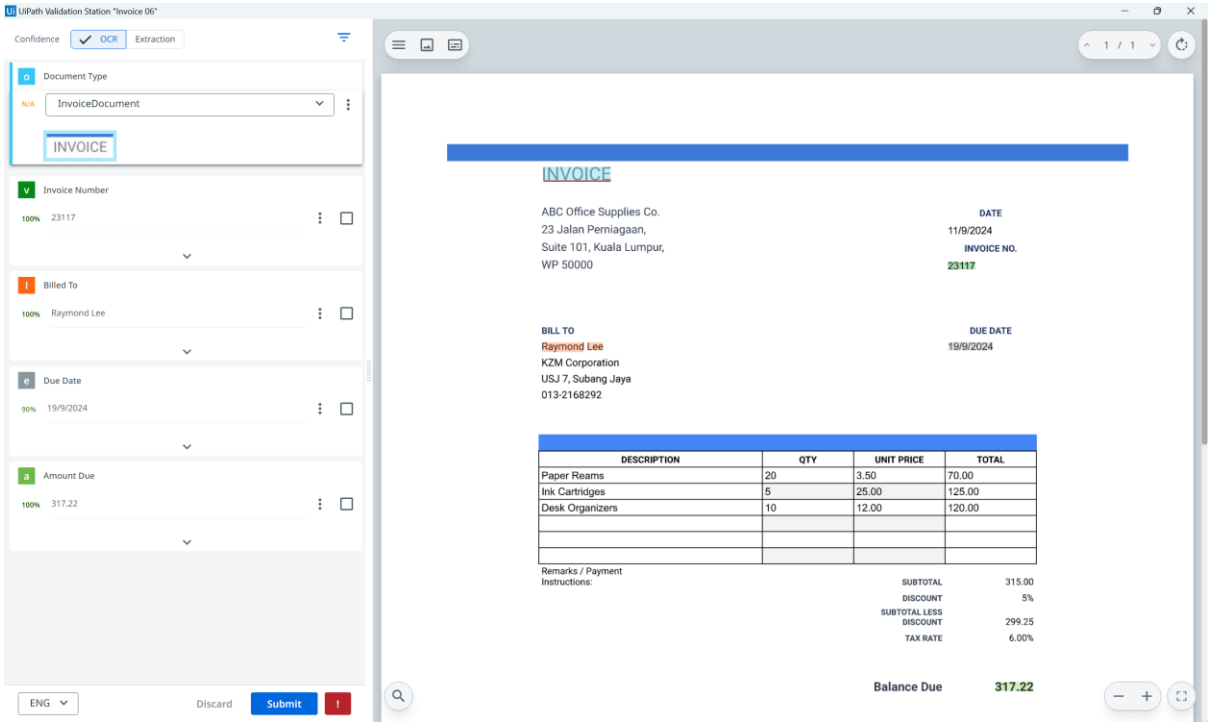
Invoice05 – OCR



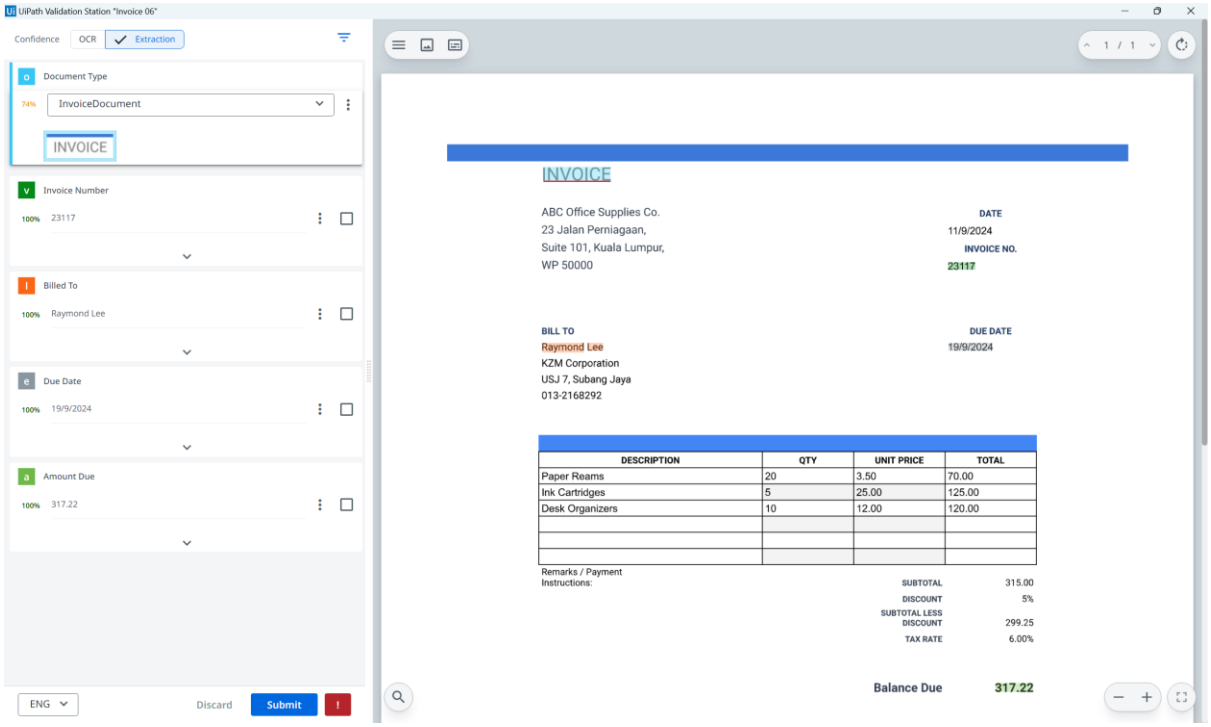
Invoice05 – Extraction



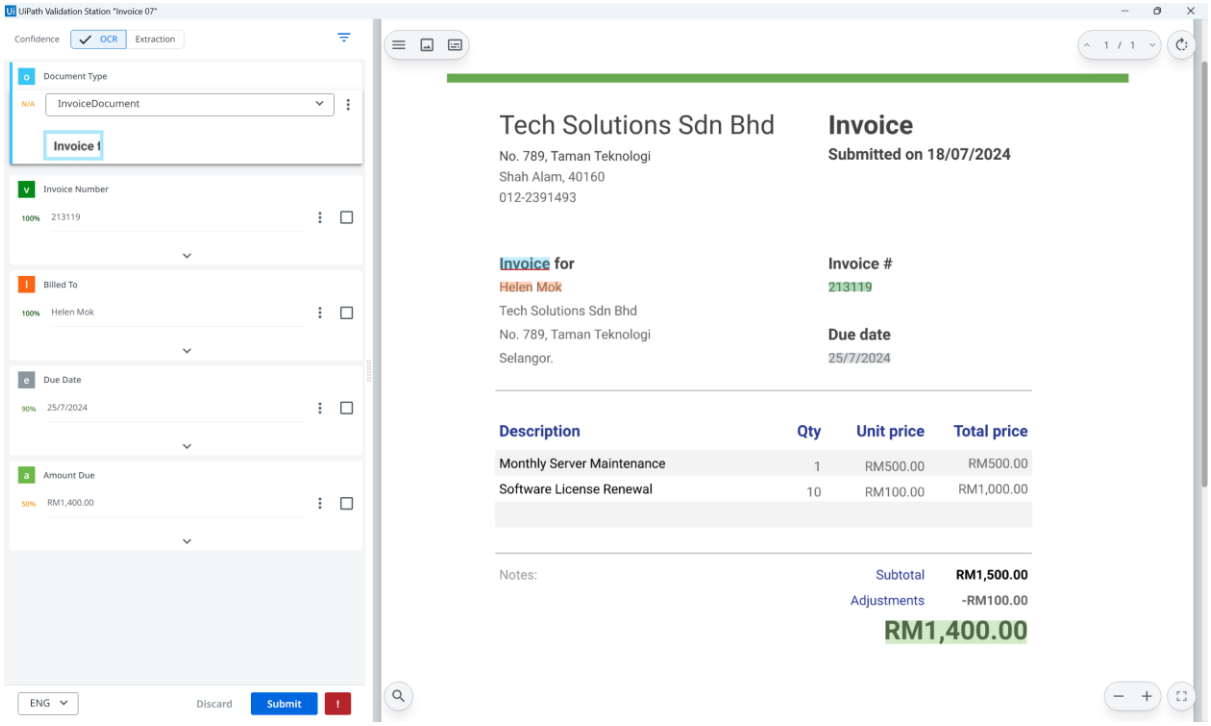
Invoice06 – OCR



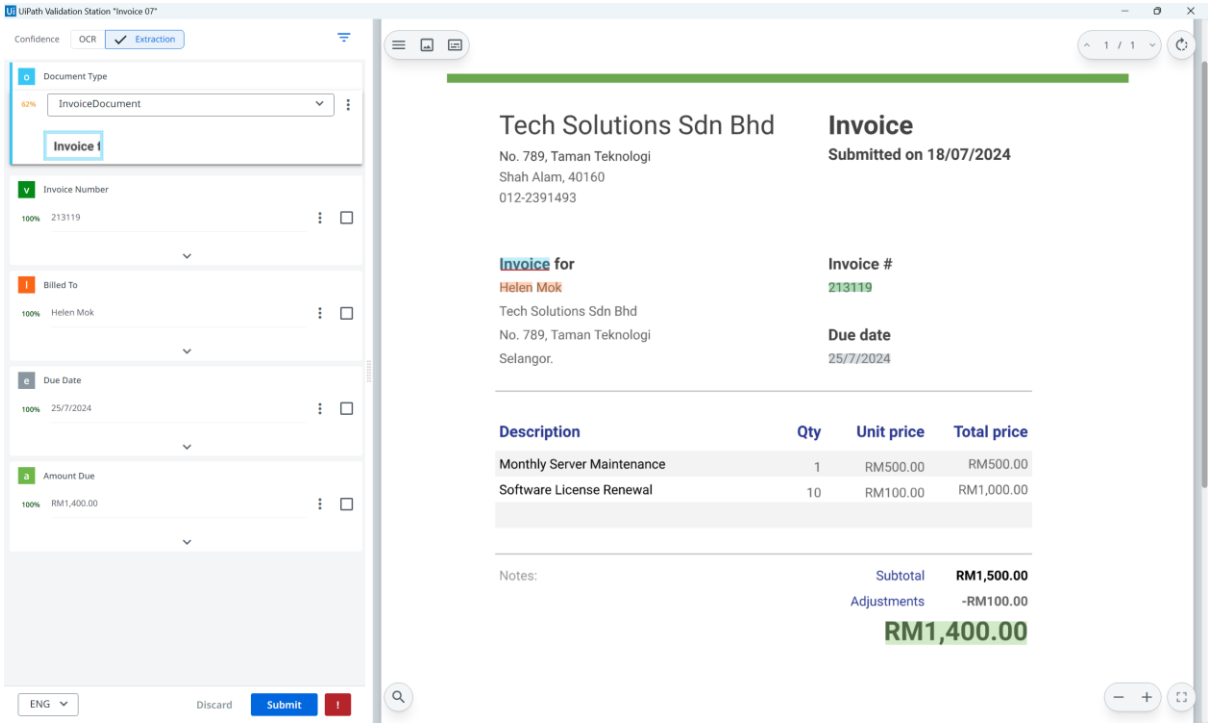
