IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 3, March 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

Enhancing Handwritten Text Recognition: A NovelApproach with MXNet and Advanced Neural Architectures

Dr. V. Muralidhar[1], D. Dharani Mahesh[2], Y. Hrudayesh[3], P. Jashwanth[4], K. Mokshitha[5]

Associate Professor, Department of CSE (AI & ML), Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, Nambur, Guntur, India[1]

Department of CSE (AI & ML), Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, Nambur, Guntur, India[2]

Department of CSE (AI & ML), Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, Nambur, Guntur, India[3]

Department of CSE (AI & ML), Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, Nambur, Guntur, India[4]

Department of CSE (AI & ML), Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, Nambur, Guntur, India[5]

ABSTRACT: In this study, we introduce an innovative approach to Optical Character Recognition (OCR), utilizing a synergistic blend of Convolutional Neural Networks (CNN), Bidirectional Long Short-Term Memory (BiLSTM), Connectionist Temporal Classification (CTC), and Single Shot Multi-box Detector (SSD) technologies, enhanced with Maximally Stable Extremal Regions (MSER). Our focus is on enhancing the recognition of handwritten texts, employing the IAM Handwriting Database and incorporating sophisticated data augmentation strategies, tailored anchor box designs, and advanced probabilistic character prediction. Our findings showcase notable enhancements in accuracy and efficiency of text recognition, supported by thoroughexperimental validations and comparative analyses with current models.

KEYWORDS: Optical Character Recognition, CNN-BiLSTM-CTC, SSD-MSER, Handwritten Text Analysis, Data Enhancement, MXNetGluon.

I. INTRODUCTION

In our research, we tackle the challenging task of handwritten text recognition using Optical Character Recognition (OCR). We innovate by combining MXNet Gluon with advanced neural architectures like CNN-BiLSTM-CTC and SSD- MSER, drawing upon significant object detection and feature extraction advancements [13, 12, 11]. Our approach extends the capabilities of OCR, especially in interpreting various handwriting styles with high accuracy.We delve into the complexities of handwritten OCR, employing sophisticateddata augmentation methods similar to "Manifold Mixup" [18]. Our work builds on previous evaluations of CNN-BiLSTM models [4] and incorporates text recovery techniques [5] and offline text extraction methods [19]. Our system sets a new benchmark in OCR technology with its ability to accurately recognize characters and text lines across diverse handwriting styles. Through extensive experimentation and comparative analysis, we demonstrate the effectiveness of our model, marking a significant progression in the OCR field.

II. LITERATURE SURVEY

1. *Introduction to the Survey*: This segment explores the evolution of OCR, focusing on deep learning architectures like CNN, BiLSTM, and CTC. It highlights the use of MXNet Gluon for handwritten text recognition, underscoring

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

significant contributions to OCR advancements.

2. *Advancements in CNN-BiLSTM-CTC:* The research delves into the development of OCR systems utilizing CNN and BiLSTM networks combined with CTC loss. Studies by Jiao et al. [5] and Kızılırmak and Yanıkoğlu [4] are highlighted for their contributions to text recovery and English handwriting recognition, respectively. Shrestha et al. [19] and the "Manifold Mixup" study [18] are noted for their innovative training processes and model generalization improvements.

3. *Role of SSD-MSER in OCR:* This section discusses the integration of SSD with MSER in OCR. Key studies by Jiang et al. [13], Zhai et al. [12], and Chen et al. [11] are examined for their advancements in text detection and segmentation, crucial for handling textual content in various formats and scales.

4. *Data Augmentation Techniques:* The importance of dataaugmentation in OCR is discussed, with emphasis on the manifold mixup method [18] and studies by Shrestha et al. [19], Jiao et al. [5], and Kızılırmak and Yanıkoğlu [4]. These approaches enhance the robustness and accuracy ofhandwriting recognition models.

5. *Comparative Studies:* This part reviews comparative analyses of different OCR frameworks, especially MXNet-based models. The advancements by Jiang et al. [13], Zhai et al. [12], and Chen et al. [11] in SSD algorithms and CNN-BiLSTM-CTCnetworks by Shrestha et al. [19] and Kızılırmak and Yanıkoğlu [4] are explored. The "Manifold Mixup" technique [18] is also noted for its impact on text recognition.

