While significant progress has been made in the image captioning task, video description is still in its infancy due to the complex nature of video data. Generatin