DOI: 10.2478/ijasitels-2025-0012

Optimization of Romanian Identity Documents Processing

Delia-Victoria Milașcon[1], Constantin Constantinescu [2], Raluca Brad [3]

[1,2,3]Lucian Blaga University of Sibiu, Department of Computer Science and Electrical Engineering

Abstract

This paper presents the development of an application based on computer vision and deep learning aimed at the automatic recognition of data from Romanian national identity documents. Its primary goal is to optimize the processing of such documents within public and private institutions. Developing such a system involves several challenges. These include the variability in the visual appearance of documents caused by wear and tear, scanning angles, camera quality, and ambient lighting conditions; the difficulty of accurately recognizing Romanian diacritics; and the correct identification of textual fields based on their positioning within the document. The system must be sufficiently robust to adapt to these variations while maintaining a high level of accuracy.

Keywords: optical character recognition, ID cards, passport, identification

1 Introduction

The processing of identity documents in Romania, encompassing optical character recognition (OCR), data extraction, authenticity verification, and integration with digital services, has undergone significant transformation as of November 2025. Driven by EU regulations like eIDAS 2.0 [1] and national initiatives under the Recovery and Resilience Plan (PNRR) [2], optimizations focus on speed, accuracy, security, and fraud prevention. These advancements address rising fraud (up 7% year-on-year in card transactions [3]) and low digital service adoption (24% for online government services [4]). Key enablers include AI-enhanced OCR, biometric chips, and modular digital ID platforms, reducing processing times from minutes to seconds while achieving over 99% accuracy in data entry [5].

Recent advances in document layout analysis, powered by deep learning, have transformed OCR-based document processing from a theoretical concept into a practical reality. Central to this progress is DocLayNet [7], a dataset comprising over 100,000 annotated pages across invoices, forms, and identity documents. With 11 semantic classes, including text blocks, titles, tables, and figures models, trained on this corpus learn to interpret documents as structured hierarchies rather than flat pixel arrays. Transformer-based architectures like LayoutLMv3 [8] with various fine tunings [9] have achieved mean Average Precision scores exceeding 0.90 on benchmark datasets of 5000 anonymized eID scans collected during a 2024–2025 pilot.

However, standard models encounter challenges when applied to Romanian identity documents, such as the standard ID or even the CEI (eID), which feature bilingual text

(Romanian and English), variable font styles, and densely packed fields. These issues are present for other countries as well. Nguyen-Trong [10] developed a complete system for Vietnamese ID card information extraction with multiple deep learning models, while authors of this paper [11] propose and algorithm for Indonesian ID cards. And the list can go on.

Accurate field detection alone is insufficient. Mobile-captured images often suffer from perspective distortion, motion blur, or poor lighting. To mitigate this, a rectification technique must be incorporated. This system [12] proposes an unsupervised geometric rectification technique, which reconstructs frontal, flat views of documents using detected text lines and border cues. This preprocessing reduces the Character Error Rate in downstream OCR from 4.2% to 0.8%.

Another major issue is the practical deployment of the solutions. Custom-trained architectures face substantial barriers: computational overhead requiring GPU infrastructure, extensive labeled training datasets (often 5,000+ documents per document type), complex fine-tuning workflows, and integration challenges with existing institutional systems. Some real-time applications exist, for example the RT-DETR detector [13], a transformer-based architecture optimized for speed, processes both sides of an ID card in under 50 ms on a mid-range smartphone GPU, enabling deployment in mobile applications like ROeID.

Building on all of the above, we propose a solution that achieves great results in localizing critical fields Romanian IDs, including the CNP (13-digit personal identifier), full names with diacritics, addresses aligned with SIRUTA codes, and validity dates.

This paper presents an intelligent hybrid deep learning architecture specifically engineered for automated processing of Romanian identity documents. Rather than training custom models from scratch, our system optimally configures and orchestrates state-of-the-art pre-trained deep learning engines—PaddleOCR's DBNet [14] for neural text detection and CRNN [15] for character recognition, complemented by Tesseract's LSTM-based [16] enhancement for Romanian diacritics—through sophisticated spatial relationship analysis and domain-specific parameter optimization. The system addresses key challenges in Romanian document processing: visual variability from wear and lighting conditions, accurate diacritic recognition (ă, â, î, ș, ț) essential for legal validity, field localization within complex layouts, and comprehensive validation through Machine Readable Zone (MRZ) integration.

The developed architecture achieves great accuracy on both identity cards and passports through intelligent configuration of complementary deep learning models, advanced geometric preprocessing, and multi-layer validation mechanisms. Processing completes in under 3 seconds on standard institutional hardware without requiring GPU infrastructure, custom model training, or cloud dependencies. Our results demonstrate that strategic optimization of pre-trained deep learning models—when engineered for domain-specific requirements—can achieve accuracy comparable to custom-trained architectures while offering the possibility of immediate deployment and minimal resource requirements, providing a production-ready solution suitable for Romanian institutional environments.

2 The Proposed System Architecture

The developed application is a complete automatic recognition system designed to extract relevant data from Romanian official identity documents. It aims to automate data entry and digitalization by processing document images and generating pre-filled official forms.

