Electronics and Computer Science
Faculty of Engineering and Physical Sciences
University of Southampton

Enhanced Generative Image Transformation (EGITT) Progress Report

Georgi Iliev
gdi1u21, 33086168

Supervisors
Supervisor: Hikmat Farhat
Second Examiner:

A project progress report submitted for the award of Computer Science Meng

December 12, 2023

Statement of Originality

- I have read and understood the ECS Academic Integrity information and the University's Academic Integrity Guidance for Students.
- I am aware that failure to act in accordance with the Regulations Governing Academic Integrity may lead to the imposition of penalties which, for the most serious cases, may include termination of programme.
- I consent to the University copying and distributing any or all of my work in any form and using third parties (who may be based outside the EU/EEA) to verify whether my work contains plagiarised material, and for quality assurance purposes.

You must change the statements in the boxes if you do not agree with them.

We expect you to acknowledge all sources of information (e.g. ideas, algorithms, data) using citations. You must also put quotation marks around any sections of text that you have copied without paraphrasing. If any figures or tables have been taken or modified from another source, you must explain this in the caption and cite the original source.

I have acknowledged all sources, and identified any content taken from elsewhere.

If you have used any code (e.g. open-source code), reference designs, or similar resources that have been produced by anyone else, you must list them in the box below. In the report, you must explain what was used and how it relates to the work you have done.

I have not used any resources produced by anyone else.

You can consult with module teaching staff/demonstrators, but you should not show anyone else your work (this includes uploading your work to publicly-accessible repositories e.g. Github, unless expressly permitted by the module leader), or help them to do theirs. For individual assignments, we expect you to work on your own. For group assignments, we expect that you work only with your allocated group. You must get permission in writing from the module teaching staff before you seek outside assistance, e.g. a proofreading service, and declare it here.

I did all the work myself, or with my allocated group, and have not helped anyone else.

We expect that you have not fabricated, modified or distorted any data, evidence, references, experimental results, or other material used or presented in the report. You must clearly describe your experiments and how the results were obtained, and include all data, source code and/or designs (either in the report, or submitted as a separate file) so that your results could be reproduced.

The material in the report is genuine, and I have included all my data/code/designs.

We expect that you have not previously submitted any part of this work for another assessment. You must get permission in writing from the module teaching staff before re-using any of your previously submitted work for this assessment.

I have not submitted any part of this work for another assessment.

If your work involved research/studies (including surveys) on human participants, their cells or data, or on animals, you must have been granted ethical approval before the work was carried out, and any experiments must have followed these requirements. You must give details of this in the report, and list the ethical approval reference number(s) in the box below.

My work did not involve human participants, their cells or data, or animals.

ECS Statement of Originality Template, updated August 2018, Alex Weddell aiofficer@ecs.soton.ac.uk

Abstract:

In recent years, Image-to-Image transformation has emerged as a fundamental component in numerous image processing and computer vision applications. Despite advancements, challenges persist, including issues such as vanishing gradients, inadequate evaluation metrics, and mode collapse, all necessitating novel solutions. This progress report outlines the ongoing dissertation project. It aims to tackle these challenges through the introduction of an adaptable and transparent Image-to-Image Transformation framework that employs a two-stream generative model and incorporates a prominent feature translation in a coarse-to-fine manner. Extensive experimentation demonstrates the model's proficiency in delivering high-quality image transformation, thereby paving the way for practical applications in diverse scenarios. The training methodology combines both Generative Adversarial Networks (GANs) and Diffusion Models, resulting in a sophisticated yet standardised tool.

This hybrid approach benefits from the strengths of both techniques, leading to improved stability during training, better initialisation and generation of high-quality samples.

The project's ultimate goal is developing a comprehensive and all-encompassing Image-to-Image transformer that aligns with recent advances in the field and furnishes practical insights into its adaptability across various datasets and scenarios.

This report presents a thorough analysis of the ongoing advancement, highlighting the project's contributions towards the dynamic and constantly investigated research arena of Image-to-Image transformation.

1. Project Description

Image-to-image transformation has recently gained significant prominence owing to its capability to convert input images from one domain to another while preserving the core content during style transfers [1]. This flexibility makes it suitable for various tasks while ensuring high-quality synthesis [2]. The prime objective of this project is to seamlessly incorporate elements of diverse Image-to-Image translation techniques, resulting in a comprehensive and unified framework that yields compelling results.

