HANDWRITTEN DIGIT RECOGNITION

BY

STEPHEN GIFT AKINPELU : 101151897

TOBECHI NWACHUKWU : 101186606

KENDALL DAZE-YACH : 101139299

GROUP 47

ABSTRACT

In this paper, we aim to develop a model that demonstrates how multi-staged AI architectures can improve overall accuracy, especially in the case of varied or unexpected datapoints. The primary AI model is designed to detect handwritten digits from the MNIST dataset, and the secondary architecture is adapted to allow for different variations within the dataset, such as flipped or blurred images. Using this two-step approach demonstrates the AI model's ability to handle different situations and improves its efficiency.

INTRODUCTION

In the digital world, a lot of individuals are transitioning from handwriting to a digital form of communication. However, we still all have those personal notes and letters that offer a sense of sincerity to us. Our handwritten project aims to bridge the gap between the traditional art of handwriting and the digital model by offering a fresh perspective for those who would like to digitally preserve their old handwritten documents, letters, or forms of any kind. Our initiative offers a cost-effective form of storage for their physical files.

A significant challenge we faced arose from the variability of handwritten digits, which can arise from factors such as noise (blurring, flipping), paper conditions, and the diversity of handwriting. To address this, we opted for an architecture that uses two datasets: the MNIST dataset and datasets that have been distorted by noise. This approach allows us to answer the question: Is it better to train an AI specifically for a dataset or can we streamline using pre-processing functions?

BACKGROUND

Our project leverages the MNIST dataset as its own benchmark. The MNIST dataset, which was created in 1994, comprises of digits handwritten by high school students and employees of the United States Census Bureau. As of 2018, researchers announced a 0.18% error rate when testing the convolutional neural network's best performance on the MNIST dataset, displaying the enduring relevance of the MNIST dataset (Kowsari, May. 2024). This breakthrough allows us to narrow down the best architecture for our tests.

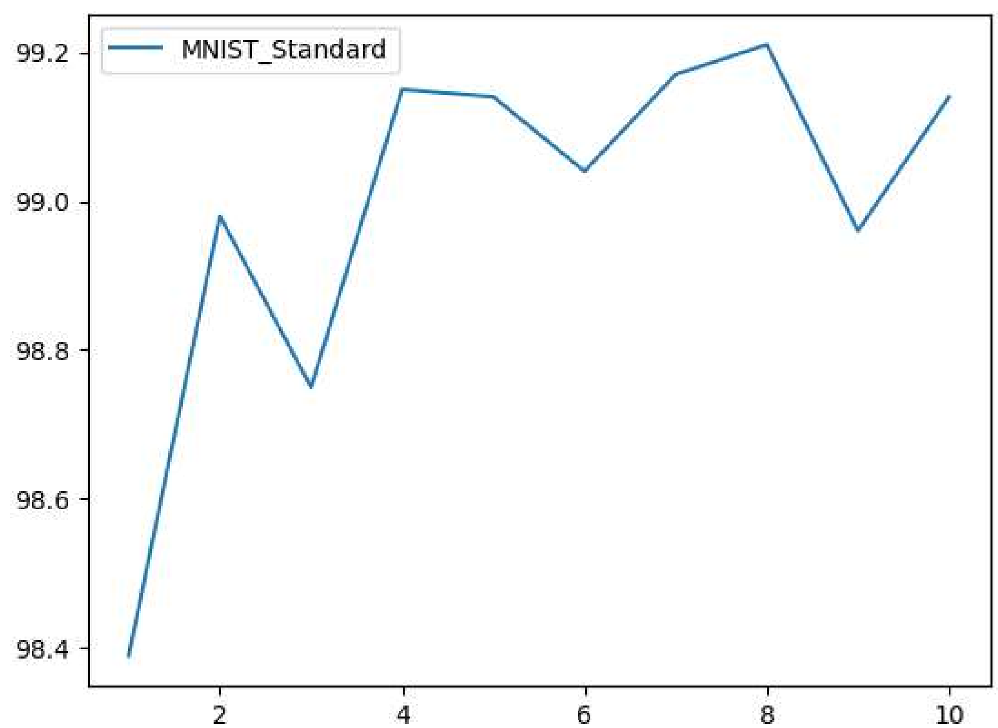
The simplicity of the MNIST dataset eliminates the need for creating a new, clean dataset of digits from scratch. The MNIST dataset test bed allows us to focus on improving the accuracy of our architecture without data pre-processing functions.

