Master Degree project proposal

Project Title

Video frame generation using General Adversarial Networks

Authors

Nikhil Nadig
Hanstavägen 49, 1302, Kista 16453
nadig@kth.se
+46 793164360
Responsibilities: Algorithms, optimization, web services, data gathering.

Óttar Guðmundsson
Trubadurvågen 1 171 69 Solna Stockholm
ottarg@kth.se
+354 6933881
Responsibilities: Architecture, write reports, loss comparison, presentation, resource gathering.

Keywords

GAN, Video, Damaged, Restore, Fix, VHS

Background

The main topic that this project will focus on is the effectiveness of neural network to generate an image X based on image frames Y and Z from a video clip. In the video, frame Y comes before Z. The new image X will be a believable image so looping through Y, X and Z respectively will seem like X is an actual video frame between Y and Z.

The second optional topic that the authors would like to look into is to use machine learning for pattern detection. This detection will be used to identify classified damages or errors in an image frame taken from a VHS video clip.

Problem statement

The inspiration and motivation for this project comes from damaged frames in VHS tapes. The problem with VHS is that they are easily damaged by magnetic forces, dust, humid or heat. There are methods of restoring VHS footage if the magnetic tape is damaged but it comes with an expensive price tag, time consumption and a lot of patience. A VHS video frame can be considered damaged if it has one or more of the following effects

- Pinking or Greening
- Screen tearing
- Chroma or Color bleeding
- Vertical or Horizontal jitter
- Rotational motion
- Interlace or Macroblock artifacts
- Snow or Analog grain
- Color loss
- Track errors

Image 1: A damaged clip from a VHS.

Problem

This project will explore the effectiveness using a GAN network to generate or fix a digital video frame that is considered damaged either by humans or by algorithm.

Here is an example of a frame that doesn't change that much and isn't severely damaged. The network might easily create a believable image so that the difference between the original and the generated one is at minimum.

Image 2: A damaged version of Fisher Price – Little people. The damaged clip is the middle one.

Image 3: The complete version of Fisher Price – Little people

The next example has a frame that is severely damaged and the difference between the frames is greater than previous example. The difference error between the generated image and the original will be greater compared to the previous example.

Image 4: A damaged version of Fisher Price – Little people. The damaged clip is the middle one.

Image 5: The correct version of Fisher Price – Little people

Purpose

The main purpose of the project is to gain better insight into GAN networks and apply it as an optional solution to a real world problem. It falls under the Neural Networks category from the course and demonstrates skills, innovation and creativity in the topic. The solution can be used a ground basis for video restoration program or a proof of concept. It can even serve as an example how GAN might be used to replace any missing or damaged data in a sequence.

Goal(s)

This project can be viewed as a personal and academic objective as well beneficial for video tape owners. By delivering and testing out GAN the authors hope to gain better insight into neural networks and how they can be applied to real world problems. For them, libraries like Keras and Tensorflow will be introduced which will hopefully serve as knowledge to solve further complex problems in the near future.

The expected results are frames considered damaged will be discarded and replaced with a frame that is generated from the images before and after the damaged part.

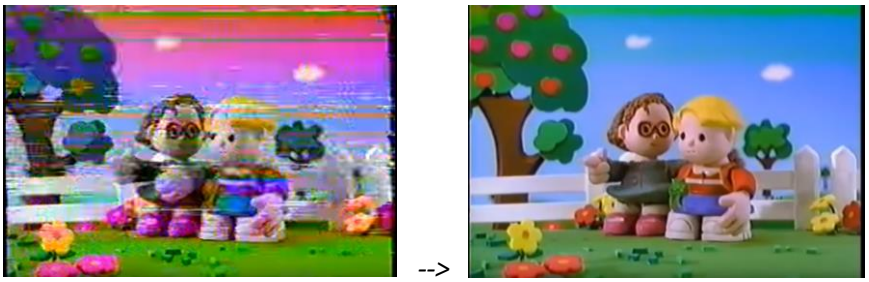

Image 6: A replaced image using the network.

If the solution will be considered as success then all businesses that restore quality in old videos might both save time and money by running the video through the algorithm instead of manually detecting and fixing every frame individually using Adobe Photoshop or other software tool. It will also help families or individuals restore quality without spending fortune on it.

Tasks

The most important task of this project will be writing and optimizing a GAN network to generate a new image. This is the only task considered to be acknowledged so the project can be marked as complete, thus it is identified as a must have. If the image generation with minimal loss is achieved in a shorter time than expected authors might consider improvements in algorithm or writing components to turn the solution into a fully-fledged software that can be deployed and used by a regular person.