The progress made so far concentrates on Image Generation, with a particular emphasis on GANs. Through trial and error, the student familiarises themselves with various techniques, contributing to the development, training, and testing of a standard GAN, a DCGAN and a Wasserstein Conditional GAN with Clipping. The research involves investigating well-defined loss functions such as cross-entropy and cosine quantization, and Wasserstein distance - all potential candidates for the adopted loss function in the final construction. The preparation and construction of Image Generation tasks accommodate both the primal and dual-domain approaches, allowing end-to-end training through back-propagation in a minimax optimisation mechanism [5]. Building upon these foundational and comprehensive processes will result in the development of a robust and adaptive framework addressing both one-to-many translation tasks in computer graphics and many-to-one translation tasks in computer vision [3] effectively.

2. Goals

The main objectives of this paper are to extensively examine different image translation techniques, with a particular focus on proposing an enhanced framework that features high image quality and synthesises images in a manner consistent with the latest advances [6] in the field. A key goal is to rigorously evaluate the generated images using quantifiable metrics in measuring image quality, thus furnishing a comprehensive insight into model performance [7].

This project aims to contribute to the advancement of image transformation by initially conducting an in-depth study and comparison of various GAN models, including standard GANs, Wasserstein GANs, conditional GANs, cyclic GANs, and deep convolutional GANs. Through this comparative analysis, valuable insights into the strengths and limitations of each approach are gained, facilitating the selection of the most suitable GAN modification for specific Image-to-Image translation tasks [8].

In pursuit with these objectives, this project ultimately seeks to develop a coherent and efficient model that aligns with recent advancements in the field. This model will operate efficiently across multiple datasets while also providing practical insights into the feasibility of diverse GAN modifications in different scenarios [9].

3. Background and Literature Review:

3.1. Backbone of Image-to-Image Transformation

The backbone of Image-to-Image Transformation serves as a foundational aspect of understanding the core principles and challenges of this transformative process. Image-to-Image translation, as a process, seeks to map an image z_X from a source domain X to a corresponding image z_Y in a target domain Y [10]. This mapping, denoted as P, necessitates indistinguishable image generation, reliant on generative models [1]. It is designed to preserve internal content while transferring external style [1] and is manifested through a specifically chosen loss function. The inclusion of the inverse mapping Q (distribution of images = Y-> X) enforces cycle consistency loss, ensuring $Q(P(X)) \approx X$ (and vice versa) [16].

The achievement of successful training is directly correlated with the selection of an appropriate loss function. In this context, four methods are identified, namely supervised, unsupervised, semi-supervised, and few-shot. Supervised training, exemplified by models like pix2pix [3], relies heavily on well-annotated paired training samples. On the contrary, unsupervised I2I translation presents inherent difficulties that require further constraints such as pixel gradients [18], cycle consistency [16], pixel values [19] and semantic features. The model, in this case, adapts to dataset-specific behaviours [2]. It processes a dataset of input-output entries through a parametric translation function that employs Convolutional Neural Networks (CNN) to cater to their local receptive fields and weight sharing [11].
Upon obtaining translated results from the generative model, the choice of suitable evaluation metrics becomes paramount. This comprehensive process emphasises the versatility of I2I transformation, which applies to various applications when provided with qualitative and quantitative satisfiable source-target ground truth data [1].

However, despite these advancements, certain limitations persist, particularly those related to feature extraction [5], computation, background changes, and non-linearity [8]. Real-world applications are grappling with challenges such as the scarcity of paired input-output samples and the potential effects of undesired content that deviate attention from the most discriminative aspects of images during translation [8]. These challenges underscore the significance of ongoing research and development aimed at imperatively addressing and mitigating them. This allows for further advancements in the field of Image-to-Image transformation.

3.2 Generative Models

Deep neural networks have revolutionised the landscape of artificial intelligence across various application domains. Generative Models play an essential role within the realm of Deep Neural

Networks [6]. Image-to-Image transformation relies heavily on image generation, which is solved immaculately by Generative Models. These models operate in semi-supervised [1] and unsupervised learning conditions by generating new samples from an unidentified data-generating distribution [7]. Their main goal is to acquire data resembling the training dataset [7].