The processing pipeline is shown in figure 1 and includes:

1. Image Upload – Users load an image of the document (scanned or photographed).
2. Document Detection – The system isolates the document's contour within the image, correcting distortions.
3. Orientation Correction – It tests four possible rotations (0°, 90°, 180°, 270°) using OCR confidence and Romanian keywords to determine the correct orientation.
4. Image Preprocessing – Noise removal, contrast enhancement, and adaptive binarization prepare the image for text recognition.
5. Text Recognition (OCR) – A hybrid approach uses PaddleOCR for speed and Tesseract for refinement—especially improving Romanian diacritics.
6. Data Association and Validation – Extracted text is mapped to document fields, verified through pattern recognition (CNP, serial number, etc.), and validated for format correctness.
7. User Review and Document Generation – The user can manually correct data before the system automatically generates pre-filled Word documents using the extracted information.

The development of the system was based on a selection of technologies, each chosen for its specific characteristics that contribute to the overall functionality of the application. Python 3.12.6 was selected as the main programming language due to its rich ecosystem of libraries for image processing and machine learning.

CustomTkinter was chosen for developing the graphical user interface. This modern library, built on top of the traditional Tkinter, offers contemporary visual components with a professional look. CustomTkinter allows for the creation of an intuitive interface that adheres to modern design principles while providing the necessary flexibility for displaying and editing complex data extracted from documents.

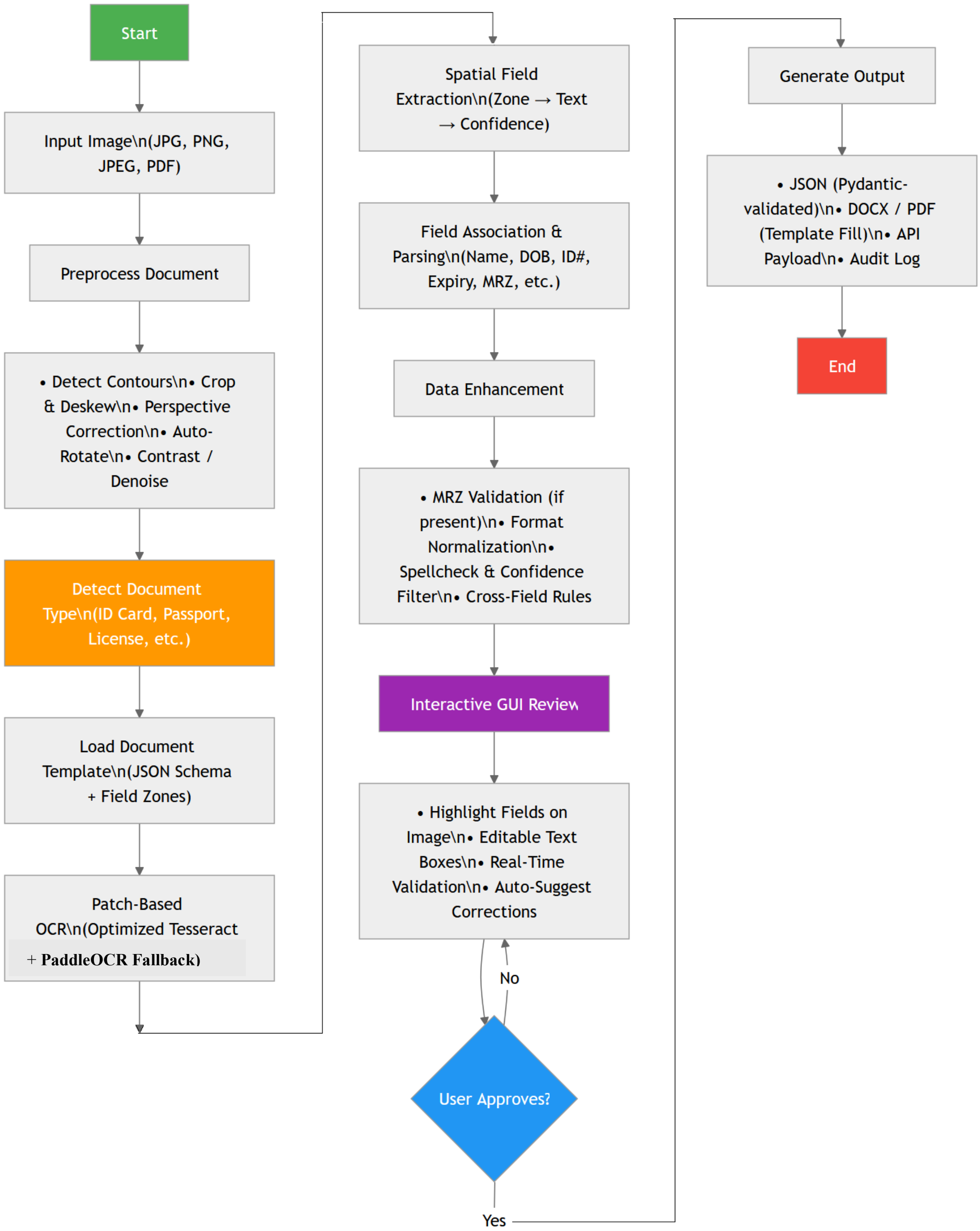

Figure 1. Processing flowchart

For image processing, the application relies on OpenCV, the reference library in the field of computer vision. OpenCV provides optimized implementations for fundamental image processing operations. Complementarily, PIL/Pillow is used for high-level image manipulation and for integration with the Tkinter graphical interface.
The central OCR component is provided by PaddleOCR, a modern framework that delivers excellent performance for text recognition in multiple languages, including Romanian. PaddleOCR uses deep neural network architectures for text detection and recognition, offering an optimal balance between accuracy and speed. To refine the

results, particularly for characters with diacritics, the system also integrates Tesseract OCR, the open-source OCR engine originally developed by HP and currently maintained by Google.
NumPy forms the foundation for efficient numerical operations, being extensively used in pixel matrix processing and in calculations required for various transformations. For document generation, the application uses python-docx, a Python library that allows programmatic manipulation of Microsoft Word documents while preserving the formatting and structure of templates.

