Shuru Wang, Alice Chu
MAIS 202

Topic: Sign Alphabet Recognition

Dataset

We will use this dataset because it is curated for convolution neural networks, which we will learn in week 8 lecture. We felt that this would be a good way to apply and practice the knowledge we obtained from the lecture. The topic (sign language recognition) is also one that we are very interested in. We planned to generate variant datasets in the following ways: rotation, flipping, resizing, cropping, and normalizing of images.

Methodology

Data Preprocessing

The dataset is very feasible. 3000 images of hand symbols are provided for each alphabet, which will ensure enough data points to train the model. The images are taken at a variety of angles, skin color of hands and background color, which provides good variation in the dataset.

The most useful information is the hand and its position. Some preprocess procedures are centering image data by subtracting mean pixel values from each pixel and using python.keras module to transform images into columns and rows.

Machine Learning model

The dataset is expected to predict the corresponding alphabet of the input sign alphabet image. The model we would like to adopt is a convoluted neural network (CNN). These models are very good at picking out patterns in images and have little dependence on pre-processing. [1] These networks are multi-layered, and as we go down the layers can recognize more complex shapes. The layers have 'filters' (convolutional kernels) to detect the image patterns, and we can think of it as sliding the filters over the image matrix and taking the dot product (as a means to apply weight and biases to the input pixel) to detect particular features (edges, corners, etc.). [1] There are various architectures, such as AlexNet, LetNet-5, [2] etc that avoid the need for feature extraction. In addition, we plan to use the softmax function as the probability function in identifying the class.

In comparison to support vector machines(SVM), they can both approximate linear and nonlinear boundaries in classification problems. [3] SVM requires more prepossessing procedures than CNN and if the number of class labels is too large, it could be problematic for SVM. [4] Moreover, CNN has been very good for image classification problems, and the examples we looked at all have very high accuracy.

Evaluation Metric

We will use a confusion matrix and accuracy score [5] to determine the performance of our algorithm, the expected accuracy is 95%. [6]

Application

User Input and Output

The user is prompted to upload an image that contains a sign alphabet with a contrastive background to the skin color. The output will be displaying the corresponding alphabet to the screen.

Reference

[1] DeepAI. 2019. Recurrent neural network. (May 2019). Retrieved February 19, 2022 from https://deepai.org/machine-learning-glossary-and-terms/recurrent-neural-network

[2] Jason Brownlee. 2019. Best practices for preparing and augmenting image data for cnns. (July 2019). Retrieved February 20, 2022 from https://machinelearningmastery.com/best-practices-for-preparing-and-augmenting-image-data-for-convolutional-neural-networks/

[3] Gabriele De Luca. 2021. SVM vs Neural Network. (August 2021). Retrieved February 19, 2022 from https://www.baeldung.com/cs/svm-vs-neural-network

[4] Nachiketa Mishra. How is CNN better than SVM and recurrent neural networks? Retrieved February 19, 2022 from https://www.quora.com/How-is-CNN-better-than-SVM-and-recurrent-neural-networks

[5]Ajitesh KumarI. 2022. Accuracy, precision, Recall & F1-Score - Python examples. (January 2022). Retrieved February 19, 2022 from https://vitalflux.com/accuracy-precision-recall-f1-score-python-example/

[6] Poulomi Chatterjee, Srishti Mukherjee, Vijaysinh Lendave, and Abhishree Choudhary. 2020. Hands-on guide to sign language classification using CNN. (October 2020). Retrieved February 19, 2022 from https://analyticsindiamag.com/hands-on-guide-to-sign-language-classification-using-cnn/