

COMPUTER VISION

(Abstract – Team 2)

Sandesh Maurya
Sanjeev Ranjan
Gaurav Dahiya

Computer vision is a field of computer science that works on enabling computers to see, identify and process images in the same way that human vision does, and then provide appropriate output. It is like imparting human intelligence and instincts to a computer. In reality though, it is a difficult task to enable computers to recognize images of different objects. Computer vision is closely linked with artificial intelligence, as the computer must interpret what it sees, and then perform appropriate analysis or act accordingly. Our seminar will briefly explain the recognition of images using Convolutional Neural Networks and how the AI describe the images , in a way like humans do , using the Midge system. Convolutional Neural Network has a series of steps to recognize an image. A matrix is associated with the image and its feature map is extracted by sliding a filter matrix over it. This feature map is rectified (made non-linear) and then the dimensionality of the image is decreased to the least possible in the pooling step. The extracted image/images we get after pooling are used to identify the actual image using fully connected layers. Now, the AI has to describe the image in a way similar to humans. Here , we are using the midge system for this purpose. Midge is a prototype system that connects computer vision to syntactic structures with semantic constraints, allowing for the automatic generation of detailed image descriptions. It identifies different objects and stuffs in the image and finds a way to connect them in a sentence , considering all possible combinations and selecting the most appropriate.