In the early stages of this project, the student familiarises themselves with Generative Models, by conducting extensive literature reviews and experimenting with different Image Generation applications, such as Deep Convolutional Generative Network (DCGANs), Conditional GANs [3], Style and Cycle GANs [16], but primarily accentuates on Generative Adversarial Networks (GANs).

Generative Adversarial Networks, introduced by Goodfellow et al. in 2014 [31], are a significant paradigm within Generative Models. They play a pivotal role in Image-to-Image Transformation. In GANs, the mapping from a random noise vector (z) to an output image (y) is learned by the Generator network (G: z -> y) [3]. This process encompasses a Minimax game between two neural networks: the Generator, which produces plausible examples from a random noise vector [3], and the Discriminator, which distinguishes between real and fake data by endeavouring to categorise the examples as authentic or synthetic [10]. The Generator strives to deceive the Discriminator by fabricating realistic samples, whereas the Discriminator aims to precisely classify and label the received data as true or false [10]. Moreover, GAN's success largely depends on the adversarial loss, which guarantees that the generated images are indistinguishable from real ones [16].

Numerous GAN variants have been introduced, finding applications in machine learning and computer vision [7] and crucially contributing to Image-to-Image translation [3], text and texture synthesis [20], hand-written font generation [21], and medical applications [22], among others.

Despite the significant advancements achieved by GANs, stable training remains a substantial challenge, impacting both the Generator and Discriminator. Persistent issues including mode collapse [23], Nash equilibrium [24], lack of proper evaluation metrics [23], and vanishing gradient [25] continue to pose challenges. Although a range of solutions have been proposed, some only partially address specific bottlenecks without comprehensively mitigating all the persistent obstacles.

4 Account of work to date

The inception of this project is characterised by careful preparation, emphasising the all-encompassing development procedure and a vigilant awareness of potential risks. A comprehensive Risk Assessment and a Gantt Chart prove to be invaluable tools during this preliminary stage, furnishing a structured framework for supervising the project. The Risk Assessment, using a methodical five-step process that includes planning, identification of hazards, evaluation of risks, precautionary suggestions, and continuous review [13], is illustrated in Figure 1. This visual representation of the risk assessment emphasises the proactive measures taken to anticipate and overcome potential hurdles.

Simultaneously, the Gantt Chart, depicted in Figure 2, serves as a visual record of the anticipated work schedule. The first semester is primarily focused on an in-depth analysis of existing literature and the formulation of the initial project draft. The dynamic nature of the development process implies inevitable modifications to the schedule, but the student remains firmly focused on the overarching project objectives. It is imperative to acknowledge that regardless of any rescheduling of development phases, the student has consistently dedicated substantial hours to this dissertation project, thereby reinforcing their commitment to achieving the final goal.

Subsequently, the project then proceeds with the identification of an appropriate dataset for constructing the experimental drafts. Several datasets are examined, with a particular emphasis on two primary sources: the CelebA dataset [12] and the Cityscapes dataset [4]. The CelebA dataset contains 202,599 face images of celebrities and includes complex backgrounds, featuring 40 binary attributes such as age, gender, and hair characteristics. Meanwhile, the Cityscapes dataset concentrates on the principle of semantic labels ↔ photographs, providing an exclusive point of view for evaluation.

Risk assessment

[14]

Problem	Loss	Probability	Risk	Mitigation
Difficulty with PyTorch Development and Image-to-Image transition	4	3	12	Gather in-depth understanding of PyTorch through early study and practice; Seek guidance from the project supervisor; Maintain regular progress monitoring and debugging
Not gaining access to a powerful enough computer for training, tuning, and testing the model	4	1	4	Investigate available computing resources from the University in advance; Consider alternative computing resources or cloud - based solutions
Other modules take significant amount of study time	2	5	10	Establish a detailed time management plan allocating time for each module; Prioritise project-related tasks and maintain flexible schedule to accommodate unforeseen academic demands; Seeking help from supervisor or reducing workload if required
Data Availability and Quality	4	3	12	Conduct extensive research on available datasets; Create strategies for supplementing data; Consider data collection as needed; Apply data quality measure
Over-ambitious Scope	3	3	9	Define clear project boundaries and objectives; Continuous reviewing of the project scope and consideration of expansion only after the initially set objectives have been achieved
User Acceptance and Feedback	2	4	8	Plan user testing and feedback acquisition to guarantee tool meets user expectations; Make incremental improvements based on user feedback; Interact with potential users and experts for guidance.

