International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Automated Validation of Korean C-3-9 Tourist Visa Applications using Object-Oriented Programming and Optical Character Recognition (OCR)

Ann Janelle Alano[1], Allen Nicole Bagaporo[2], Bernard Bullicer[3], Lois Renzo Sarmiento[4], Carmelita H. Benito[5], Jesus Paguigan[6]

[1,2,3,4]Graduate Student, College of Computing and Information Technology, Adamson University
[5]Professor, College of Computing and Information Technology, Adamson University
[6]Professor, Adamson University and Eulogio "Amang" Rodriguez Institute of Science and Technoĺogy

Abstract
This research presents a prototype application deployed on Google Cloud Platform (GCP) App Engine to automate the validation process for Korean C-3-9 Tourist Visa applications. By leveraging Object-Oriented Programming (OOP) principles and Optical Character Recognition (OCR) technology, the system efficiently scans PDF application forms to ensure completeness of data required. This automated method will boost applicants' confidence and encourage them to analyze their own applications, which could shorten processing times and increase overall effectiveness. This research offers time-saving solution in validating the required fields especially with the fluctuating numbers of tourist in South Korea in 2023 which is 349,956 (tourism.gov.ph, 2023).Given the potential for widespread use not only for applicants but also for officials, this research also explores various network efficiency techniques for future improvements following cloud infrastructure via GCP. These techniques seek to improve the system's overall efficiency, high performance, security, manageability, and adaptability to various visa kinds.

Keywords: Optical Character Recognition, Object Oriented Programming, Visa application validation system, Google Cloud Platform

1. Introduction

This research focuses on the development of a prototype system aimed at automating the validation process of Korean Visa applications, specifically targeting the C-3-9 Tourist Visa. The prototype makes use of the principles of Object-Oriented Programming (OOP) to efficiently process and validate application forms, which would be in PDF format, to ascertain that the fields required are well completed. The prototype integrates Optical Character Recognition technology; where data from application forms is digitally extracted and checked to see whether the required fields are fully completed. This research presents a promising solution to address the challenges of manual visa validation. By combining automation, OOP principles, OCR technology, and cloud infrastructure, the

proposed system offers a time-saving and efficient solution that benefits both applicants and officials, ultimately contributing to a smoother and more streamlined visa processing experience. To address this issue, this research proposes a novel solution: an automated visa validation system deployed on Google Cloud Platform (GCP) App Engine. This system leverages the power of Object-Oriented Programming (OOP) principles and Optical Character Recognition (OCR) technology to efficiently scan PDF application forms and ensure data completeness.

1.1 Objectives

The primary objective of this research is to develop a prototype system for Korean visa validation process utilizing Object-Oriented Programming (OOP) principles with Optical Character Recognition (OCR) Technology and how it can help with optimizing system's design.

- To design a system applying the principles of OOP
- To check the visa document for data accuracy and data validation
- To automate data validation and accuracy utilizing Optical Character Recognition (OCR)

1.2 Problem Statement

The current process for validating Korean Visa applications, particularly for C-3-9 Tourist visas, is manual and time-consuming, often resulting in incomplete applications and delays in processing. Human error is inevitable, leading to incomplete applications and the need for resubmission, further delaying processing. The manual review process is time-consuming, especially with the fluctuating number of tourist visa applications, as evidenced by the 349,956 applications received in South Korea in 2023 alone (tourism.gov.ph, 2023).

This research aims to address these challenges by developing a prototype system that automates the validation process of Korean Visa applications, specifically for C-3-9 Tourist visas. This system will leverage Object-Oriented Programming (OOP) principles and integrate Optical Character Recognition (OCR) technology to extract and confirm the completeness of all required fields from the PDF application form.

2. Literature/Theoretical Framework

The proposed system is envisioned as a concurrence of these theories, providing the OOP structural backbone, OCR ensures data extraction correctness, and automation drives efficiency. These components are supported by network optimization for scalability and tourism-related frameworks that contextualize the system's relevance to South Korea's visa application processes thus, this framework allows the prototype to address the immediate problem of the completeness of applications while remaining adaptable for broad application in document validation. Optical Character Recognition (OCR) Technology OCR is a pivotal component for automating the data extraction process from PDF forms, ensuring that fields are read and validated accurately.