6. *Research Context:* The research integrates MXNet Gluon with CNN-BiLSTM-CTC and SSD-MSER, focusing on gaps identified in existing studies. It leverages MXNet's efficiency, CNN-BiLSTM-CTC's accuracy, and SSD-MSER's text segmentation capabilities. Insights from studies like Shrestha et al. [19] and the manifold mixup approach [18] are used to enhance the model's adaptability to handwriting variations. This novel integration aims to advance practical OCR applications.

III. METHODOLOGY

A. PAGE SEGMENTATION

Data Compilation: Our research utilized the IAM Handwriting Database [1], which includes 1,539 pages of texts written by 657 distinct individuals. This diverse collection provided a comprehensive range of handwriting samples, including sentences, lines, and words, serving as the primary dataset for our study.

Initial Image Enhancement: The first step involved standardizing the size and pixel values of input images to create a uniform dataset suitable for neural network training. We implemented data augmentation to introduce variety, accommodating different handwriting styles. Additionally, denoising techniques were applied to enhance text clarity, thereby improving the overall quality and effectiveness of the images for further processing.

Precise Page Segmentation: Utilizing the MSER algorithm [10], our approach focused on accurately segmenting pages. MSER was instrumental in detecting stable regions within the images, enabling the isolation of textual components like words and lines. This detailed segmentation is essential for the effective application of CNNs in later stages of character recognition.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

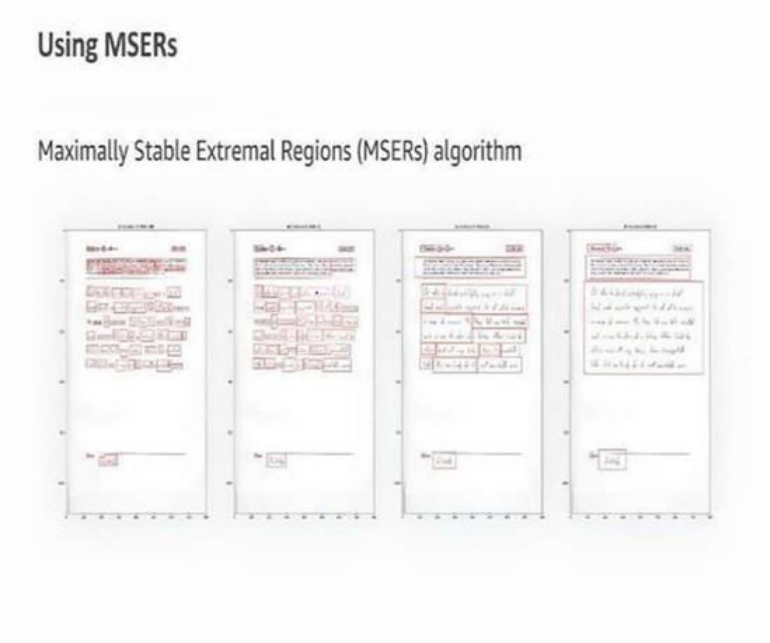


Fig. 1: MSER Algorithm in Action for Text Segmentation- This figure showcases the Maximally Stable Extremal Regions (MSER) algorithm's phases in segmenting handwritten texts, with each step visually represented. Textual regions identified by the algorithm are marked in red, aiding in the preparation of data for CNN-based character recognition, as detailed in [10].

Character Recognition with CNN: Our study utilizes Convolutional Neural Networks (CNN) to decode characters from scanned documents. The multi-layered CNN architecture, as shown in Figure 2, processes visual data through a series of specialized layers. These include convolutional layers for feature extraction, batch normalization for learning consistency, and pooling layers for data condensation and computational efficiency. This architecture is pivotal in accurately locating character coordinates, highlighting CNN's pattern recognition strength.

