VISUAL UNDERSTANDING USING DEEP LEARNING

Raviteja Pothala

ECE-GMRIT

ravitejapothala333@gmail.com

Nikhitha Pithani

CSE-GMRIT

nikhitha.cherry@gmail.com

Abstract

Describing the contents of images is a challenging task for machines to achieve. Deep learning has been widely adopted to determine the key tasks for image understanding. Deep learning uses the hierarchical architectures for learning high-level of abstractions in data, a subcategory of machine learning which deals with the use of neural networks to improve the state-of-the-art in many traditional Artificial Intelligence domains like computer vision, speech recognition, natural language processing and much more. We aim to review the state-of- the- art in a deep learning approach, Convolutional Neural Networks (CNN) in computer vision. In this, the applications of deep learning in some of the computer vision tasks such as human pose estimation and image classification are discussed. In the human pose estimation, we mainly tried to summarize deep learning schemes to estimate the human articulation from still images by using holistic processing method and part-based processing methods. The approaches based on bags of visual words (BoW) for image classification are discussed. For these two applications in the computer vision domain, using of the advancements of CNN based schemes is done, as it is the most extensively utilized and most suitable for images. Finally, we aimed to look at the future trends and challenges in training deep neural networks.

Keywords: Deep learning, Computer vision, CNN, Applications, Holistic processing, Part-based processing, BoW.