3 Image Pre-processing Steps

The process of detecting a document within an image represents the first major technical challenge of the system. In practice, users may provide images in which the document appears under various conditions: it may be photographed on a complex background, occupy only a portion of the image, or be captured at an oblique angle. The function detect_document_contour() implements a robust algorithm that automatically identifies and extracts the document from this variable context.
The algorithm begins by converting the image to grayscale, an essential step that reduces computational complexity and removes color variations that are irrelevant for contour detection. This conversion simplifies the image from three color channels (BGR) to a single intensity channel, while preserving essential information about edges and contrasts.
The next step involves applying a Gaussian filter [17] to reduce noise. This filter is essential because image noise can generate false contours that interfere with the correct detection of the document edges. A 5x5 Gaussian kernel performs smoothing that preserves important edges while eliminating minor intensity variations. Actual edge detection is performed using the Canny algorithm [18], applied with three different threshold sets to cover a wide range of image conditions. This multi-threshold approach ensures robust detection in the presence of varying contrast and lighting.
For each configuration, the algorithm applies morphological operations to connect fragmented edges. Dilation followed by a morphological closing operation helps form complete contours even when the initial detection produces broken edges. The evaluation of detected contours represents the most sophisticated aspect of the algorithm. For each identified contour, the system calculates a score based on multiple criteria reflecting the likelihood that the contour corresponds to an identity document:

- Aspect ratio criterion evaluates how close the detected shape is to the standard proportions of a Romanian ID card (85.6 mm × 54 mm). Contours with aspect ratios near this value receive higher scores.
- Rectangular approximation criterion checks how well the contour can be approximated by a polygon with few vertices. A real document should be approximately rectangular.

After selecting the optimal contour, the algorithm adds generous margins (10% of the document’s dimensions) to ensure that no important content is lost during cropping. This conservative approach prioritizes preserving the integrity of the information over achieving a perfectly tight crop. Figure 2 shows a Romanian ID card in different phases of preprocessing.

Figure 2. Document detection and extraction from an image. Original image (1); Document detection (2); Edge detection (3, 4, 5); Final result (6)

4 Optical Character Recognition Optimizations with Tesseract

The development of optimizations for Tesseract has undergone a significant evolution, culminating in the current implementation that prioritizes the preservation of detail over aggressive processing. This section describes the refined approach that maximizes Tesseract's performance for Romanian documents.

Early versions of the system implemented a complex preprocessing pipeline that included intensive bilateral filtering, adaptive histogram equalization (CLAHE), morphological opening and closing operations, and even image sharpening techniques. While theoretically sound, this approach produced disastrous results in practice.

The fundamental problem was that each processing step, although improving the visual appearance of the image for the human eye, actually degraded the information useful for OCR. Excessive sharpening transformed diacritics into unrecognizable artifacts. Aggressive morphological operations incorrectly merged or split character components. Extreme binarization completely eliminated the subtle nuances necessary to distinguish between similar characters.

The final implementation adopts the principle of elegant simplicity. Rather than attempting to "enhance" the image, the system focuses on providing Tesseract with a representation that is as faithful as possible to the original text, with only the minimal cleaning necessary to remove obvious noise. This minimalist approach not only produces superior results but also reduces processing time and code complexity. Removing unnecessary preprocessing steps makes the system more robust and predictable.

The system implements a decision algorithm designed to select the optimal OCR result by intelligently combining outputs from two complementary engines: the primary OCR module and Tesseract. This mechanism goes beyond a simple "best match" approach,

instead using multiple criteria to evaluate the quality of the extracted text and determine which version should be retained for each field.
At the core of this decision process is the diacritics criterion. Romanian documents often contain characters with diacritics (ă, â, î, ș, ț), which are critical for accurate recognition of names, locations, and other key fields. The system counts all variants of diacritics in both the primary OCR and Tesseract outputs, considering uppercase, lowercase, and older cedilla-based forms. This allows the algorithm to prioritize results that preserve these essential linguistic details.
Complementing the diacritics analysis is a similarity criterion. To avoid selecting results that deviate substantially from the original, the algorithm calculates textual similarity between the two outputs. Both texts are normalized to lowercase, ensuring that differences in capitalization do not influence the comparison. This step prevents the system from mistakenly favoring a result that introduces unrelated or erroneous text.
The decision logic itself follows a clear hierarchy. Results from Tesseract that contain more diacritics than the original are given the highest priority, reflecting the importance of preserving linguistic accuracy. Secondary consideration is given to outputs with at least 60% similarity to the original and a diacritic count equal to or higher than that of the primary OCR. For certain fields, such as residence, additional rules apply: if both outputs have the same number of diacritics, the system selects the more complete result, recognizing that Tesseract may capture extra details—such as multi-line addresses—that the primary OCR engine might miss.
To maintain high data quality, the system incorporates several safety validations. Short text segments (fewer than two characters), text resembling MRZ zones, segments containing noise characters, or outputs with similarity below 50% are automatically rejected. This ensures that the algorithm does not introduce errors while attempting to improve results.
The decision-making process is further refined with field-specific handling. Different types of fields have unique requirements: for instance, the place of birth field uses a relaxed similarity threshold and employs fallback logic for detecting counties, while the residence field is processed with multi-line support and additional address-specific validations.
Finally, the algorithm includes an implicit safety mechanism. In cases where none of the improvement criteria are clearly met, the original result from the primary OCR engine is retained. This conservative approach prevents quality degradation in ambiguous scenarios, ensuring that the system avoids introducing new errors while striving to enhance the accuracy of extracted information.
Through this hybrid, criterion-driven process, the system achieves a careful balance between precision and reliability, leveraging the strengths of Tesseract for fine-grained details such as diacritics, while maintaining the robustness and consistency of the primary OCR engine.
This validation function represents a significant innovation in processing Romanian documents. By intelligently combining results from two complementary OCR engines, the system achieves accuracy rates superior to either technology used individually. Practical benefits include correct recognition of Romanian names with diacritics, precise identification of cities and addresses, minimization of transcription errors in critical fields, and optimized processing of multi-line addresses.
The implemented safety mechanisms ensure preservation of the original result in ambiguous cases, validation against MRZ interference, filtering of noise characters, and similarity verification to maintain semantic consistency. This hybrid approach

