MCAST IICT Research Paper

Kledi Azzopardi
Institute of Information & Communication Technology
Malta College of Arts, Science & Technology
Corradino Hill
Paola PLA 9032
{kledi.azzopardi.d56725}@mcast.edu.mt

***Abstract*—This dissertation investigates how artificial intelligence (AI) models compare to traditional tools in the context of image editing. Specifically, it evaluates Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) against Adobe Photoshop for two common editing tasks: hair colour transformation and image inpainting. The research addresses the growing demand for faster and more scalable photo editing solutions without compromising realism. Over 530,000 facial images were used to train the models, and results were assessed using both quantitative metrics (MSE, SSIM, PSNR) and qualitative user ratings based on realism, blending, and detail preservation. The findings show that GANs produce results close to Photoshop in terms of visual quality while significantly improving speed and automation. VAEs, although efficient and interpretable, produced blurrier outputs and scored lower across all evaluation metrics. This research concludes that GANs offer a promising alternative to manual editing for specific tasks, and it encourages further investigation into combining the strengths of both AI models to enhance realism and control in automated editing workflows.**

***Index Terms*—GAN, VAE, Photoshop, Image Editing, Visual Realism**

I. Chosen Research

This dissertation aims to bring more realism to photo editing through the comparison of Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) with classical image manipulation techniques, i.e., Photoshop. The dissertation discusses how generative artificial intelligence-based models can be leveraged to automate and improve some of the photo-editing tasks, i.e., hair colour change and inpainting an image, with extremely high visual realism. Manual photo editing with Photoshop is traditional and time-consuming, while AI approaches provide efficient, automated, and scalable solutions. AI editing is by no means without its share of problems with regard to computational requirements, ethical issues, and simplicity in dealing with realism.

A. Research Onion

1) **Philosophical Approach**: The ideal approach is adopted in this research, whereas the objective is to explore and compare quantitative findings based on AI-generated images.
2) **Research Approach**: A deductive approach that applies previous hypotheses on GANs, VAEs, and traditional photo editing methods.
3) **Strategy**: A comparative experimental study using AI models trained on a dataset of facial images.

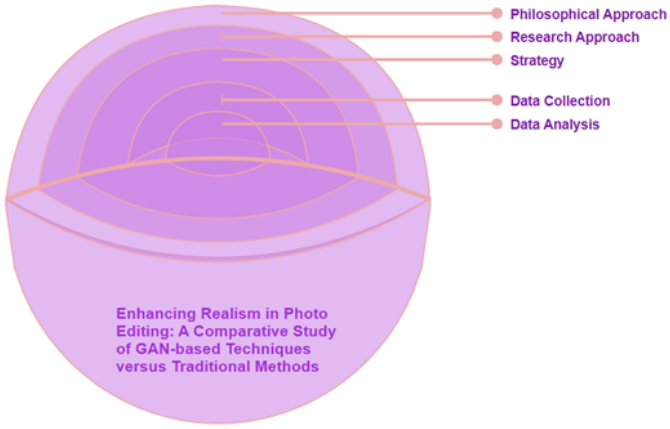

Fig. 1. Research Onion showing the layered methodology used to compare GAN-based and traditional photo editing techniques.

4) **Data Collection**: Empirical evaluation of the quality of images based on metrics such as MSE (Mean Squared Error), SSIM (Structural Similarity Index), and PSNR (Peak Signal-to-Noise Ratio).
5) **Data Analysis**: Quantitative measurement utilising AI-powered outputs, and qualitative evaluation by visual verification.

B. Background to this Research Theme

Image editing plays a vital role in creative industries, digital marketing, and even forensic analysis. Traditional tools like Photoshop are highly effective but require time and skill. In contrast, artificial intelligence has introduced new ways to automate and speed up the editing process. Techniques like GANs and VAEs have changed the way images are enhanced, allowing for automatic tasks like super-resolution, inpainting, and style transfer. GANs are made up of two parts: a generator that tries to produce realistic images, and a discriminator that checks how real the image looks. VAEs use something called a latent space to apply clear, structured changes to images. Both methods have their strengths and weaknesses. Conditional GANs can handle many types of image-to-image translation, such as converting labels into realistic photos, which shows how useful they can be in different editing tasks [1].

C. Hypothesis

Previous research, including [2] on StyleGAN, has shown that AI models can produce realistic image edits while keeping the subject's identity clear. However, earlier work has not compared AI-based editing methods with standard manual tools like Photoshop, which is the main focus of this study. This research is based on the idea that using GANs to edit images can make them look more realistic than Photoshop. VAEs, while useful for making controlled changes, often result in blurrier and less detailed images. The study also suggests that combining GANs and VAEs could offer the best of both approaches: high-quality visuals and better control. Lastly, it assumes that AI editing is much faster and easier to use than doing everything manually in Photoshop.

D. Research Aim and Purpose Statement

The aim of this research is to evaluate how well AI-based image editing techniques perform compared to traditional Photoshop editing. It focuses on measuring the realism, structure preservation, and overall efficiency of approaches like GANs and VAEs, while also considering their individual strengths and weaknesses. The study further explores whether AI-driven editing tools are practical and effective for everyday use in fields such as digital marketing, fashion, and forensic imaging.

