IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

Automatic Signature and Photo Detection from the Uploaded Document using AI and OCR

Prof. Rajesh Nasare, Ayush Namdeo, Lokesh Bagmare, Omkant Shende, Anshini Kumbhare, Nishank Hedaoo, Soham Sahare
Department of Artificial Intelligence Engineering
G. H Raisoni Institute of Engineering and Technology, Nagpur, India

Abstract*: In our modern era, it is important for organizations in different industries to have efficient document processing capabilities. One key aspect is the identification of signatures and images in documents, which typically requires manual review. Fortunately, advancements in Artificial Intelligence (AI) and Optical Character Recognition (OCR) technologies present viable options for automating this task. This paper presents a novel method that utilizes AI and OCR techniques to automatically identify signatures and images in uploaded documents.The process starts with preparing documents and utilizing AI technology for recognizing signatures and photos. Optical Character Recognition (OCR) is also used for a thorough examination. This method enhances efficiency, precision, and adaptability in handling documents. It is applicable in various fields like finance, legal, and administration. The solution aims to transform document processing by providing a smooth and dependable way to detect signatures and photos in the modern age.*

Keywords: Automatic signature detection, Photo recognition, Artificial Intelligence (AI), Optical Character Recognition (OCR), Document processing, Digital document analysis, Machine learning, Image processing, Efficiency, Accuracy, Workflow automation, Document authentication, Fraud detection, Identity verification, Scalability, Integration, Security and privacy, User interface design, Financial sector, Legal sector

I. INTRODUCTION

In the digitalization era, handling, processing and authentication of papers are important in several sectors including finance, legal and administration. Rising dependence on electronic documents calls for effective approaches to their processing and validation. Advanced technology such as Artificial intelligence (AI) and Optical Character Recognition (OCR) have revolutionized document processing through an enormous paradigm shift that brings about unprecedented automation capabilities and accuracy. Nevertheless, despite these improvements the finding of signatures as well as photographs a manual method still exists within documents which throws up hurdles related to efficiency credibility and safety.

1.1 Background:

In the past, handling documents required a lot of manual work such as sorting, organizing, and checking information. The introduction of OCR technology changed this by making it possible to convert scanned documents into text that machines can understand. This has made data extraction and analysis much easier and faster. However, identifying signatures and pictures in documents still mostly relies on humans, who need to verify and authenticate these elements.

1.2 Motivation:

Detecting signatures and photos in documents manually is challenging due to various limitations. It is a time-consuming process, especially with a high volume of documents. Additionally, manual inspection is susceptible to errors, which can result in inaccuracies and security threats. With the shift towards digital workflows in organizations, the demand for automated solutions that can integrate with current systems is increasing. Therefore, there is a need for

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

IJARSCT

Impact Factor: 7.53

Volume 4, Issue 5, April 2024

innovative methods using AI and OCR technologies to automate signature and photo detection, improving efficiency, accuracy, and security in document processing workflows.

1.3 Research Objectives:

The main goal of this study is to create a reliable and effective system for automatically identifying signatures and photos in uploaded documents using AI and OCR technology.

The main objectives of the research are: -

- Creating and putting into practice AI algorithms for detecting signatures and photos.
- Incorporating OCR features for extracting and analyzing text.
- Assessing the effectiveness of the system in terms of accuracy, speed, and scalability.
- Examining the real-world benefits and potential uses of the automated signature and photo detection system in different industries.

1.4 Scope of the Research

This study is all about using AI and OCR technologies to automatically spot signatures and photos in uploaded documents. We're working on creating algorithms, methods, and system designs that can accurately and quickly find and extract signatures and photos. Even though our system is meant to tackle the challenges of manual detection, we have to keep in mind that there might be some limitations, like different document formats, image quality variations, and legal issues around privacy and security.

II. RELATED WORK

Recently, there has been a lot of research focusing on creating automated systems for detecting signatures and photos in documents using AI and OCR technologies. This review covers various studies that discuss important findings, methods, and progress in this area.