OCR converts scanned images or documents into machine-readable text by identifying patterns and characters. Advances in OCR algorithms, such as deep learning-based OCR, enhance accuracy even with non-standardized form layouts. OCR enables the system to analyze visa application forms by extracting data from scanned PDF and verifying completeness of fields, thus limiting human errors.

Wang (2023) discuss OCR technology represents the core of modern digitization processes since it provides both immediate practical benefits and potential future development opportunities. Addressing its current limitations through further advancements in machine learning, OCR stands to remain an important instrument for the ongoing digital transformation of industries worldwide.

Prototype extraction and adaptive OCR by Yihong Xu and G. Nagy (1999) This research tackles the challenge of maintaining OCR accuracy when dealing with low-quality scanned documents. Instead of relying on a single, universal OCR system, they propose a way to create custom OCR systems tailored to each specific document. his allows them to create a custom OCR system that works better on that specific document, even if the scan quality is low. Essentially, they are teaching the OCR system to recognize the specific style and imperfections of a particular document, leading to more accurate results.

2.1 Object-Oriented Programming

Object-Oriented Programming (OOP) is a programming paradigm based upon objects having both data and methods that aims to incorporate the advantages of modularity and reusability (Habtamu Keno, 2020).
Thus, the prototype was design using the OOP programming paradigm and can be reuse to other document validating transactions.

2.2 Optical Character Reader (OCR)

Optical Character Reader (OCR) converts scanned images or documents into machine-readable text by identifying patterns and characters. Advances in OCR algorithms, such as deep learning-based OCR, enhance accuracy even with non-standardized form layouts. OCR enables the prototype to analyze visa application forms by extracting data from scanned PDF and verifying completeness of fields, thus limiting human errors.
OCR technology represents the core of modern digitization processes since it provides both immediate practical benefits and potential future development opportunities. Addressing its current limitations through further advancements in machine learning, OCR stands to remain an important instrument for the ongoing digital transformation of industries worldwide (Wang, 2023).
According to Method and System for Accessing Table Content in a Digital Image of the Table [3] with US Patent No. 11,966,758 B2 enables interoperability may be the ability of a computer system or application to work with other computer systems or applications, typically without special effort on the part of the users (Vaindiner A, 2024).
The prototype will be working the same way as the above system but the difference is a standard document will be utilized for easier and structured extraction of information from the C-3-9 visa application document.

3. Methodology

In Figure 1, the conceptual framework that discusses the flow of the usage of the prototype.
The user will have the C-3-9 Tourist Visa application documents filled out and saved in pdf file format. Then it will be fielded-in in the prototype using a button. Once the prototype was able to receive the pdf file; the user will then click another button to scan the pdf file for its completeness and validity.
Messages will then appear on the screen indicating if all fields were properly filled out and if not the prototype will be indicating the field of the documents which were omitted.

Figure 1. Discusses the flow of the usage of the prototype through a conceptual framework