Invoice06 – Extraction



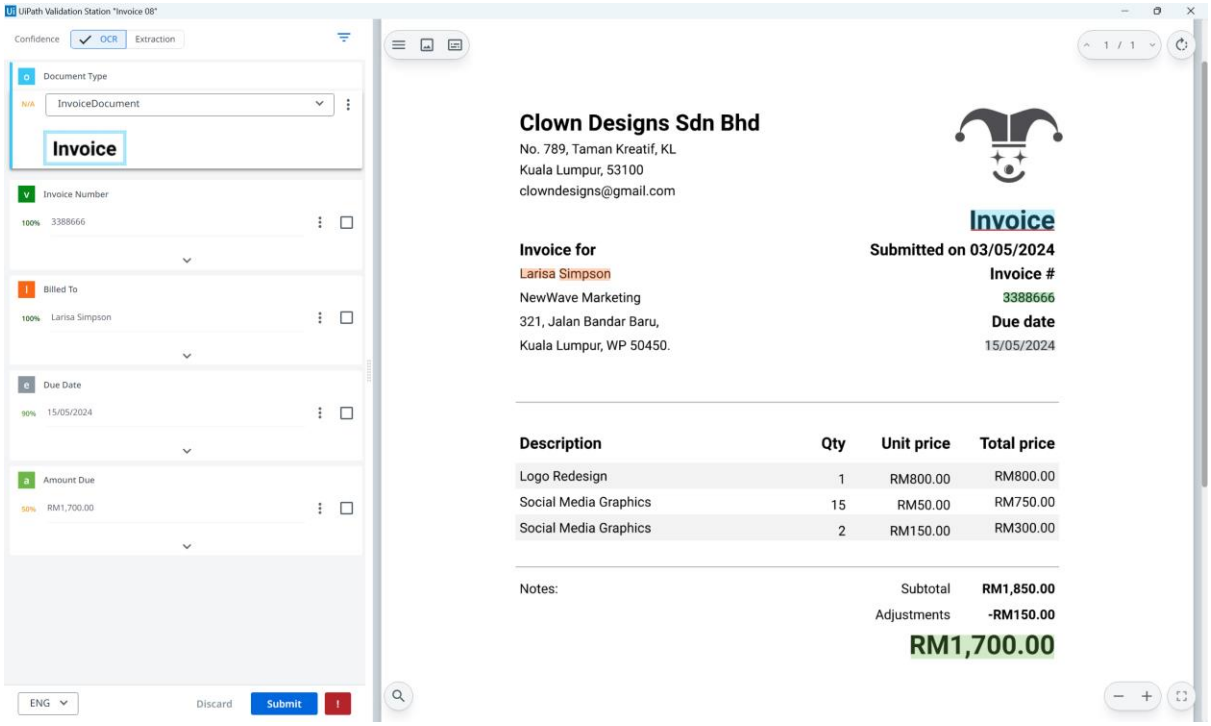
Invoice07 – OCR



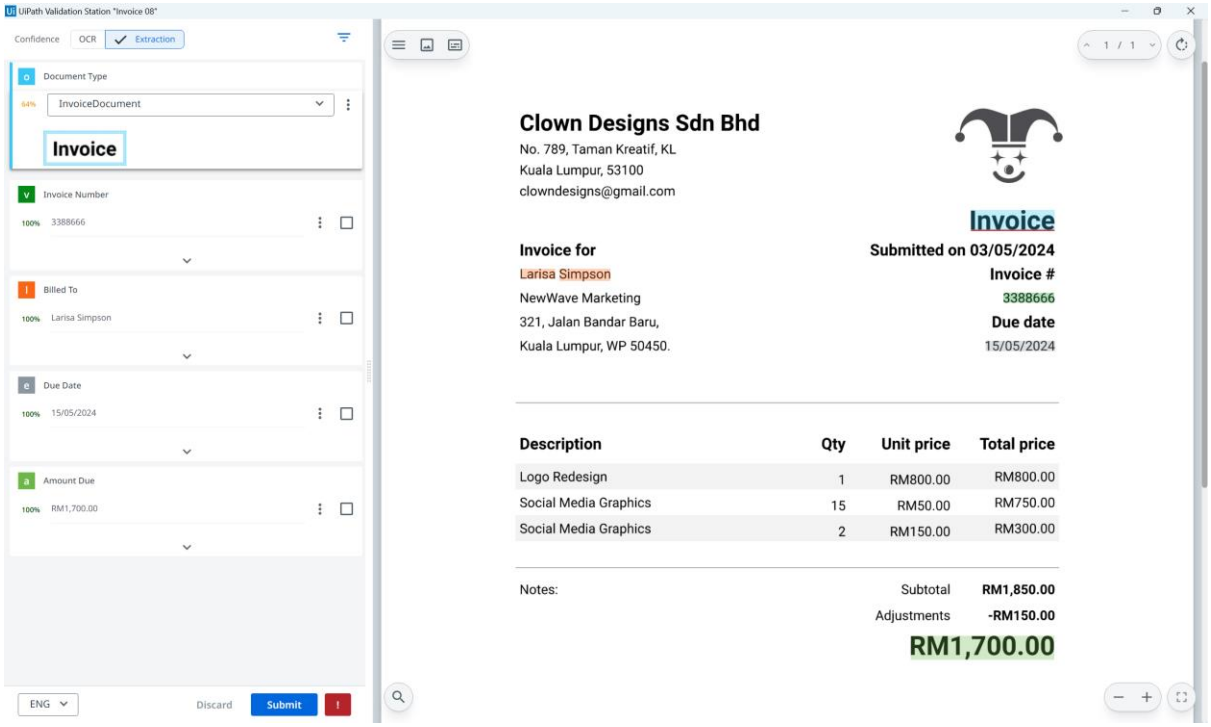
Invoice07 – Extraction



Invoice08 – OCR

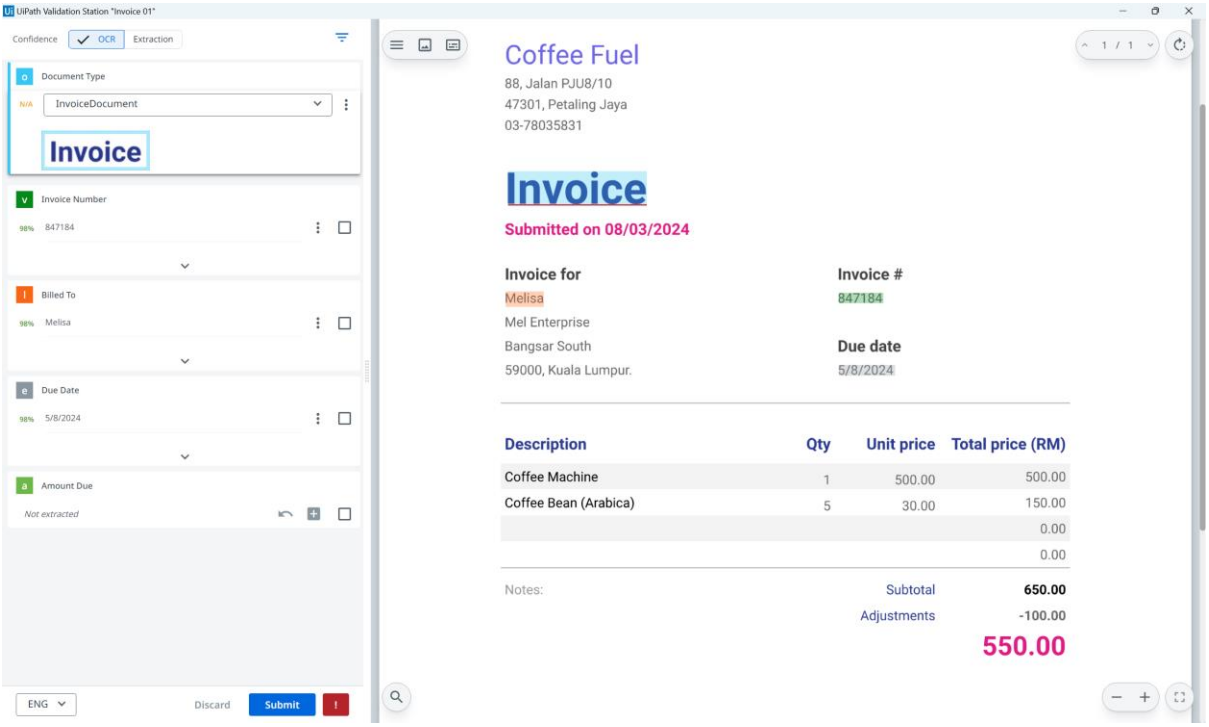