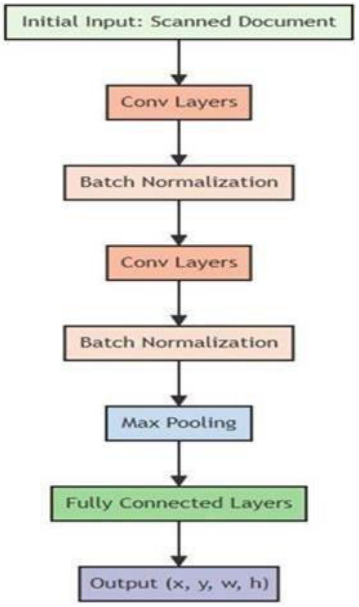


Fig.2 Convolutional Neural Network (CNN) Architecture for Character Recognition.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

Augmenting Data for Generalization: To ensure our model's effectiveness across varied handwriting styles, we implemented data augmentation techniques like rotation and scaling. This approach, inspired by Jiao et al. [5] and Kızılırmak and Yanıkoğlu [4], broadens the model's exposure to different handwriting styles, enhancing its adaptability and generalization. Such methods are critical for the model's capacity to process diverse handwriting, aligning with our goal of improved generalization.

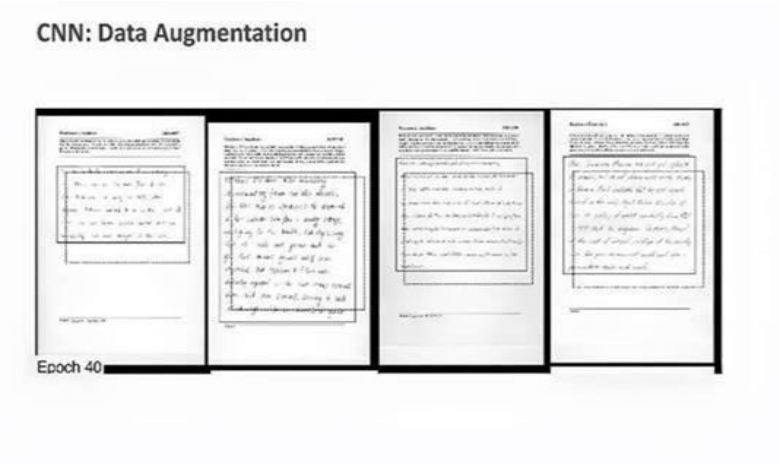


Figure 3, "Data Augmentation in CNN for OCR" [10], showcases how varying rotation and scale during training improves our CNN's ability toaccurately interpret diverse handwriting styles, essential for real-world OCR applications.

Model Training and Refinement: The model's training began with optimizing the Mean Square Error (MSE) loss function, refining predictions of character bounding boxes. This was followed by enhancements using the Intersection over Union (IoU) metric to improve text outlining accuracy within images, as illustrated in Figures 4, 5, and 6. This iterative training method, focusing first on MSE and then on IoU, significantly boosted the model's performance and precision, as evidenced by the trends in the figures.

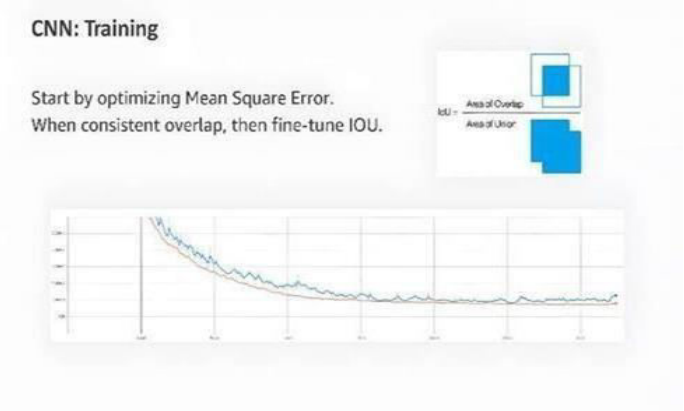


Fig. 4 Early Stage MSE Optimization: This figure, adapted from the methodology proposed by Apache MXNet [10], depicts the initial training phase where minimizing the Mean Square Error (MSE) was paramount for the accurate prediction of character bounding boxes.