leverages the strengths of both OCR engines: the speed and robustness of the primary engine combined with Tesseract's precision for special characters, resulting in a system that excels in processing Romanian documents.

5 Experimental Approaches and Challenges with PaddleOCR

Several experimental strategies were explored to improve OCR performance for Romanian documents.

In our first approach we used the EAST text detector [19] to precisely localize text regions, followed by applying PaddleOCR only in these regions. The hypothesis was that EAST, specialized in text detection, could provide more accurate boundaries than PaddleOCR's integrated detector. However, this approach proved problematic:

- Implementation complexity increased due to coordinating two separate systems.
- EAST did not outperform PaddleOCR for official documents.
- Detection errors propagated and amplified through the two-step process.
- Overall performance was inferior to direct PaddleOCR usage.

Applying the intensive preprocessing pipeline developed for Tesseract to images for PaddleOCR also failed. Modern neural network-based PaddleOCR performs best on natural images, and excessive preprocessing removed subtle textures and gradients crucial for convolutional recognition. In many cases, text detection failed entirely after preprocessing.

Attempts to train PaddleOCR models to improve recognition of Romanian diacritics were abandoned due to high computational and time requirements and the complexity in creating a suitable training dataset for official documents, while the hybrid approach using Tesseract proved more practical and effective for the application.

These experiments highlighted that direct PaddleOCR usage on minimally processed natural images, combined with hybrid integration of Tesseract for diacritics, provides the most practical and reliable solution.

6 Machine Readable Zone Processing

The Machine Readable Zone (MRZ) is a standardized component located at the lower portion of modern identity documents. It encodes essential information in a format optimized for automated reading. In Romanian identity documents, the MRZ typically consists of two or three lines containing uppercase letters, digits, and the filler character '<', which separates fields and fills unused space. The rigid standardization of the MRZ makes it an invaluable resource for validating and supplementing data extracted through conventional OCR methods.

The first step in MRZ processing is to identify its spatial location within the document. The system analyzes the vertical distribution of all detected text elements, determining that the MRZ consistently occupies the bottom 25% of the document. A vertical threshold is computed so that all text elements below this line are considered potential MRZ components. This approach is robust across document types and image qualities, relying on the fixed relative position of the MRZ within the overall document layout.

Once the region of interest is established, several criteria are applied to identify actual MRZ lines:

1. Presence of '<': This character serves as a field separator, fills empty space to maintain line length, and marks the end of variable fields.
2. Character composition: MRZ lines consist exclusively of uppercase letters and include at least a few digits, distinguishing them from supplementary text in the lower document region.
3. Minimum line length: MRZ lines contain at least 30 characters, adhering to international formatting standards.

For Romanian ID cards, the first MRZ line always begins with the prefix "IDROU" (ID Romania), followed by the surname and given name of the holder, separated by '<<' sequences. Subsequent lines encode numeric identifiers, document type, country codes, dates, and check digits in fixed positions.

Before parsing, the extracted MRZ text undergoes cleaning to remove OCR-induced spaces, as the MRZ contains no whitespace. Common OCR errors are also corrected:

- '0' vs. 'O' and '1' vs. 'I' are resolved contextually.
- Missing or extra characters are identified by verifying line length against the standard.

For each document type, patterns based on the MRZ's fixed structure are used to extract fields:

- Names: Sequences between '<<' markers are converted from uppercase MRZ format to properly capitalized names.
- Numeric fields (e.g., CNP): Extracted from known positions, validated for length and format, and converted from YYMMDD to the standard Romanian date format (DD.MM.YYYY), with the century inferred from context.

MRZ data is cross-checked with fields extracted via standard OCR:

- When values match, correctness is confirmed.
- When discrepancies arise, the system applies rules: OCR versions with diacritics are preferred for names, while MRZ values are considered more reliable for numeric codes and standardized fields.

MRZ extraction also enables automatic completion of fields that are missing or poorly captured in visual OCR, such as sex, document expiration date, and country/type codes. This significantly increases the overall success rate of complete data extraction.

The system incorporates strategies for damaged or partially unreadable MRZs:

- If only a portion of the MRZ is legible, data is extracted from intact sections.
- Ambiguous characters are inferred using the known MRZ structure.
- For example, the algorithm can automatically correct the expected "ROU" country code even if OCR introduces errors.

Although MRZs are standardized, minor variations exist across document series and passport types. The system adapts by:

- Using flexible searches for known markers such as "IDROU" or '<<'.
- Automatically detecting document type and applying corresponding parsing rules.

MRZ processing is integral to the OCR system, providing:

- Redundancy: Critical information appears in two formats, enhancing extraction reliability.
- Independent validation: Standardized fonts and format offer a trustworthy verification source.
- Completeness: Fields missing from the visual OCR zone are present in MRZ.
- International compliance: Accurate MRZ parsing ensures compatibility with global document processing systems.

