Sign Language Digits Classification with Neural Networks

Daniel-Domițian Iuonac, Politehnica University Timișoara

I.INTRODUCTION

Sign languages (also known as signed languages) are languages that use manual communication to convey meaning. This can include simultaneously employing hand gestures, movement, orientation of the fingers, arms or body, and facial expressions to convey a speaker's ideas.

Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano. It was developed with a focus on enabling fast experimentation. Being able to go from idea to result with the least possible delay is key to doing good research.

It is propose to create a neural network (NN) that predicts sign language digits. The NN is written in Python 3 using Keras. The source code is available in [1].

II.DATASET

The dataset used is an open-source dataset prepared by Turkey Ankara Ayrancı Anadolu High School's students[2].

Details of dataset:

- Image size: 100x100 pixels
- Color space: RGB
- Number of classes: 10 (Digits: 0-9)
- Number of participant students: 218
- Number of samples per student: 10

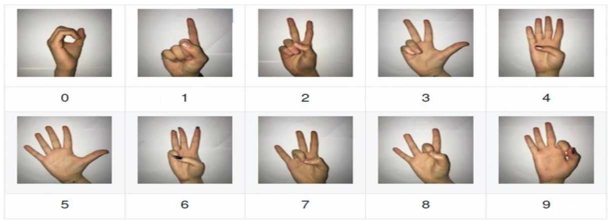

Figure 1. Dataset preview

Data preprocessing**:** A few images from some digit classes where removed because the number of images for each digit class varied. The new dataset has 2040 images (204 images for each class). The images were resized to 64x64 pixels.

The dataset was splited in two parts, 80% for training and 20% for testing.

III. CLASSIFIER COMPARISON

Multilayer Perceptrons, or MLPs for short, are the classical type of neural network.

They are comprised of one or more layers of neurons. Data is fed to the input layer, there may be one or more hidden layers providing levels of abstraction, and predictions are made on the output layer.

MLPs are suitable for classification prediction problems where inputs are assigned a class or label.

Convolutional Neural Networks, or CNNs, were designed to map image data to an output variable.

They have proven so effective that they are the go-to method for any type of prediction problem involving image data as an input.

The benefit of using CNNs is their ability to develop an internal representation of a two-dimensional image. This allows the model to learn position and scale invariant structures in the data, which is important when working with images.

The MLP has 6 Dense layers. The first 5 layers have ‘relu’ activation and units as 128, 64, 64, 32 ,16. The sixth layer has 10 units and ‘softmax’ activation. The loss function used is ‘categotical_crossentropy’ and ‘adam’ optimizer.

The CNN has 8 layers in total: 3 Conv2D layers, 2 MaxPooling2D layers and 3 Dense layers. The convolutionals layers have ‘relu’ activation and 256, 128, 64 units. The MaxPooling2D layers have the pool size (2, 2). The first 2 Dense layers have ‘relu’ activation and 64, respectively 32, units. The last Dense layer has ‘softmax’ activation and 10 units. The loss function used is ‘categorical_crossentropy’ and ‘rmsprop’ optimizer. The results of the classifiers on test set are depicted in Table 1. Training and validation loss for CNN are depicted in Figure 2.

Table 1. Classifier Accuracy.

Classifier	**Accuracy [%]**
MLP	61.24
CNN	94.85

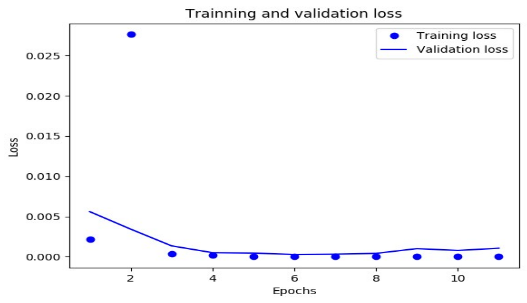

Figure 2. Training And Validation Loss for CNN

IV. Convolutional Network Architecture

Conv2D (64, 64 ,256) ‘relu’	MaxPooling2D (2, 2) ‘relu’	Conv2D (32, 32 , 128) ‘relu’	MaxPooling2D (2, 2) ‘relu’	Conv2D (32, 32 , 64) ‘relu’	Dense (64) ‘relu’	Dense (32) ‘relu’	Dense (10) ‘softmax’

V. FUTURE WORK

The next step will imply the development of a Graphical User Interface (GUI) application in order to predict sign language digits in real time using a camera. The software will have the capability to create a custom dataset by capturing the user gestures. The hyperparameters can be tunned through the application and the CNN can be retrained on the new dataset by pressing a button.

References

[1] https://github.com/dd-iuonac/sign-language-digits-classifier

[2] https://github.com/ardamavi/Sign-Language-Digits-Dataset