Invoice08 – Extraction

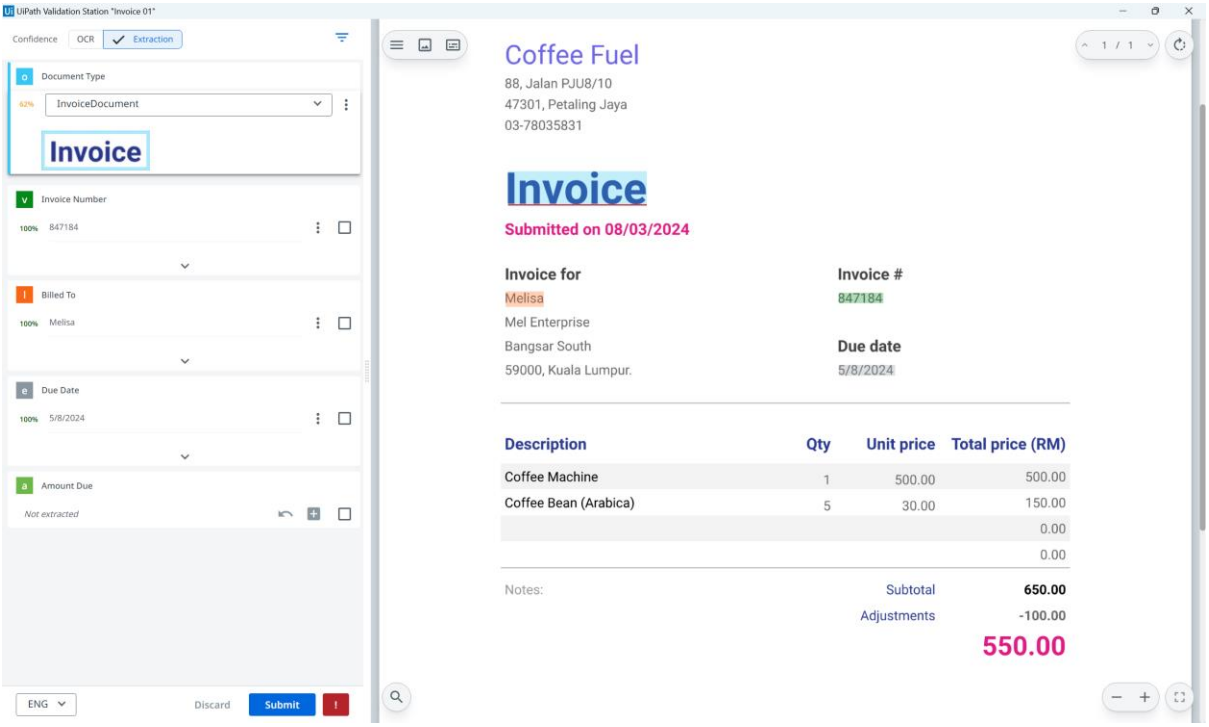


Test Results of Invoice Processing with Tesseract OCR and Extraction

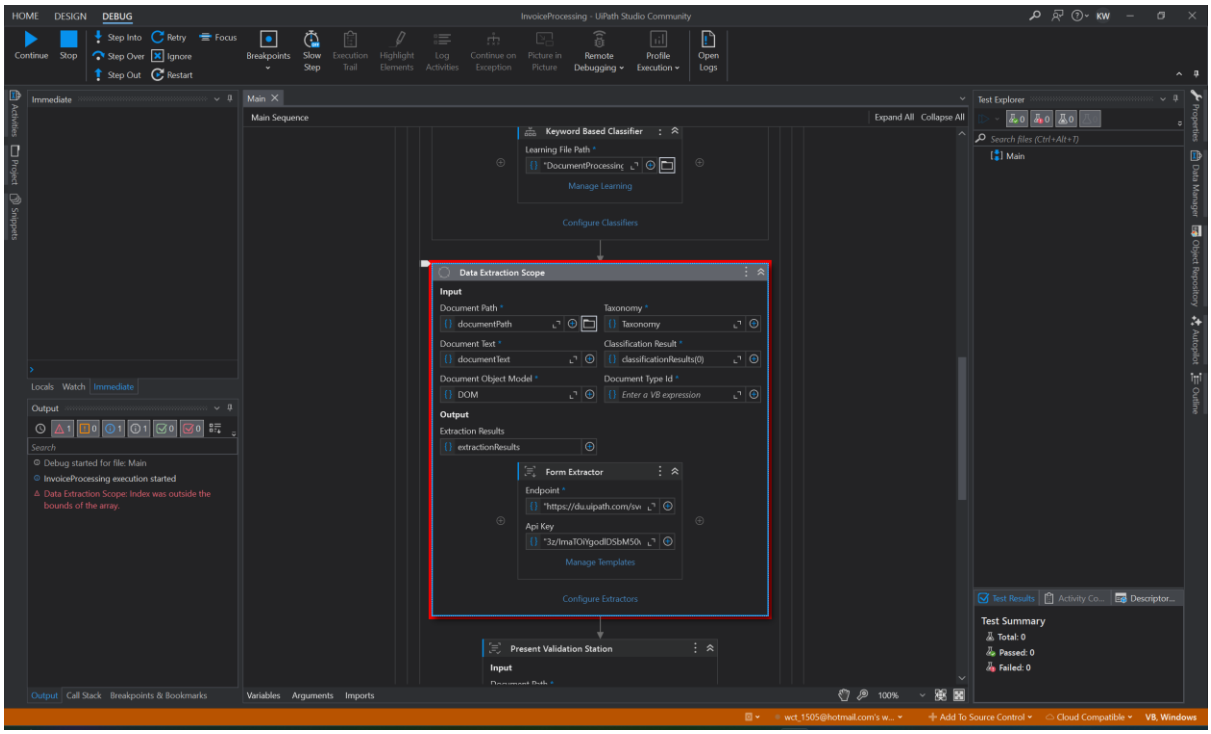
Invoice01 – OCR



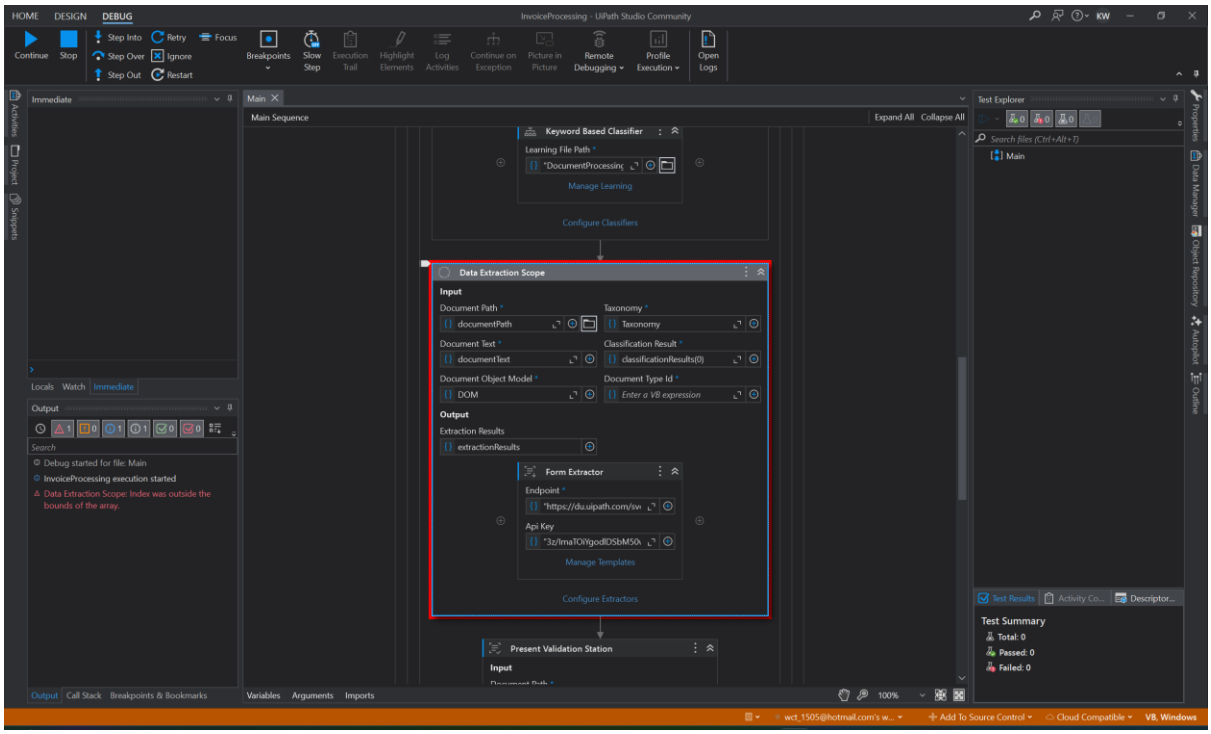