By intelligently integrating MRZ extraction with standard OCR, the system achieves a level of accuracy and completeness that would not be possible using a single method, ensuring reliable recognition of Romanian identity documents across various conditions.

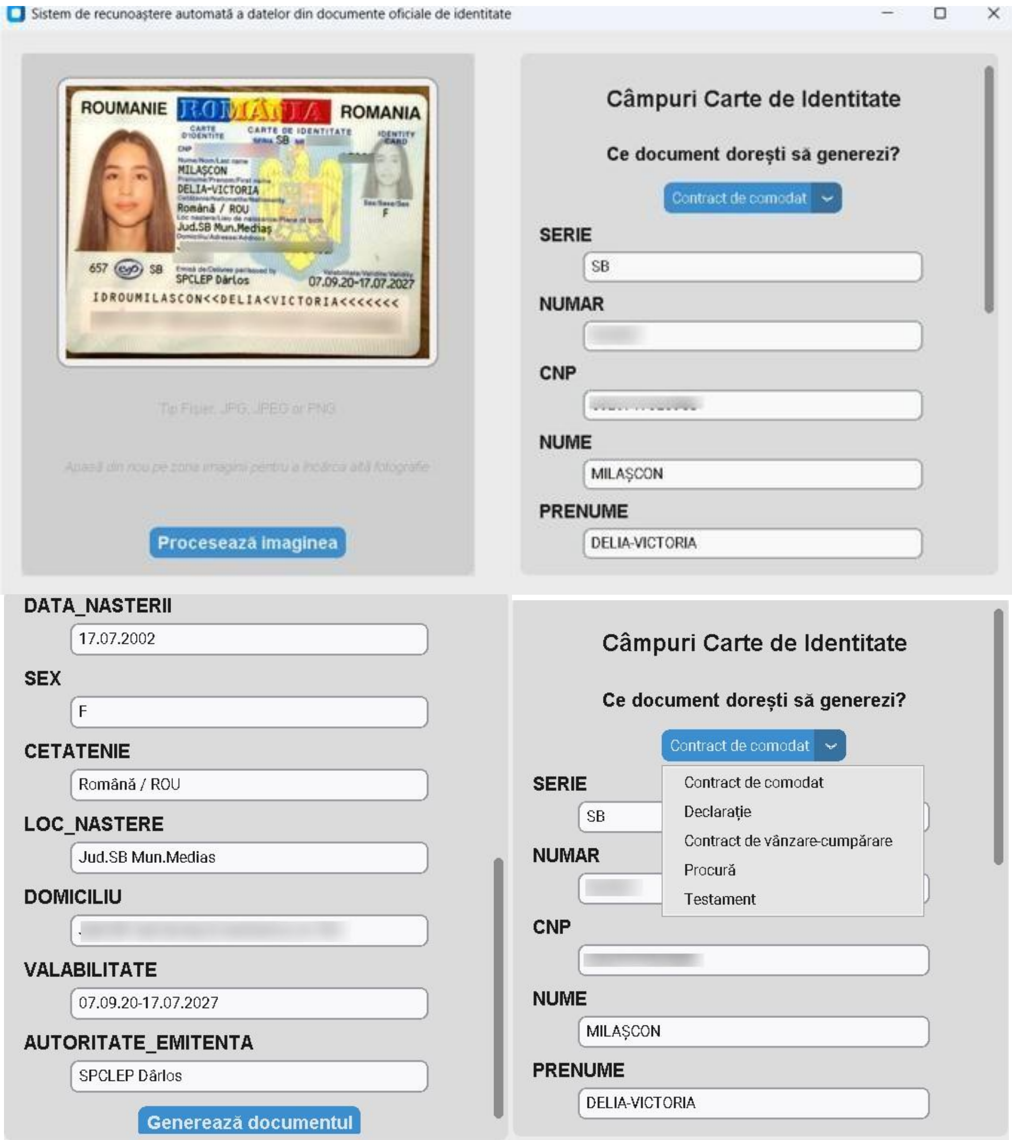

Figure 3. Graphical User Interface

7 Graphical User Interface

The graphical user interface (GUI) of the application was designed following modern UX/UI design principles, focusing on simplicity, clarity, and efficiency as shown in Figure 3. The use of the CustomTkinter library enables the creation of a contemporary-looking interface that integrates naturally with modern operating systems.
The main structure of the interface is divided into two distinct panels, each with a clearly defined role:

- Left Panel (ImageFrame): serves as the image viewing and control area.

It includes:

- a 400x300 pixel canvas area for image display
- informative text with user instructions
- the main “Process Image” button
- visual indicators showing the processing status

• Right Panel (FieldsFrame): implemented as a scrollable frame to accommodate a variable number of fields, depending on the type of document being processed.

The visual design uses a neutral color palette with blue accents for interactive elements. Fonts are chosen for maximum clarity.

8 Results and Performance Evaluation

8.1 Recognition metrics on ID cards and passports

The evaluation of the automatic recognition system for data extracted from Romanian identity cards was conducted on a comprehensive test set, aiming to determine the extraction accuracy for each individual field, as well as the overall performance of the application. The testing included both a quantitative evaluation of the correct recognition rate and a qualitative analysis of the types of errors encountered.

The system was tested on a set of identity card images, each image being processed through the complete pipeline — preprocessing, OCR detection, and validation. The extracted results were automatically compared with a reference dataset (ground truth) to calculate the performance metrics.