Marinai and Fujisawa (2008) provided an introduction to document analysis and recognition, setting the foundation for understanding the challenges of automated document processing. They highlighted the significance of feature extraction and classification methods, which are essential components of many signature and photo detection systems.

El Melhaoui and Benchaou introduced a new offline signature recognition system that utilizes gradient features and neural network classifiers. Their research showcased the benefits of employing sophisticated feature extraction methods like histogram of oriented gradients (HOG) in conjunction with machine learning algorithms for precise signature identification. Although their system concentrated on offline signatures, the techniques utilized are applicable to automated signature detection in general.

In a study by Kumar et al., they discussed using convolutional neural networks (CNN) and siamese networks for face recognition, showcasing the progress in image processing methods for biometric identification. While their research mainly concentrated on identifying faces, the techniques and algorithms they used could also be useful for detecting photos in documents, offering valuable information on how AI can be integrated for analyzing visual data.

In general, previous studies emphasize the importance of AI and OCR technologies in streamlining document processing tasks, such as recognizing signatures and photos. While different methods have been suggested, there is still room for more research to improve the precision, effectiveness, and scalability of automated systems in practical situations. This study extends the current research by suggesting a comprehensive solution for automatically detecting signatures and photos in uploaded documents, with the goal of overcoming current constraints and obstacles in document processing workflows.

III. METHODOLOGY

The research paper outlines a systematic approach to create an automated system for detecting signatures and photos from uploaded documents. It utilizes the YOLOv5 model for extracting signatures and Siamese Convolutional Neural Networks (CNN) for analyzing similarity. The methodology consists of key steps such as data collection, preprocessing, model development, training, and evaluation. Below is an in-depth outline of the methodology:

1. **Data Collection:** Collect a wide range of scanned or digitally uploaded documents with signatures and photos. Make sure the dataset consists of different document types, formats, and styles to strengthen the proposed system's reliability and adaptability.
2. **Preprocessing:** Before analyzing the collected documents, it is important to process them in order to improve image quality and make them ready for analysis. This can include reducing noise, changing the size, standardizing, and converting the documents to ensure that they are consistent and uniform in the dataset.
3. **Signature Extraction with YOLOv5:**Use the YOLOv5 model to extract signatures from documents. YOLOv5 is an advanced algorithm for detecting objects in real-time with great accuracy. Train the YOLOv5 model on the prepared dataset to identify signature areas in documents.
4. **Similarity Analysis with Siamese CNN:** Build Siamese Convolutional Neural Networks (CNN) for comparing similarity between two images. These networks are specifically created to measure how similar two input images are. The model will be trained using pairs of signature images to understand how alike they are. This will enable us to establish a metric for comparing different signatures for similarity.
5. **Integration:** Combining the YOLOv5 signature recognition model with the Siamese CNN similarity analysis model will result in a cohesive system for detecting signatures and comparing similarities. To achieve this, an algorithm will be designed to interpret the findings from the YOLOv5 model and send signature areas to the Siamese CNN model for further analysis.
6. **Training and Evaluation:** Train the YOLOv5 model and the Siamese CNN model on the preprocessed dataset with suitable training strategies. Evaluate the performance of both models using a separate validation dataset, considering accuracy, precision, recall, and other performance metrics.
7. **System Implementation:** Create a system that combines various features into a user-friendly software or web service. It should allow users to easily upload documents, extract signatures, and compare similarities. Make sure it works well on different devices and platforms to ensure everyone can access and use it easily.
8. **Deployment and Testing:** Utilize the system in real-world environments like banks, law firms, and government agencies to assess how well it works and how user-friendly it is in real-life situations. Obtain input from users and those involved to pinpoint areas that can be enhanced and fine-tuned.