Invoice01 – Extraction



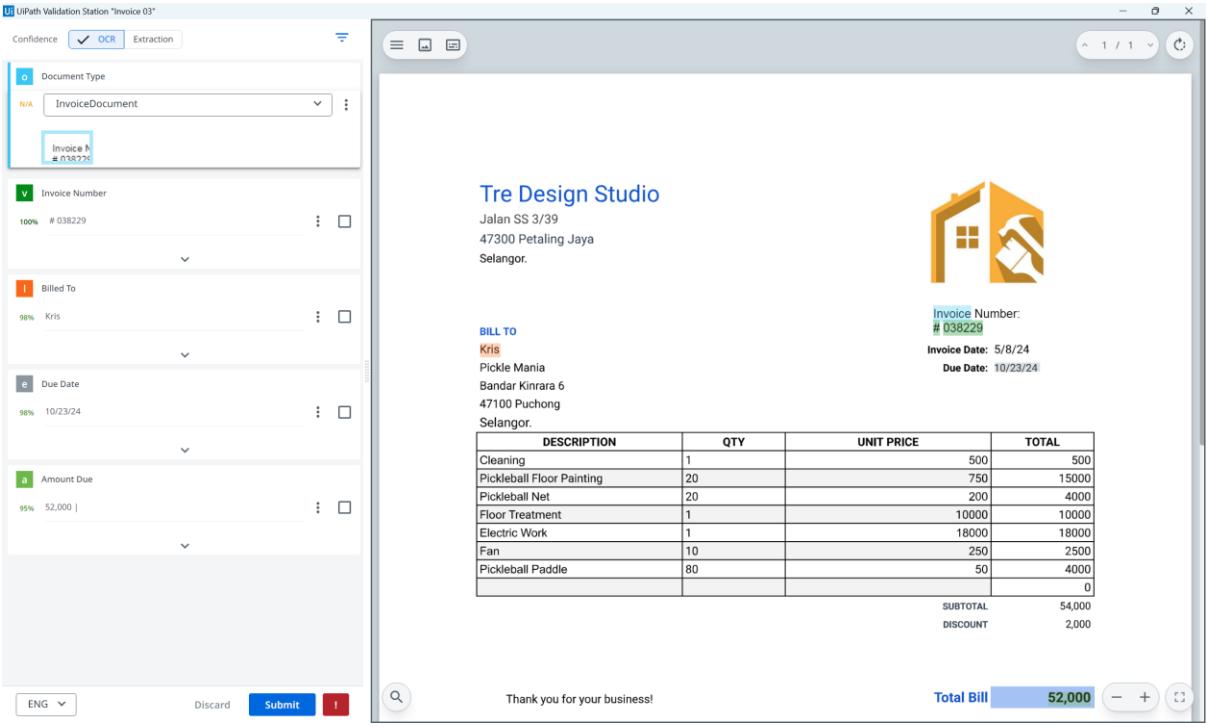
Invoice02 – OCR



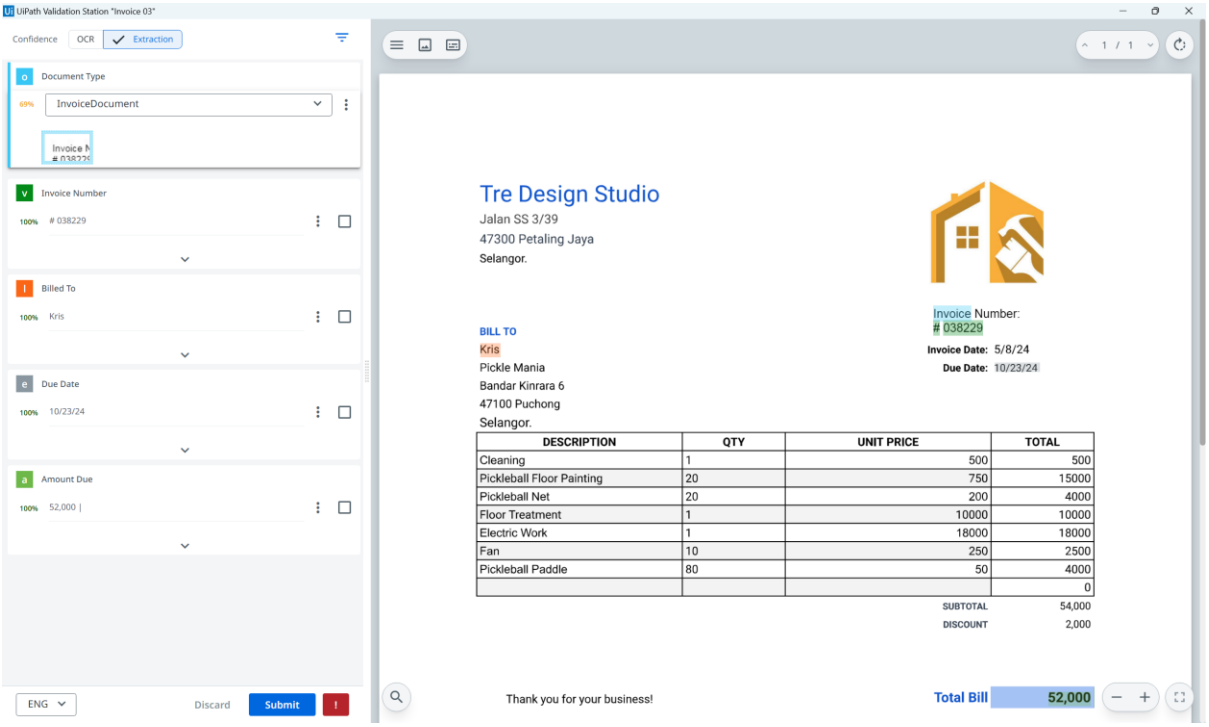
Invoice02 – Extraction



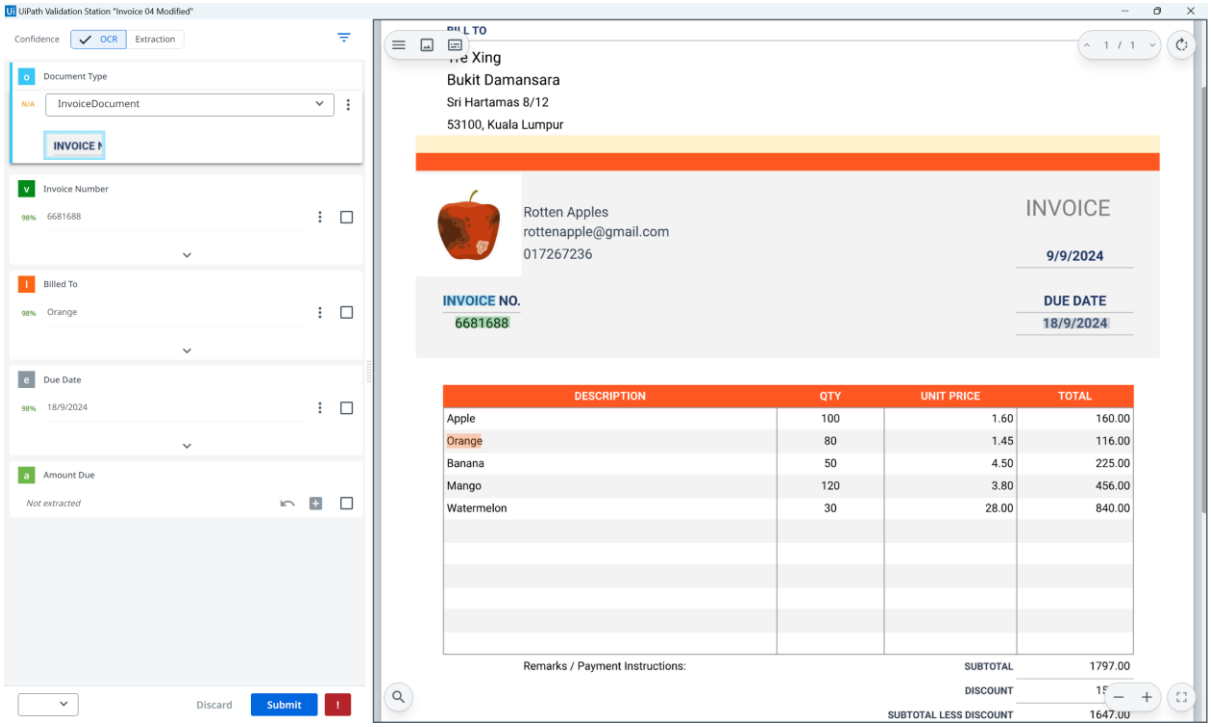
