WROCŁAW UNIVERSITY OF SCIENCE AND TECHNOLOGY
FACULTY OF ELECTRONICS

FIELD: Computer Science
SPECIALIZATION: Internet Engineering (INE)

MASTER OF SCIENCE THESIS

Research on methods of changing objects in images using Deepfake technology

Badania metod zmiany obiektów na obrazach z wykorzystaniem technologii Deepfake

AUTHOR:
Michał Zendran

SUPERVISOR:
Dr inż. Andrzej Rusiecki

GRADE:

WROCŁAW 2020

Contents

Chapter 1

Abstract

Abstract. To be written at the end of the works.

Chapter 2

Introduction

2.1 Motivation

Machine learning has found many, different applications in the field of image data processing and computer vision. From picture classification to image denoiseing and resolution enhancement, artificial neural networks has gained the opinion of exceptionally useful tool. But for some time, a new, controversial use-case has been getting more and more attention in both media and research circles. So called "deepfake" technology has opened doors to many new possibilities of picture generation but also raised many issues of moral and legal matter.

Deepfake is a technology from the field of machine learning designed to combine and overlay objects in images or videos creating deceptively realistic counterfeits. The name comes from combination of two terms: ”deep learning” and ”fake”, and has its origins in a Reddit user named ”deepfakes”. Initially the term was associated only with face-swapping technology, but with time it was extended to all deep learning implementations of changing objects in images.

Deepfake technology has already found multiple applications such as changing seasons in the landscapes images, transforming horses into zebras or "repainting" images in styles of different artists. But the most controversial and impactful use-case so far is already mentioned face-swapping. In times of overwhelming amount of news it's getting harder and harder to filter out fake ones from valuable peaces of information. People generally tend to not check sources of information but rather blindly follow hot stories in social medias and television. Such environment combined with capabilities of deepfake gives possibilities of influencing elections by misrepresenting politicians in forged videos to defame or blackmail theme. Another popular use-case of described technology is creating erotic videos by replacing faces of porn actress with faces of well-known celebrities. This application might have less dangerous consequences then influencing world politics but may be hurtful to people that became objects of such act.

Although there are many malicious ways of using deepfake technology it might also be used for good reasons such as helping people to cope with the loss of the loved once or in entertainment filed by de-ageing actors to play younger-selves. Besides, to be able to detect and fight harmful applications of deepfake it might be vital to deeply understand algorithms and techniques behind it. Therefore, conducting research on that part of machine learning field seems to have great meaning in incoming times.

2.2 Objective and assumptions

This project aims to implement and compare four different methods of changing objects in images with application of artificial neural networks. For sake of this research, human faces were chosen as an object of replacement, as it rises a complex issue of simultaneous color, texture and shape modification.

As there are no numerical methods of measuring quality of images obtained from deepfake algorithms, the only way of appraising results of methods discussed in this research is visual evaluation. To be able to fairly rate each implemented technique, the same set of images will be used as a learning dataset for all cases. Therefor, effects of all approaches will be visually evaluated and compared with each other, which will result in the final assessment. This rating of methods is the expected outcome of the research.

There are two main factors that will be taken into a consideration during a results evaluation process. First of them is a resemblance of the faked image, to the appearance of the imitated person. The more striking similarity, the better. The other crucial aspect is preservation of original facial expression and pose, as the believable deepfake must capture the source material movements. Resultant of those two factors will be the main feature to be rated. It is assumed that neither of mentioned characteristics should outweigh the other one, but rather, the final effect should be well balanced composition of both aspects.

2.3 State of the art

What was done in this area so far. Four different methods with great success of GANs.

2.4 Naming conventions and terminology

Below, all abbreviations and naming conventions used in the research are listed and explained:

- Deepfake – name of the deep learning technology of swapping objects in images or an end result generated witch such technology
- ANN – artificial neural network
- CNN – convolutional neural network
- VAE – variational autoencoder
- GAN – generative adversarial network
- VAE-GAN – variational autoencoder-Generative adversarial network

Chapter 3

Theoretical background

3.1 Artificial neural network

Explain what are ANN, main idea, training process and so on.

3.2 Convolutional neural networks

Explain how it works, what are main use-cases and so on.

3.3 Supervised vs unsupervised training

Description of both and what are main differences and when to use which.

Chapter 4

Deepfake methods

4.1 Variational auto encoder

Idea behind deepfake generated by VAE without CNN

4.2 Convolutional variational auto encoder

Idea behind deepfake generated by VAE with CNN

4.3 VAE-GAN

Idea behind deepfake generated by GAN actually "VAE-GAN".

4.4 CycleGAN

Describe what is it, what it consists of, what are its applications, why I thought it should work for deepfake. Explain how it works exactly. Show learning process and results (good ones: horses to zebras and bad ones: face to face). Idea behind deepfake generated by CycleGAN. Explain why I'm assuming it should it work?

Chapter 5

Datasets

5.1 Datasets description

How dataset for deepfake learning should look like Used Datasets: VoxCeleb (description)

5.2 Data pre-processing

how I prepared my own datasets. All operations from videos to npz files

Chapter 6

Technologies

6.1 Software and Libraries

As in title...

6.2 Hardware

As in title ... (My hardware, Google colab, Google cloud?)

Chapter 7

Networks implementation

Detailed description of implementation of each method. What are the topologies, what callbacks were used, why those parameter, why those batches itp

7.1 Variational auto encoder

7.2 Convolutional variational auto encoder

7.3 VAE-GAN

7.4 CycleGAN

Chapter 8

Results

Presentation and discussion of results for each method

Chapter 9

Conclusions

Bibliography

[1] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The LaTeX Companion*. Addison-Wesley, Reading, Massachusetts, 1993.

List of Figures

List of Tables