Abstract

**Researchers**:

Sidharth Pal (Matrikel Nr. 00503319)

Siddharth Anil Naik (Matrikel Nr. 00616219)

**Research focus**:

Aid for visually impairment using Computer vision and other AI tools

**Professor**:

Prof. Dr. Ralph Lano

**Course**:

Master’s in Software engineering for Industrial Applications

# Project Title:

Visual to Audio Aid for Visually Impaired

# Problem Description:

Visual impairment can halfway or completely halt the daily routine life and the reliance rate can go up to 100%.

# Objective of Project:

Our project aims to help Visual impaired with the help of Computer vision based Visual to Audio Aid. The field of computer vision has been transformed by the introduction of deep learning.

We aim to advance computer vision analysis from static scenes and images toward dynamic scenes and to integrate audio-visual perception, eventually enabling these systems to understand visual inputs.

# Rationale of the Project:

We are using Image captioning which generates meaningful caption upon understanding the given image. The video stream is frames of images. We aim to process the image frames at a certain interval for feature matrix with a CNN. The feature matrix which is then fed to LSTM based RNN to generate meaningful caption.

The caption is compared to the previous caption for semantic text similarity using another RNN. In case of significant difference, the newly generated caption is than given as an input to a text to speech conversion model which than delivers an audio output.

This system employs non-invasive human-machine interface which in the case of visually impaired, transforms visual information into audio aid.

# Keywords:

Computer Vision, Image Processing, Semantic Text, Convolutional Neural Network, Recurrent Neural Network, Artificial Intelligence, Deep Learning, Machine Learning.