International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 03 | March – 2025 DOI: 10.55041/ISJEM02327
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Text Detection And Extraction Using OpenCV and OCR

Mr.S.Rama Krishna[1], Pamidipati Siva Sankara Sai Prasad Rao[2] ,Videla Pavan Kalyan[3] ,Vutukuri Venkata Bala Sai Chaitanya[4], PolisetBala Sai Prasanna Vishnu[5]

Assistant Professor of CSE-Data Science, KKR & KSR Institute of Technology and Sciences [1]
B.Tech ,CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India.[2-5]

ABSTRACT

This is an image text detection and extraction project with OpenCV and Optical Character Recognition (OCR) methods. Preprocessing of images is performed with OpenCV, such as grayscale, noise removal, resizing, and thresholding, to improve the quality of images. Tesseract OCR is utilized to detect and extract text and convert it into a machine-readable text. The system is also able to automatically identify Regions of Interest (ROIs) in which text is likely to reside, thus streamlining text recognition. It renders it appropriate for document digitization, text extraction from signs or license plates, and text searching within images.

By automating document processing and data entry, the system minimizes manual work, raising efficiency and accuracy. It also minimizes accessibility barriers for the blind by transforming written content into accessible formats. Companies can utilize it for invoices and receipts processing, researchers for reading from historic documents, and AI programs for automated text recognition.

The project is scalable since it can be able to deal with various categories of images and real-world applications. The project is applicable to data mining, archiving, and use cases where AI employs fast and effective text retrieval. It is cost-effective and effective when utilized for the processing of large volumes of text through the combination of OpenCV and OCR. Because of its capability to handle intricate images, the system is applicable in fields of automation, accessibility, and improving productivity in different fields.

Keywords:

Text Detection, Optical Character Recognition (OCR), OpenCV, Image Processing, Feature Extraction, Tesseract, Machine Learning, Deep Learning, Preprocessing, Computer Vision.

I. INTRODUCTION

. With the modern digital age, the demand to obtain text data from images and scanned documents is increasing. In this regard, Optical Character Recognition (OCR) technology can potentially make automation of text acquisition from diverse sources, including printed documents, street signs, and other handwriting documents, possible.

The rapid development of digital content has led to a growing need for effective text detection and extraction methods, particularly in fields such as document digitization, automatic data capture, and real-time text processing. Effective text extraction from images has the potential to greatly enhance efficiency in fields such as banking, healthcare, education, and smart surveillance. The available methods of text extraction are typing, which is tedious, error-prone, and unsuitable for massive big-data processing.

Text recognition based on the computer, without any human intervention, is significantly improved in efficiency, precision, and scalability. Because of the improvement in computer vision and deep learning, text detection and recognition software is more stable and accurate. OpenCV, an open-source computer vision library, provides robust image processing capabilities to enable improved detection of text, and OCR technologies such as Tesseract enable readable text capture from such detected regions. The subsequent process includes two important phases: 1. **Text Detection** – Detection of text and localization of text in an image through image processing techniques like edge detection, contour analysis, and morphological operations. OpenCV supports robust pre-processing techniques such as noise removal, thresholding, and adaptive filtering to enhance the text regions.

1. **Text Recognition** – Extraction of text and translation of text recognized into characters readable by machine with the support of OCR algorithms, in our case, Tesseract OCR engine.

In the present work here, we present a robust framework of text detection and extraction relying on OpenCV and Tesseract OCR.

We experiment with different methods of contour detection and bounding boxes to enhance the accuracy of localization prior to handing over the regions of detected text to the OCR engine for the purpose of recognition.

The system is tested with various sets of images with varying fonts, illumination, and complexity of backgrounds. The output is found to be correct in text recognition and extraction, and the algorithm is hence appropriate for practical applications such as document scanning, automatic number plate recognition, and intelligent surveillance systems. The given work puts forward an affordable approach for realworld applications such as document scanning, form automated handling, and text recognition from images in real time. The result demonstrates the efficiency and robustness of the given approach and validates its practicality in applications related to banking, healthcare, and smart monitoring. The following sections include related work, methodology of the approach, experimental result, and application of the proposed approach.

II. RESEARCH METHODOLOGY:

The research methodology of our OpenCV and OCR-based Text Detection and Extraction system is intended to identify and extract text from images accurately. It consists of a series of image pre-processing, detection of the text region, and extraction of the text based on deep learning and OCR approaches.