Invoice03 – OCR



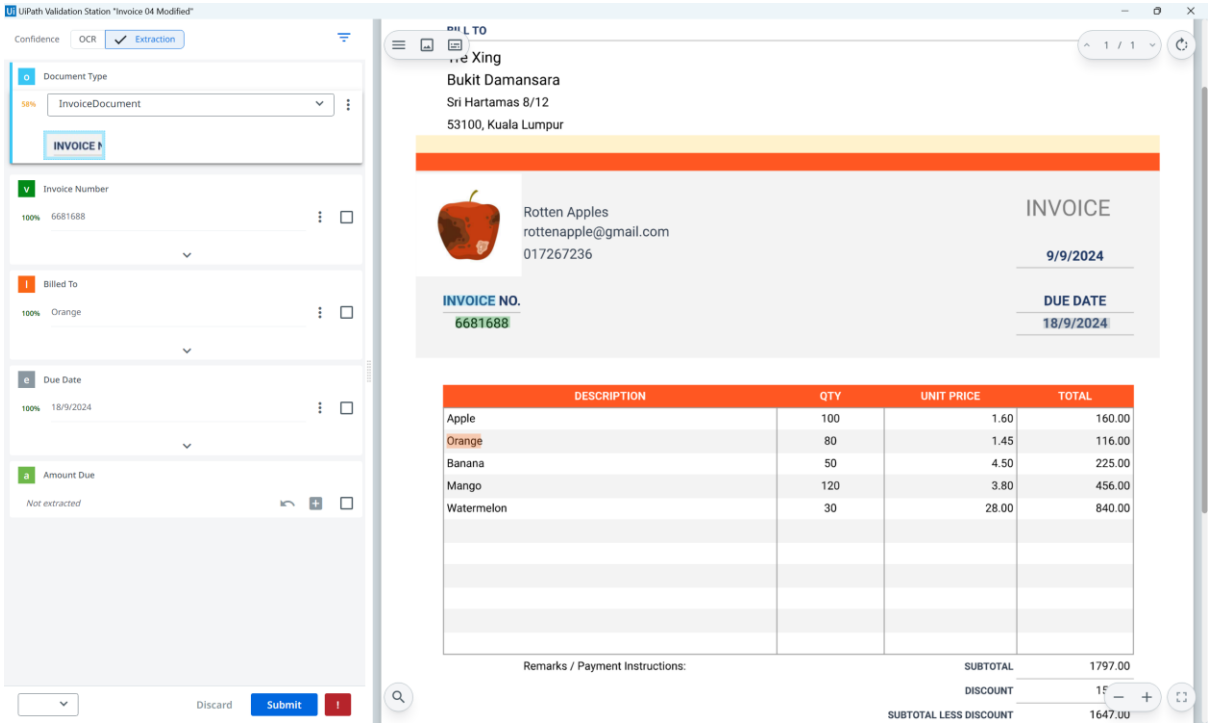
Invoice03 – Extraction



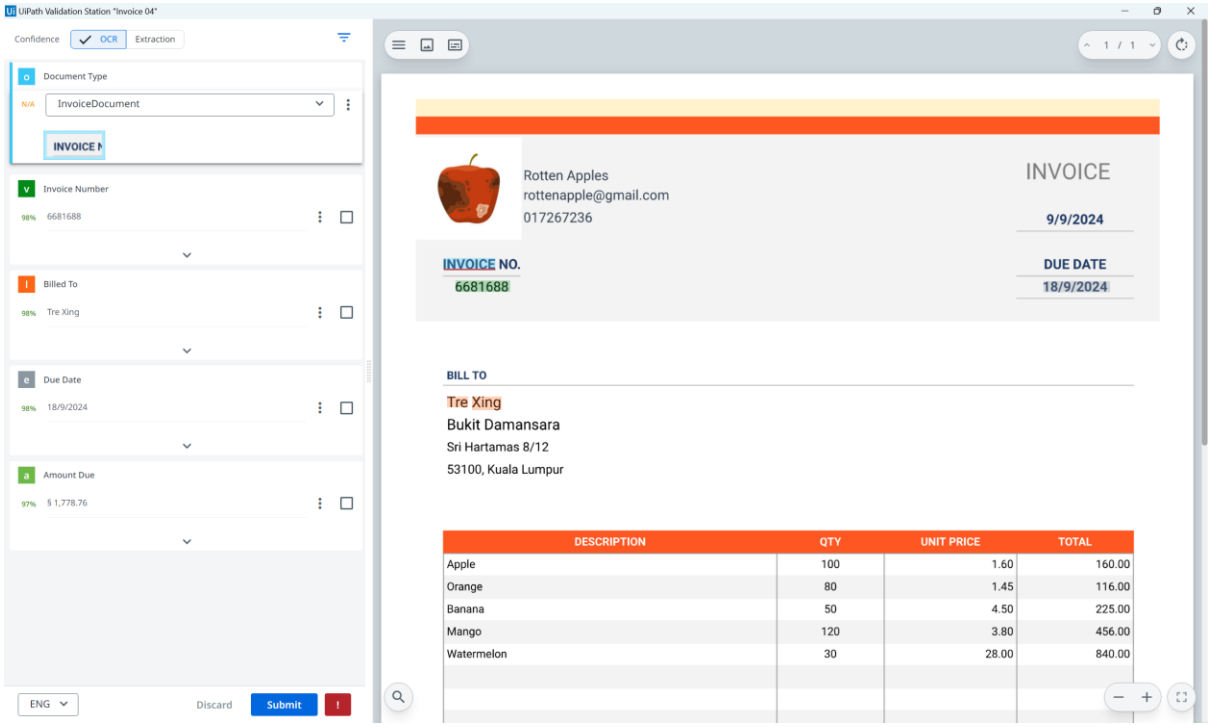
Invoice04 Modified – OCR



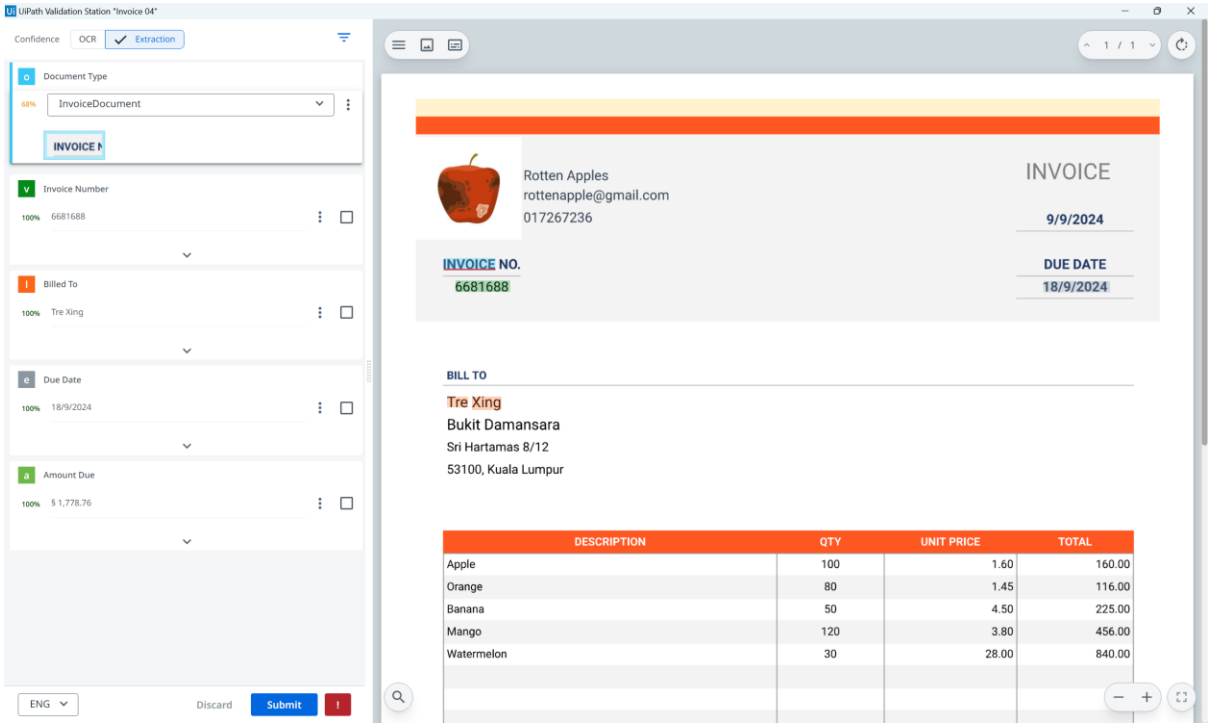
Invoice04 Modified – Extraction



