ECE/CS/ME 539 Introduction to Artificial Neural Networks

Project Progress Report

# Multi-Language Alphabetical Classifier

Team # : 11

Team Members:

Caleb Federman, Department of Electrical and Computer Engineering Undergraduate

Elias Cassis, Department of Electrical and Computer Engineering Undergraduate

Connor Braun, Department of Computer Sciences Undergraduate

Tate Waugh, Department of Computer Sciences Undergraduate

[Github](https://github.com/calebfederman/handwritten-alphabet-language-classifier)

ABSTRACT

Our project successfully initiated the development of a Multi-Language Alphabetical Classifier using a convolutional neural network (CNN). We processed a dataset of 6,831 handwritten English characters, converting them into a uniform grayscale format and resizing them for neural network compatibility. The CNN model was trained and tested, achieving an impressive accuracy of 90.17% on the test set. This foundational phase sets the stage for future expansion to include multiple languages, aiming to recognize characters from diverse alphabets and identify their corresponding languages.

**Introduction**

Overview

Our objective is to develop a machine learning framework adept at recognizing and classifying characters from diverse alphabets, subsequently identifying the associated language. By employing a convolutional neural network, the system will be trained on a comprehensive dataset of written characters, striving for proficient classification that can accurately discern the alphabet and language. Furthermore, the framework will encompass a component capable of analyzing words, where it leverages the certainty levels of individual character classifications, along with contextual cues like word length and dictionary references, to ascertain the language, and thus the individual characters, of the entire word.

Motivations

The project is driven by the challenge of recognizing diverse alphabetic characters and the broader goal of facilitating automated translation and text processing in multiple languages.

Significance

Our work contributes to the field of OCR, particularly in enhancing the accuracy of multilingual character recognition.

Related Works

#### Relevant works related to our implementation:

OCR Systems:

* [Comprehensive overview of OCR](https://ieeexplore.ieee.org/document/9151144#:~:text=Given%20the%20ubiquity%20of%20handwritten,intelligence%2Fmachine%20learning%20tools%20to)1
* [Survey on OCR systems](https://arxiv.org/pdf/1710.05703.pdf#:~:text=A%20Survey%20on%20Optical%20Character,the%20capability%20to%20very%20easily)2
* [OCR System Mechanism](https://link.springer.com/chapter/10.1007/978-3-319-50252-6_2#:~:text=Optical%20character%20recognition%20%28OCR%29%20,increasing%20attention%20in%20both)3
* [OCR for text recognition](https://ieeexplore.ieee.org/document/9935961/#:~:text=From%20this%20comes%20a%20need,accuracy)4

Cross-Language Character Recognition

* [Cross-Lingual Handwritten character recognition](https://www.sciencedirect.com/science/article/pii/S0167865522001490#:~:text=,31%2C%2034)5
* [Cross-Lingual Learning for Text Processing](https://www.sciencedirect.com/science/article/pii/S0957417420305893#:~:text=%23%20%E3%80%901%E2%80%A0Cross,it%20was%20not%20possible%20previously)6
* [Script Recognition in Multi-Script Documents](https://dl.acm.org/doi/10.1145/3396167#:~:text=This%20literature%20examines%20the%20Script,handwritten%2C%20Nandinagari%2C%20and%20Hebrew%2C%20wh)7
* [Cross-Lingual Text Image Recognition](https://www.researchgate.net/publication/361356717_Cross-Lingual_Text_Image_Recognition_via_Multi-Hierarchy_Cross-Modal_Mimic#:~:text=Cross,Institute%20of%20Automation%20of%20Chinese)8

Advanced Neural Network Architectures

* Our lecture notes from class
* Specifically, Exercise 25 & 26 based on CNNs
* Specifically, Exercise 12 based on One-Hot Encoding

**Method**

Data

We utilized a dataset of handwritten English characters. Additionally, we preprocessed the 6,831 images by converting them to grayscale and resizing them to 28x28 pixels.We then performed some affine transformations on some images in order to broaden the scope of the model.

Algorithm and Program

Developed a CNN with three convolutional layers, ReLU activations, pooling, and dropout layers. Adjusted the model to optimize the character recognition task. Additionally, we implemented one-hot encoding for character classification.

Platform

Employed Google Colab for development due to its collaborative features and access to GPU resources.

Experiments

Conducted training and validation experiments, focusing on model accuracy and loss metrics.

Results

The model demonstrated effective learning capabilities with an increasing accuracy trend over training epochs. As it improved consistently, we were able to peak accuracy at 90.17% on the test set after 30 epochs.

Evaluation

Evaluated the model using accuracy metrics and visual inspection of its predictions. Both of these are demonstrated through our printed output as well as the created graphs (shown below).