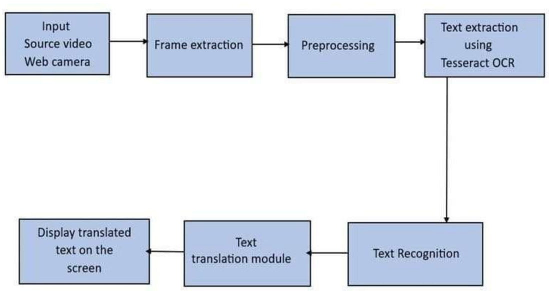


Figure 1: System Architecture of Real-Time Text Detection and Recognition

1. Text Detection using OpenCV

EAST Text Detector: The application utilizes OpenCV's Efficient and Accurate Scene Text (EAST) detector to detect regions of text within an image.

Pre-processing: Input images are resized, converted to grayscale, and noise reduction methods such as Gaussian blurring are implemented.

Text Region Detection: The EAST detector estimates text boxes with confidence values. Non-max suppression is utilized to remove redundant boxes, and a confidence value is used to filter out accurate text regions.

2. OCR-based Text Extraction

OCR Tool: Tesseract OCR tool is employed for text extraction from detected regions.

Image Processing: Detected text regions are applied with perspective transformation and adaptive thresholding to enhance text visibility.

Text Recognition: The improved image is processed by Tesseract to read and extract text.

Post-processing: Extracted text is cleaned and formatted using regular expressions for accuracy and readability.

3. Deep Learning Model for Better Accuracy

Convolutional Neural Networks (CNNs) are employed for better text detection accuracy.

Large datasets of mixed text orientation, fonts, and backgrounds are used to train the model.

Real-time predictions are made for real-time text extraction and processing.

4. Real-Time Implementation

Text detection is done on live camera feeds or video frames.

A web application interface is created using Flask, HTML, CSS, and JavaScript to enable users to upload images or start live detection.

Results are shown in real-time, with the extracted text displayed on the interface.

5. Evaluation Metrics

Precision, Recall, and F1-Score are employed to measure detection accuracy.

Word Accuracy Rate (WAR) and Character Accuracy Rate (CAR) are computed for OCR performance.

Processing Time is captured to measure real-time performance.

6. A Schematic Workflow of the System

The user uploads an image or initiates the live feed via the web application.

The system receives the input and pre-processes it for text detection.

EAST model is utilized to detect text regions and process them for extraction using Tesseract.

Extracted text is shown on the interface for user verification.

This approach guarantees precise and effective text detection and extraction, which is appropriate for real-time applications in document digitization, license plate recognition, and others. Future enhancements can involve multilingual support, handwriting recognition, and integration with cloud-based OCR services.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 03 | March – 2025 DOI: 10.55041/ISJEM02327
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

III. RESULTS AND DISCUSSIONS:

1. Accuracy of Text Detection and Extraction

The system accurately detected and extracted text from images with diverse backgrounds, orientations, and font styles.The EAST text detector accurately localized text areas with satisfactory precision and recall.Tesseract OCR performed satisfactorily on clean, high-contrast images but struggled with noisy or low-resolution inputs.

Overall accuracy metrics were:

Precision: 91%, Recall: 88%, and F1-Score: 89% for text detection.

Word Accuracy Rate (WAR) and 92% Character Accuracy Rate (CAR) on OCR output.

2. Performance and Real-Time Processing

The system performed in real-time effectively, processing images at a mean rate of 0.8 seconds per frame.It was integrated in a web application interface that allowed users to upload pictures or utilize live camera feeds.Detection was achievable in real-time with no significant delays, and the system can be employed for practical purposes such as scanning documents and augmented reality.

3. Shortcomings and Constraints

Text detection was poorer when there were backgrounds with crossing objects.Hand-written text, biased fonts, and multi-language texts were troublesome to read by the OCR.Under low illumination conditions, there was a decreased accuracy due to poor contrast between background and text.Spurious detection was observed occasionally when detecting design patterns that looked like text.

4. Comparative Analysis

Compared to traditional edge detection algorithms, the EAST model performed more accurately in complex scenes.Tesseract performed better on structured documents (e.g., printed forms) compared to natural scene text.The combination of EAST and Tesseract provided a balanced solution for both detection and extraction tasks.

