VRIJE
UNIVERSITEIT
BRUSSEL

COMPUTER GENERATED CAR DESIGN

Assignment 3 - Computational Creativity

Lennert Bontinck

March, 2020-2021

Student number: 568702

Computer Science: AI

Abstract

As per requirement of the Computational Creativity course, a creative system will be designed and constructed. This will be done by completing multiple smaller milestones. This report focuses on the third of those milestones. Some more information is given about which data will be used for the system. The expected flow of the system, a schedule for development and some technical aspects are discussed as well.

Section 1.1 discusses what data is needed, how it could be collected and which pre-bundled datasets exist. Section 1.2 highlights the four most important components for the system and what technology should be used for them. In section 2.1 it is explained that some issues where encountered this milestone resulting in a sub-optimal report. Finally a new, less error-prone, schedule is given in section 2.2.

All source files for this project are available on GitHub (Bontinck, 2021). It is noted that this report is written by modifying the VUB based LaTeX template from De Smet (2020).

Contents

I More technical details **1**

1.1 Data to be used 2

1.2 Required components and flow of the system 3

II Conclusions **5**

2.1 Issues this assignment 6

2.2 Expected roadmap 6

References **7**

Part I

More technical details

1.1 Data to be used

In an ideal scenario, the GAN would make use of custom training data collected for this assignment. Some technical detail about the required training data, which are images, is given below:

- Should be of JPG format.
- Should have a maximum resolution of 512x512, preferably even lower (limited computational power).
- Should be of cars from a European brand, Peugeot and Mercedes in particular.
- Should contain the same car in diverse angles.
- Would ideally be labelled with brand, model, colour, body type and release year. The more meta-data the merrier.
- From the literature study it seems over 50 thousand images per brand should be needed, *as a minimum.*

The images could be scraped from the web. Thus a custom scraper would need to be written to collect images. The scraper should work by using popular car auction websites since these would allow for collecting meta-data as well.

While collecting training data for the project would be ideal, this would require a considerable amount of time. Luckily, if it would turn out unfeasible to do this, many existing pre-bundled training sets exist. The LSUN-Stanford Car Dataset by Kramberger and Potočnik (2020) could be used, as discussed in the previous assignment. This dataset consists of over 2 million images.

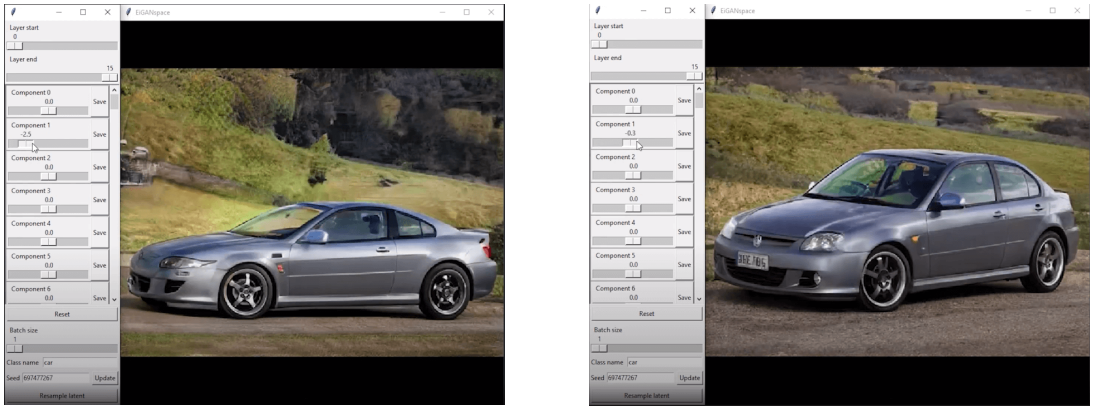
The only drawback of using the LSUN-Stanford Car Dataset is that it consists of images from multiple, mostly American, car brands. My personal knowledge of these is lesser than my knowledge of European car brands. This could impose some issues for the evaluation portion of this project since it would be harder to assess if the generated image resembles an existing car. However, this can be minimized by having enough juries with expertise to evaluate the images. This will be discussed in greater detail in the next assignment.