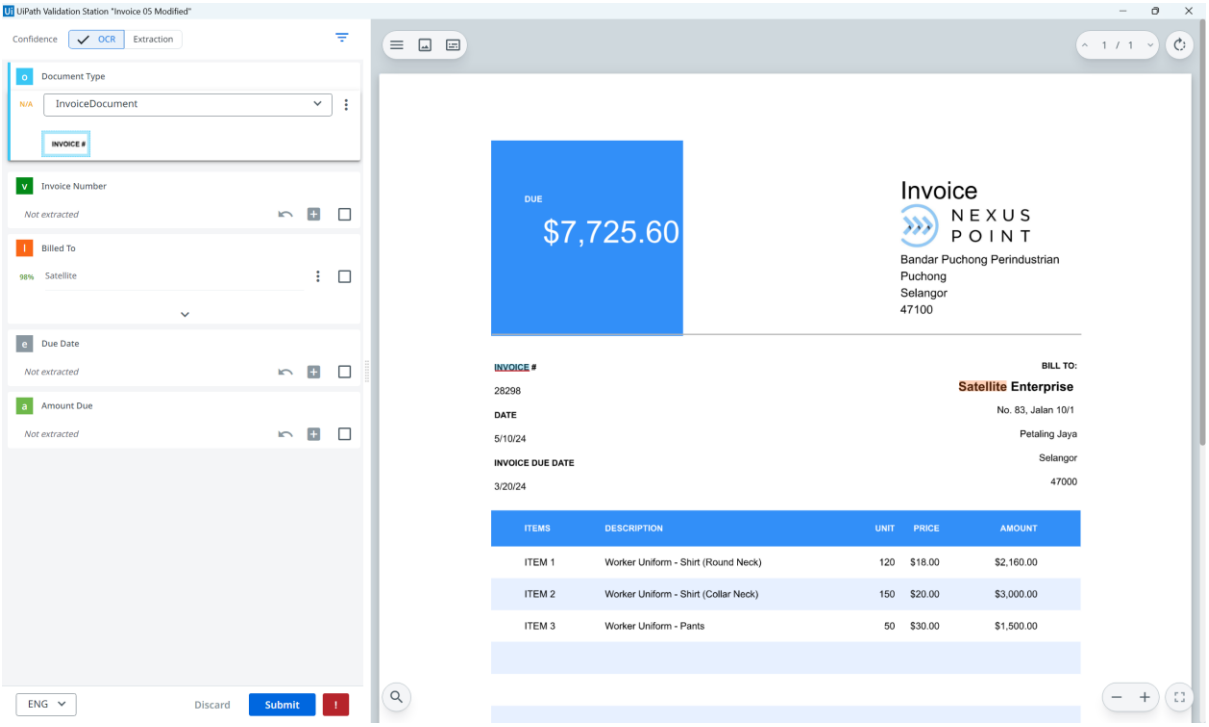
Invoice04 – OCR



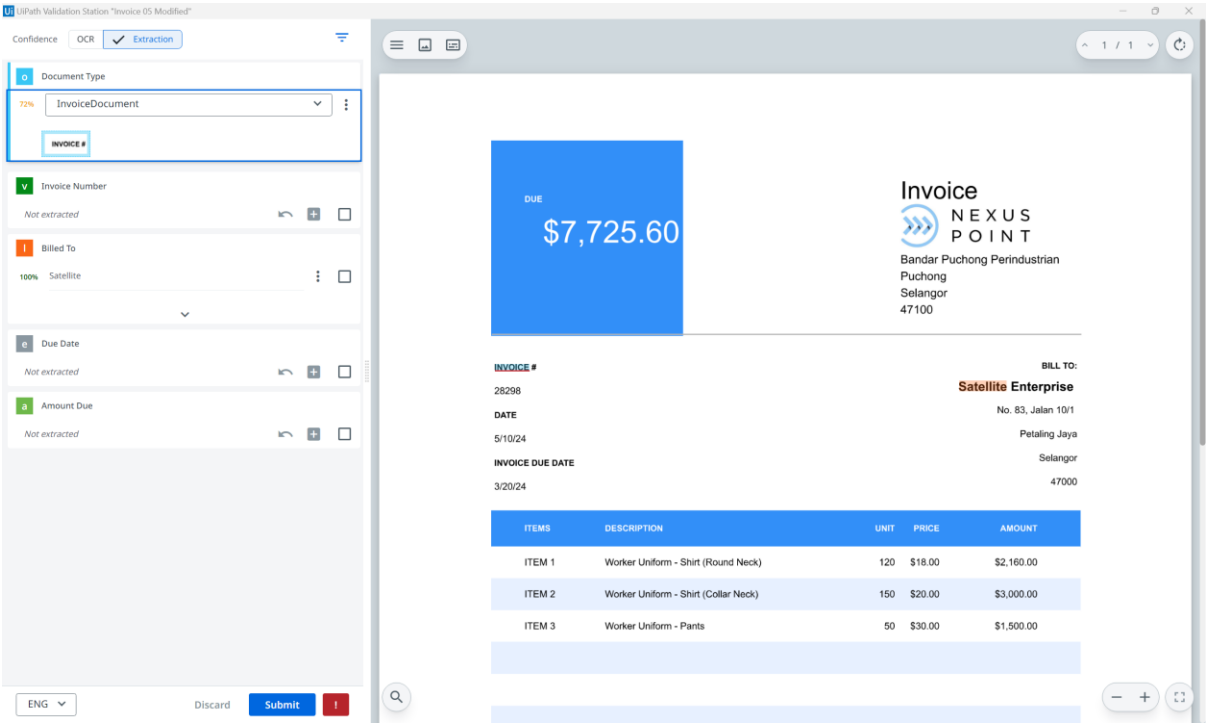
Invoice04 – Extraction



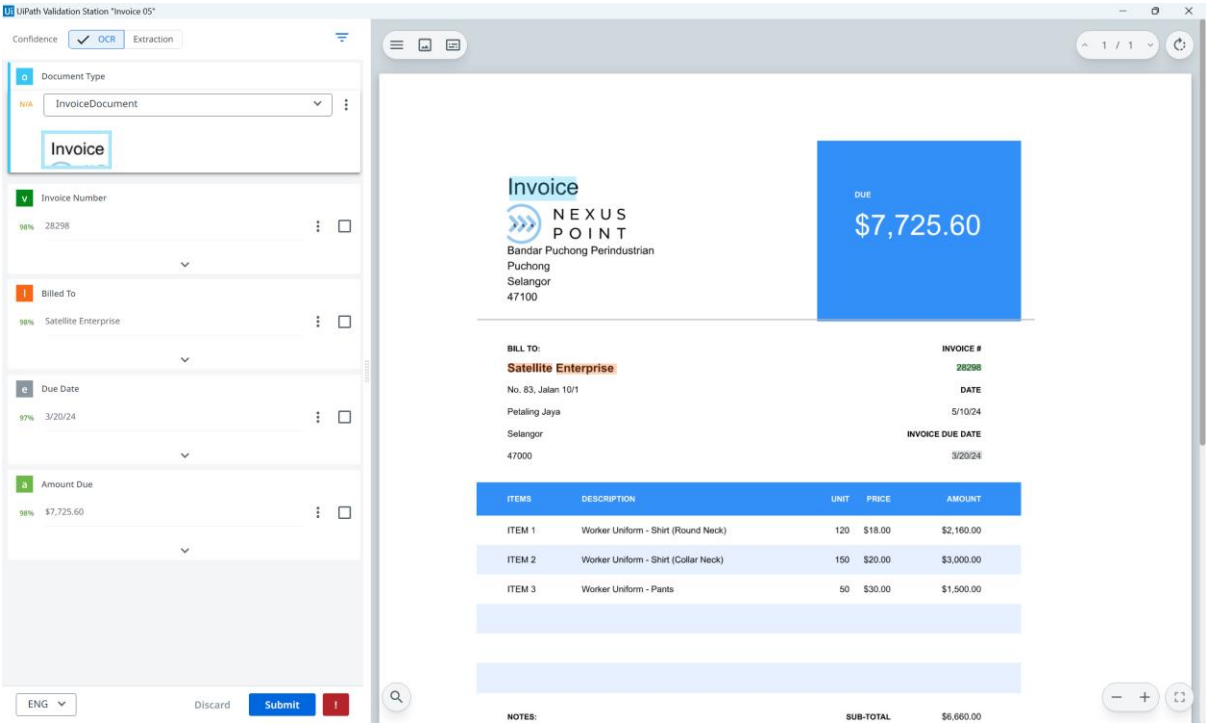
Invoice05 Modified – OCR