Figure 1: Risk Assessment of the project

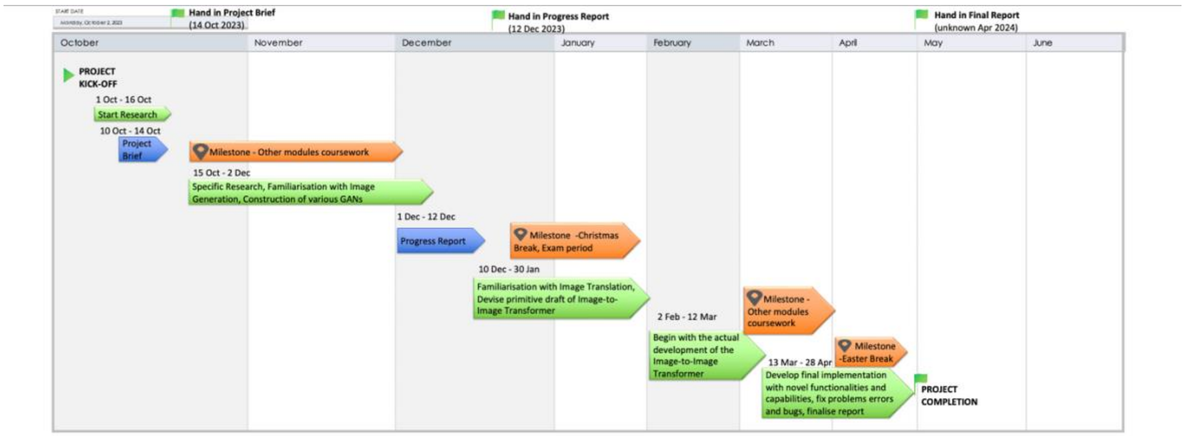

Figure 2: GANTT Chart of the project [14]

These datasets are strategically chosen to support Unsupervised multi-domain Image-to-Image Generation and Transformation. Their compatibility and adaptability to GANs influence this selection process. The datasets yield both qualitative and quantitative results, thus validating their effectiveness in the project's context. The inclusion of the datasets in the project is further justified by their ability to produce tangible results, which can be compared against ground truth images, as demonstrated by Figures 6,7 and 8. This also reinforces the project's basis and aligns with the objectives outlined in the project description and goals.

Following the dataset selection, the subsequent phase involves the creation and implementation of two crucial components in both the Image Generation and Image-to-Image transformation processes. These components, namely the generator and the discriminator, are shown in Figure 3. They engage in a Minimax game, which is a fundamental concept within GANs [31]. The architecture of both the Generator and Discriminator incorporates multiple dimensions, expanding from features multiplied by two to four, six, eight, and sixteen dimensions. The Generator excels at producing synthetic yet plausible images, contributing to the generation process [1]. Conversely, the Discriminator effectively distinguishes between generated (fake) and authentic data [1]. Notably, the designed components demonstrate their versatility by being adaptable and implementable across various GAN models.

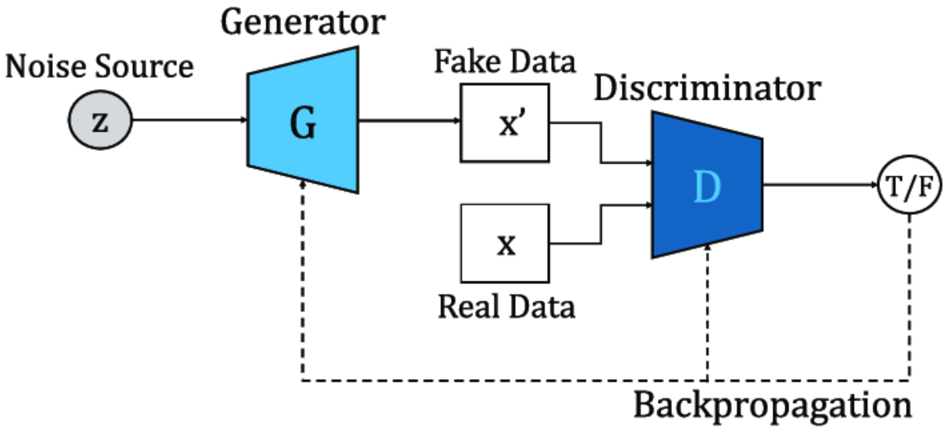

Figure 3: Generator and Discriminator modus operandi [17]

