Name	Sufyan Ahmad
Roll No	201980013
Assignment	Survey Paper
Course	Computer Vision
Section	A

GIFT UNIVERSITY , GUJRANWALA

SUMMARY

"A Survey on Automatic Image Caption Generation"

Shuang Bai and Shan An

Introduction

Image caption generation is the task of automatically generating a natural language description of an image. It is a challenging task due to the need to capture the semantic content of the image, generate grammatically correct sentences, and be creative and informative.

Approaches to Image Caption Generation

There are two main approaches to image caption generation: retrieval-based and generation-based methods.

- Retrieval-based methods first retrieve a set of captions from a database of captions, and then select the caption that is most relevant to the given image. These methods are relatively simple to implement, but they can be limited by the quality of the captions in the database.
- Generation-based methods directly generate a caption from the image, without using a database of captions. These methods are more complex to implement, but they can potentially generate more creative and informative captions.

Neural Network-based Methods

Neural network-based methods have been shown to be effective for both retrieval-based and generation-based image caption generation. These methods use deep learning to learn the relationship between images and captions.

- Attention-based methods are a type of neural network-based method that have been shown to be particularly effective for image caption generation. These methods allow the model to focus on different parts of the image when generating the caption.

Future Research Directions

There are still many challenges in image caption generation, such as:

- Generating captions for images with multiple objects
- Generating captions that are creative and informative
- Generating captions in different languages

Future research in this area is likely to focus on addressing these challenges.

Conclusion

Image caption generation is a rapidly developing field with the potential to have a wide range of applications. Neural network-based methods have shown great promise for this task, and future research is likely to further improve the performance of these methods.