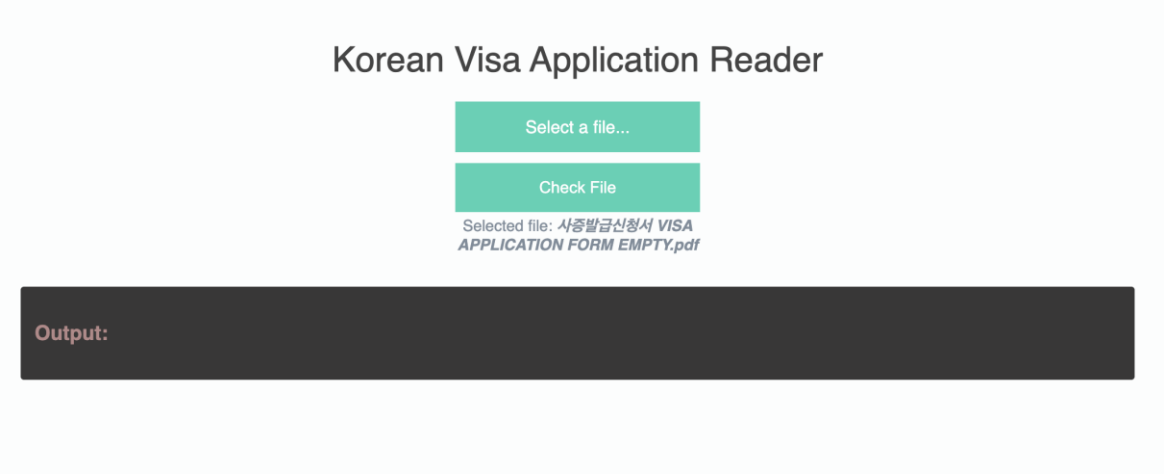

Figure 2. User Interface after selecting the PDF file

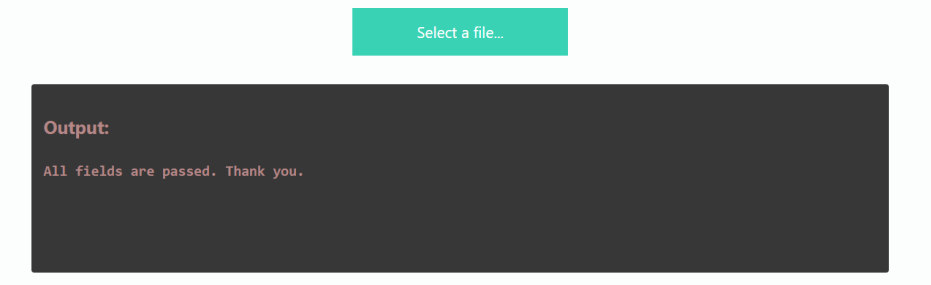

Figure 3. User Interface showing successful return message that fields are validated and passed

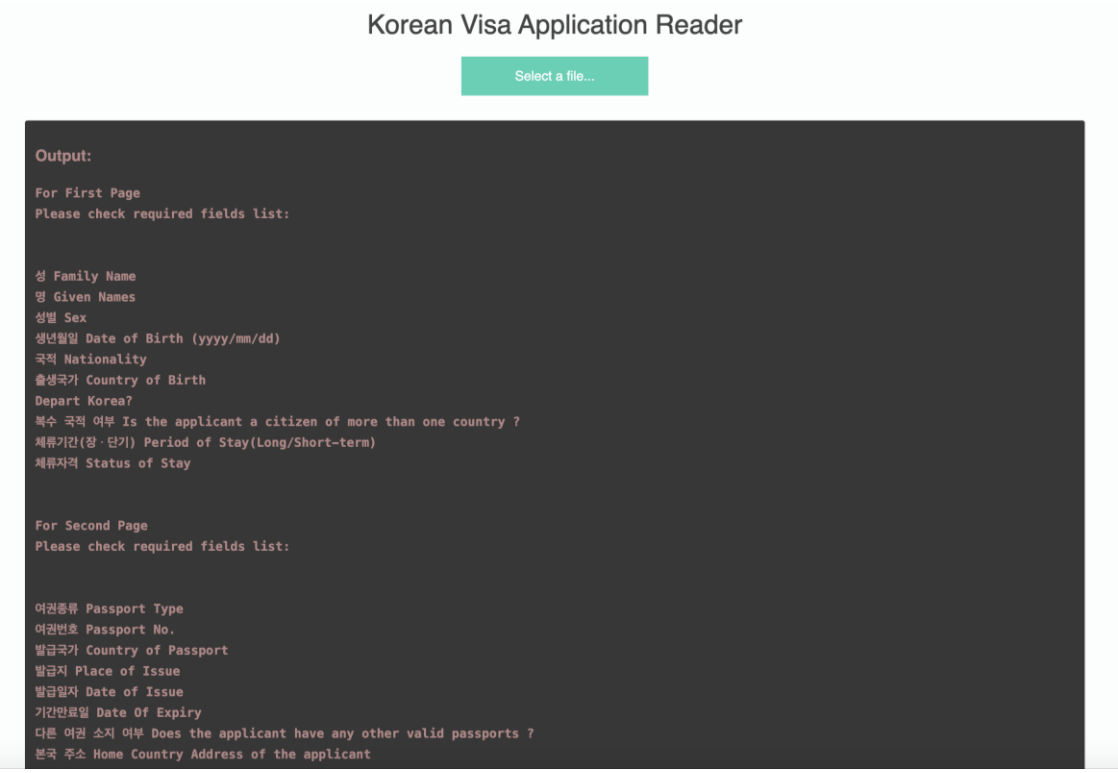

Figure 4. User Interface showing display message if the uploaded file has incomplete details filled out or has empty details

4. Discussion of Results

The prototype of the visa application validation system demonstrated promising results, aligning with the proposed methodology and addressing the identified needs of visa applicants.

