1 Reflections

I found a great interest working on this project on as I'd never created an image to image model, considered the memorability of images, or even read a research paper before starting. Luckily my project supervisor, Adrian Bors, suggested a range of great research papers on image memorability, and that allowed me to start with a good footing.

There is a lot of openly available research into machine learning and computer vision, AlexNet sparked somewhat of a machine learning renaissance in 2012, and in the decade since we have seen countless incredible papers released. The industry as a whole is very open, not only do academics produce research but large companies with huge funding typically post their research online for anyone to read. This abundance of research makes getting up to speed on the state of the art very easy. Not only is there a lot of free research available, but there are also countless articles and videos produced with the aim of making learning concepts easier.

A big challenge for me was reading and understanding lots of these maths heavy research papers, I would often have to read a paper multiple times until I truly got what was happening, this was difficult and required perseverance but now I'm more comfortable with having to spend a long time figuring out what a paper is showing.

My dissertation is written using LaTeX, and bar a few small hiccups, it was very easy to learn to use as there are lots of online resources available. I typically write notes in markdown, which syntactically are very similar to TeX files, I imagine this transition was easier than going from a traditional WYSIWYG editor to LaTeX. I had never used LaTeX before but I have grown very fond of it.

I find it difficult to work at home, and prefer to work in the University Library or any of the other various study spaces available. Unfortunately, my laptop is not computationally powerful enough to train most deep networks, my workflow became a cycle of writing code at the library and running it on my computer overnight. This work structure did take a while to get used to but it didn't pose much of a hassle in the end. I could have theoretically set up a Jupyter remote server to run on my computer at home, and connected to that from my laptop, but I was concerned about SSH vulnerabilities. If

I were to do this project again I would spend the time learning how to set up a Jupyter server as I could have been far more productive. I considered training on Google Colab but unfortunately I hit the free GPU usage limit very quickly and didn't want to pay for credits.

I would like to pursue a career in computer vision and what I have learned working on this project will be immensely useful. My goal was to learn to predict memorability maps from images but while trying to achieve that I have learned so much in the surrounding areas of computer vision and deepened my knowledge of deep learning.

I've taken 3 courses at the University of York on intelligent systems, and in one of them implement my own unconditional GANs, because of that the transition to creating my own conditional image to image networks wasn't too foreign. One of my strengths is that I'm really able to stick to a problem for a long time, if something wasn't working I would happily spend hours fixing it without needing a break.

Unfortunately diffusion models take a long time to train, and even longer to evaluate over the approximately 300 images in the VISCHEMA validation set, to evaluate one of my diffusion models with 2000 noise steps over the entire validation dataset would take roughly 2 hours, compared to the seconds for an autoencoder or GAN. This has meant that my third experiment took significantly longer to run than the first two, and has been the cause of many headaches. I regret that I was unable to achieve great results using a diffusion model and wish I had spent more time getting them to work, I spent too much time perfecting my Autoencoder and GAN results and should have spent more time on the diffusion, I naively assumed that it would take a similar amount of time to get good results with all three.

Overall I am proud of the work produced in this project, I have learned a lot about computer vision and memorability and I have produced, what I would consider, great results.