**MSc. in Computing**

**Practicum Approval Form**

# Section 1: Student Details

|  |  |
| --- | --- |
| Project Title: | Low quality facial image restoration using generative models |
| Student ID: | 22267441, 21261888 |
| Student name: | Christian Dave Cobalida, Yamini Pravin Karpe |
| Student email | [christian.cobalida2@mail.dcu.ie,](mailto:christian.cobalida2@mail.dcu.ie,) [yamini.karpe2@mail.dcu.ie,](mailto:yamini.karpe2@mail.dcu.ie,) |
| Chosen major: | Artificial Intelligence, Data Analytics |
| Supervisor | Prof. Hossein Javidnia |
| Date of Submission | 28/11/2022 |

# Section 2: About your Practicum

Please answer all questions below. Please pay special attention to the word counts in all cases.

**What is the topic of your proposed practicum? (100 words)**

The main objective of the proposed practicum is to reconstruct the missing part of an image by learning feature representations from a set of images. The main objective is achieved by training a CNN with the missing part and making it possible to predict what is in the missing part using its features. Given a corrupted image with parts of an image missing or distorted, our objective is to provide a seamless and plausible replacement for a random region of pixels in the image with the help of available visual data. Then we train the model to regress to the missing pixel values. Our objective is to explore ways to improve image reconstruction methods using CNNs and deep learning models. We will use ensemble methods that includes CNN, binomial logistic regression and other neural networks.

**Please provide details of the papers you have read on this topic (details of 5 papers** expected).

1. Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A. Efros, “Context Encoders: Feature Learning by Inpainting,” Proc. International Conference on Computer Vision and Pattern Recognition (CVPR), 2016
2. Chao Yang, Xin Lu, Zhe Lin, Eli Shechtman, Oliver Wang, and Hao Li, “High-Resolution Image Inpainting using Multi-Scale Neural Patch Synthesis,” Proc. International Conference on Computer Vision and Pattern Recognition (CVPR), 2017
3. Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S. Huang, “Generative Image Inpainting with Contextual Attention,” Proc. Computer Vision and Pattern Recognition (CVPR), 2018.
4. Zhang, W. et al., 2021. Context-Aware Image Inpainting with Learned Semantic Priors},. CoRR, Volume abs/2106.07220.
5. Fujii, R., Hachiuma, R. & Saito, H., 2021. RGB-D Image Inpainting Using Generative Adversarial Network with. CoRR, Volume abs/2110.07413

**How does your proposal relate to existing work on this topic described in these papers?** (200 words)

The previous existing works were either based on

1. GANs which is formed of a single Encoder/decoder CNN generator (designed in different ways) in addition to local and/or global discriminator and the loss function was using both the adversarial loss and construction loss.
2. Multi-stage network with 1st stage(generator) in-paint the image with rough estimate and a 2nd stage (texture network) to refine the in-painted part to match the valid pixel and the loss was based on construction and texture loss

Our proposal will be based on using Conditional GANs, the architecture will be consisting of 2 networks:

* Context capture: its job is to be able to capture the feature and the context of the image and be able to generate a set of features. This will be achieved through using a publicly available pre-trained network such as ResNet or inception and applying a transfer learning strategy.
* Conditional GANs: its job is to take the set of features generated by the context capture as a condition to limit the latent space when generating the output image and be able to in-paint the images based on this condition. The generator will be based on encoder/decoder and single/multiple discriminators will be deployed.

The loss shall be based on using Construction loss and Adversarial loss.

What are the research questions that you will attempt to answer? (200 words)

* In-paint image with random missing regions with reasonable pixels that match the context and the texture of the original image.
* Use conditional GANs + Context capture network.
* Transfer learning used for context capture network.
* GANs is based on CNN Generator and discriminator.
* Multiple losses to be deployed (Construction and adversarial).
* Multiple evaluation methods: Qualitative and Quantitative.
* Compare results with existing state of the art results.

How will you explore these questions? (Please address the following points. Note that three or four sentences on each will suffice.)   
  
- What software and programming environment will you use?

* Python will be the main language used, the project will be developed using a combination of Jupyter notebooks and within VS Code in addition to Google Collab.

- What coding/development will you do?

* Image processing on the images of the dataset
* Apply masking on the images to generate the images with missing area
* Train the context network with the images or use transfer learning to extract features need for the conditional GAN
* Train and test the conditional GAN for image inpainting
* Visualize the inpainting results
* Evaluate the inpainting results using qualitative and quantitative metrics.

- What data will be used for your investigations?

* Flickr-Faces-HQ (FFHQ) is a high-quality image dataset of human faces, originally created as a benchmark for generative adversarial networks (GAN).

- Is this data currently available, it not, where will it come from?

* The dataset is available on Kaggle consists of 52,000 high-quality PNG images at 512×512 resolution and contains considerable variation in terms of age, ethnicity and image background. It also has good coverage of accessories such as eyeglasses, sunglasses, hats, etc.

- What experiments do you expect to run?

The datasets will be divided into train/validation/test sets and below experiments are expected to take place:

* try different ways to extract the context features of the images
* try different ways to apply the condition to the GAN
* try different implementations of the CNN generator
* try different adversarial loss functions
* What output do you expect to gather?

- How will the results be evaluated?

Results will be evaluated by using an unseen test set of images and allowing the network to in-paint the missing gaps which will be evaluated:

* Qualitatively (manually judged)
* Quantitatively (L1/L2 Loss, PSRN)