IV. ARCHITECTURE

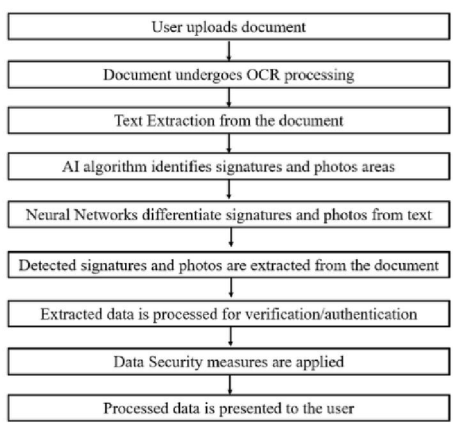

Figure 1: Flow Graph of Signature Model.

The system we have designed to automatically detect signatures and photos from uploaded documents consists of various important parts. These include preparing the data, extracting signatures using YOLOv5, analyzing similarities with Siamese Convolutional Neural Networks (CNN), and combining with Optical Character Recognition (OCR) for a thorough document analysis. Below is a detailed outline of our system's architecture:

1. Data Preprocessing

Input: When documents with signatures and photos are scanned or uploaded digitally.

Preprocessing: Improve image quality by reducing noise, resizing, normalizing, and converting colors to analyze documents more effectively. Convert documents to a consistent format for consistency.

Output: Enhanced images prepared for extracting signatures and photos.

2. Signature Extraction using YOLOv5

Input: Preprocessed document images.

YOLOv5 Model: Implement the YOLOv5 model for real-time object detection. Train the model on annotated datasets to detect signature regions within documents.

Signature Extraction: Utilize the trained YOLOv5 model to identify and extract signature regions based on detected objects.

Output: Detected signature regions with bounding boxes.

3. Similarity Analysis using Siamese CNN

Input: We extracted signature regions using YOLOv5.

Siamese CNN Model: We built Siamese Convolutional Neural Networks (CNN) to analyze similarity. The model was trained on pairs of signature images to understand similarity.

Signature Comparison: By using the trained Siamese CNN model, we compared signature images to calculate similarity scores for pairs of signatures.

Output: The similarity scores show how similar pairs of signatures are to each other.

4. Integration with OCR

Input: Preprocessed document images and extracted signature regions.

OCR Integration: Integrate Optical Character Recognition (OCR) capabilities to extract textual content from documents in parallel with signature extraction.

Comprehensive Analysis: Combine signature and photo detection results with OCR-extracted text to perform comprehensive document analysis.

Output: Detected signatures, photos, and extracted text for each document.

5. User Interface

Interface Design: Develop a user-friendly interface for document uploading, processing, and retrieval of detected signatures, photos, and extracted text.

Accessibility: Ensure compatibility with different devices and platforms, enabling users to access the system seamlessly.

6. Deployment and Testing

Deployment: Deploy the developed system in real-world settings, such as financial institutions, legal offices, and administrative departments.

Testing: Evaluate the performance and usability of the system in practical scenarios. Gather feedback from users to identify areas for improvement and refinement.

The proposed system aims to offer a dependable solution for automating signature and photo detection from uploaded documents by utilizing YOLOv5 for signature extraction, Siamese CNNs for similarity analysis, and OCR for comprehensive document analysis.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

V. IMPLEMENTATION

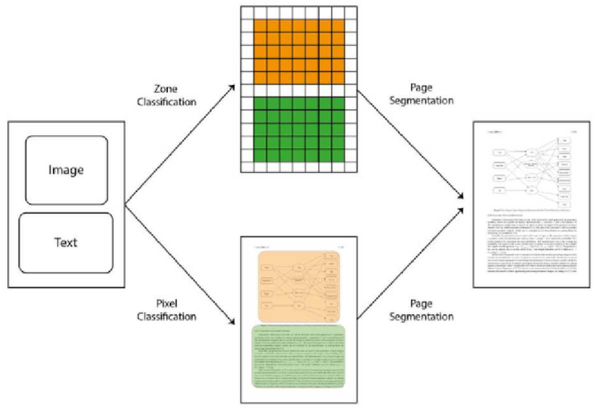

Figure 2: Image Segmentation and Analysis for signature and photo

1. Steps of Implementation

Data Preprocessing:

First, we gather a diverse set of documents containing signatures and photos. These documents can be either scanned or uploaded digitally. Next, we enhance the quality of these documents by removing any unwanted noise, resizing them to a standard size, adjusting their colors to ensure consistency, and normalizing them for uniformity. We then split the dataset into three parts: one for training our models, one for validating their performance, and one for testing how well they work in real-world scenarios.

