Journal of Informatics, Computer Science, Data Science And Arificial Intelligence
Volume 2 No. 2 Juni 2025, ISSN : 3031-9145

IMPLEMENTATION OF DEEP LEARNING ALGORITHM IN HANDWRITING TO TEXT DOCUMENT CONVERSION APPLICATION

Alif Lufti Santri [1], Khairuddin Nasution [2], Tasliyah Haramini [3]
[1]Universitas Islam Sumatera Utara, aliflutfi1245@gmail.com.
[2]Universitas Islam Sumatera Utara, khairuddin_nst@uisu.ac.id. [3]Universitas Islam Sumatera Utara, tazlie@gmail.com.

ARTICLE INFO

ARTICLE HISTORY:
Received :22 Juni 2025
Revised :25 Juni 2025
Accepted : 27 Juni 2025

Keywords: *Handwriting, recognition, Tesseract.js, OCR, Web Application, Digitalization*

ABSTRACT

The development of information and communication technology has driven the need for systems capable of efficiently converting handwritten text into digital text. This study aims to develop a web-based application capable of real-time handwriting recognition using Tesseract.js, a JavaScript library for optical character recognition (OCR). The application is designed to assist users in converting handwritten documents into editable text formats, thereby enhancing productivity and information accessibility.

The methods used in this study include uploading handwritten images, preprocessing the images to improve input quality, and applying OCR algorithms using Tesseract.js to recognize characters and words. The recognized results are then displayed on the user interface, with an option for manual correction if needed. The study also evaluates the accuracy of the text recognition produced by the application by comparing the recognition results with the original text.

The results show that the developed application is capable of recognizing handwriting with a satisfactory level of accuracy, despite variations in handwriting styles. This application is expected to make a significant contribution in the field of document digitization and data processing, and serve as a reference for the development of similar systems in the future.

INTRODUCTION

In the rapidly advancing era of digitalization, the need to convert information from analog to digital formats has become increasingly important. The conventional process of converting handwritten text into digital text still relies on manual retyping, which is not only time-consuming but also prone to human error.

Handwriting is the most natural and personal form of written communication for humans. Although keyboard technology and digital input devices have advanced rapidly, handwriting still plays an important role in various aspects of life, such as in education, medicine, law, and administration. Important documents such as medical records, application forms, written examinations, and historical archives are often still in handwritten form, requiring digitization for storage, retrieval, and data processing purposes (Memon et al., 2020).

The development of Artificial Intelligence (AI) and Machine Learning technologies has opened up significant opportunities to automate the handwriting recognition process. Techniques such as Deep Learning, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) have shown promising results in recognizing complex patterns in human handwriting (Li, 2024).

In Indonesia, the need for handwriting recognition technology is increasing in line with digitalization efforts across various sectors. Government institutions, educational organizations, and private companies possess thousands—even millions—of handwritten documents that need to be digitized. The manual process that has been used so far not only requires substantial human resources but also results in varying levels of accuracy and lengthy processing times (Ash, 2024).

An effective and accurate handwriting recognition application can significantly enhance the efficiency of document digitization processes. Moreover, this technology can be integrated with various information systems to support business process automation, such as form processing, archive digitization, and the development of mobile applications capable of recognizing handwriting in real time.

An application is a software program designed to run on a specific system, aimed at facilitating various activities carried out by humans (Huda & Priyatna, 2019). An image is a simple two-dimensional medium displayed on an opaque surface, and it can be developed in various forms through a combination of visual elements (images) and verbal elements (words) (Hermawan, 2013).

Image is a two-dimensional artifact that either records the visual appearance of physical objects, like photographs, or provides a visual representation of concepts or artificial data, like graphics or synthetic images (Ten Brinke et al., 2020).

Handwriting recognition stands out as a prominent field of study within artificial intelligence. OCR is a field that empowers the translation of various document types

and images into data that can be analyzed, edited, and searched. Therefore, it offers users a convenient means of preserving handwritten content without manual text input, ultimately saving time and effort in storing and retrieving documents (Alhamad et al., 2024).

Tesseract.js is a WebAssembly-based JavaScript library that serves as a port of the Tesseract OCR (Optical Character Recognition) engine. This library enables text recognition from images directly within browser environments or on the server side (Node.js), without the need for additional software installation. Tesseract.js supports more than 100 languages, automatic text orientation detection, and provides a simple programming interface for reading words, paragraphs, and character boundaries within an image (Rakshit et al., 2020).

Deep learning is a particular kind of machine learning that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts… Deep learning is based on artificial neural networks, a type of computing system roughly mimicking the biological neural networks of animal brains. Representation learning sets deep learning apart from other techniques in ML. Unlike manual feature engineering, feature learning enables computers to automatically find the representations required for classification from raw data (Mishra et al., 2021).

Deep learning is a specialized form of machine learning that uses several processing layers to learn complex abstractions (concepts) in data by building a hierarchy of abstractions, wherein each level of abstraction is created using the lower-level abstractions learned by the preceding layer in the hierarchy… Unlike traditional machine learning, deep learning does not require manual feature engineering prior to model development because of their capability to perform hierarchical feature learning, where higher-level features are defined in terms of lower-level features (Misra & Li, 2019).