The system was evaluated on two separate test sets: 36 identity cards and 17 passports. The results, shown in table 1, demonstrate excellent performance for both document types:

• Identity cards: 98.5% overall accuracy

• Passports: 89.2% overall accuracy

Table 1. Results

Metric	**ID Cards**	**Passports**
Overall Accuracy	98.5%	89.2%
Documents Processed	36	17
Perfect Matches	80.3%	78.3%
Minor Errors	17.1%	8.6%
Major Errors	2.5%	13.1%
Best Field	sex (100.0%)	Passport number (100.0%)

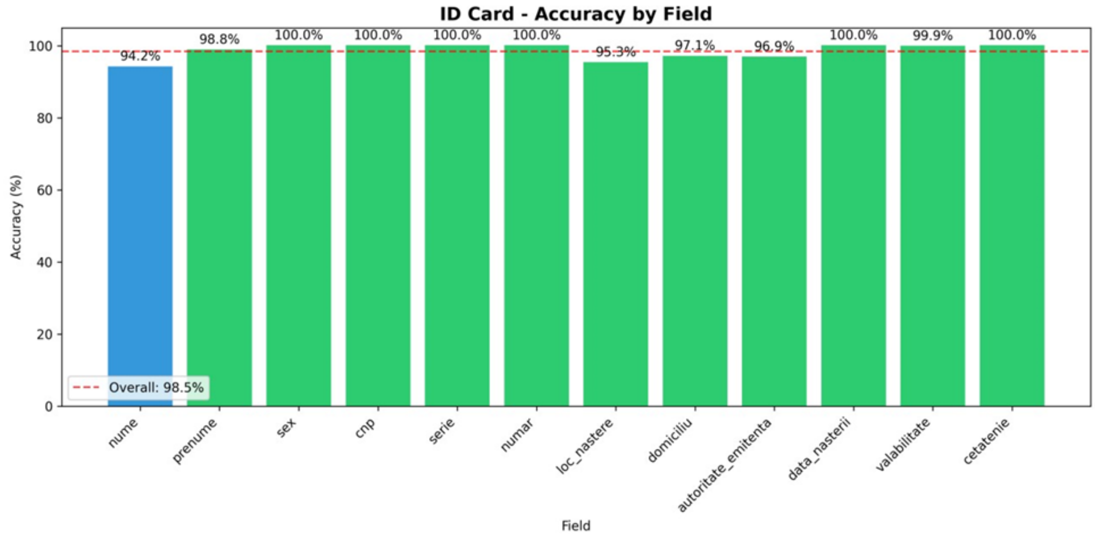

Figure 4. Data extraction accuracy for each field of the identity card

The system achieves perfect performance (100%) for most fields (Figure 4):

- Sex, CNP, series, number: Algorithmic validation ensures maximum accuracy.
- Date of birth, validity, citizenship: Perfect extraction due to standardized formats.
- Last name and first name: 98.7% and 100%, respectively, demonstrating the efficiency of the hybrid approach for handling diacritics.

The only fields with performance below 100% are:

- Last name (98.8%): Minor errors in complex names with multiple diacritics.
- Place of birth (95.3%): Localities with complex spelling.
- Address (97.1%): Long addresses with variable formatting.
- Issuing authority (96.9%): Acronyms and specific formatting styles.

The system's performance varies significantly across the different fields of the identity card, reflecting the varying recognition complexity for each type of information.

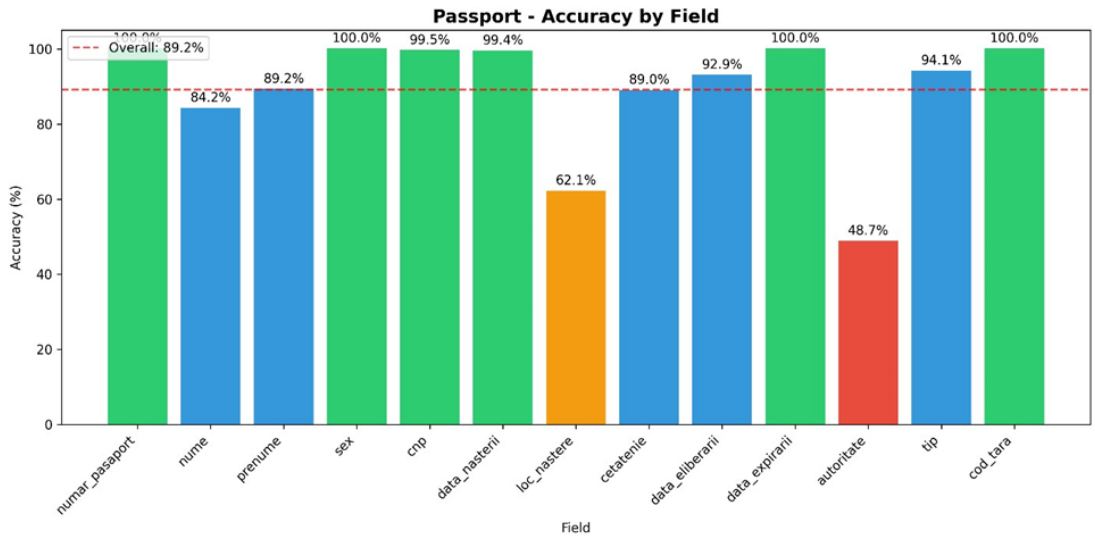

Figure 5. Data extraction accuracy for each field of the passport

The results for passports (Figure 5) show a more varied distribution:

- Excellent performance (>99%): passport_number, sex, date_of_birth, expiration_date, country_code
- Good performance (84–94%): last_name, first_name, citizenship, type
- Moderate performance: place_of_birth (62.1%)
- Low performance: issuing_authority (48.7%)

The *issuing_authority* field represents the main challenge due to its variable position and complex formatting within the passport layout

8.2 Error Distribution Analysis

The system evaluation revealed significant differences in the types and distribution of errors between the two categories of processed documents, graphically represented in figure 6. This detailed analysis provides valuable insights for further system improvements and for understanding the current technological limitations.