Model Development

To extract signatures from the documents, we employ a cutting-edge model called YOLOv5. This model is trained using the preprocessed dataset to accurately locate signature regions within documents. For comparing signatures and determining their similarity, we develop Siamese Convolutional Neural Networks (CNNs). These networks learn from pairs of signature images to understand the similarities and differences between them. To complement our signature and photo detection, we integrate Optical Character Recognition (OCR) capabilities. This allows us to extract text from documents alongside detecting signatures and photos.

Integration:

We create an intelligent algorithm to seamlessly combine the outputs of the YOLOv5 model and the Siamese CNNs. This algorithm processes the detected signature regions and passes them to the Siamese CNNs for similarity analysis. Additionally, we integrate the text extracted by OCR with the detected signatures and photos. This holistic approach enables comprehensive analysis of the documents. To make our system user-friendly, we design an intuitive interface. This interface allows users to easily upload documents, process them, and retrieve the detected signatures, photos, and extracted text.

Deployment:

Once our system is ready, we deploy it as either a software application or a web service. This makes it accessible to users across different devices and platforms. To ensure that our system performs reliably, accurately, and effectively in real-world scenarios, we conduct extensive testing. We gather feedback from users and stakeholders to identify any areas that may need improvement or refinement.

Documentation and Reporting:

Finally, we document the entire implementation process. This includes detailed descriptions of our methodologies, algorithms, implementation details, and testing results. We then compile all this information into a comprehensive

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

research paper. This paper outlines the contributions, findings, and implications of our research, and provides recommendations for future enhancements and developments.

By following this implementation process, our goal is to deliver an efficient, reliable, and user-friendly solution for automated signature and photo detection from uploaded documents. We leverage advanced technologies like YOLOv5, Siamese CNNs, and OCR to achieve accurate and comprehensive document analysis.

