

Computer Vision Approaches based on Deep Learning and Neural Networks: Deep Neural Networks for Video Analysis of Human Pose Estimation

This paper mainly focuses on three sections. Firstly how to improve the existing solution in the field of human pose estimation through CNNs. Secondly how to implement a deep learning algorithm on Recurrent Neural Networks (RNN) so that human pose estimation can be done through RNN. Lastly, learn how to use an unsupervised learning paradigm to train NNs. One of the main motivations behind the paper is to improve gesture recognition in sign language videos. As deaf people cannot hear anything. They have to understand everything seeing sign language. But everyone doesn't understand sign language equally. So if there is a subtitle or text version of the signs of the sign language, then it will be very easy to understand. So in order to do it gesture recognition and detection of the upper body part are needed.

So for solving this problem Deep Neural Network (DNN) is proposed to solve it in a new way. DNN is successfully used to object localization. They suggested a cascade of DNN-based pose predictors to solve the problem. Another way to solve this problem is by using CNN. Using CNN is used to recognize a person in an image to estimate the pose and classify the action. NNs are used to learn conditional probabilities. Learning this the presence of the parts and their spatial relationships within image patches can be found.

Through this paper, we learned that the remaining solution should be improved and other architecture can be approached to solve this problem. By solving this problem a system can be designed to convert sign language to text. This will help people to understand sign language better.

The paper could have shown other ways apart from CNN and RNN approaches to estimate human pose. So that it could be more clear which the best approach to solve this problem.

This paper will help us to learn about how to use CNN and computer vision for the gesture and pose estimation. As our project is mainly based on computer vision and CNN, so learning about these things and their applications is very important.

Human face detection in a complex background

The paper attempts to show a way how to develop an automated face recognition system that can identify a face in a complex background where the number of the face, size of the face, and the location of the face is unknown. So without knowing these pieces of information the system can identify faces, how many are there, and so on. Using this system we can identify faces in black and white images. But the system can also be used for colorful images and videos as well by using some color information or video characteristics. So using the system we will be able to identify a person or count the number of people present in an image.

The paper shows a hierarchical knowledge-based pattern recognition system that consists of three levels. In the first level, the image is taken as an input and then the whole image is scanned to find all possible candidates that can be human faces with all possible sizes and locations. In the second level, the basics face location is done. Which is done on the basis of a window of 8 x 8 cells over each candidate's face, and according to these rules the candidates are screened. For level one and two analysis mosaic images is used. On basis of the mosaic images, the rules are established for the first two levels. These two level's search work is based on the quartet and octet of mosaic images. Using the grey box the different parts of the face are detected. In the final and third level, the face detected in the second level is processed and improved. To improve these images edge detection method for eye and mouth regions is proposed. If all detection fits well then they are declared as images.

Using this system we can identify faces in the images, the number of the faces, and identify the person as well.

For the face, they have ignored some matter. According to their theory in order for an object to be detected as a face, there must be eyes, mouth, and nose present. There can be many exceptional cases. A person can be wearing glasses or wearing a mask. So in these cases, eyes, mouth, or nose will be hidden behind these objects. But in these cases, faces won't be detected as faces.

This paper will help us how to find or detect a face in a complex background. As our project is to identify faces so detecting faces is a vital part of the project.