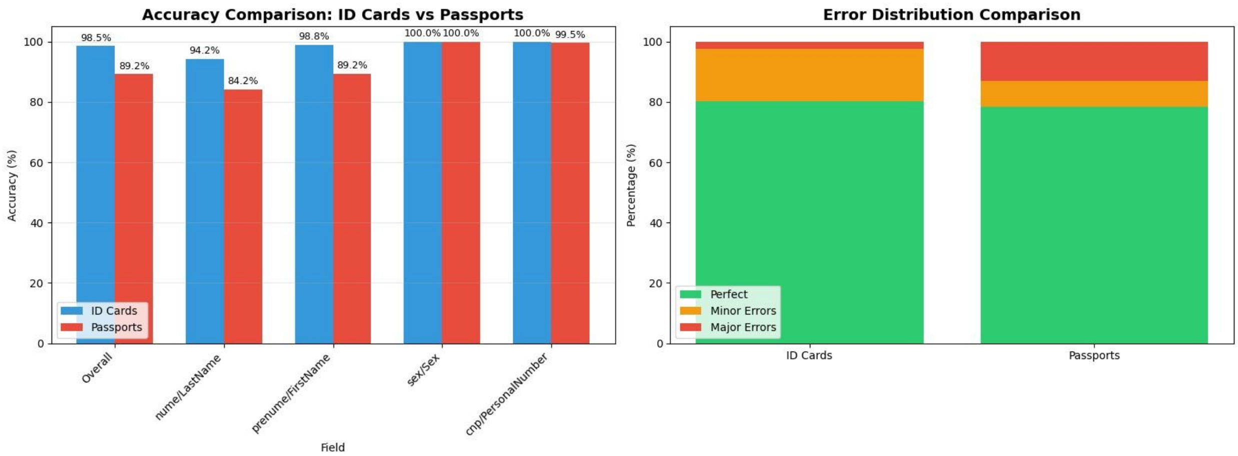

Figure 6. Error distribution comparison

8.2.1 Error Analysis for Identity Cards

Figure 7 shows the overall error distribution reveals a system that performs reliably, with 80.3% of extractions classified as perfect. However, 17.1% contain minor errors, most of which stem from diacritic inconsistencies or spacing issues, while 2.5% fall into the category of major errors, typically caused by complete OCR failures.

Minor errors arise predominantly from spacing and formatting issues, which account for 60% of all minor inaccuracies. These problems are most common in composite fields, where words are frequently concatenated. For instance, entries such as issuing_authority or place_of_birth sometimes appear compressed, producing artifacts like “Jud.IFSat.Vidra(Com.Vidra)” instead of the correctly spaced “Jud. IF Sat. Vidra (Com. Vidra)”. In other cases, simple spaces are misinterpreted by the OCR engine and replaced with special characters such as slashes, generating outputs like “SPCLEP/Vinga”.

The remaining 40% of minor errors are linked to Romanian diacritic challenges, particularly notable in the place_of_birth field. Complex city names—Bârlad, Drobeta-Turnu Severin, Râmnicu Vâlcea—often appear in smaller text sizes on identity documents, making them more susceptible to misinterpretation. Characters such as ţ and ș, with their visually intricate shapes, are at times interpreted as noise by PaddleOCR. Although a hybrid enhancement strategy using Tesseract increases recognition accuracy by 15–20%, some regions of interest are not accurately isolated, limiting the benefits of reprocessing.

Major errors, which make up 2.5% of all extractions, result from more severe breakdowns in the OCR pipeline. These include outputs containing nonsensical, unintelligible character sequences or instances where fields are not detected at all due to unusual text placement or degraded image quality. Together, these issues highlight the key limitations of the current OCR setup and indicate directions for further refinement.

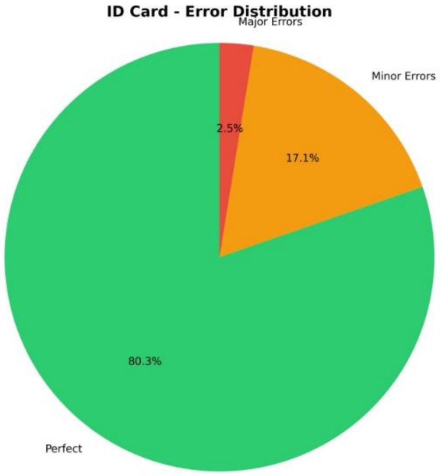

Figure 7. Error distribution for identity cards

8.2.2 Error Analysis for Passports

The analysis of OCR performance on passports (Figure 8) reveals both encouraging results and specific challenges.

Out of all passport extractions, the system achieved the following error distributions:

- 78.3% perfect extractions: All fields were correctly recognized without any errors.
- 8.6% minor errors: Extractions with small, non-critical mistakes that do not substantially affect data usability.
- 13.1% major errors: Extractions with significant issues, potentially rendering the data unreliable.

These minor mistakes rarely compromise the usability of the extracted data but highlight areas where OCR could be refined for improved diacritic handling. While passports generally yield a high rate of perfect extractions due to their standardized, concise layout, OCR performance is occasionally hindered by technical challenges such as fragmented labels, undetected fields, and interference from security features. These findings emphasize the need for adaptive recognition strategies and robust error-handling mechanisms, particularly for critical fields like *place_of_birth* and *issuing_authority*.