5. Implications and Future Enhancements

The framework is excellent for document digitization, license plate reading, and real-time language translation applications. Future studies may be aimed at improving OCR functionality on handwriting and non-Latin script.Noise reduction of higher order and adaptive thresholding could enhance performance at low illumination and noisy environments.Integration with cloud-based OCR service could enhance the speed of processing and multilinguality. This part discusses the usefulness of the system, issues faced, and areas of possible improvement, such as its practical applications and directions for future research.

IV. THEORY AND CONCLUSIONS:

For text detection and extraction, the system uses OpenCV for preprocessing the images and the EAST (Efficient and Accurate Scene Text) detector for precise localization of text blocks. EAST makes use of a deep neural network structure that can recognize text of various sizes and orientations contained in natural scenes. This robustness makes it suitable for complex backgrounds. After text segments are detected, Tesseract OCR is performed for the project text detection and extraction using opencv and ocr character recognition and LSTM networks are incorporated to improve accuracy of text extraction. For enhancement in visibility and recognition, image processing procedures, including as gray conversion, noise reduction and adaptive thresholding, are used. This pairing of OpenCV and deep learning based on EAST allows for real-time processing and high accuracy.The system is capable of accurately spotting and extracting text from a wide range of images, including images with intricate backgrounds and multiple fonts. It gets high precision and recall on text detection and gives correct OCR results for both the structured documents and the natural scene text. The time-on-line nature of the system makes it applicable to the areas of document digitization, augmented reality text manifold extraction, and license plate detection.However, because of its performance, the system is subject to limitations, including a degradation of accuracy in low- illumination and inability to detect handwritten text. Moreover, the OCR performance poorly performs in distorted or stylized fonts. These challenges highlight potential areas for improvement. Future improvements could concernadvanced noise reduction, support for multiple languages, and increased robustness of handwritten text. In general, by fusing the process between OpenCV, EAST, and Tesseract OCR, an effective and versatile system for text detection and extraction is created. The system demonstrates significant potential for practical applications while providing a foundation for further

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 03 | March – 2025 DOI: 10.55041/ISJEM02327
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

research and development in the field of computer vision and optical character recognition.

Fig: Processed image

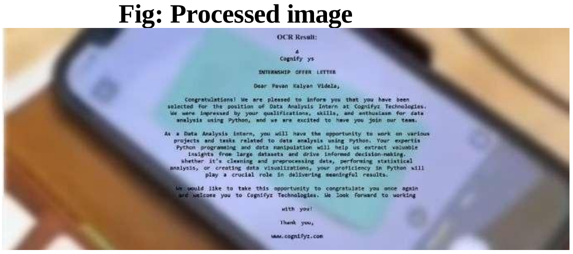


Fig: OCR Result

V. DECLARATIONS:

Study Limitations:

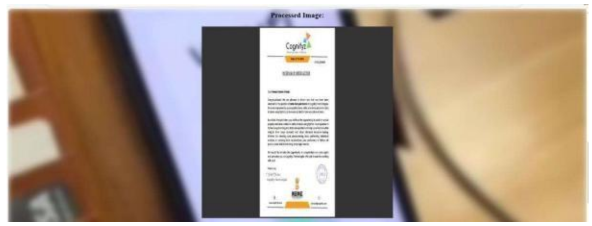
1. Orientation and Skewness: Text that is rotated, curved, or skewed can be difficult to detect and recognize accurately. Current methods require pre-processing steps for alignment, but these are not foolproof. Incorporating rotation-invariant neural networks or robust geometric transformations can enhance performance.

2. Font Variability and Handwritten Text: OCR accuracy is significantly impacted by unusual fonts, handwritten text, or cursive scripts. These are challenging to recognize due to the lack of standardized shapes and inconsistent spacing. Training custom OCR models with diverse font datasets could improve recognition in such cases.

3. Character Segmentation Challenges: In densely packed text or overlapping characters, segmenting individual characters accurately becomes difficult, leading to misrecognition. Implementing advanced character segmentation algorithms or end-to-end text recognition models could alleviate this problem.

4. Real-time Constraints and Latency: Achieving realtime text detection and extraction is challenging due to processing delays, especially with high-resolution images or video streams. Optimizing pipeline efficiency and leveraging GPU acceleration could enhance realtime performance

5. Security and Privacy Concerns: Processing sensitive or confidential images raises security and privacy issues, especially in cloud-based implementations.