Selecting an appropriate loss function is crucial to the developmental trajectory. A well-chosen loss function has the potential to provide a unified approach that applies to tasks with various loss formulations [3]. During the examination of various loss functions across tasks, the student prioritises Wasserstein GAN due to its use of similarity-preserving learning and quantization error control [27]. Nevertheless, introducing critic functions into WGANs sometimes necessitates the imposition of a 1-Lipschitz constraint, which can be addressed using heuristics like weight clipping [7], spectral normalisation [1], and gradient penalty [1]. Overall, this methodology offers two significant benefits: enhanced stability for optimisation and correlation with the generator's convergence, resulting in improved sample quality and a meaningful loss metric [27].

While the Loss Function can be compared as an internal measure that evaluates a model's accuracy in approximating a desired output, performance metrics are employed post-training to evaluate the model's accuracy in predicting unseen data sets [28]. Following extensive experimentation with loss functions using Wasserstein GANs, the student navigates towards using the Wasserstein/ Earth Mover's Distance, defined as a metric in Figure 4 [5]. However, although aligned with the loss function experimentation, its effectiveness in developing the Image-to-Image transformer is surpassed by a preference for the Fréchet Inception Distance (FID) [39] with the formulation presented in Figure 5. FID is conceptualised as the distance between two instances of the same graph immersed in a metric space [39]. Its features incorporate reduced susceptibility to noise, with a lower value indicating overall better performance.

$$W(Pr, Pg) = \inf_{\gamma \in \pi(Pr, Pg)} E_{(x, y) \sim \gamma} [\, ||x - y|| \,]$$

Figure 4: WGAN suggested replacement distance derived from Earth-Mover (EM) or Wasserstein distance [20]

$$FID(r, g) = ||\mu_r - \mu_g||_2^2 + Tr(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g})$$

Figure 5: Fréchet Inception Distance equation evaluating the quality of generated samples [7]

5 Proposed Design

All experiments and implementations are conducted using Jupyter Notebooks within the PyTorch framework on the Google Collab Pro software, which is purchased using the allocated budget for the Third-Year Project. This software yields quicker results than running on a personal machine. The total number of training epochs depends on the complexity of the solution. If multiple generators are used, the number of epochs may need adjustment to accommodate the increased computational demand. The generated samples are three-channel images with a resolution of 64x64 pixels, providing a tangible output for evaluation and further analysis. This implementation aligns with the objectives outlined in the project brief and contributes to the foundational aspects of the Image-to-Image Transformation tool. The utilisation of the PyTorch framework and the structured architectural progression of the Generator and Discriminator networks assist in achieving the project's objective of developing a cohesive and robust framework.

This research concentrates on supervised and unsupervised methods, each with its own set of pros and cons. The experimental phase encompasses the application of different GAN models, each with specific objectives outlined in the project description.

A sequential protocol guides the standard GAN experiment, comprising data loading and preprocessing, generator and discriminator, weight initialisation, training loop execution, loss function, optimiser application, and visualisation of outcomes. Graphical representations of the initial, intermediate, and final training outcomes are illustrated in Figure 6, Figure 7, and Figure 8, respectively. The ultimate goal is to train the image generation process to produce authentic images similar to those present in the CelebA dataset [12].