Continuous Assessment and Enhancement: Our research implemented a dynamic evaluation process during the CNN training, where the model was assessed at defined intervals using key metrics like character accuracy and text localization precision. These metrics were instrumental in fine-tuning the model's architecture and training methods. Following industry benchmarks, such as the protocols by Apache MXNet [10], our training started with a focus on reducing Mean Square Error (MSE) to establish a fundamental accuracy level. We then advanced our assessment with the Intersection over Union (IoU) metric for a deeper understanding of the model's text localization abilities. Figure 7, inspired by Apache MXNet's work [10], illustrates our evaluation technique at specific epochs, showing how MSE and

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

IoU guided our training refinements. Through this cyclical assessment and adjustment process, our CNN model achieved significant accuracy and versatility.

Fig. 5 Illustration of Model Evaluation at Epoch 20 (Adapted from [10]) Depicted here is the model's performance at Epoch 20, showcasing the accuracy in character bounding box predictions and text localization, which were key to the enhancements made throughout the training regimen.

B. LINE SEGMENTATION WITH SINGLE SHOT MULTI BOX DETECTOR (SSD)

Data Augmentation for Diverse Handwriting Representation: To ensure our model could handle the wide range of human handwriting, we expanded our dataset through data augmentation. Techniques like rotation, scaling, and shearing were utilized to mimic the variability encountered in real- world scenarios. Drawing from the work of Jiang et al. [13], we incorporated context- aware enhancements to refine our model's recognition capabilities in complex scenes. We also adapted multiscale analysis methods from Chen et al. [11] and dense feature fusion from Zhai et al. [12], enriching our approach to handle diverse handwriting sizes and intricate text features.

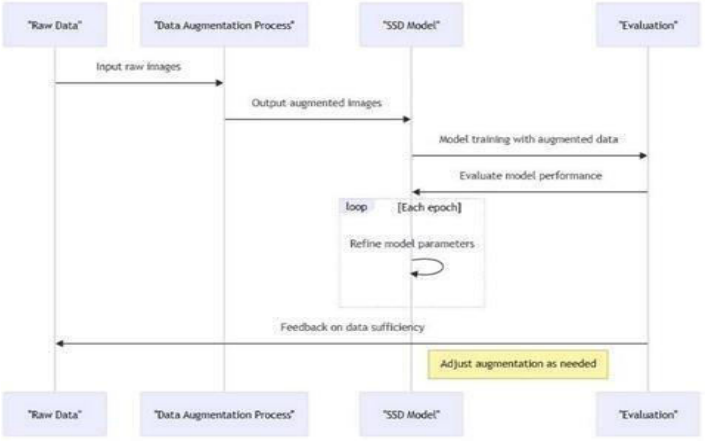


Fig. 6 Schematic representation of the data augmentation workflow within our SSD-based model training process, highlighting the iterative refinement of model parameters and adjustment of data augmentation in response to feedback on data sufficiency.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

Customized Anchor Box Design for Line Detection: Recognizing the limitations of standard anchor boxes for text line segmentation, we introduced custom-designed anchor boxes. These were tailored to align with the unique dimensions of textlines, improving the precision of our SSD-based OCR system. By modifying the anchor boxes using the **mxnet.ndarray.contrib.MultiBoxPrior** function, as inspired by Apache MXNet [10], our model achieved enhanced accuracy in detecting text lines, especially within documents with complex layouts. This custom configuration significantly boosted the performance metrics of our OCR system, demonstrating the effectiveness of our specialized approach to anchor box design.

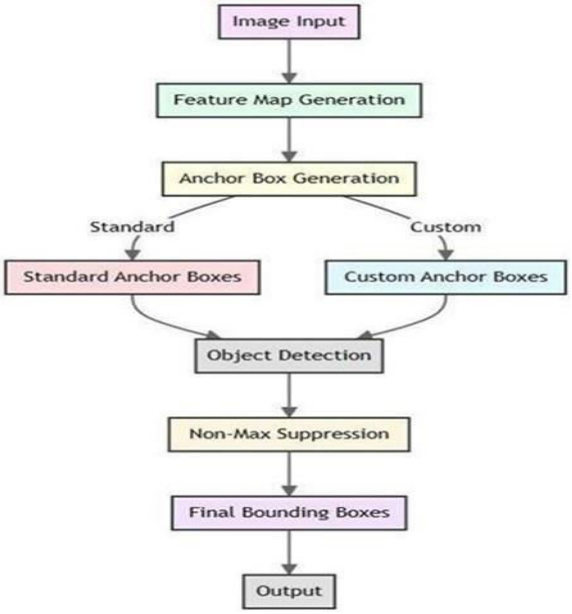


Fig. 7 A flowchart depicting our OCR framework's process from image input to the output, highlighting the bifurcation at the anchor box generation stage where both standard and custom anchor boxes are generated.