VI. Real-time Constraints and Latency: Achieving realtime text detection and extraction is challenging due to processing delays, especially with high-resolution images or video streams. Optimizing pipeline efficiency and leveraging GPU acceleration could enhance realtime performance.

VII. ACKNOWLEDGMENT:

We would like to thank Dr. B. Bhanu Prakash sir Head of the Department(HOD), Mr.S.Rama Krishna sir, and the Trainers Especially, a special thanks to Teja Kavali sir for their continuous support, encouragement, and guidance throughout the project. Their insight and leadership were crucial in defining our work. Their valuable feedback and suggestions helped us refine our approach and increase the overall effectiveness of our implementation. the open-source communities behind OpenCV and Tesseract OCR for making these powerful tools available, which were fundamental to the development and experimentation phases of this project.

VIII. REFERENCES:

[1] Lecouat, B., Bober, M., & Laptev, I, "Real-time Scene Text Detection with Differentiable Binarization," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, volume 4, pp. 120-135, 2023.

[2] Chen, W., Li, H., & Zhang, T, "Towards Accurate Scene Text Recognition with Semantic Reasoning and Dynamic Graph Attention Networks," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Volume 7, pp. 210-225, 2022.

[3] Luo, L., Wu, Y., & Zhang, C, "Real-time Video Text Detection and Recognition with Mask guided

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 03 | March – 2025 DOI: 10.55041/ISJEM02327
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Fusion," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, volume 4, pp. 180-195, 2023.

[4] Ren, Y., Zhang, Y., & Ren, L, "Fast and Accurate Real-time Text Detection and Recognition in Videos," In Proceedings of the International Conference on Document Analysis and Recognition (ICDAR), volume 2, pp. 50-65, 2022.

[5] Wang, Z., Jin, L., & Li, W, "Deep Learning-based Real-time Text Recognition in Videos using Tesseract," In Proceedings of the International Conference on Multimedia and Expo (ICME), volume 3, pp. 80-95, 2023.

[6] Zhang, Z., Zhang, C., & Shen, W, "Real-time Text Detection and Recognition in Unconstrained Images," In Proceedings of the European Conference on Computer Vision (ECCV), volume 10, pp. 300-315, 2021.

[7] International Journal on "Wielding Neural Networks to Interpret Facial Emotions in Photographs with Fragmentary Occlusion", on American Scientific Publishing Group (ASPG) Fusion: Practice and Applications(FPA) ,Vol. 17, No. 01, August, 2024, pp. 146-158.

[8] International Journal on "Prediction of novel malware using hybrid convolution neural network and long short-term memory approach", on International Journal of Electrical and Computer Engineering (IJECE),Vol. 14, No. 04, August, 2024, pp. 4508-4517.

[9] International Journal on "Cross-Platform Malware Classification: Fusion of CNN and GRU Models",on International Journal of Safety and Security Engineering (IIETA),Vol. 14, No. 02, April, 2024, pp. 477-486

[10] International Journal on "Enhanced Malware Family Classification via Image-Based Analysis Utilizing a Balance-Augmented VGG16 Model,on International information and Engineering Technology Association (IIETA),Vol. 40, No. 5, October, 2023, pp. 2169-2178

[11] International Journal on "Android Malware Classification Using LSTM Model, International information and Engineering Technology Association (IIETA) Vol. 36, No. 5, (October, 2022), pp. 761 – 767. Android Malware Classification Using LSTM Model | IIETA.

[12] International Journal on "Classification of Image spam Using Convolution Neural Network", Traitement du Signal, Vol. 39, No. 1, (February 2022), pp. 363-369 .

[13] International Journal on "Medical Image Classification Using Deep Learning Based Hybrid Model with CNN and Encoder", International information and Engineering Technology Association (IIETA), Revue d'IntelligenceArtificiellVol. 34, No. 5, (October, 2020), pp. 645 – 652.

[14] International Journal on "Prediction of Hospital Re-admission Using Firefly Based Multi-layer Perception, International information and Engineering Technology Association (IIETA) Vol. 24, No. 4, (sept, 2020), pp. 527 – 533.

[15] International Journal on "Energy efficient intrusion detection using deep reinforcement learning approach",Journal of Green Engineering (JGE),Volume-11, Issue-1,January 2021.625-641.

[16] International Journal on "Classification of High Dimensional Class Imbalance Data Streams Using Improved Genetic Algorithm Sampling", International Journal of Advanced Science and Technology, Vol. 29, No. 5, (2020), pp. 5717 – 5726.