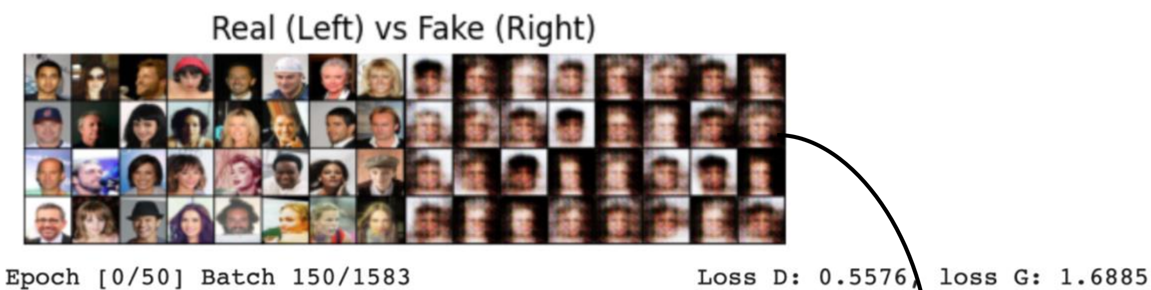

Figure 6: Start of training standard GAN

Figure 7: Progress of training standard GAN

Figure 8: End of training standard GAN

The Deep Convolutional GAN (DCGAN) adds complexity with a Generator that employs transpose convolutional layers to upscale and batch normalisation. Meanwhile, the Discriminator features convolutional layers, leaky ReLU activation, and batch normalisation for effective distinction. The code for this task is trained on both the MNIST [15] and CelebA [12] datasets, demonstrating the adaptability of the GAN model to diverse datasets.

The implementation of the Conditional Wasserstein GAN with Gradient Penalty extends the experiment by incorporating model weight initialisation, training loops, data loading and configuration, generator-discriminator interaction, logging, visualisation, and consideration of hyperparameters, and is depicted in Figures 9 and 10. This model mitigates certain limitations inherent in standard GANs such as mode collapse and training instability, by introducing the gradient penalty mechanism.