II. Review of Research Methodology

Multiple research investigations explore the utilisation of AI technology for image modifying processes. The two major AI approaches which exist today are GANs and VAEs. Research has shown GANs to become celebrated for their ability to manufacture exceptional photo-realistic images. The work of [2] represents an excellent example of StyleGAN. The popularity of VAEs stems from their capacity to provide image editing structure, although their final images have lower clarity compared to those generated by GANs. GANs generate high-quality images with known training issues, yet VAEs produce organised results, although they introduce image blur [3]. GANs can achieve high-resolution outputs even when upscaling single images, reinforcing their ability to maintain clarity during complex transformations [4].

A. Literature Map

B. Distinguishing Between Journals

Associate professors review academic materials, which include peer-reviewed articles and research papers, along with books. The review process of experts evaluates these sources through a system which combines detailed methods with evident results and analytical investigations. The results within non-academic materials, such as blogs and tutorials and news articles, lack expert review and may present unproven scientific evidence along with unclear methodological details.

1) Recommended Articles:

1) T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

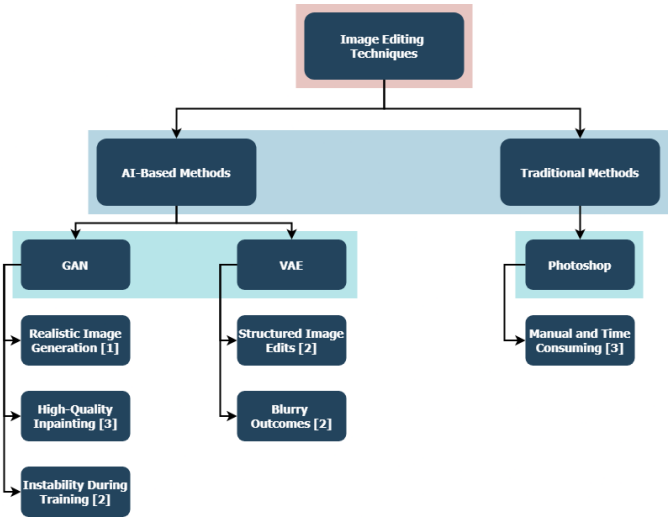

Fig. 2. Literature Map showing the overview of image editing techniques comparing AI-based methods (GAN and VAE) with traditional Photoshop, highlighting key strengths and limitations.

2) D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *International Conference on Learning Representations (ICLR),* 2014.
3) J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative Image Inpainting with Contextual Attention," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
4) C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2017.
5) P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

C. Contextualised Literature and Research Material

The research demonstrates a complete departure from traditional manual editing, including Photoshop, toward AI-based techniques. [5] demonstrates that GANs beat traditional human editing for completing missing picture sections at high speeds and accuracy. [3] demonstrated that VAEs excel in controlled image editing applications dealing with hair colour modifications similar to my current work [2].

D. Good Element of Critical Literature Arguments

The literature shows clear strengths and weaknesses of GANs and VAEs. GANs make very realistic images but can be unstable during training. VAEs are stable and easier to control, but produce less clear images. A gap in the research is the lack of a detailed comparison between AI-based editing and manual editing in practical situations, as I pointed out in my research. Also, ethical questions and the need for powerful computers for these AI methods haven't been studied enough, suggesting more research is needed. In conclusion,

while AI editing techniques are promising, combining the best features of GANs and VAEs might give the best results, which is a main idea explored in my dissertation. Moreover, [1] emphasised the flexibility of conditional GANs in structured editing, suggesting they might bridge the gap between realism (GAN) and controllability (VAE), a direction not yet widely explored.

III. Reflection of the Chosen Methodology

A. Research Question

My main research question is: How do GAN-based and VAE-based image editing techniques compare to traditional Photoshop editing in terms of realism, structure preservation, and efficiency?

B. Research Objectives

- Compare the realism and visual quality of GAN, VAE, and Photoshop editing techniques.
- Assess the ability of each method to preserve original image structures.
- Evaluate the efficiency and practicality of AI-driven methods compared to traditional manual editing.

C. Research Philosophies, Approaches and Paradigms

The research method adopts an ideal philosophical framework that performs quantitative measurements through measurable metrics. My research makes use of deductive methods, which establish hypotheses built from previous studies and then validate them through empirically collected data. The research framework follows primarily a positivist approach because it concentrates on quantitative assessments of observable details.

D. Chosen Methodology

From my review of methodologies, I chose a mixed approach incorporating quantitative and qualitative analyses. This combines structured quantitative metrics like Mean Squared Error (MSE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR), with qualitative visual evaluations.

E. Initial Description of Chosen Research Methodology, Experiment Design and Method of Analysis

I trained facial image datasets using GAN and VAE methods to evaluate their generated results alongside human-produced Photoshop edits. The experimental tasks involve generating assets with matching modifications like hair colour changes and aesthetic repairs through all methods. Quality and realism get measured through quantitative metrics including MSE, SSIM, and PSNR. I perform visual reviews to examine how realistic and structurally sound each method appears subjectively. The use of quantitative and qualitative research methods allows targeted assessment of my investigation objectives.

