



BLIVIZ - A Blind's Vision

Vasanth P - CSE C

Saadhana G - CSE C



OBJECTIVE OF BLIVIZ



To enable the vision of blind people to come to reality by avoiding the need of external person or other factors such as stick



INTRODUCTION

BLIVIZ is an AI powered application which generates a caption for a given image. This application consists of three main components.

1. Feature Extractor - ViT model - Extract features from image
2. Caption generator - GPT2 - Generate captions based on the features extracted
3. TexttoSpeech - gTTS - Convert caption in text to voice output



EXISTING SYSTEM AND SOLUTIONS TO EXISTING SYSTEM

We can't imagine a world which we can't see where the blind persons lead their daily life in this circumstance which is pretty hard to live. There are two main existing solutions for it now

- 01 They can take the help of an another person to describe what is happening around them or carry around a stick

- 02 AI created using CNN and RNN



DRAWBACKS OF EXISTING SYSTEM

1. Dependent on external factors such as stick or an external person
2. Risk of getting hurt while touching objects that could cause harm
3. RNN can't remember long information

BLOCK DIAGRAM OF PROPOSED SYSTEM





ADVANTAGES OF PROPOSED SYSTEM

1. Not dependent on external factors such as stick or an external person
2. No risk of getting hurt while touching objects that could cause harm
3. Transformer models are SOTA models which are very accurate than traditional models



HARDWARE AND SOFTWARE REQUIREMENTS

HARDWARE REQUIREMENTS:

- Laptop or Desktop

SOFTWARE REQUIREMENTS:

- Python
- Pytorch
- Transformers
- OpenCV
- PIL
- Streamlit



REFERENCES

1. <https://arxiv.org/pdf/1502.03044v3.pdf>
2. https://openaccess.thecvf.com/content/ACCV2020/papers/He_Image_Captioning_through_Image_Transformer_ACCV_2020_pAper.pdf



Project timeline





Thank you!

Lorem ipsum dolor sit amet,
consectetur adipiscing elit. Curabitur
eleifend a diam quis suscipit.