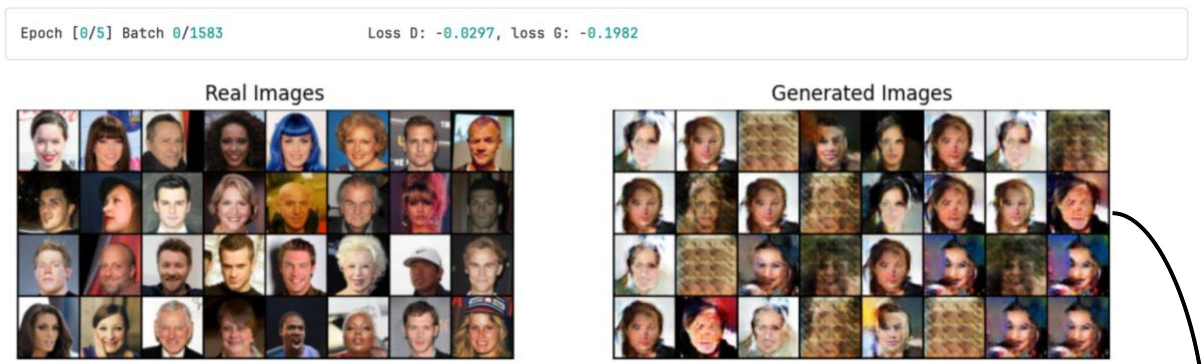

Figure 9: Start of training Wasserstein GAN with Gradient Penalty

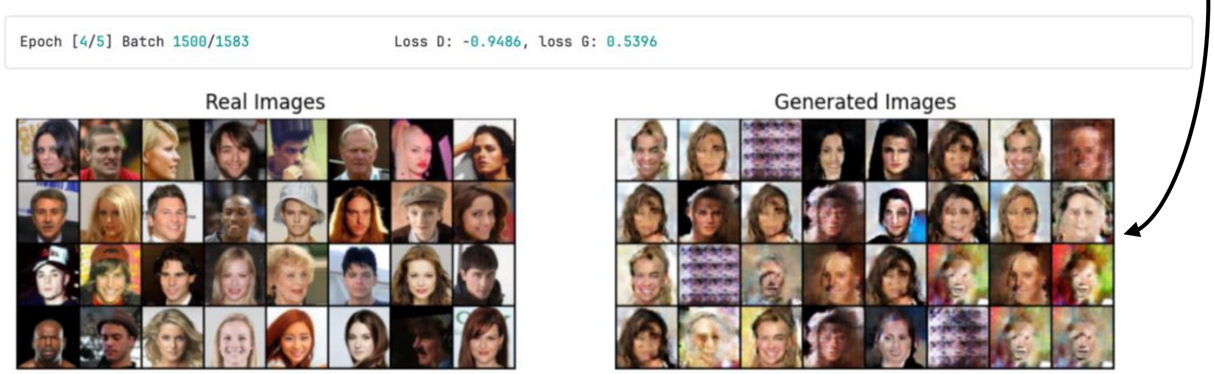

Figure 10: End of training WassersteinGAN with Gradient Penalty

Building upon the foundational knowledge gained from experimental GAN implementations, the development focus shifts towards the construction of a mock Image-to-Image Transformer. This transformative process draws inspiration from the pix2pix [3] architecture, undergoing modifications to enhance its capabilities. Expanding beyond the conventional generator and discriminator, this transformer incorporates an encoder or variational autoencoder component. The implemented loss function comprises a GAN loss, a feature matching loss, a VGG loss (which measures the Euclidean distance between the generated image and the real image maps), and a KL divergence loss. The training process is fine-tuned using Adam optimisers, each with distinct learning rates and betas.

The preprocessing phase involves the migration of tensors to the GPU and the transformation of label maps into one-hot encoding. Furthermore, the utility methods are implemented to extract edges from tensors, re-parameterise the variational autoencoder, and verify GPU availability. A class was created to generate HTML pages and facilitate visualisation and summarisation of the results. A dedicated testing loop ensures thorough evaluation, enhancing the model's comprehensiveness. Despite the significant strides made in the development of the Image-to-Image mock transformer, its development requires additional refinement and modification. Presently, the model has not produced any visual output since various merging components are vital for successful training. Therefore, there is still room for improvement. Future efforts will be directed towards addressing these limitations and enhancing the capabilities of the model in the upcoming stages of the project.

6 Plan of the Remaining Work

The project's initial stages concentrate on examining various Image Generation models. As attention moves towards the practical application of image transformation, a significant turning point emerges. The forthcoming work schedule will encompass the remainder of the first semester and the entire second semester, with an emphasis on the actual image transformation. Feature extraction is a crucial process that will be used to extract fundamental image components that ensure seamless transfer. The plan involves reviewing the existing mock transformer and refining it until the desired outcome is achieved. Subsequently, a new transformer will be constructed and equipped with advanced capabilities to yield comprehensive and satisfactory results. A key phase of this process entails a comparative analysis of all the constructed models. The version demonstrating the most favourable outcomes shall be chosen to progress further.

In accordance with the project's objectives, the integration of Diffusion Models would constitute a significant accomplishment owing to their demonstrated excellence in producing superior images and contemporary adoption in the field. This necessitates the development of a hybrid algorithm that seamlessly merges GANs and Diffusion Models. Both techniques, renowned for their competency in generative tasks, are presumed to synergistically enhance image synthesis, facilitate adaptation to diverse datasets, and foster more stable training environments. The introduction of Diffusion Models is thought to strengthen training stability by contributing to better initialisation and the production of high-quality samples. However, it is essential to acknowledge that, despite their potential, the combination of these models presents challenges and requires further exploration, given the limited existing research in this domain. The upcoming phases of the project will delve into overcoming these challenges and refining the hybrid model so that optimal performance is achieved.

7 Support Required to Complete the Project

Further support may be required for the successful development of the model. The current subscription to Google Collab Pro fails to meet the student's expectations. The limited availability of faster GPUs on Google Collab Pro obstructs continuous usage, prompting the student to explore alternatives. While considering the utilization of the Idiridis Supercomputer at the University of Southampton, the student is informed that it may not reach its optimal performance for Image-to-Image Transformation tasks.

Due to the significant amount of time consumed by training and running the models, the student is currently searching for viable alternatives to tackle the present challenges.