SSD Architectural Refinements for Text Recognition: Our enhanced SSD framework is specifically optimized for text recognition, featuring a modified ResNet-34 for feature map generation and a dual-pathway anchor box generation for improved line detection. Custom anchor boxes are designed to capture textual nuances, influenced by recent advances in multi-scale and contextual detection [11][12][13], leading to superior performance in text line accuracy.

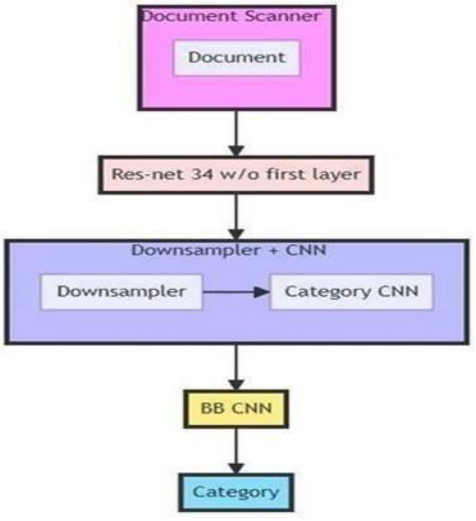


Fig. 8 A flowchart representing our refined SSD architecture, from image input through feature map generation, anchor box bifurcation, objectdetection, and final classification.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

Selective Refinement with Non-Maximum Suppression: We've fine-tuned our SSD model with Non-Maximum Suppression (NMS) todifferentiate between text lines effectively and reduce overlap. Our architecture's object detection is further refined by a bifurcated anchor box approach and dense feature fusion, ensuring precise text categorization.

Training and Optimization Strategy: The training of our model was meticulously planned, starting with MSE reduction and evolving towards IoU optimization. Advanced techniques like cyclical learning rates and early stopping were employed to enhance learning efficiency and prevent overfitting.

Enhanced Word-Level Detection: Our model's ability to detect individual words within lines has been significantly improved by introducing a 'Word SSD' adaptation. This specialized enhancement is critical for effectively isolating words in complex text arrangements.

Benchmarking through Iterative Evaluation: Our model underwent iterative benchmarking, using Intersection over Union (IoU) for quantitative evaluation and visual reviews for qualitative insights, leading to its ongoing enhancement. It achieved a mean Character Error Rate (CER) of 8.4 on pre- segmented lines, demonstrating robust text recognition, and a full pipeline mean CER of 11.6, guidingfuture refinements [2].

Results and Performance Insights: Our model reached a Training IoU of 0.593 and a Test IoU of 0.573. Qualitative analysis, particularly from NMS visuals (Figure adapted from [10]), indicated areas for improvement like correcting mislabeled smudges andmisaligned boxes, to boost text detection accuracy.

C. HANDWRITTEN TEXT RECOGNITION SYSTEM

Our system begins with an image of handwritten text, utilizing a CNN-BiLSTM model to produce a string of recognized text by extracting character probabilities. This model adeptly converts the complexities of handwritten script into digital text, significantly outperforming traditional methods in offline English handwriting recognition [4].

Enhanced Handwriting Detection with CNN-BiLSTM: The system's core utilizes a hybrid CNN-BiLSTM architecture, where CNN extracts image features and BiLSTM processes the sequential nature of handwriting, enhancing the accuracy of character sequence predictions.

