**CONCLUSION**

In this paper, a hybrid encoder-decoder based model to generate the effective caption of the image by using the Flickr8k dataset. During the encoding phase, the proposed model used transfer learning-based model like VGG16 and ResNet5o and YOLO for extracting the image features. A concatenate function is used to combine the feature and removes the duplicate one. For the decoding, BiGRu and LSTM are used to get the complete caption of the image. Further BLEU value is evaluated of both the captions generated by BiGRU and LSTM. Final caption is considered whose METEOR value is high. The proposed model is also evaluated by METEOR and ROUGE. The proposed model achieved score BLUE-1: 0.67, METEOR: 0.54 and ROUGE: 0.31 on Flickr8k dataset. The experimental results show the better results through BLUE, METEOR and ROUGE when compared to another state-of-art models. The model is also helpful in generating the captions at real time.