METHODOLOGY

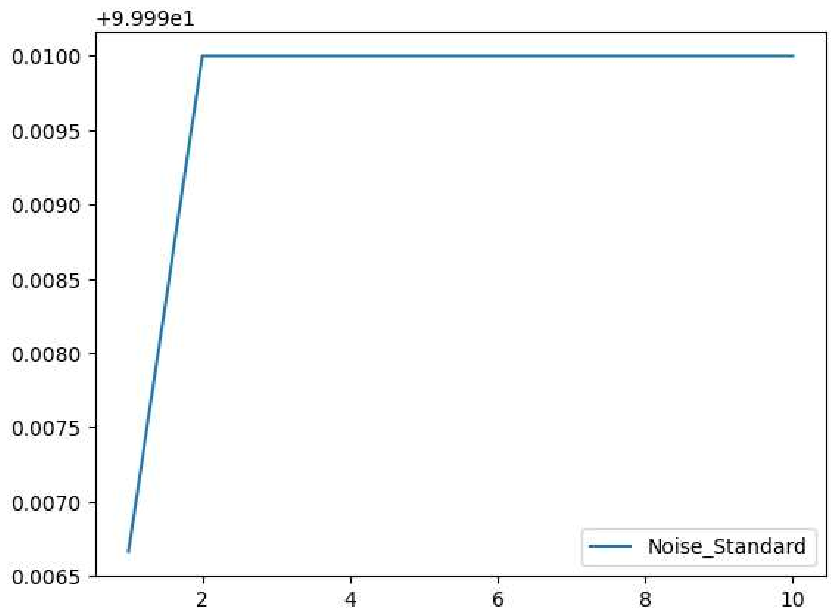
The approach for implementing this project had several stages. The first stage consisted of building an architecture that was able to correctly identify the digits in the MNIST dataset. This started with training the model against the dataset, and then thorough testing was carried out to evaluate the accuracy of the model. For this stage, an accuracy average of 99.301% was attained. The second stage involved building an architecture that could determine whether an image had some noise. Noise in the context of this project, simply means an initial image being altered, either by flipping or blurring the image. The MNIST dataset was preprocessed before testing this architecture. Different images in the dataset were either flipped or blurred at random, to create a new dataset, and this is what the second architecture was trained on. Putting these two architectures together, we had "created" a third architecture capable of identifying a more diverse set of images than in the original MNIST dataset. The steps for identifying an image would consist of the second architecture determining whether the image was flipped or blurred, applying an appropriate filter to reverse the noise, and then passing the resulting image into the first architecture for recognition. From tests performed, this multi-stage architecture performed better at recognizing a more diverse data set than each of the pieces of the architecture.

RESULTS

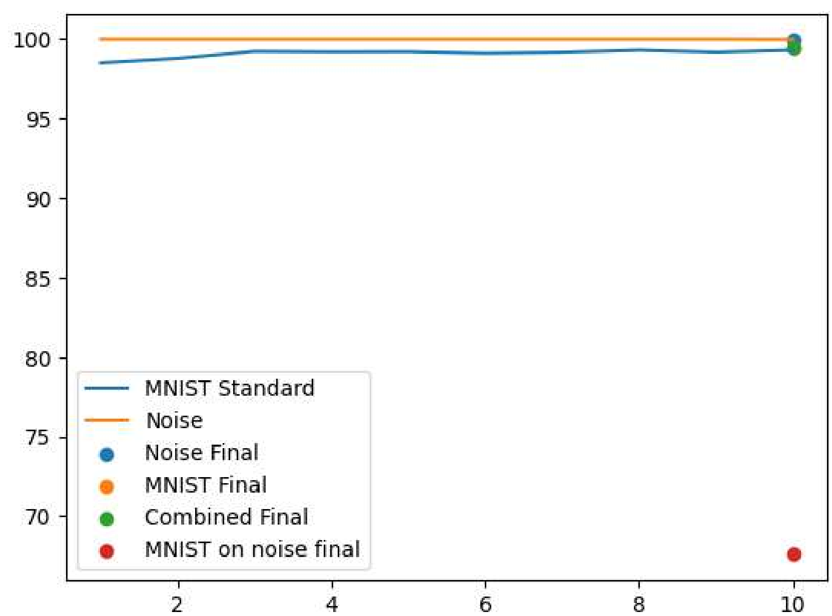
The results from testing the multi-staged architecture and comparing its accuracy to that of the standalone architectures proved that having a multi-staged architecture, especially in this context, improved accuracy. The plot below shows the accuracy of the first architecture, which was strictly for identifying digits from the MNIST dataset. This performed with a high accuracy of 99.301%.