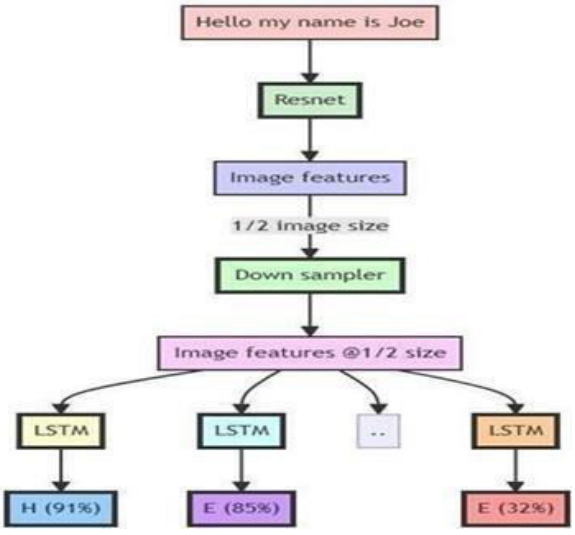


Figure 9: This figure shows the feature extraction and reduction workflow using ResNet and a down-sampler, followed by LSTM networks for character prediction with confidence scores.

Probabilistic Character Prediction: LSTM networks yield probabilistic predictions for each character, crucial in handling the variability of handwriting styles and enhancing the system's robustness [6][9].

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

Challenges in Handwritten Text Recognition: We address several challenges, including complex alignment modeling and the need for streamlined transcription. Our method employs Connectionist Temporal Classification (CTC) loss to align input images with target text,enhancing model alignment accuracy.

Sequence Decoding Using Beam Search: Our methodology incorporates beam search during inference, efficiently exploring potential sequences and improving transcription quality. This approach is particularly effective with models trained using CTC loss.

Lexicon-Enhanced Text Decoding: We introduce Lexicon- Enhanced Text Decoding, which begins with neural network output and includes steps like addressing contractions, tokenization, and lexicon-driven insight for word alternatives, significantly augmenting transcription accuracy.

Language Model-Enhanced Text Decoding: Incorporating a pre-trained language model, Our approach integrates the 'awd_lstm_lm_1150' pre-trained language model into our handwriting recognition system, Language Model-Enhanced Text Decoding. This method starts with tokenization, advancing to token ID conversion.

It then leverages the language model for a deeper understanding of the text, using perplexity assessment to choose the sequence with the lowest value. This incorporation enhances both accuracy and contextual comprehension, adeptly interpreting complex handwritten nuances and uniting language modeling with recognition precision.

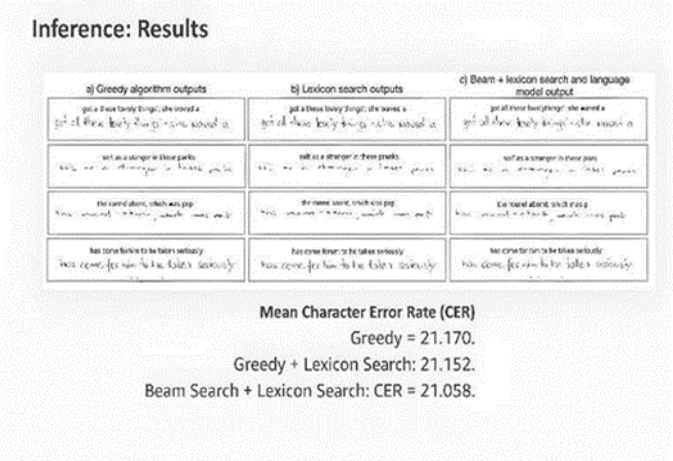


Fig. 10: Text Decoding Method Comparison Compares Greedy Algorithm, Lexicon Search, and Beam Search with Language Model in handwritten recognition. Beam Search with Language Model achieves the lowest CER (21.058), based on AWS Labs and Apache MXNet methods [2, 10].

IV. EXPERIMENT AND RESULT

Model Training and Precision Enhancements: Utilizing the IAM Handwriting Database, our model underwent rigorous training, focusing on Mean Square Error (MSE) and Intersection over Union (IoU) optimizations, achieving a refined Training IoU of 0.593 anda Test IoU of 0.573.

Advanced Line and Word Detection: The SSD model, enhanced for line detection with custom anchor boxes, and the 'Word SSD' for precise word localization, set new accuracy standards in OCR technology.