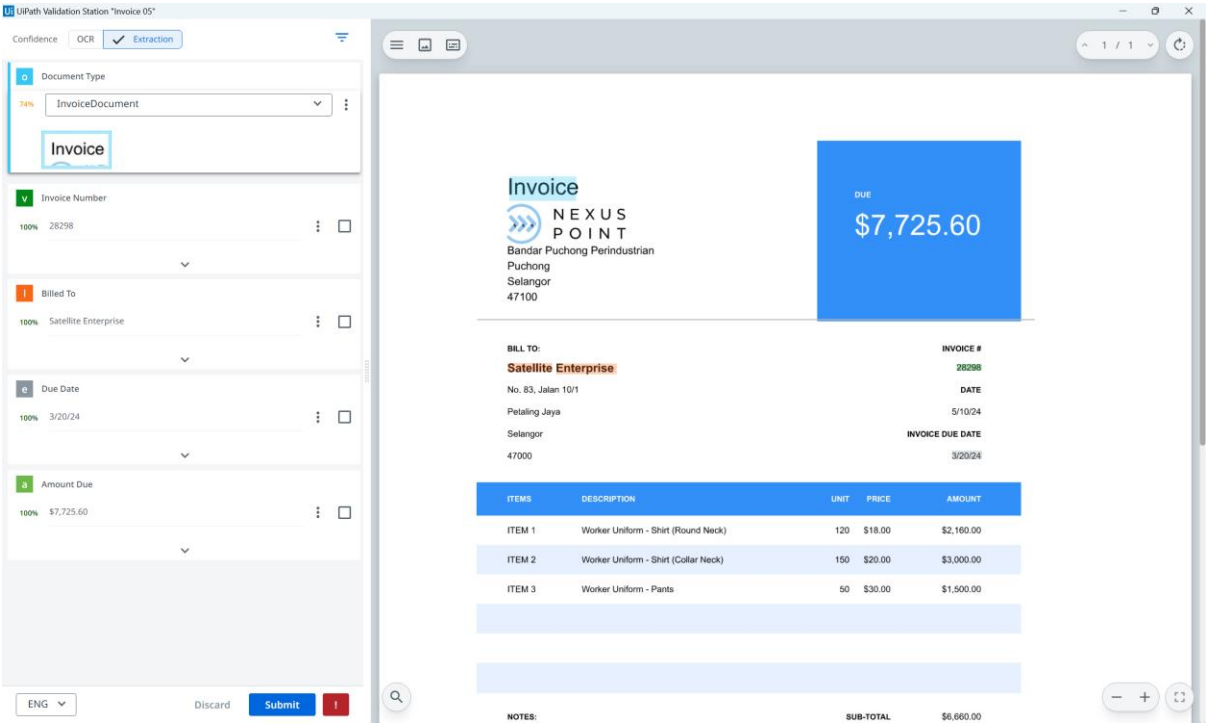
Invoice05 Modified – Extraction



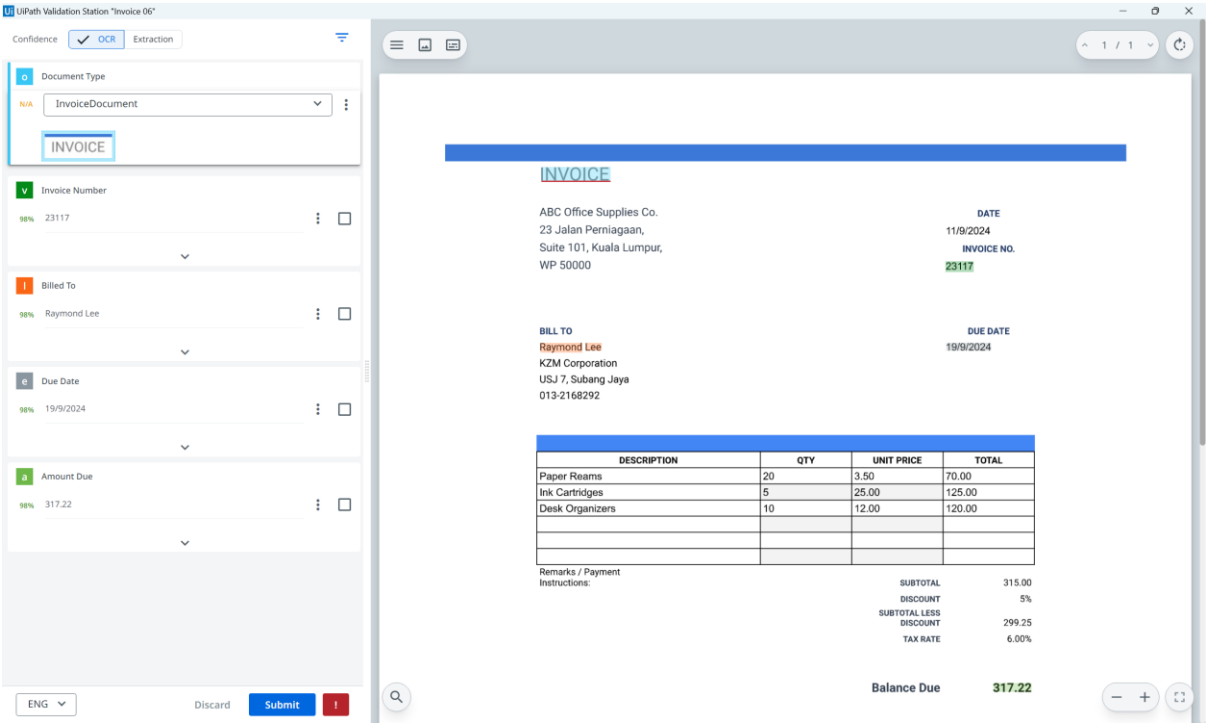
Invoice05 – OCR



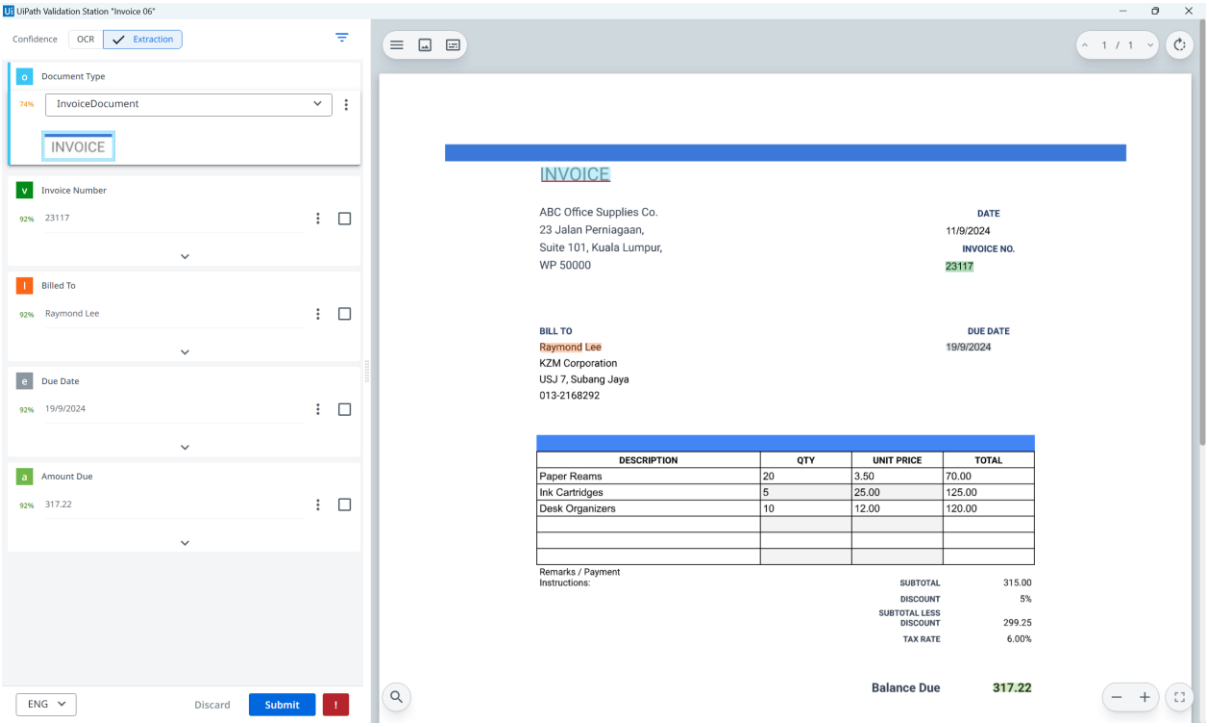
Invoice05 – Extraction



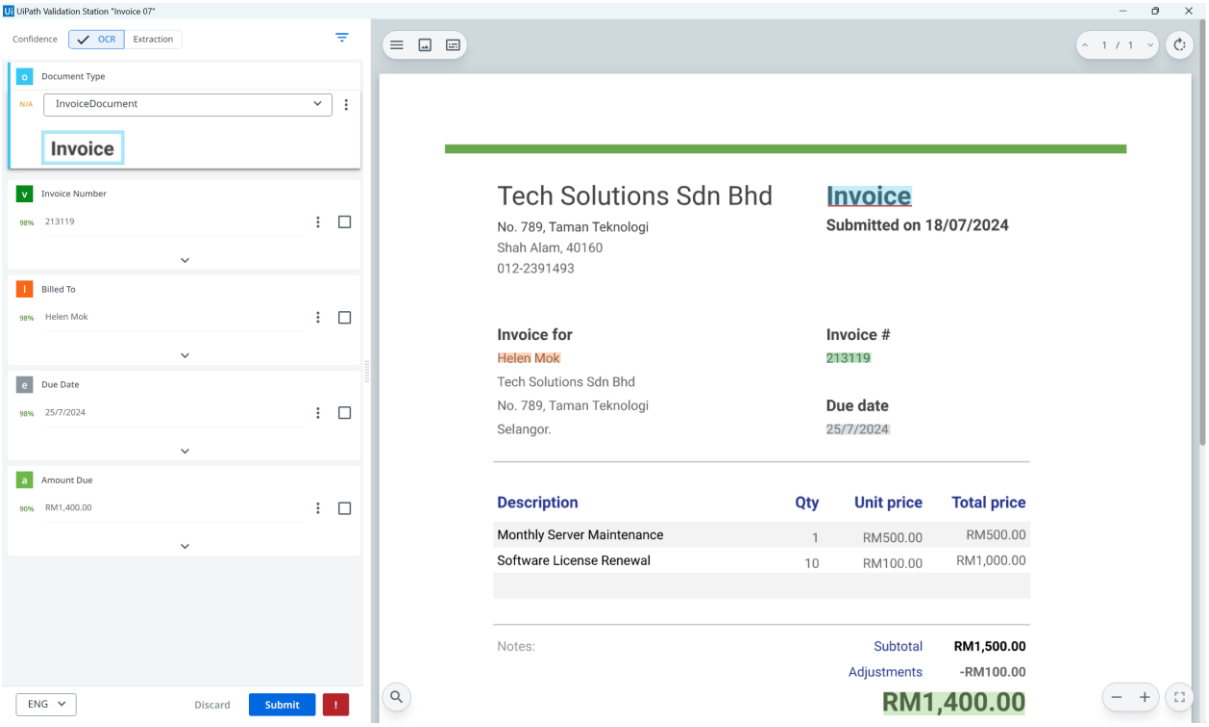