8 References

[1] Pang, Y., Chen, Z., Qin, T., & Lin, J. (2021). Image-to-Image Translation: Methods and Applications. Arxiv. https://arxiv.org/pdf/2101.08629.pdf
[2] Jiang, L., Zhang, C., Huang, M., Liu, C., Shi, J., & Loy, C. C. (2020, July 25). Tsit: A simple and versatile framework for image-to-image translation. arXiv.org. https://arxiv.org/abs/2007.12072
[3] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2018, November 26). Image-to-image translation with conditional adversarial networks. arXiv.org.x
[4] Cordts, M., Omran, M., & Ramos, S. (2020, October 17). The cityscapes dataset. Cityscapes Dataset. https://www.cityscapes-dataset.com/
[5] Doan, K. D., Manchanda, S., Wang, F., Keerthi, S., Bhowmik, A., & Reddy, C. K. (2020, March 30). Image generation via minimizing Fréchet distance in discriminator feature space. arXiv.org. https://arxiv.org/abs/2003.11774
[6] Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017, June 21). CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. Arxiv. https://arxiv.org/pdf/1706.07068.pdf
[7] Li, X., Omar, B., & Jabbar, A. (2020, June 9). A Survey on Generative Adversarial Networks: Variants, Applications, and Training. Arxiv. https://arxiv.org/pdf/2006.05132
[8] Tang, H., Xu, D., Sebe, N., & Yan, Y. (2019, August 27). Attention-guided generative adversarial networks for unsupervised image-to-image translation. arXiv.org. https://arxiv.org/abs/1903.12296
[9] Waite, J. (2020). Generative Adversarial Networks (dissertation) - Electronics and Computer Science Faculty of Physical Sciences and Engineering University of Southampton
[10] Jiang, S., Tao, Z., & Fu, Y. (2019, January 23). Segmentation guided image-to-image translation with Adversarial Networks. arXiv.org. https://arxiv.org/abs/1901.01569
[11] Peters, T., & Farhat, H. (2023, November 21). High-resolution image-based malware classification using multiple instance learning. arXiv.org. https://arxiv.org/abs/2311.12760
[12] Liu, Z., Luo, P., Wang, X., & Tang, X. (2021, September 10). Large-scale celeb faces attributes (celeba) dataset. CelebA Dataset. https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html
[13] Champion, J. (2000). Risk assessment: A Five-step process - Sage journals. Sage Journals. https://journals.sagepub.com/doi/abs/10.1177/175045890001000701
[14] Ramos, D. (2020, September 29). Free monthly Gantt chart templates. Smartsheet. https://www.smartsheet.com/content/monthly-gantt-chart-templates
[15] LeCun, Y., Cortes, C., & Burges, C. (1998, November). The mnist database. MNIST handwritten digit database. http://yann.lecun.com/exdb/mnist/
[16] Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2020, August 24). Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv.org. https://arxiv.org/abs/1703.10593

[17] TONELLO, A. M., LETIZIA, N. A., RIGHINI, D., & MARCUZZI, F. (2019, June 17). Machine learning tips and tricks for power line communications. IEEE Access. https://ieeexplore.ieee.org/document/8737766
[18] Bousmalis, K., Silberman, N., Dohan, D., Erhan, D., Krishnan, D.: Unsupervised pixel-level domain adaptation with generative adversarial networks.
[19] Shrivastava, A., Pfister, T., Tuzel, O., Susskind, J., Wang, W., Webb, R.: Learning from simulated and unsupervised images through adversarial training.
[20] N. Jetchev, U. Bergmann, and R. Vollgraf, "Texture synthesis with spatial generative adversarial networks," arXiv preprint arXiv: 1611.08207, 2016.
[21] Y. Tian. zi2zi: Master Chinese calligraphy with conditional adversarial networks, 2017.
[22] Y. Xue, T. Xu, H. Zhang, L. R. Long, and X. Huang, "Segan: Adversarial network with multi-scale l1 loss for medical image segmentation," Neuroinformatics, vol.16, pp.383–392, 2018.
[23] I. Goodfellow, "NIPS 2016 tutorial: generative adversarial networks," arXiv preprint arXiv:1701.00160, 2016.
[24] L. Ratliff, S. Burden, and S. Sastry, "Characterization and computation of local Nash equilibria in continuous games," in Annual Allerton Conference on Communication, Control and Computing, pp. 917–924, 2013.
[25] A. Borji, "Pros and cons of gan evaluation measures," in Computer Vision and Image Understanding, vol. 179, pp. 41–65, 2019.
[26] I. Goodfellow, J. Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengi, "Generative adversarial nets," in Neural Information Processing Systems, pp. 2672– 2680, 2014.
[27] Arjovsky, M., Chintala, S., & Bottou, L.´eon. (2017, December 6). Wasserstein Gan. arXiv. https://arxiv.org/pdf/1701.07875.pdf
[28] Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A., & Chavez-Urbiola, E. A. (2023, September 6). Loss functions and metrics in deep learning. arXiv.org. https://arxiv.org/abs/2307.02694