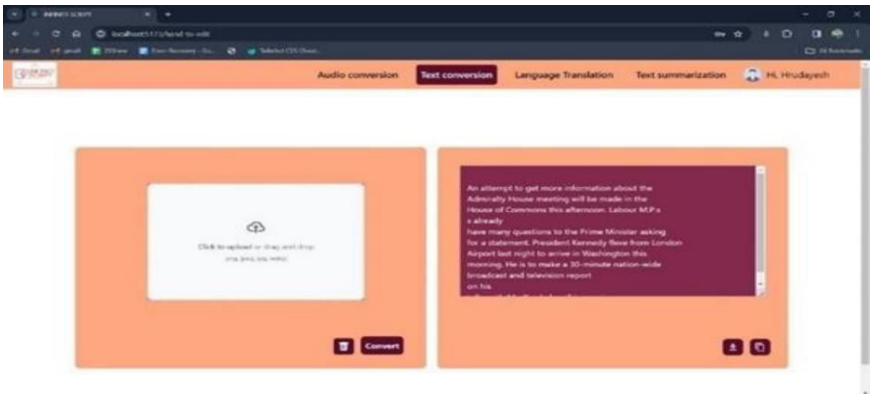
Accuracy Metrics and System Performance: Our system demonstrated strong text recognition capabilities, with a mean Character Error Rate (CER) of 8.4 on pre-segmented lines. Full pipeline testing indicated a higher mean CER of 11.6, directing us toward specific improvement areas.
*Innovative Decoding and Alignment Techniques:*The CNN- BiLSTM system's use of LSTM for probabilistic predictions and CTC loss for sequence alignment, as in Figure 17, contributed significantly to accuracy. A comparative analysis of decoding methods showed the Beam Search with Lexicon Search and Language Model method to be most effective, achieving the lowest CER at 21.058.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

Interface Integration: The interface, essential for user experience in OCR processing, will be illustrated in the research paper, showingthe practicality and efficiency of the system in handling user inputs for text conversion tasks.



V. CONCLUSION

Our research marks a notable advancement in Optical Character Recognition (OCR) technology, uniquely combining Convolutional Neural Networks (CNN), single- shot multi-box detector (SSD) methodologies, and innovative techniques in handwriting recognition. This study has successfully addressed the complex challenges of OCR by implementing a comprehensive approach that includes data compilation, tailored image enhancement, intricate line and page segmentation, and meticulous model training.

The integration of the IAM Handwriting Database enriched our dataset with diverse handwriting styles, crucial for the model's adaptability. The application of advanced algorithms like MSER and CNN significantly improved the accuracy of character recognition and page segmentation.

Our research's cornerstone was the introduction of custom anchor box designs in the SSD framework, which greatly enhanced line segmentation and text localization accuracy. The model's training and refinement process, involving iterative tuning of Mean Square Error (MSE) and Intersection over Union (IoU) metrics, ensured high accuracy in character bounding box predictions and text localization. Our innovations, including down-sampling for computational efficiency and the use of LSTM networks for probabilistic character prediction, have markedly improved the system's performance.

Furthermore, the inclusion of Connectionist Temporal Classification (CTC) loss and sequence decoding with beam search has refined transcription accuracy in our handwriting recognition system. Lexicon-enhanced and Language Model Enhanced Text Decoding methodologies have been crucial in enhancing the precision and contextual understanding of the recognized text. These methods, alongside advanced decoding techniques and linguistic insights, have significantly reduced the Character Error Rate (CER) in our evaluations.

In conclusion, this research presents a highly refined, efficient, and accurate OCR system, setting a new benchmark in handwriting recognition technology. The integration of CNN, SSD, and advanced neural network architectures, combined with our systematic approach to data handling and model training, has not only improved text recognition capabilities but also widened the scope for practical applications in various sectors. This study, underlining the importance of continuous innovation in AI and pattern recognition, paves the way for future advancements in OCR and contributes significantly to Our study represents a significant leap forward in Optical Character Recognition (OCR) by integrating Convolutional Neural Networks (CNN) with single-shot multi-box detector (SSD) methods and novel handwriting recognition techniques. We tackled OCR's challenges through comprehensive data compilation from the IAM Handwriting Database and sophisticated image processing strategies.
The outcome is a refined, efficient, and high-accuracy OCR system that establishes a new standard in handwriting recognition. Our research underscores the critical role of ongoing innovation in AI, providing substantial contributions to natural language processing and computer vision, and opening avenues for practical OCR applications across various sectors.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 3, March 2024 ||