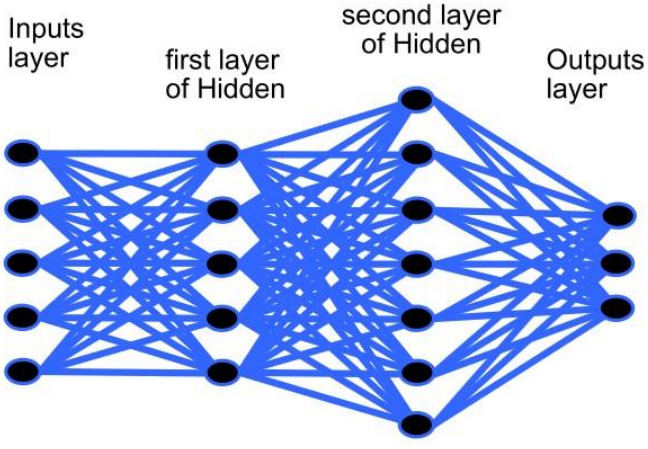


Figure 1. Architecture Illustration in Deep Learning

Convolutional Neural Network (CNN) is a classification method in machine learning that assigns labels to data during the training process, thus falling under the category of supervised learning. Generally, the architecture of CNN consists of two main components: the feature extraction part and the classification part. The feature extraction is performed through convolutional layers, while the classification is carried out using multilayer perceptrons (Ghandi & Ramadhan, 2024).

Convolutional Neural Network (CNN) was developed as an extension of the multilayer perceptron (MLP), specifically designed to handle two-dimensional data such as images. CNN has proven to be highly efficient in recognizing patterns in images and has been applied in various fields, including facial recognition, object detection, and handwriting recognition. In CNN, the convolutional layer is responsible for extracting important features from images by applying convolution operations on the image pixels. This operation uses filters or kernels that slide across the image to generate feature maps. These feature maps are then used as input to the classification layer, which may consist of fully connected layers or multilayer perceptrons to perform classification based on the extracted features. During the training process, the model's weights and parameters are updated iteratively using backpropagation algorithms and gradient optimization to minimize prediction errors (et al., 2021).

RESEARCH METHODS

This study employs a quantitative experimental approach using software engineering methods. The main objective of this research is to develop and test a deep learning-based application capable of automatically converting handwritten text into digital text.

The study implements a Convolutional Neural Network (CNN) model for feature extraction from handwritten images, combined with a Recurrent Neural Network (RNN) such as LSTM to recognize character sequences, or utilizes an end-to-end model such as a Convolutional Recurrent Neural Network (CRNN) or a Transformer-based model.

Example:

Implementation of Deep Learning Algorithm in a Handwriting-to-Text Document Conversion Application

Input Image: “NAMA SAYA ALIF”

Convolutional Layer: Detects features such as edges, corners, and patterns.

Feature Extraction: Uses shape analysis, contour detection, or CNN models to capture relevant patterns.

Fully Connected Layer: Classifies the extracted features into characters or words.

CNN Processing:

- Frame 1: “N”
- Frame 2: “A”
- Frame 3: “M”
- Frame 4: “A”
- Frame 5: “S”
- Frame 6: “A”
- Frame 7: “Y”
- Frame 8: “A”
- Frame 9: “A”
- Frame 10: “L”
- Frame 11: “I”
- Frame 12: “F”

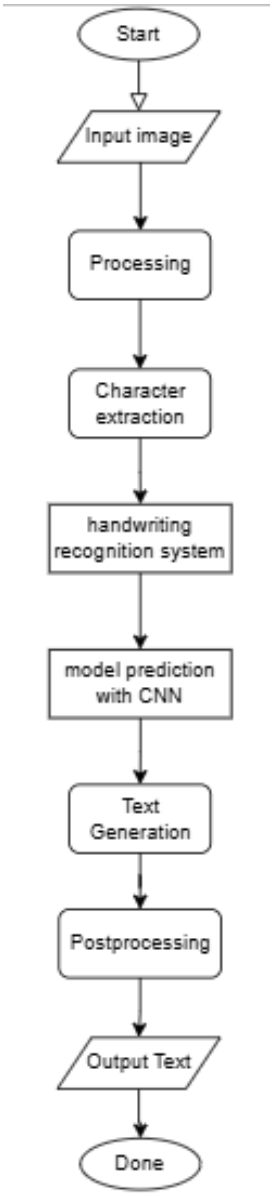
Decoding: Processes the conversion of “N”, “A”, “M”, “A”, “S”, “A”, “Y”, “A”, “A”, “L”, “I”, “F” and the character sequence.

Post-Processing: Combines and corrects the text using a language model or spelling rules.

Output: Displays the final recognized text from the application.

The system developed in this study is an artificial intelligence (AI)-based application capable of automatically converting handwritten text into digital text. This system uses deep learning algorithms to perform the extraction, identification, and translation of handwritten characters into digital text that can be stored or further edited by the user.