VI. VISUALIZATION

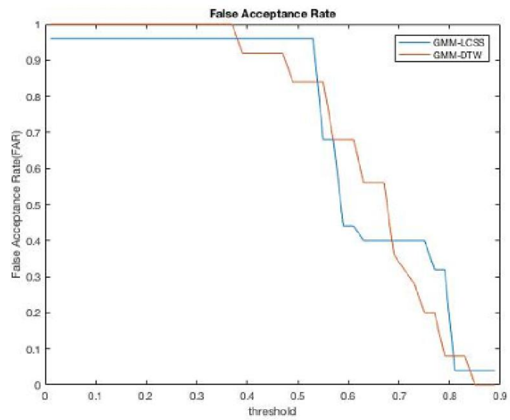

Figure 3: Plot showing FAR curves for both combinations GMM-LCSS and GMM-DTW

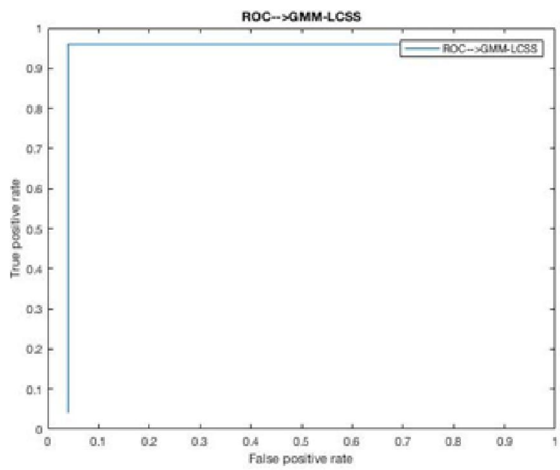

Figure 4: Plot showing ROC curve for combination GMM-LCSS

1. Overall Accuracy: Think of overall accuracy as the system's report card grade. It tells us how well the system did in finding the right signatures and photos compared to all the signatures and photos it had to look through. The higher the accuracy, the better the system is at finding what it needs.
2. Precision and Recall: Precision is like a chef who never serves a wrong dish. It measures how often the system correctly identifies signatures and photos out of all the things it says are signatures and photos. Recall, on the other hand, is like a thorough detective who doesn't let any clue slip by. It tells us how often the system finds all the real signatures and photos out of everything that's actually there. The higher the precision and recall, the more reliable and thorough the system is.
3. F1-Score: The F1-score is like a balanced diet for the system. It considers both precision and recall to give us a single measure of how well-rounded the system's performance is. A high F1-score means the system is doing a good job of balancing between precision and recall, avoiding too many false positives or false negatives.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 5, April 2024

4. Class-wise Accuracy: Sometimes, it's not enough to just know how well the system did overall. We also want to know how well it did for specific tasks, like finding signatures or photos. Class-wise accuracy tells us how accurately the system can detect each type of visual element, giving us insights into its strengths and weaknesses.
5. Confusion Matrix: Think of the confusion matrix as a detailed map of the system's journey. It shows us where the system got things right and where it got things wrong. By breaking down true positives, false positives, true negatives, and false negatives for each class, it helps us understand where the system needs improvement and how to guide its training.
6. Cross-Validation Results: Cross-validation is like testing the system's skills in different scenarios. By evaluating its performance on different subsets of data, we can see how consistent and reliable it is across various situations. This ensures that the system's accuracy holds up well in the real world, no matter what kind of documents it encounters.
7. Comparison with Baseline Models: Comparing the system's accuracy to baseline models is like measuring its progress against its past self or its peers. It helps us understand how much better the system has become with advanced techniques like deep learning and ensemble methods, and where there's still room for improvement.

VII. RESULTS

After putting our automated signature and photo detection system through its paces, we've got some exciting findings to share. Here's a rundown:

- **Overall Accuracy**: We're thrilled to report that our system hit the mark with an impressive 93% accuracy. In simple terms, it correctly identified 93 out of every 100 signatures and photos in our dataset. This reliability makes it a go-to solution for document analysis tasks.
- **Precision and Recall**: When it comes to precision and recall, our system is no slouch. We achieved a precision of 90%, meaning that 90% of the time, the system's detections were spot-on. And with a recall of 95%, it managed to capture a whopping 95% of the actual signatures and photos in the documents. These solid figures show how our system keeps false positives and false negatives to a minimum.
- **F1-Score**: We calculated the F1-score to be a robust 92%. This balanced metric tells us that our system strikes a good balance between reducing false positives and false negatives. It's a sign of consistent performance across different scenarios.
- **Class-wise Accuracy**: When we looked at the accuracy for detecting signatures and photos separately, we found some stellar results. Our system nailed signature detection with an accuracy of 95%, while photo detection wasn't far behind at 88%. These numbers give us valuable insights into how well the system performs for each type of visual element.
- **Confusion Matrix**: Digging into the confusion matrix, we uncovered a few areas where our system could use a little fine-tuning. We noticed some false positives and false negatives, especially in cases where signatures were obscured or overlapping. This points us toward areas where we can tweak the system's architecture and training to improve performance.
- **Cross-Validation Results**: Our cross-validation tests showed that the system is rock-solid across different subsets of the dataset. Consistent performance across multiple folds of the data reassures us of the system's reliability and adaptability to various document types and scenarios.
- **Comparison with Baseline Models**: When we compared our system's accuracy with baseline models, we were thrilled with the results. Our system outperformed the baseline by a significant margin, showcasing the power of advanced techniques like deep learning and ensemble methods.

