**Sign Language Recognition using Convolutional Neural Networks**

Very few people understand sign language. Moreover, contrary to popular belief, it is not an international language. Obviously, this further complicates communication between the Deaf community and the hearing majority. The alternative of written communication is cumbersome, because the Deaf community is generally less skilled in writing a spoken language [17].

17 - Van Herreweghe, M.: Prelinguaal dove jongeren en nederlands: een syntactisch onderzoek. Universiteit Gent. Faculteit Letteren en Wijsbegeerte (1996)

In our work, we build on the results of Roel Verschaeren [18]. He proposes a CNN model that recognizes a set of 50 different signs in the Flemish Sign Language with an error of 2.5%, using the Microsoft Kinect. Unfortunately, this work is limited in the sense that it considers only a single person in a fixed environment.

18 - **Verschaeren, R.: Automatische herkenning van gebaren met de microsoft Kinect (2012)**

Kinect camera

The Microsoft Kinect is also used in [2] that proposes a recognition system for 239 words of the Chinese Sign Language (CSL). Here, the 3D movement trajectory of the hands are used besides a language model to construct sentences. This trajectory is aligned and matched with a gallery of known trajectories. The top-1 and top-5 recognition rates are 83.51% and 96.32% respectively.

2 - **Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., Chen, X., Zhou, M.: Sign Language Recognition and Translation with Kinect (2013), http://vipl.ict.ac.cn/sites/default/\_les/papers/\_les/2013 FG xjchai Sign Language Recognition and Translation with Kinect.pdf**

CNNs (based on [13]) are feature extraction models in deep learning that recently have proven to be to be very successful at image recognition [12], [3], [20], [7].

**13 - Lecun, Y., Bottou, L., Bengio, Y., Ha\_ner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE 86(11) (1998)**

**12 - Krizhevsky, A., Sutskever, I., Hinton, G.: Imagenet classi\_cation with deep con- volutional neural networks. Advances in Neural Information pp. 1{9 (2012),** [**http://books.nips.cc/papers/\_les/nips25/NIPS2012 0534.pdf**](http://books.nips.cc/papers/_les/nips25/NIPS2012%200534.pdf)

**3 - Cire\_san, D., Meier, U., Schmidhuber, J.: Multi-column deep neural networks for image classi\_cation. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 3642{3649. IEEE (2012)**

**20 - Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional neural net- works. arXiv preprint arXiv:1311.2901 (2013)**

**7 - Goodfellow, I.J., Bulatov, Y., Ibarz, J., Arnoud, S., Shet, V.: Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv preprint arXiv:1312.6082 (2013)**

**American Sign Language Character Recognition Using Convolution Neural Network**

Paulo Trigueiros, Fernando Ribeiro, and Luís Paulo Reis [2] have proposed a real time vision based system whose purpose is to recognise Portuguese sign language. They used Kinect Camera to extract hand features. For model training and gesture classification open source Dlib library was used, a general purpose cross platform C++ library capable of SVM multiclass classification.

2 - **Trigueiros, P., Ribeiro, F. and Reis, L.P. “Vision-based Portuguese sign language recognition system”. In New Perspectives in Information Systems and Technologies, 2014 Volume 1 (pp. 605-617). Springer International Publishing.**

Neha V. Tavan, Prof. A.V. Deorankar [3] in their work implemented an algorithm to extract HOG features. These features were then used to train an artificial neural network which was later used for the purpose of gesture recognition.

3 - **Tavari, Neha V., and A. V. Deorankar. “Indian Sign Language Recognition based on Histograms of Oriented Gradient.” International Journal of Computer Science Information Technologies 5 (2014).**

Haitham Hasan, S. Abdul-Kareem [4] have proposed a technique for hand gesture recognition based on shape analysis. They used neural network based approach to classify among six static hand gestures namely open, close cut, paste, maximize and minimize. They have used a unique multi-layer perception of neural network for classification using back-propagation learning algorithm. They were able to achieve an accuracy of 86.38%.

4 **- Hasan, Haitham, and S. Abdul-Kareem. “Static hand gesture recognition using neural networks.” Artificial Intelligence Review 41, no. 2 (2014): 147-181.**

Bhumika Gupta, Pushkar Shukla, Ankush Mittal [6] They have used HOG and SIFT to extract feature for image. These features are then combined into a single matrix. Correlation is computed for these matrices and is fed to a K-Nearest Neighbour Classifier. Out of 200 gestures 179 were identified correctly.