The application aims to accelerate the digitization process of handwritten documents, such as lecture notes, manual forms, archives, and others, which are traditionally processed manually (by retyping). By leveraging deep learning technology, the system is designed to accurately and efficiently recognize various styles of handwriting. In this Handwriting To Text Online application, a flowchart is used to represent the workflow of the system. The following is the design flowchart for the handwriting recognition application:



RESULTS AND DISCUSSION

On the main screen of the Handwriting To Text application, the user clicks the “Choose File” button to upload an image file in JPG or JPEG format. In addition to the “Choose File” button, the application also features “Start” and “Refresh” buttons, which are used to begin the process and to restart the application, respectively.

After that, when the user selects an image file in JPG or JPEG format from a folder on their PC or laptop—for example, I will choose one of the image files.

On the screen, once the user has selected an image file, the file name and the chosen image or photo to be converted will appear next to the “Choose File” button. A preview of the image will also be displayed before it is processed by the application for conversion.

Next is the display when the user clicks the “Start” button located below the image preview. When the “Start” button is pressed, the application will automatically process the image, and the progress bar will run from 0% to 100%. Once it reaches 100%, the result will be displayed in the “Result” section.

Next is the result display using the example handwritten image above. Once the process is complete, the converted text will appear in the “Result” section. The text can then be copied and pasted into Word or other applications.

When the user clicks the Refresh button, the display will return to its initial state.

CONCLUSION

The conclusion of the Handwriting To Text application is that it is an online application designed to convert scanned handwriting into text on a computer.

This application uses Handwriting Recognition, which is a part of OCR (Optical Character Recognition), to automatically convert handwriting into text. OCR greatly assists in scanning characters from images and makes the process more versatile.

In the development of this handwriting conversion application, a Deep Learning algorithm is also used. This algorithm includes several methods, one of which is Convolutional Neural Networks (CNN). CNN is highly suitable for handwriting conversion applications due to its ability to identify character and number patterns with high accuracy. Convolutional Neural Networks (CNN) are used to convert handwritten images into machine-readable text.

REFERENCES

Alhamad, H. A., Shehab, M., Shambour, M. K. Y., Abu-Hashem, M. A., Abuthawabeh, A., Al-Aqrabi, H., Daoud, M. S., & Shannaq, F. B. (2024). Handwritten Recognition Techniques: A Comprehensive Review. *Symmetry*, *16*(6), 1–25. https://doi.org/10.3390/sym16060681

Ash, G. (2024). *Perancangan Website Text Scanner Untuk Konversi Handwritten ke Teks Digital Dengan Menggunakan Optical Character Recognition*. *10*(3), 854–

860.

Elngar, A. A., Arafa, M., Fathy, A., Moustafa, B., Mahmoud, O., Shaban, M., & Fawzy, N. (2021). Image Classification Based On CNN: A Survey. *Journal of Cybersecurity and Information Management*, *6*(1), PP. 18-50. https://doi.org/10.54216/jcim.060102

Ghandi, S., & Ramadhan, Y. R. (2024). *PENERAPAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN) DALAM APLIKASI PENDETEKSI PENYAKIT DAUN TANAMAN KENTANG BERBASIS ANDROID*. *8*(5), 8701–8708.

Hermawan, B. (2013). Multimodality: Menfsir Verbal, Membaca Gambar, dan Memahami Teks. *Jurnal Pendidikan Bahasa Dan Sastra*, *13*(2), 102–117. https://doi.org/10.17509/bs

Huda, B., & Priyatna, B. (2019). Penggunaan Aplikasi Content Management System (CMS) Untuk Pengembangan Bisnis Berbasis E-commerce. *Systematics*, *1*(2), 81. https://doi.org/10.35706/sys.v1i2.2076

Li, R. (2024). *A review of neural networks in handwritten character recognition*. *0*, 169–174. https://doi.org/10.54254/2755-2721/92/20241736

Memon, J., Sami, M., Khan, R. A., & Uddin, M. (2020). Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR). *IEEE Access*, *8*, 142642–142668. https://doi.org/10.1109/ACCESS.2020.3012542

Mishra, R. K., Reddy, G. Y. S., & Pathak, H. (2021). The Understanding of Deep Learning: A Comprehensive Review. *Mathematical Problems in Engineering*, *2021*. https://doi.org/10.1155/2021/5548884

Misra, S., & Li, H. (2019). Deep neural network architectures to approximate the fluid-filled pore size distributions of subsurface geological formations. In *Machine Learning for Subsurface Characterization*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-817736-5.00007-7

Rakshit, S., Kundu, A., Maity, M., Mandal, S., Sarkar, S., & Basu, S. (2020). *Recognition of handwritten Roman Numerals using Tesseract open source OCR engine*. 1–5. http://arxiv.org/abs/1003.5898

Ten Brinke, W., Squire, D. M. G., & Bigelow, J. (2020). The Meaning of an Image in Content-Based Image Retrieval. *CEUR Workshop Proceedings*, *240*(January 2020), 710–719.