

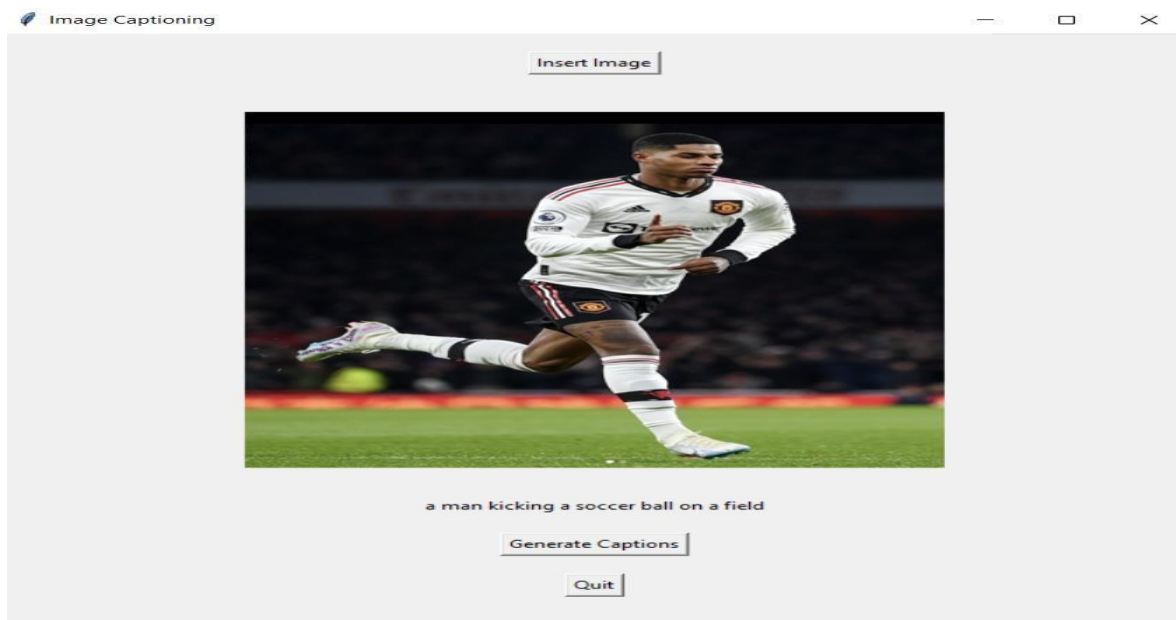
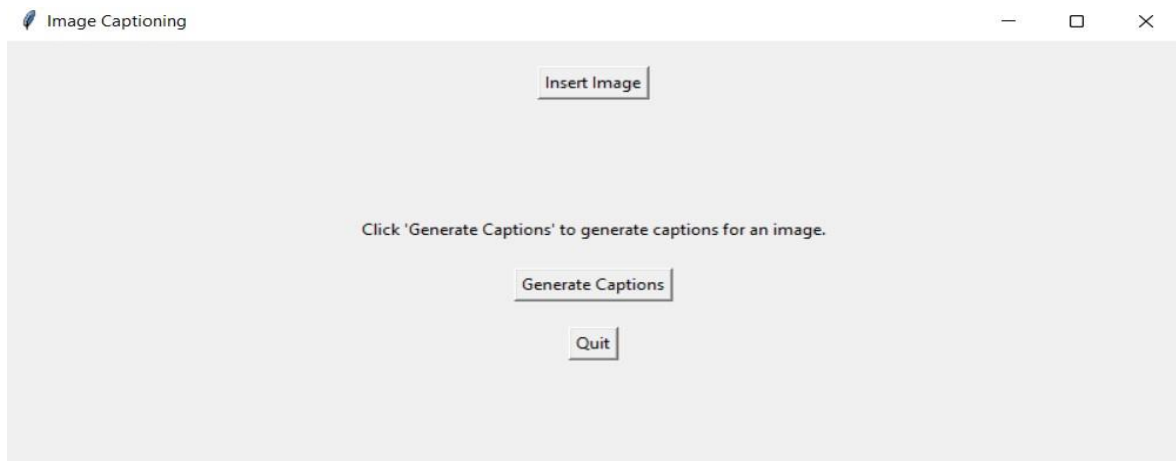
ImgCapGen

This application is build on a pretrained model,

<https://huggingface.co/nlpconnect/vit-gpt2-imagecaptioning/tree/main>

The interface of model is built using tkinter library

Run ImgCapGen.py then click on insert image and your caption will be generated below the picture.



There are far better models for image captioning which give more accurate results but as time is of the essence here I have built this model as finding a pre-trained usable model is harder than it seems.

If required we could build a model using ResNet or autoencoders or LSTM which would give better results than this, however I don't have enough computing power myself to build it as I found while training flickr8k dataset.

The interface can also be improved much more and more function could be added.

As for the multiple captions for image, I didn't find any pre-trained model but we could build one in the future