World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:13, No:01, 2019

Using Convolutional Neural Networks to Distinguish Different Sign Language Alphanumerics

Authors : Stephen L. Green, Alexander N. Gorban, Ivan Y. Tyukin

Abstract : Within the past decade, using Convolutional Neural Networks (CNN)'s to create Deep Learning systems capable of translating Sign Language into text has been a breakthrough in breaking the communication barrier for deaf-mute people. Conventional research on this subject has been concerned with training the network to recognize the fingerspelling gestures of a given language and produce their corresponding alphanumerics. One of the problems with the current developing technology is that images are scarce, with little variations in the gestures being presented to the recognition program, often skewed towards single skin tones and hand sizes that makes a percentage of the population's fingerspelling harder to detect. Along with this, current gesture detection programs are only trained on one finger spelling language despite there being one hundred and forty-two known variants so far. All of this presents a limitation for traditional exploitation for the state of current technologies such as CNN's, due to their large number of required parameters. This work aims to present a technology that aims to resolve this issue by combining a pretrained legacy AI system for a generic object recognition task with a corrector method to uptrain the legacy network. This is a computationally efficient procedure that does not require large volumes of data even when covering a broad range of sign languages such as American Sign Language, British Sign Language and Chinese Sign Language (Pinyin). Implementing recent results on method concentration, namely the stochastic separation theorem, an AI system is supposed as an operate mapping an input present in the set of images $u \in U$ to an output that exists in a set of predicted class labels $q \in Q$ of the alphanumeric that q represents and the language it comes from. These inputs and outputs, along with the interval variables $z \in Z$ represent the system's current state which implies a mapping that assigns an element $x \in \mathbb{R}^n$ to the triple (u, z, q). As all xi are i.i.d vectors drawn from a product mean distribution, over a period of time the AI generates a large set of measurements xi called S that are grouped into two categories: the correct predictions M and the incorrect predictions Y. Once the network has made its predictions, a corrector can then be applied through centering S and Y by subtracting their means. The data is then regularized by applying the Kaiser rule to the resulting eigenmatrix and then whitened before being split into pairwise, positively correlated clusters. Each of these clusters produces a unique hyperplane and if any element x falls outside the region bounded by these lines then it is reported as an error. As a result of this methodology, a self-correcting recognition process is created that can identify fingerspelling from a variety of sign language and successfully identify the corresponding alphanumeric and what language the gesture originates from which no other neural network has been able to replicate.

Keywords : convolutional neural networks, deep learning, shallow correctors, sign language
Conference Title : ICNN 2019 : International Conference on Neural Networks
Conference Location : Tokyo, Japan
Conference Dates : January 07-08, 2019