**Results**

Conditions

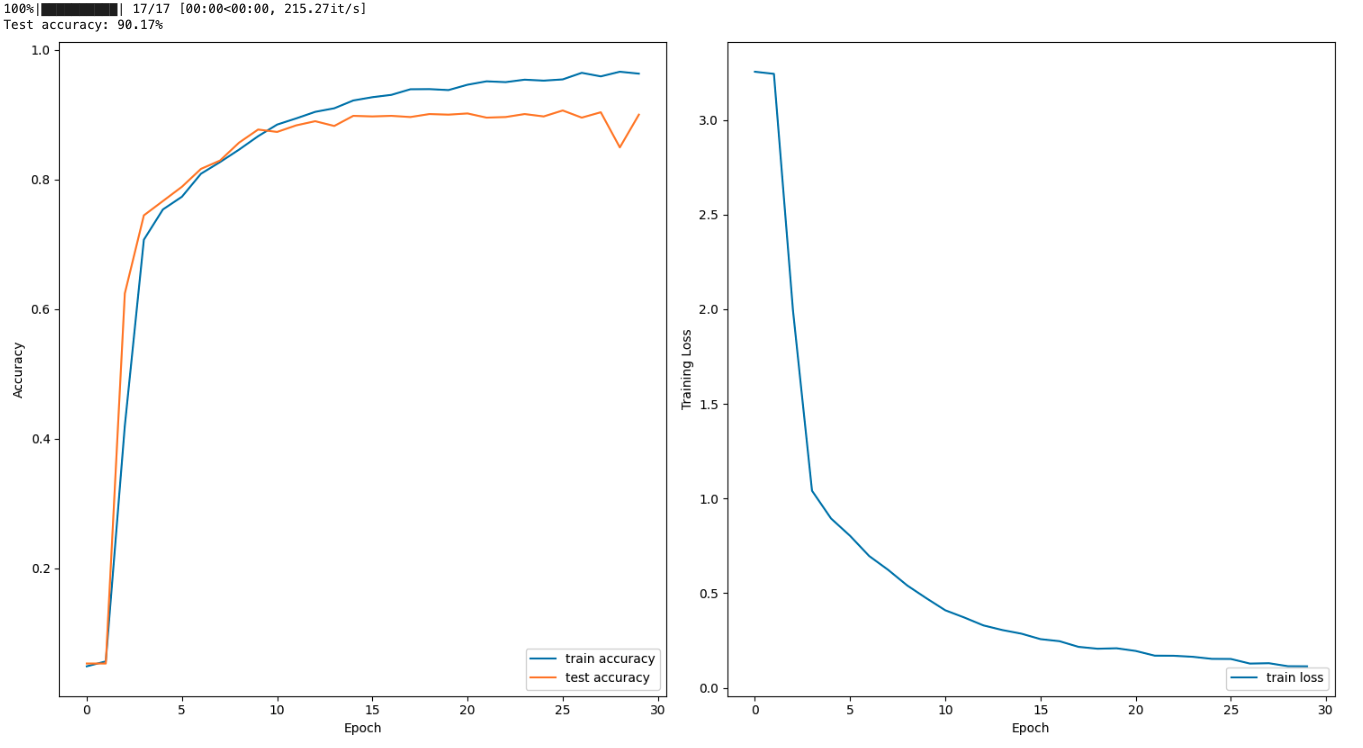
We employed the Google Colab environment with Python and PyTorch. Our dataset consisted of the English alphabet with 6,831 images. We additionally used templates of CNN training based on our in-class exercise from 12 (One-hot encoding), 25 & 26 (CNN training).

