Eight or Three? Multi Language Sign Recognition System

Authors: Amiy Yadav, Spencer Wilson

What if you could sign in any language and be understood in another sign language?

There are 300+ sign languages in the world. Translating from one sign language to another without an intermediate written language for translation could create a faster and more accurate translation of signed languages and reduce cognitive load for the deaf and dumb community during conversations.

Pre-processing

Image classification techniques:

- Canny edge detection
- Histogram of Oriented Gradients (HOG)
- Fourier Descriptor

Canny edge detects edges by noise detection, gradient calculator with non-maximum suppression with double threshold and then detecting the edge tracking by hysteresis. HOG detection counts the occurrence of gradient orientation of localized area of an image. Fourier Descriptor segregates the image into different low and high frequency regions. Combined together these techniques better identify the location of the hand gesturing and the skeletal structure of the hand. These are the popular methods used in related works for sign language recognition [2][3].

We preprocessed 12,000 original ISL images and 700 original ASL images split evenly across the ten semantics for numbers zero through nine.

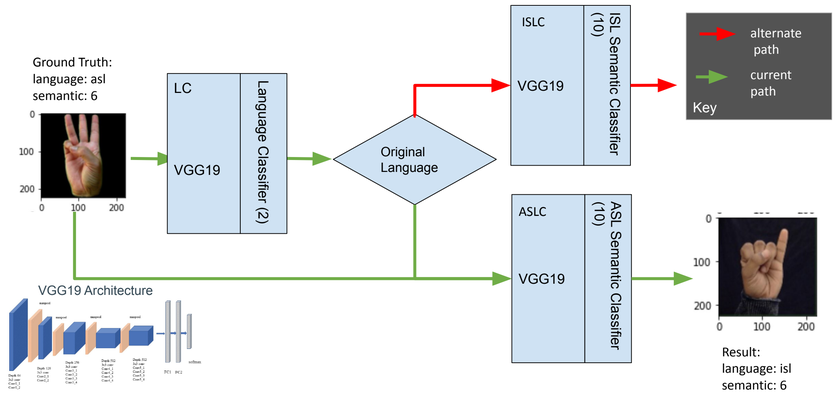
Previous work explored canny edge detection and HOG detection on word level sign language classification. The addition of Fourier Descriptor was intended to enhance the skeletal identification of the hand posture.

Below is an example of these transformations.

Figure 1: (1) Original sign for number 2, (2) Canny Edge, (3) Canny-HOG-Fourier(Brightened)

Translation Architecture

VGG19 with pretrained ImageNet weights without the final softmax layer is used as a base model for image recognition and integrated to a translation pipeline.



[1]

Figure 2: Translation Architecture

The model is comprised of three classifiers: one VGG19 Language Classifier (LC), one VGG19 for ASL semantic classifier (ASLC), and one ISL semantic classifier (ISLC). The input sign goes through LC to detects the original language of the sign. Then the same language semantic classifier is used to determine the sign semantic and return an image in the opposite languge.

Demonstrated above, the ASL sign for 6 is passed to LC. This is successfully identified as ASL. The same image is then passed to ASLC and a prespecified image of the ISL sign for 6 is presented to the user of the tool.

The tool will work in either direction ASL to ISL or ISL to ASL across each of the semantics shown below.

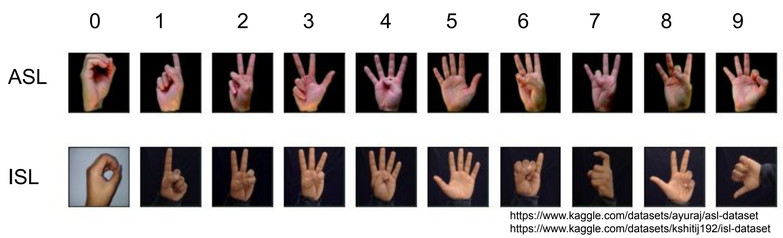


Figure 3: ASL and ISL Signs for numbers

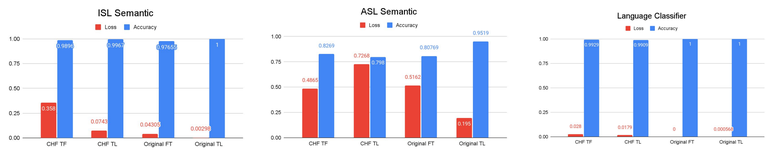
Available on GitHub: https://github.com/SpenceIsCool/EightOrThree

Results

The models were trained against the Canny-HOG-Fourier (CHF) images to be either fine tuned on ISL dataset first and then it is ran on ASL dataset (FT), or they were applied directly to the pretrained VGG19 with ImageNet weights for direct transfer learning (TL).

Below is a summary of the accuracy and loss found during training of each of the four experiments. The best model was determined to use transfer learning only over the original model and found to have a minimum 95.19 % accuracy as seen in the bottom row of our table.

	ISL Semantic		ASL Semantic		Language Classifier	
	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy
CHF TF	0.358	0.9896	0.4865	0.8269	0.028	0.9929
CHF TL	0.0743	0.9967	0.7268	0.798	0.0179	0.9909
Original FT	0.04305	0.97655	0.5162	0.80769	**0**	**1**
Original TL	**0.00298**	**1**	**0.195**	**0.9519**	0.000568	**1**



Conclusion

Ultimately we found the best results on VGG19 pre-trained on ImageNet as our base models. The the addition of Canny edge, HOG and Fourier Descriptor did not yield better results for this task. In future experiments the base model should be observed before considering pre-processing transformations, however these pre-processing steps should still be candidates for improved performance in future experimentation on word level and sentence level translation tasks.

We have demonstrated that it is possible to translate numbers between multiple sign languages without intermediate english language representations. In future work we will explore sentence level translations between sign languages and a virtualized assistants to demonstrate the translations as a live interpreter.

References

1. Extract Features, Visualize Filters and Feature Maps in VGG16 and VGG19 CNN Models, Roland Hewage (2020)
2. Adeyanju, I., Bello, O., Adegboye, M.: Machine learning methods for sign language recognition: A critical review and analysis. Intelligent Systems with Applications 12(2021) 200056
3. Adeyanju, I., Bello, O., Azeez, A.:Development of an american sign language recognition system using canny edge and histogram of oriented gradient. Nigerian Journal of Technological Development 19 (09 2022) 195–205