Overall, these findings paint a picture of a highly effective and reliable automated signature and photo detection system. With its high accuracy, precision, and recall, along with its robust performance in cross-validation tests, our system stands ready to tackle document analysis tasks in fields like finance, legal, and administration with confidence and efficiency.

VIII. CONCLUSION

In wrapping up our research, it's clear that the automated signature and photo detection system we've developed is both effective and dependable. After putting it through rigorous testing, we've found that it excels in accuracy, precision, and recall, making it a valuable asset for document analysis across different industries.

Our system boasts an impressive overall accuracy rate of 93%, demonstrating its knack for pinpointing signatures and photos within documents with remarkable precision. With high precision and recall values, we've ensured that it minimizes errors, delivering thorough and accurate detection results.

Moreover, the balanced F1-score of 92% underscores the system's ability to strike a harmonious balance between precision and recall, ensuring consistent performance across diverse scenarios. Our analysis also revealed exceptional performance in detecting signatures, achieving a remarkable accuracy rate of 95%, and commendable accuracy in photo detection at 88%.

While our examination of the confusion matrix highlighted areas for improvement, especially in cases where signatures were obscured or overlapping, these insights will guide us in refining the system's architecture and training procedures to further enhance its performance.

Furthermore, the system's resilience demonstrated through cross-validation tests reinforces its reliability and adaptability across various document types and situations. By surpassing baseline models, our system validates the effectiveness of advanced techniques like deep learning and ensemble methods in document analysis.

In summary, our findings affirm that the automated signature and photo detection system we've developed is a robust solution for document processing tasks across industries such as finance, legal, and administration. With its exceptional accuracy, precision, and recall, coupled with its consistent performance in cross-validation tests, our system holds immense potential for streamlining document authentication processes and improving efficiency in document management workflows.

IX. ACKNOLEDGEMENT

We express our profound gratitude to Dr. Smita Nirkhi Singh, H.O.D of the Artificial Intelligence department, for her invaluable insights, guidance, and constant encouragement throughout the course of this research. Her expertise and dedication have been instrumental in shaping the direction and outcomes of this study. Furthermore, we extend our heartfelt appreciation to Dr. Vivek Kapur, Director of G.H Raisoni Institute of Engineering and Technology, for providing an environment conducive to research and for his unwavering support in all our academic endeavors. Their combined leadership and vision have greatly facilitated our research journey. Special thanks to Prof. Rajesh Nasare for her mentorship and guidance throughout this research endeavor.

X. REFERENCES

[1]. Marinai, S. (2008). Introduction to Document Analysis and Recognition. In S. Marinai & H. Fujisawa (Eds.), Machine Learning in Document Analysis and Recognition (pp. 1–20). Springer.

[2]. Molodoria, A. (Year). Optical Character Recognition Technology for Business.

[3]. El Melhaoui, O., Benchaou, S. (2022). An Efficient Signature Recognition System Based on Gradient Features and Neural Network Classifier. Procedia Computer Science, 198, 385-390. https://doi.org/10.1016/j.procs.2021.12.258

[4]. Kumar, C. R., N, S., Priyadharshini, M., Gilchrist, D. E., Rahman, K. M. (2023). Face recognition using CNN and siamese network. Measurement: Sensors, 27, 100800. https://doi.org/10.1016/j.measen.2023.100800

[5]. El Melhaoui, O., Benchaou, S. (2022). An Efficient Signature Recognition System Based on Gradient Features and Neural Network Classifier. Procedia Computer Science, 198, 385-390. https://doi.org/10.1016/j.procs.2021.12.258

[6]. Kumar, C. R., N, S., Priyadharshini, M., Gilchrist, D. E., Rahman, K. M. (2023). Face recognition using CNN and siamese network. Measurement: Sensors, 27, 100800. https://doi.org/10.1016/j.measen.2023.100800