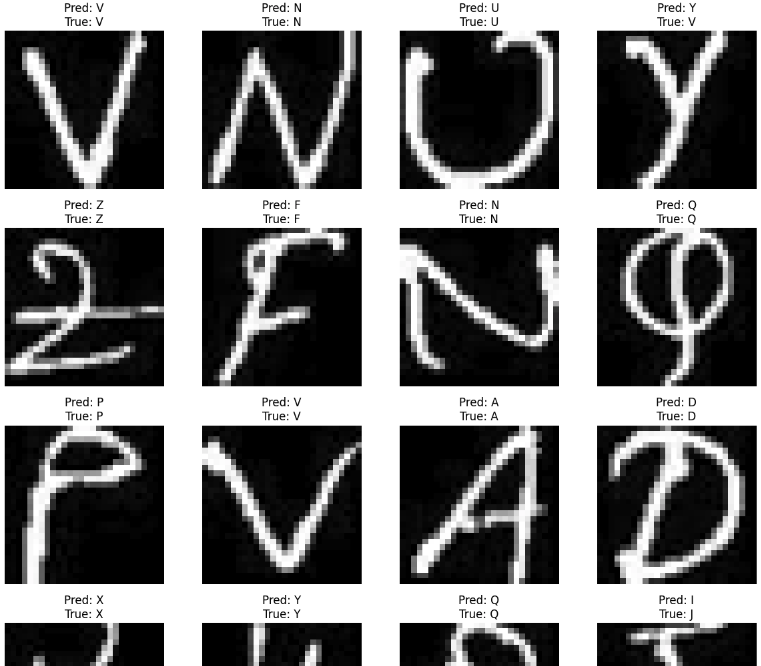
Results

Our program yielded an accuracy that increased from about 4.9% to 90.17% over 30 epochs.. Additionally, the training loss decreased from 3.225 to 0.114, indicating effective learning and model convergence

Quality

Our program yielded high accuracy in character recognition, with correct classifications in 15/16 visual test examples as shown below. Additionally, we experienced consistent performance improvement, validating the model’s design and training approach.





(\*Note: the dataset indicated that the true value of “Y” was “V”. However, our training still proved accurate as it predicted it’s value to be “Y”)

**Discussion**

Based on our initial results, the model shows promise in recognizing English characters, aligning with our objective of developing a multilingual character classifier. Challenges included resolving technical issues and optimizing the CNN architecture. Specifically, we had to work to ensure the shapes of the layers were being translated correctly; this took up most of our troubleshooting time. Additionally, other challenges may occur as we work with different types of datasets that aren’t as well prepared as the English dataset. This may lead to more preprocessing work before we can get to the training and testing phase. Future work involves expanding the database to include other languages and enhancing the model’s ability to discern language context and improve recognition accuracy. We are excited to continue with our training as we will be able to use our english training as an effective template for the rest of the languages that we will be training on. Once our letter classification is performing significantly well enough, we can add in next word prediction possibly utilizing a RNN. Thus our final result should be able to predict letters/language, then predict possible next words.

**References**

1 J. Memon, M. Sami, R. A. Khan and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," in *IEEE Access*, vol. 8, pp. 142642-142668, 2020, doi: 10.1109/ACCESS.2020.3012542.

2 Islam, Noman, et al. “A Survey on Optical Character Recognition System.” Rnal of Information & Communication Technology-JICT, vol. 10, no. 2, Jan. 2016, pp. 4–7. jms.ilmauniversity.edu.pk/index.php/JICT/article/download/641/362.

3 Chaudhuri, A., Mandaviya, K., Badelia, P., Ghosh, S.K. (2017). Optical Character Recognition Systems. In: Optical Character Recognition Systems for Different Languages with Soft Computing. Studies in Fuzziness and Soft Computing, vol 352. Springer, Cham. <https://doi.org/10.1007/978-3-319-50252-6_2>

4 R. Avyodri, S. Lukas and H. Tjahyadi, "Optical Character Recognition (OCR) for Text Recognition and its Post-Processing Method: A Literature Review," *2022 1st International Conference on Technology Innovation and Its Applications (ICTIIA)*, Tangerang, Indonesia, 2022, pp. 1-6, doi: 10.1109/ICTIIA54654.2022.9935961.

5 Nirmala S Guptha, V. Balamurugan, Geetha Megharaj, Khalid Nazim Abdul Sattar, J. Dhiviya Rose,

Cross lingual handwritten character recognition using long short term memory network with aid of elephant herding optimization algorithm,

Pattern Recognition Letters,

Volume 159,

2022,

Pages 16-22,

ISSN 0167-8655,

https://doi.org/10.1016/j.patrec.2022.04.038.

(<https://www.sciencedirect.com/science/article/pii/S0167865522001490>)

6 Dhivya S and Usha Devi G. 2021. Study on Automated Approach to Recognize Characters for Handwritten and Historical Document. ACM Trans. Asian Low-Resour. Lang. Inf. Process. 20, 3, Article 37 (May 2021), 24 pages. https://doi.org/10.1145/3396167

7 Dhivya S and Usha Devi G. 2021. Study on Automated Approach to Recognize Characters for Handwritten and Historical Document. ACM Trans. Asian Low-Resour. Lang. Inf. Process. 20, 3, Article 37 (May 2021), 24 pages. https://doi.org/10.1145/3396167

8 Chen, Zhuo & Yin, Fei & Yang, Qing & Liu, Cheng-Lin. (2022). Cross-Lingual Text Image Recognition via Multi-Hierarchy Cross-Modal Mimic. IEEE Transactions on Multimedia. PP. 1-13. 10.1109/TMM.2022.3183386.