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, 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. Y. Bengio. Learning deep architectures for AI. Foundations and trends in Machine Learning, 2009.
2. G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. Science, 2006.
3. C. Doersch, A. Gupta, and A. A. Efros. Context as supervisory signal: Discovering objects with predictable context. In ECCV, 2014.

4. A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. ICLR, 2016.
5. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In NIPS, 2014.
6. J. Hays and A. A. Efros. Scene completion using millions of photographs. SIGGRAPH, 2007.

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

Computer vision has made tremendous progress on semantic image understanding tasks such as classification, object detection, and segmentation in the past decade. Convolutional Neural Networks (CNNs) [1, 2] have greatly advanced the performance in these tasks [3, 4, 5]. The success of such models on image classification paved the way to tackle harder problems, including unsupervised understanding and generation of natural images. Some of the earliest work in deep unsupervised learning are autoencoders [1, 2]. Doersch, et al. [3] used the task of predicting the relative positions of neighbouring patches within an image as a way to train an unsupervised deep feature representations. Radford et al. [4] proposed new convolutional architectures and optimization hyperparameters for Generative Adversarial Networks (GAN) [5] producing encouraging results.In computer graphics, filling in large holes is typically done via scene completion [6], involving a cut-paste formulation using nearest neighbours from a dataset of millions of images. We intend to utilize the research available to approach the problem with a more advanced model which may lead to a better understanding of facial image restoration.

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

1. Can we use Deep Learning techniques to restore corrupted facial images generated from the dataset?

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?

- The environment will revolve around the Anaconda framework. It supports Python packages and has RStudio which will be used in detecting and visualising facial images. Visual Studio Code might also be used for writing and editing code.

- What coding/development will you do?

- Training the model with various algorithms such as GAN, CNN, loss functions, etc.

- 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?

- We train our context encoders by regressing to the ground truth content of the missing (dropped out) region and then using two loss functions to determine the errors between the output and the given target value.

- What output do you expect to gather?

- The trained model will be able to determine whether the facial image is real or fake.

- How will the results be evaluated?

- It will be probability based on various tests performed during the training on a facial image output such as inference on the effect of epochs in training, using only one loss function and using additional loss function, etc. It will also be evaluated how our model performed compared to other models using L1 to L2 loss metrics.