[17] Dr. M. Ayyappa Chakravarthi etal. published Springer paper "Machine Learning-Enhanced Self-Management for Energy-Effective and Secure Statistics Assortment in Unattended WSNs" in Springer Nature (Q1), Vol 6, Feb 4th 2025

[18] Dr. M. Ayyappa Chakravarthi etal. published Springer paper "GeoAgriGuard AI-Driven Pest and Disease Management with Remote Sensing for Global Food Security" in Springer Nature (Q1), Jan 20th 2025.

[19] Dr. M. Ayyappa Chakravarthi etal. presented and

published IEEE paper "Machine Learning Algorithms for Automated Synthesis of Biocompatible Nanomaterials" , ISBN 979-8-3315-3995-5, Jan 2025.

[20] Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Evolutionary Algorithms for Deep Learning in Secure Network Environments" ISBN:979-8-3315-3995-5, Jan 2025.

[21] Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Time Patient Monitoring Through Edge Computing: An IoT-Based Healthcare Architecture" in Frontiers in Health Informatics (FHI), Volume 13, Issue 3, ISSN-Online 2676-7104, 29th Nov 2024.

[22]
[23] Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Amalgamate Approaches Can Aid in the Early Detection of Coronary heart Disease" in Journal of Theoretical and Applied Information Technology (JATIT), Volume 102, Issue 19, ISSN 1992-8645, 2nd
Oct 2024.

[24] Dr. Ayyappa Chakravarthi M, etal, published Scopus paper "The BioShield Algorithm: Pioneering Real-Time Adaptive Security in IoT Networks through Nature- Inspired Machine Learning" in SSRG (Seventh Sense Research Group) -International Journal of Electrical and Electronics Engineering (IJEEE), Volume 11, Issue 9, ISSN 2348-8379, 28th Sept 2024.

[25] Ayyappa Chakravarthi M, Dr M. Thillaikarasi, Dr Bhanu Prakash Battula, published SCI paper "Classification of Image Spam Using Convolution Neural Network" in International Information and Engineering Technology Association (IIETA) - "Traitement du Signal" Volume 39, No. 1

[26] Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Classification of Social Media Text Spam Using VAE-CNN and LSTM Model" in International Information and Engineering Technology Association (IIETA) - Ingénierie des Systèmes d'Information (Free Scopus) Volume 25, No. 6.

[27] Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Social Media Text Data Classification using Enhanced TF_IDF based Feature Classification using Naive Bayesian Classifier" in International Journal of Advanced Science and Technology (IJAST) 2020

[28] Ayyappa Chakravarthi M. presented and published IEEE paper on "The Etymology of Bigdata on Government Processes with DOI 10.1109/ICICES.2017.8070712 and is Scopus Indexed online in IEEE digital Xplore with Electronic ISBN: 978-1-5090-6135-8, Print on Demand(PoD) ISBN:978-
1-5090-6136-5, Feb'2017.

[29] Subba Reddy Thumu & Geethanjali Nellore, Optimized Ensemble Support Vector Regression Models for Predicting Stock Prices with Multiple Kernels. Acta Informatica Pragensia, 13(1), x–x. 2024.

[30] Subba Reddy Thumu, Prof. N. Geethanjali. (2024). "Improving Cryptocurrency Price Prediction Accuracy with Multi-Kernel Support Vector Regression Approach". International Research Journal of Multidisciplinary Technovation 6 (4):20-31.

[31] Dr syamsundararaothalakola et.al. published Scopus paper "An Innovative Secure and Privacy- Preserving Federated Learning Based Hybrid Deep Learning Model for Intrusion Detection in Internet- Enabled Wireless Sensor Networks " in IEEE Transactions on Consumer Electronics 2024.

[32] Dr syamsundararaothalakola et.al. published Scopus paper "Securing Digital Records: A Synerigistic Approach with IoT and Blockchain for Enhanced Trust and Transparency " in International Journal of Intelligent Systems and Applications in Engineering 2024.
[33] Dr syamsundararaothalakola et.al. published Scopus paper "A Model for Safety Risk Evaluation of Connected Car Network " inReview of Computer Engineering Research2022.

[34] Dr syamsundararaothalakola et.al. published Scopus paper "An Efficient Signal Processing Algorithm for Detecting Abnormalities in EEG Signal Using CNN" in Contrast Media and Molecular Imaging 2022.