**Title of the Project :** Visual Audio For Visionless Using Deep Learning

**Name of the Students :**  Preethi. R

Renuka. G

Veerapareddy Divya

**Register Number(s)    :** 211419104202

211419104220

211419104298

**Name of the Guide      :**  Mrs. D. Jennifer, M.E. (Ph.D.)

**ABSTRACT**

The sense of sight is incredibly important for human beings, and its absence can significantly impact an individual's quality of life. Blindness can limit an individual's ability to perform everyday tasks independently, including navigating unfamiliar environments, identifying objects, and reading. According to the World Health Organization (WHO), approximately 285 million people worldwide are visually impaired, with 39 million of them being completely blind. Visual audio for Visionless combines the concept of image captioning and TTS that involves converting an image into textual and audio description using deep learning. This task is challenging as it requires the system to understand the content of an image and generate natural language captions that accurately describes the image content in both spoken and written form. In the proposed system, CNN (Inception V3) serves as the encoder and LSTM with attention serves as the decoder and generates the caption word by word, conditioning the prediction on the previous generated. Inception V3 and LSTM for image captioning can achieve good results due to the complementary strengths of these two models. Ultimately, captions and audio description for images in real-time are generated as soon as the system captures the scenario as an image.