The second plot shows the results from the second architecture, which has the sole responsibility of determining what type of noise an image has (flip or blur). It also had a high accuracy, and the results are indicative of that. (Pay close attention to plot scale).

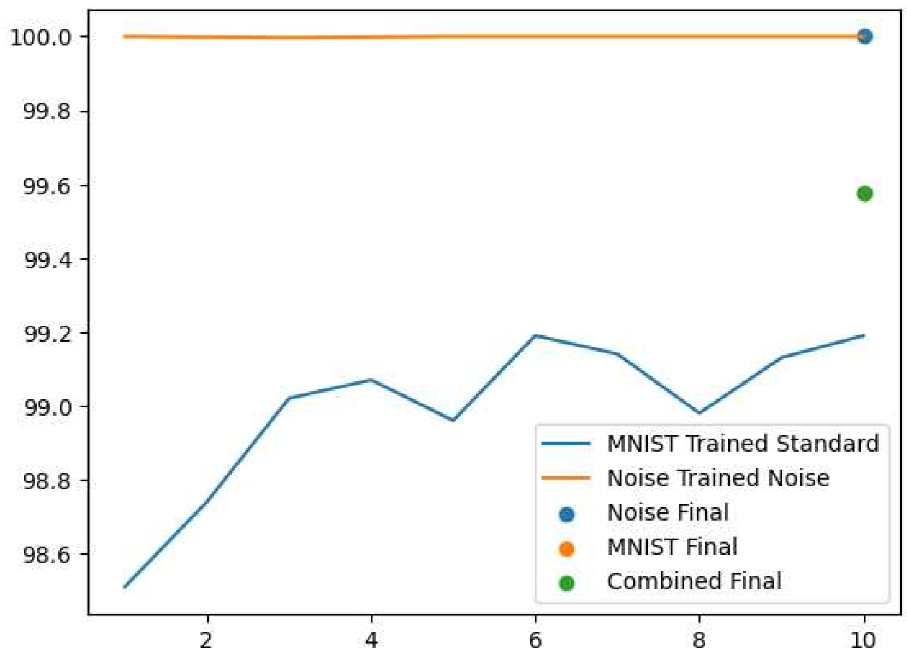


The next plot shows the results from putting the first architecture, which was strictly for recognition of images in the MNIST dataset, up against the preprocessed dataset consisting of images with noise. As we can see, the performance dropped off significantly compared to the other tests, and it had approximately 2 out of every 3 correct (~67% accuracy). *(Red Dot)*



FIGURES CONT'D

The last plot then shows the results from combining both architectures, and this architecture was able to accurately predict whether the images contained noise and recognize the digits, achieving high accuracy in both tasks with an overall accuracy of approximately 99.58%.



CONCLUSION

This project represents a significant milestone in the recognition of handwritten texts or extracts into digital form, as mentioned in the introduction section. By successfully varying the initial dataset and developing an architecture capable of handling additional edge cases such as noise, we have been able to lay a strong foundation for further progress. Handwritten texts contain a combination of digits and letters, making it essential to extend our capabilities beyond digit recognition alone. This project serves as a crucial first step towards achieving that goal.

To further advance towards converting handwritten texts into digital form, the next logical step would be to implement handwritten letter recognition. Recognizing handwritten letters poses unique challenges due to diverse writing styles. Building upon the architecture developed in this project, future efforts can focus on training models specifically designed to accurately identify and classify handwritten letters.