In both of these cases, the used training data would have been already publicly available so there are no expected issues for licensing.

1.2 Required components and flow of the system

As was already discussed in the previous assignments, the following components are required and will be used from top to bottom:

- A crawler that collects input training images from the web.
  - Can be bypassed by using LSUN-Stanford Car Dataset.
- A DCGAN that generates images of cars based on the collected training images.
  - The official TensorFlow implementation of StyleGan2 will be used for this (Karras et al., 2020).
  - The computational requirement for this is rather gigantic. Luckily a pre-trained car GAN exists as already discussed in the previous assignment. Those could be used as-is or they can be further fine-tuned.
- A framework that allows for control over the GAN.
  - Since the resulting GAN is of the black-box principle, control over the AI is hard, however, not impossible.
  - Some research was already done in the previous assignment surrounding control over hidden layers. Since then, a paper that talks about the development of a GUI for this purpose has caught my attention, GANSpace (Härkönen et al., 2020).
  - GANSpace is compatible with StyleGan2 (Härkönen et al., 2020) and supplies an easy-to-use interface.
  - GANSpace has a demo over the pre-trained StyleGan2 car-related GAN, shown in figure 1.
- An online questionnaire.
  - Allows getting human feedback on whether or not the resulting cars look like actual cars.
  - A tool previously created by me can be modified to hold this survey in a randomised fashion (Bontinck, 2019).
  - More details about this tool will be given in the next assignment.



(a) Two door coupe

(b) Four door sedan

Figure 1: Fragments of the GANSpace demo demonstrating the control over the AI to generate similar looking cars with different properties (Härkönen, 2020).

Part II

Conclusions

2.1 Issues this assignment

As was discussed in the previous assignment the goal was that a basic version of the GAN would already be implemented by now. The training process over a subset of the LSUN-Stanford Car Dataset was initiated (using the first 200.000 images) but crashed prematurely. It wasn't directly clear why my computer crashed and restarting this process and get a working demo on time for this assignment would not have been possible. Because of this, this report was shorter and less fine-tuned than the previous ones. However, a clear understanding of what is needed and how it can be achieved is clearly in place.

2.2 Expected roadmap

The following forms an expected roadmap for the remainder of this project:

- 23/03 - 02/04: Implementation of the evaluation tool and writing of the evaluation assignment.
- 02/04 - 11/04: Development of the GAN.
- 11/04 - 15/04: Control over the GAN and collection of images/videos for evaluation.
- 15/04 - 29/04: Collection of evaluation data.
- 29/04 - 18/05: Writing of the research paper.

References

Bontinck, L. (2019). *Je kijkt er naar, maar ziet het niet: Datacompressie principes — ontstaan en uitdagingen — implementaties — afbeeldingscompressie: Png — jpeg — jpeg2000 — webp — heif — videocompressie: H.264-avc — h.264-svc — h.265/hevc — av1* [Available under "evaluatietool"]. Retrieved March 21, 2021, from https://github.com/pikawika/bachelorproef-compressie

Bontinck, L. (2021). *Computational creativity project* [GitHub commit: 8286859...]. Retrieved March 7, 2021, from https://github.com/pikawika/VUB-CC-Project

De Smet, R. (2020). *Vub latex huisstijl* [GitHub commit: d91f5579...]. Retrieved November 2, 2020, from https://gitlab.com/rubdos/texlive-vub

Härkönen, E. (2020). *Ganspace: Discovering interpretable gan controls.* https://www.youtube.com/watch?v=jdTICDa_eAI

Härkönen, E., Hertzmann, A., Lehtinen, J., & Paris, S. (2020). Ganspace: Discovering interpretable gan controls.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. *Proc. CVPR*.

Kramberger, T., & Potočnik, B. (2020). Lsun-stanford car dataset: Enhancing large-scale car image datasets using deep learning for usage in gan training. *Applied Sciences*, *10*(14). https://doi.org/10.3390/app10144913