Invoice06 – OCR



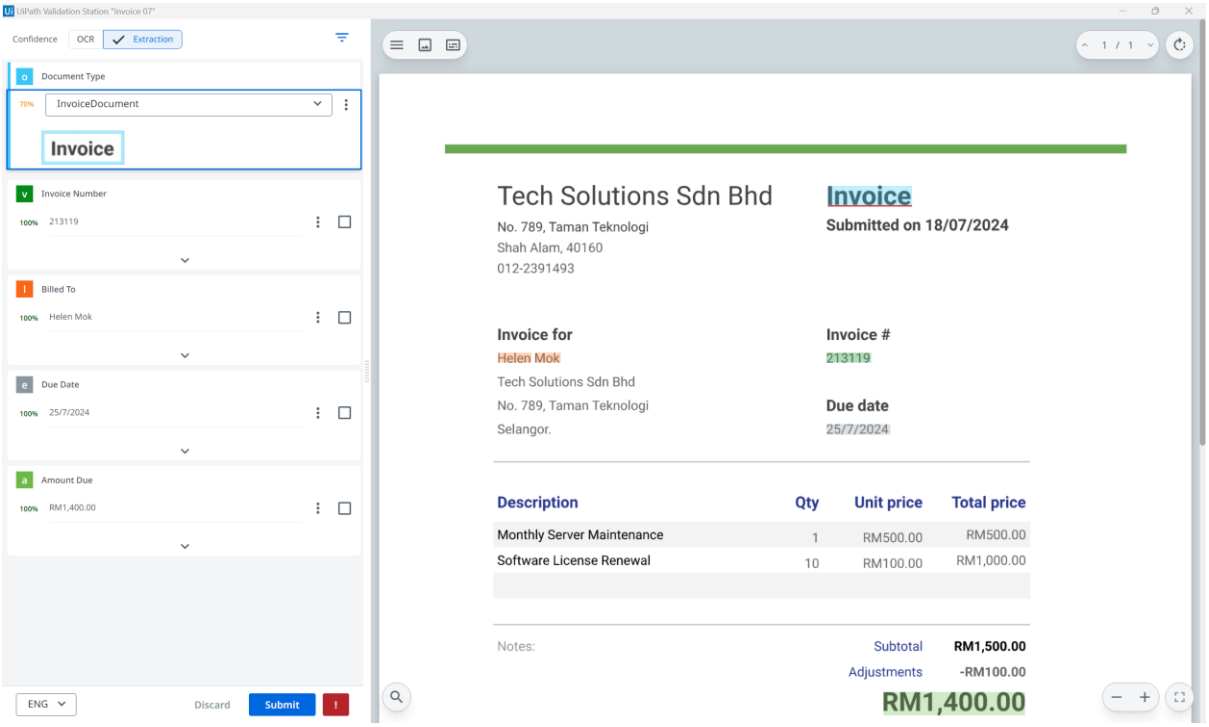
Invoice06 – Extraction



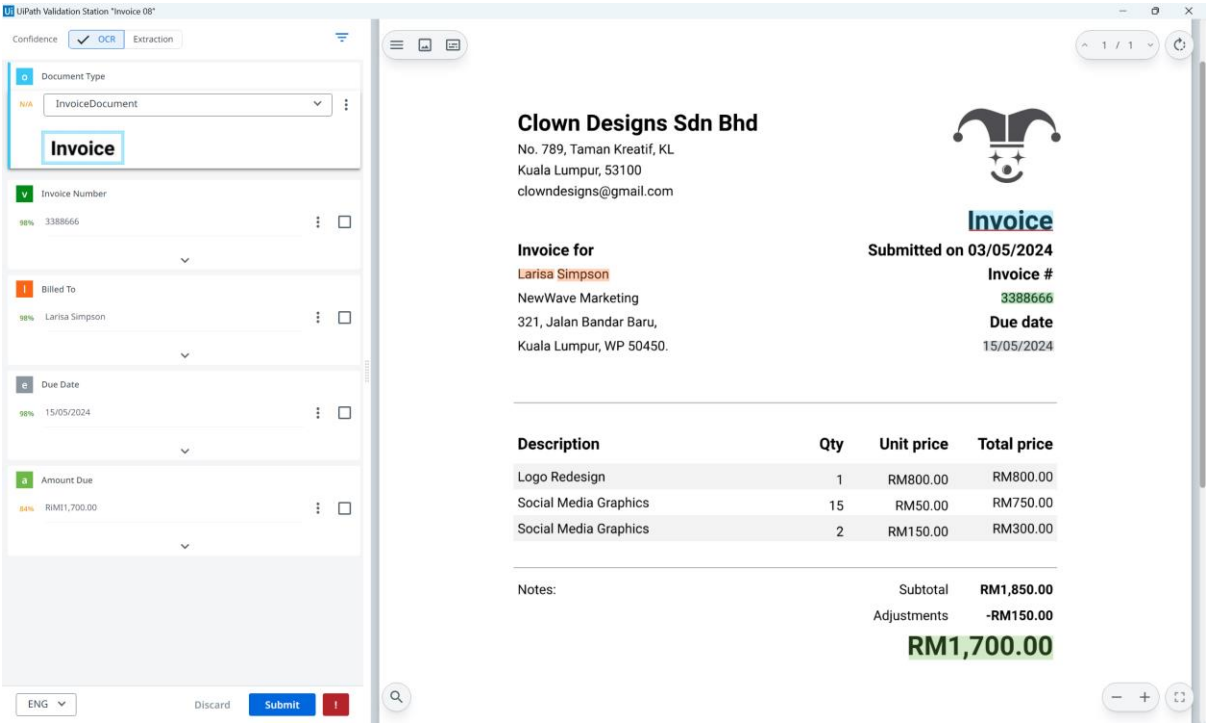
Invoice07 – OCR



Invoice07 – Extraction



Invoice08 – OCR



Invoice08 – Extraction