- User Friendliness: A significant majority (80%) of respondents found the system easy to use, indicating a successful user interface design.
- Validation Accuracy: While the system achieved a 60-50% accuracy rate in validation, it's important to note that some users may have uploaded incorrect PDF files, affecting the overall accuracy.
- Field Validation: A substantial 80% of respondents reported that the system accurately validated the required fields, demonstrating its effectiveness in identifying and verifying essential information.
- Future Application: All respondents (100%) expressed confidence in the system's potential to aid future visa applications, highlighting its perceived value.

Overall, the prototype's performance, coupled with its user-friendly interface and accurate field validation, suggests that it can significantly streamline the visa application process and reduce the risk of errors and

rejections.

4.1 Summary of Findings

Below are the key findings based on user feedback

- **User Interface:**
  - Users expressed a desire for a more intuitive and user-friendly interface, especially for those unfamiliar with complex PDF forms.
  - The current system's limitations in handling various visa types and specific requirements were highlighted.
- **Technical Performance:**
  - Concerns were raised about the system's security measures, particularly in protecting sensitive personal information.
  - Scalability issues were identified, especially during peak usage periods.
  - Performance optimizations were suggested to improve response times and overall system efficiency.

4.2 Recommendation:

- **Enhance User Interface:**
  - Intuitive Design: Implement a more intuitive and visually appealing interface that guides users through the application process with clear instructions and visual cues.
  - Error Handling and Validation: Improve error handling mechanisms to provide clear and actionable feedback to users, preventing data entry mistakes and ensuring accurate submissions.
  - Accessibility: Ensure the system is accessible to users with disabilities by adhering to web accessibility standards.
- **Strengthen Security:**
  - Data Encryption: Implement robust encryption techniques to protect sensitive personal information both during transmission and storage.
  - Secure Authentication: Employ strong authentication methods, such as multi-factor authentication, to safeguard user accounts and prevent unauthorized access.
  - Regular Security Audits: Conduct regular security audits and vulnerability assessments to identify and address potential security risks.
- **Improve Scalability and Performance:**
  - Load Balancing: Implement load balancing techniques to distribute traffic across multiple servers, optimizing resource utilization and preventing system overload.
  - Performance Optimization: Continuously monitor system performance and optimize database queries, network communication, and other critical components.
  - Caching Strategies: Employ caching mechanisms to reduce server load and improve response times for frequently accessed data.
- **Expand System Functionality:**
  - Support Multiple Visa Types: Extend the system's capabilities to handle a wider range of visa types, including tourist, business, student, and work visas.
  - Integrate with Document Verification: Integrate with document verification services to automate the verification process and reduce manual effort.

4.3 Final Remarks

By addressing the identified issues and implementing the proposed recommendations, the Korean Visa

Application PDF Reader can significantly enhance user experience, improve system security, and optimize performance. This will ultimately lead to a more efficient and reliable visa application process for both applicants and processing authorities.

5. Bibliography/ References/Appendices:

1. "THE PHILIPPINES RECORDS 1.2M INTERNATIONAL TOURISTS IN FIRST TWO MONTHS OF 2024" https://beta.tourism.gov.ph/news_and_updates/the-philippines-records-1-2m-international-tourists-in-first-two-months-of-2024/#:~:text=In%20terms%20of%20international%20visitor,28.50%20percent%20of%20the%20total%2C
2. Junmiao W., "A Study of The OCR development History and Directions of Development" https://www.researchgate.net/publication/377803605_A_Study_of_The_OCR_Development_History_and_Directions_of_Development
3. Habtamu K., Object Oriented System Analysis and Design (OOSAD). 2020 (Online) http://ndl.ethernet.edu.et/bitstream/123456789/90366/3/Revised%20OOSAD%20Module2020.pdf
4. Vaindiner A. et al. "Method and System for Accessing Table Content in a Digital Image of the Table," US 11966758 B2. [Online]. Apr 23, 2024. Available: https://patents.google.com/patent/US11966758B2/en?oq=us11966758