**Review on paper:** Bangla Sign language Detection using SIFT and CNN- 2018

**Abstract**

This paper discusses the research paper Bangla Sign Language Detection using SIFT and CNN. In the research, the authors have used SIFT feature extraction and CNN classification to compare the accuracy increase compared to just using CNN for Bangla Sign Language detection.

**Introduction**

Most researchers have used SVN, ANN and KNN for the classification of Bangle sign language. Some researchers have used CNN with minimal preprocessing but in the paper the researchers have used a preprocessing SIFT (Scale-Invariant feature Transform) technique to eliminate the problems of rotation and scaling in images before using CNN to check how well it performs compared to only using CNN.

**Literature Review**

Most researchers use minimal preprocessing techniques when using CNN. But in their research they have used skin masking techniques and SIFT technique to preprocess the images before classifying the images using CNN. There is a problem in image classifications which is there can be rotation, scaling and resolution and illumination in the images that can be difficult for the CNN to identify but using SIFT they have addressed the issue.

**Methods**

For the detection of Bangla hand signs they have taken the images manually. They took a total of 7600 images of 51 Bangla letters for training and testing. Before they used the images on CNN they pre-processed the images by Skin Masking and SIFT (Scale-Invariant feature Transform).

The images are first in RGB color space which is converted to HSV. The upper and lower boundary was set to do skin masking so all the colour except the skin colour is removed from the images. Then these images are converted to grayscale. The skin masking techniques was used to crop the Region of Interest that has only the image of the hand.

Before passing the images to CNN they address another major problem in training that is the rotation, scaling and resolution and illumination changes that can hamper training. They used SIFT (Scale-Invariant feature Transform) for solving it. SIFT was implemented using the following steps:

1. Key point detection
2. Key point descriptor generation

Then the feature vectors from SIFT was clustered using K-means clustering because CNN takes a set amount of features as input. Then they converted each images as a histogram where in x- axis they put the features and in y-axis they put the frequency of each features in an image. This is what they call Bag of Feature set and this was used it to train and test the CNN.

**Results**

Comparing the results between only using CNN and using SIFT alongside CNN their results show that there is a higher accuracy for using SIFT with CNN for classifying each Bangla letter. They have shown that SIFT works well with CNN.

**Discussion**

The results show that it is effective to use preprocessing techniques since it can increase the accuracy of classification by a large amount. In Bangla sign language there are many letters and so it is difficult to get a high accuracy and there is an absence of a large dataset with images that have good quality. Since, in our project we haven’t done enough preprocessing but adding preprocessing such as SIFT we can get a higher accuracy for Bangla sign detection.

**Conclusion**

Using SIFT with CNN has shown to increase the accuracy of detecting the hand signs. So researchers should use preprocessing technique to increase the accuracy. There is another preprocessing technique SURF which may perform better with CNN and further research is required on it.

**Reference-**

S. S. Shanta, S. T. Anwar and M. R. Kabir, "Bangla Sign Language Detection Using SIFT and CNN," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1-6, doi: 10.1109/ICCCNT.2018.8493915.

URL: https://ieeexplore.ieee.org/document/8493915