UNIVERSITI TEKNOLOGI MARA

PHOTO-REALISTIC IMAGE GENERATION USING GENERATIVE ADVERSARIAL NETWORK WITH MULTIPLE TEXTUAL DESCRIPTION FOR FORENSIC SKETCHLESS RECOGNITION

NUR NABILAH BINTI BAHRUM

MSc

January 2024

ABSTRACT

The process of identifying photos from a sketch has been explored by many researchers, and the performance of the identification process is almost perfect, particularly for viewed sketches. Suspect identification based on sketches is one of the applications in forensic science. To identify the suspect using these kinds of methods, a face sketch is required. Hence, the methods require skilled artists to sketch the suspect based on descriptions provided by eyewitnesses. However, the skills of these artists are different from one another, which results in different rendered sketches. Therefore, this work attempts to propose a new identification method based only on forensic face-written descriptions. To investigate the feasibility of the proposed method, this study has evaluated the performance of some text-to-photo generators on both viewed and forensic datasets using three different models of GAN which are SAGAN, DFGAN, and DCGAN. Then, the generated images are compared to the real photo contained within those datasets to evaluate how well the GAN model recognizes the faces. The results demonstrated that the generated photos by the DCGAN models is better than the other two models which are achieve better value of FID, Clean-FID and KID. Then by using the DCGAN model, this study attempts in developing multi-text-to face GANs pseudo-photo generator for forensic sketch recognition that consist three analysis which are by using single description text-to-face GAN Generator, multiple description text-to-face GAN generator and concatenated description text-to-face GAN generator. From the result obtain from this analysis it shows that, the generated photo that had been generated by DCGAN model using multiple concatenation description text-to-face GAN using three eyewitnesses is more closer to the real photo compared to the generated photo that had been generated using single and multiple description. Therefore, this method had been chosen to generate the mugshot photo of the suspect person from the PRIP-HDC dataset and from the result obtained the FID, Clean-FID and KID value for the generated mugshot photo is 131.769, 131.128 and 0.083, respectively. Other than that, by using the generated mugshot photo that had been generated using proposed method is able to recognize 9 correct identities at rank-5. Furthermore, this study performed a qualitative analysis on the generated mugshot photo using the proposed method and compared it to other methods. According to the qualitative analysis results, a significant majority of the respondents indicated a preference for the proposed method, as it was found to generate mugshot photos that closely resembled real mugshot photos when compared to other methods. This finding demonstrates that the study's implementation of the proposed method successfully improved the quality of the generated mugshot photo.

ACKNOWLEDGEMENT

TABLE OF CONTENT

CHAPTER 1
INTRODUCTION

1.1   Background of Study

Face sketch recognition systems have grown in popularity worldwide due to technological advancements. Sketch identification of a suspect is one of the most frequently used methods in applying face recognition algorithms for forensic science [1]. The face sketch recognition system begins with artists manually drawing a sketch, which is then match with mugshot photos in the database to determine the most similar image to the sketch and thus identify the person. Matching methods for photo and sketch images can be implemented in two ways. The first one involves matching the images in the sketch modality by first converting the photo to a sketch, and the second way involves first converting the sketch to a photo and then matching in the photo modality. This is because both of the sketch and photo images are better to be matched in the same modality (intra-modality) and this method has been proven to has great performance as compared to the different modality images (inter-modality).

In 2015, learned features methods for face sketch recognition became popular and replaced the traditional techniques based on hand-crafted features [2]. Due to the substantial gains in system accuracy over other approaches, deep learning-based method become a standard approach in face recognition systems. However, there are limitations of deep learning for face sketch recognition. One of the primary reasons is the need for a large number of examples to train deep neural networks robustly to avoid problems such as overfitting and local minima [3]. Apart from that, generative adversarial networks (GANs) have been introduced, and GANs have become popular in image-to-image translation tasks and have made a lot of progress in recent years. In a typical GAN model, the discriminator tries to figure out if the inputs are real or fake. At the same time, the generator learns to make sharper and more realistic samples that can't be notified apart from the real samples. For example, Pix2Pix uses conditional GAN to do the task of translating image styles in a supervised way. In order to reduce the difficulty of obtaining image pairs, CycleGAN preserves key attributes between the input and translated images by employing cycle consistency loss [4]. However, all of