Project

- GAN network that generates a new image in between two different images
- Estimate and calculate loss between original and generated image

Algorithm improvement (optional)

- Select how many images before and after undamaged part that the network will train on
- Basic feature
    - Automatic pattern detection using machine learning to find frames that are damaged
- Advanced feature

- Detect bounding box sections of image considered damaged
- Generate only damaged parts of an image and replace them

Software implementation (optional)

- Deliver a software where user can select a video to be fixed. The video will be analyzed, restored and rendered back again
- Select types of VHS errors that should be looked for and fixed
- Implement a client / server architecture that will upload video to a server, run network and calculations there and send fixed video back to client.

Method

The method used will be GAN (Generative adversarial network). The reason for this pick as a solution for the project is that the algorithm is unsupervised so no labels are needed to find which entities should be fixed. The network is based on two networks, namely a Generator and Discriminator, which compete with each other in a zero sum game framework. The generator will try to create an image based on the image before the damaged part that will fool the Discriminator into accepting the image as the predecessor of it. This method is often used to generate images that are authentic to human observers so authors think that the network is a perfect fit for this implementation.

These methods will be implemented in Python 3.0 using Tensorflow and Keras. Jupyter notebook, Floydhub and Amazon web services will be used for development.

Milestone chart (time schedule)

1. (2017-11-06) Gather resources and information regarding GAN networks and optimization
2. (2017-11-06) Find two versions of a demo clip that will be used in this project. One that is damaged which represents the input noise for the generator and another one that isn't which represents the real data set for discriminator to value
3. (2017-11-07) Split videos into images for each frame and select sections to test out
4. (2017-11-10) Setup environment and necessary packets to run Tensorflow, panda, Keras
5. (2017-11-16) Write a mockup code for Generator
6. (2017-11-23) Write a mockup code for Discriminator
7. (2017-11-30) Implement a training loop for competition between Generator and Discriminator
8. (2017-12-04) Save generated image
9. (2017-12-07) Search and apply an algorithm to measure the loss of generated image compared to the real expected image
10. (2017-12-10) Improve and refactor code
11. (2017-12-11)Finalize project conclusion and write a report
12. (2017-12-12) Prepare presentation
13. (2017-12-14) Present a working prototype

Risks, Consequences and Ethics

Risks of this project is that we might not be able to deliver before deadline. Both authors of this project are familiar with GAN networks and Tensorflow but neither of them have ever implemented their own solution in the framework. The other risk is simply that the solution might not be good enough and people would rather have their VHS damaged for that old 90's feeling.

So far, authors have only found one factor which might be a questionable regarding both ethics and consequence. The aim of the solution is to replace damaged part of videos as it was never damaged. The network will do its best based on the trained data to predict what was originally at the damaged part, but that does not mean it will be correct. In the case of a human watching a repaired section which is incorrectly generated, it might confuse the person and distort its memories.

An example might be one of the authors watching a video of himself in a red jacket, but the footage of the jacket is damaged so the GAN might generate a green jacket instead. That might make the author wonder where this green jacket came from and why there is no footage of the red one.

Otherwise, there are no other ethical dilemmas or consequences regarding the solution since this solution is only meant to restore and improve damaged parts of a video.

Summary

The project will mainly explore if a GAN network can be used to replace or restore frames considered damaged from a video that is copied digitally to a computer. Instead of using human capital to manually fix anomalies for a hefty price tag the network can be used as a substitute for a cheaper and possibly faster restoration.

Expected results is a generated image that will fit in a sequence of continues images that can be viewed in a row so they will represent a video. If time is not of the essence an interesting approach for analyzing results could be rendering the images together to make a new video. If a random group of people prefer watching a fixed video, rather than watching it with the damaged part, the solution can be considered as an optional solution for video fixing.

Authors have found couple of ways to both improve and optimize the solution to let it possibly create a better damage restoration than only using a single GAN network to replace a whole frame. Narrowing the scope for this project to only using the network seems like a reasonable approach that can meet deadline.

Reference(s)

[1] Photo restoration, web site, 2016, https://github.com/WenchenLi/PhotoRestoration

[2] Image Completion with Deep learning in Tensorflow, web site, 2016, http://bamos.github.io/2016/08/09/deep-completion/

[3] Generating videos with Scene Dynamic, web site, 2016, http://carlvondrick.com/tinyvideo/

[4] Introduction to Restoring Video Quality: How to Improve Video with Filters, web site, 2016, http://www.digitalfaq.com/guides/video/introduction-restore-video.htm