6 - **Gupta, Bhumika, Pushkar Shukla, and Ankush Mittal. "K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion." In 2016 International Conference on Computer Communication and Informatics (ICCCI), pp. 1-5. IEEE, 2016.**

S.Nagarajan and S.Nagarajan [7] used Edge Oriented Histogram features and multiclass SVM. The edge histogram count of input sign language alphabets is extracted as the features and applied to a multiclass SVM for classification. The average accuracy of the system was 93.75 %.

7 **- Nagarajan, S., and T. S. Subashini. "Static hand gesture recognition for sign language alphabets using edge oriented histogram and multi class SVM." International Journal of Computer Applications 82, no. 4 (2013).**

Cnn – cnn hoyagatta research paper 1

**LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86, no. 11 (1998): 2278-2324.**

**Real-time American Sign Language Recognition with Convolutional Neural Networks**

CNN -

[11] L. Pigou et al. Sign Language Recognition Using Convolutional Neural Networks. European Conference on Computer Vision 6-12 September 2014

[10] J. Atwood, M. Eicholtz, and J. Farrell. American Sign Language Recognition System. Artificial Intelligence and Machine Learning for Engineering Design. Dept. of Mechanical Engineering, Carnegie Mellon University, 2012.

While linear classifiers are easy to work with because they are relatively simple models, they require sophisticated feature extraction and preprocessing methods to be successful [2, 3, 4]. Singha and Das obtained accuracy of 96% on 10 classes for images of gestures of one hand using Karhunen-Loeve Transforms

**2 - Singha, J. and Das, K. “Hand Gesture Recognition Based on Karhunen-Loeve Transform”, Mobile and Embedded 232 Technology International Conference (MECON), January 17-18, 2013, India. 365-371.**

**3 - D. Aryanie, Y. Heryadi. American Sign Language-Based Finger-spelling Recognition using k-Nearest Neighbors Classifier. 3rd International Conference on Information and Communication Technology (2015) 533-536.**

**3 - R. Sharma et al. Recognition of Single Handed Sign Language Gestures using Contour Tracing descriptor. Proceedings of the World Congress on Engineering 2013 Vol. II, WCE 2013, July 3 - 5, 2013, London, U.K.**

Some neural networks have been used to tackle ASLtranslation [8, 9, 10, 11]. Arguably, the most significant

advantage of neural networks is that they learn the mostimportant classification features. However, they require

considerably more time and data to train. To date, most have been relatively shallow. Mekala et al. classified video of ASL letters into text using advanced feature extraction and a 3-layer Neural Network [8]. They extracted features in two categories: hand position and movement. Prior to ASL classification, they identify the presence and location of 6 “points of interest” in the hand: each of the fingertips and the center of the palm. Mekala et al. also take Fourier

Transforms of the images and identify what section of the frame the hand is located in. While they claim to be able to correctly classify 100% of images with this framework, there is no mention of whether this result was achieved in

the training, validation or test set.

Admasu and Raimond classified Ethiopian Sign Language correctly in 98.5% of cases using a feedforward

Neural Network [9]. They use a significant amount of image preprocessing, including image size normalization, image background subtraction, contrast adjustment, and image segmentation. Admasu and Raimond extracted features with a Gabor Filter and Principal Component Analysis.

**Real-time sign language recognition based on neural network architecture**

One important means of communication method for the hearing impaired community is the use of sign language, as in [1]

{introduction}

1 - Paulraj M P, Sazali Yaacob, Mohd Shuhanaz bin Zanar Azalan, Rajkumar Palaniappan, “A Phoneme based sign language recognition system using skin color segmentation”, Signal Processing and Its Applications (CSPA), pp: 1 – 5, 2010.

Signing takes place in a 3D space, called signing space close to the trunk and the head, as in [7]. {intro}

7 - Bauer, B.; Hienz, H., “Relevant features for video-based continuous sign language recognition”, Automatic Face and Gesture Recognition, 2000. Proceedings, pp: 440 – 445, 2000.

Occlusion plays a crucial role factor in real time as while signing; some fingers or even a whole hand can be occluded, as in [12].

12 - P. Mekala, R. Salmeron, Jeffrey Fan, A. Davari, J. Tan, "Occlusion Detection Using Motion-Position Analysis", IEEE 42nd Southeastern Symposium on System Theory (SSST'10), Tyler, TX, pp. 197-201, March 7-9, 2010.

In order to satisfy the memory requirements and the environmental scene conditions, preprocessing of the raw video content is highly important [14]. {Methodology}

14 - Brian L. Pulito, Raju Damarla, Sunil Nariani, " 2-D Shift Invariant image Classification Neural Network, which overcomes Stability, Plasticity Dilemma", Vol 2, International Joint Conference on Neural Network, San Deigo, June 17-21,1990.