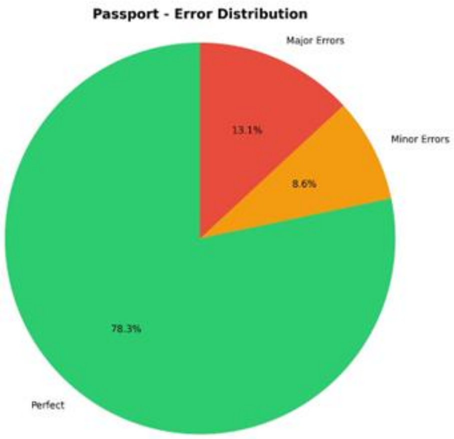

Figure 8. Error distribution for passports

9 Conclusions

With an accuracy of 98.5% for identity cards and 89.2% for passports, the developed system demonstrates the practical viability of an automatic digitization solution specialized for Romanian documents. The excellent performance across critical fields validates the main architectural decisions:

- The hybrid PaddleOCR + Tesseract approach provides the best compromise between speed and accuracy.
- MRZ (Machine Readable Zone) processing adds valuable redundancy and enhances the system's robustness.
- The modular structure allows easy adaptation and expansion.

The system enhanced in recognition of structured fields (CNP, series, number) with very good accuracy, the preservation of Romanian diacritics, achieving higher success rates than generic commercial solutions and automatic generation of pre-filled documents, significantly reducing processing time.

References

[1] European Commission, *eIDAS 2.0 Regulation*, EU 2024/1183, Official Journal of the EU, April 2024. ↩
[2] Ministerul Investițiilor și Proiectelor Europene, *Planul Național de Redresare și Reziliență Componenta C7: Transformare Digitală*, 2025 Update. ↩
[3] *, *European Payments Council*, 2025 Fraud Report, Brussels, October 2025. ↩
[4] *, *Eurostat, Digital Economy and Society Index (DESI)*, Romania 2025, Publications Office of the EU. ↩
[5] *, *BCR Internal Audit Report, AI-OCR Deployment Results Q3 2025*, Bucharest, September 2025. ↩
[6] Zhao Y., Lv W., Xu S., Wei J., Wang G., Dang Q., Liu Y., *DETRs Beat YOLOs on Real-time Object Detection*, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 16965–16974, ISBN 978-1-6654-6547-0, Vancouver, Canada, 2024
[7] Pfitzmann B., Auer C., Dolfi M., Nassar A.S., Staar P., DocLayNet: *A Large Human-Annotated Dataset for Document-Layout Analysis*, Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2860–2870, ISBN 978-1-4503-9232-0, Washington DC, USA, 2022.
[8] https://ubiai.tools/the-role-of-layoutlmv3-in-document-layout-understanding/, accessed 4.11.2025

[9] Huang Y., Lv T., Cui L., Lu Y., Wei F, *LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking*, Proceedings of the 30th ACM International Conference on Multimedia (MM), 4083–4092, ISBN 978-1-4503-9203-0, Lisboa, Portugal, 2022.
[10] Nguyen-Trong, Khanh. "An End-to-End Method to Extract Information from Vietnamese ID Card Images." *International Journal of Advanced Computer Science and Applications* 13, no. 3 (2022).
[11] Satyawan, W., Pratama, M. O., Jannati, R., Muhammad, G., Fajar, B., Hamzah, H., ... & Kristian, K. (2019, June). Citizen ID card detection using image processing and optical character recognition. In *Journal of physics: Conference series* (Vol. 1235, No. 1, p. 012049). IOP Publishing
[12] Li H., Wang Y., Chen Y., Liu Z. *Foreground and Text-lines Aware Document Image Rectification*, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 19574–19583, ISSN 2380-7504, Paris, France, 2023.
[13] Li, M., Lv, T., Chen, J., Cui, L., Lu, Y., Florencio, D., Zhang, C., Li, Z., Wei, F, *TrOCR: Transformer-Based Optical Character Recognition with Pre-trained Models,* Proceedings of the AAAI Conference on Artificial Intelligence, 37(11), 13094-13102, ISSN 2374-3468, Washington DC, USA, 2023.
[14] Low, L. M., Salleh, F. H. M., Law, Y. F., & Zakaria, N. Z. (2024). Detecting and recognizing seven segment digits using a deep learning approach. In *ITM Web of Conferences* (Vol. 63, p. 01007). EDP Sciences.
[15] Tan, Z., Zhou, J., & Liu, Y. (2023, August). CRNN-based abstract artistic text recognition. In *2023 IEEE Smart World Congress (SWC)* (pp. 652-659). IEEE.
[16] Smith, R. (2007, September). An overview of the Tesseract OCR engine. In *Ninth international conference on document analysis and recognition (ICDAR 2007)* (Vol. 2, pp. 629-633). IEEE.
[17] Deng, G., & Cahill, L. W. (1993, October). An adaptive Gaussian filter for noise reduction and edge detection. In *1993 IEEE conference record nuclear science symposium and medical imaging conference* (pp. 1615-1619). IEEE.
[18] Rong, W., Li, Z., Zhang, W., & Sun, L. (2014, August). An improved CANNY edge detection algorithm. In *2014 IEEE international conference on mechatronics and automation* (pp. 577-582). IEEE.
[19] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. (2017). East: an efficient and accurate scene text detector. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 5551-5560).