| DOI:10.15680/IJIRSET.2024.1303243 |

VI. ACKNOWLEDGMENT

I am deeply thankful for the substantial support provided by a few key organizations, which was vital to the progression of our research. First, a heartfelt acknowledgment to AWS Labs, whose open-source endeavor in Handwritten Text Recognition within Apache MXNet [2] has been a foundational element in both our understanding and methodology for this research. Equally, my gratitude is extended to the Apache MXNet community. Their informative 'OCR with MXNet Gluon' presentation [10] has been aguiding force in our research process, providing essential insights that have significantly influenced our work additionally, my sincere appreciation goes to the University of Bern. Their generosity in allowing us access to the IAM Handwriting Database [1] has been fundamental in the training and evaluation of our Handwritten Text Recognition models. We have steadfastly adhered tothe principles of academic honesty, meticulously acknowledging all referenced works and stringently steering clear of any plagiaristic practices. The depth and quality of our study have been significantly augmented by the essential support and contributions from AWS Labs, the Apache MXNet community, and the University of Bern. Their aid has been instrumental in elevating our research to its current stature, for which we express our heartfelt gratitude.

REFERENCES

1. IAM Handwriting Database, University of Bern. [Online]. Available:http://www.fki.inf.unibe.ch/databases/iam-handwriting-database.
2. AWS Labs. Handwritten Text Recognition for Apache MXNet. [Online]. Available: https://github.com/awslabs/handwritten- text- recognition-for-apache-mxnet/tree/master.
3. Shonenkov, D. Karachev, M. Novopoltsev, M. Potanin, D. Dimitrov, and A. Chertok, "Handwritten text generation and strikethrough characters augmentation," Computer Optics, vol. 46, no. 3, pp. 1-13, June 2022. doi: 10.18287/2412-6179-co- 1049.
4. F. Kızılırmak and B. Yanıkoğlu, "CNN-BiLSTM model for English Handwriting Recognition: Comprehensive Evaluation on the IAM Dataset," Research Square, Nov. 17, 2022. [Online]. Available: https://doi.org/10.21203/rs.3.rs- 2274499/v1.
5. L. Jiao, H. Wu, H. Wang, and R. Bie, "Text Recovery via Deep CNN- BiLSTM Recognition and Bayesian Inference," in IEEE Access, vol. 6, pp. 76416-76428, 2018, doi: 10.1109/ACCESS.2018.2882592.
6. N. Bhardwaj, "A Research on Handwritten Text Recognition," International Journal of Scientific Research in Engineering and Management, vol. 06, no. 04, Apr. 2022, doi: 10.55041/ijsrem12649.
7. L. Kumari, S. Singh, V. V. S. Rathore, and A. Sharma, "Lexicon and attention-based handwritten text recognition system," Machine Graphics and Vision, vol. 31, no. 1/4, pp. 75–92, Dec. 2022, doi: 10.22630/mgv.2022.31.1.4.
8. Ansari, B. Kaur, M. Rakhra, A. Singh, and D. Singh, "Handwritten Text Recognition using Deep Learning Algorithms," 2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST), Dec. 2022, Published, doi: 10.1109/aist55798.2022.10065348.
9. R. G. Khalkar, A. S. Dikhit, and A. Goel, "Handwritten Text Recognition using Deep Learning (CNN & RNN)," IARJSET, vol. 8, no. 6, pp. 870–881, Jun. 2021, doi: 10.17148/iarjset.2021.861
10. Apache MXNet. (2017). OCR with MXNet Gluon [PowerPoint slides].Available: https://www.slideshare.net/apachemxnet/ocr- with-mxnet- gluon#11.
11. Z. Chen, K. Wu, Y. Li, M. Wang, and W. Li, “SSD-MSN: An Improved Multi-Scale Object Detection Network Based on SSD,” IEEE Access, vol. 7, pp. 80622–80632, 2019, doi: 10.1109/access.2019.2923016.
12. S. Zhai, D. Shang, S. Wang, and S. Dong, “DF-SSD: An Improved SSD Object Detection Algorithm Based on DenseNet and Feature Fusion,” IEEE Access, vol. 8, pp. 24344–24357, 2020, doi: 10.1109/access.2020.2971026.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details