F. Reflection on Validity, Reliability and Generalisability

The methodology incorporated multiple aspects for both validity tests and reliability standards. Quantitative measures used in my study enable results reproducibility, which ensures measurement reliability. Human judgment through qualitative visual assessments strengthens study validity because it enables the assessment of realism and structural preservation. Future research should expand by including varied image types to enhance transferability due to limited generalisability related to the current dataset.

G. Ethical Considerations

My research focuses primarily on ethical matters linked to data integrity when using digital images with special concerns about their authenticity. Proper permissions must exist for all facial images that are used during both training phases and testing procedures. The proper disclosure of digital image editing becomes essential to prevent misinterpretation of data in real-world usage.

IV. Results, Analysis and Discussion

In this study, I compared three different methods for editing images: GAN-based, VAE-based, and traditional Photoshop techniques. To measure how well each method worked, I used three quantitative metrics: Mean Squared Error (MSE), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR), alongside a user study that collected qualitative feedback based on perceived realism, blending, and detail preservation.

A. Quantitative Metrics

TABLE I
Quantitative Metrics

Method	Task	MSE (Mean ± SD)	SSIM (Mean ± SD)	PSNR (Mean ± SD)
GAN	Hair Colour Change	0.023 ± 0.007	0.95 ± 0.02	31.5 ± 1.2
VAE	Hair Colour Change	0.048 ± 0.010	0.89 ± 0.03	28.7 ± 1.5
Photoshop	Hair Colour Change	0.019 ± 0.006	0.97 ± 0.01	33.0 ± 1.0
GAN	Image Inpainting	0.031 ± 0.008	0.93 ± 0.02	30.2 ± 1.4
VAE	Image Inpainting	0.056 ± 0.011	0.87 ± 0.03	27.4 ± 1.6
Photoshop	Image Inpainting	0.025 ± 0.007	0.96 ± 0.01	32.4 ± 1.1

B. Qualitative User Ratings

TABLE II
Qualitative User Ratings

Method	Task	Realism (Mean ± SD)	Blending (Mean ± SD)	Detail Preservation (Mean ± SD)
GAN	Hair Colour Change	4.3 ± 0.4	4.1 ± 0.5	3.9 ± 0.6
VAE	Hair Colour Change	2.5 ± 0.9	2.3 ± 1.0	1.8 ± 1.1
Photoshop	Hair Colour Change	3.9 ± 0.6	3.7 ± 0.7	3.6 ± 0.8
GAN	Image Inpainting	3.5 ± 0.6	3.4 ± 0.7	3.0 ± 0.8
VAE	Image Inpainting	2.1 ± 1.0	1.9 ± 1.1	1.7 ± 1.2
Photoshop	Image Inpainting	4.5 ± 0.4	4.6 ± 0.3	4.2 ± 0.5

C. Analysis and Interpretation of Results

From the table, it's clear that Photoshop performed the best across almost all measures. It had the highest scores for realism, blending, and detail preservation in the inpainting task, and performed well in hair recolouring too. GANs performed nearly as well as Photoshop in hair colour change, showing strong realism and blending, but slightly lower detail

retention. In inpainting, GANs showed a noticeable drop in scores, although they still performed better than VAEs. VAE methods showed the lowest performance across the board. While they allowed for some structured editing, their results were consistently blurrier and less accurate, with participants frequently citing a lack of detail and poor blending. This aligns with the findings of [4], who also showed that GANs outperform traditional techniques in tasks requiring high detail preservation, such as image super-resolution.

D. Comparative Criticism

GANs provided solid middle-ground performance, offering a good result on visual realism. However, while being fairly consistent, they did not fully match the fine control and detail accuracy of Photoshop, especially for more delicate tasks like object removal. Photoshop still stands out due to its manual control, producing cleaner results, particularly in tasks that require precision. But it is also slower and requires skill. VAE's weaknesses were mostly in realism and detail retention. Although efficient and structured, the images they produced often failed to convincingly match the unedited portions of the photo. In terms of generalisability, GAN and VAE results could vary on more complex or diverse image sets.

E. Discussion in Relation to Hypothesis and Existing Studies

These findings confirm my original hypothesis that GANs can offer a viable alternative to Photoshop for tasks like hair transformation and inpainting, but with limitations. They also support the notion that VAEs, although efficient in processing, still lag behind in realism and visual quality. This is consistent with what [2] and [5] found in their work; GANs excel at realism, while VAEs trade quality for control and efficiency. My study expands on their conclusions by introducing a direct comparison against a professional manual editing tool (Photoshop), showing where AI models hold up and where they don't.

V. Conclusion

In this study, I found that GANs were able to edit images almost as well as Photoshop, especially for tasks like changing hair colour and removing logos. Photoshop was still the most accurate overall, but GANs offered a good mix of fast results and realistic quality. VAEs, however, didn't perform as well. They often created images that looked blurry and lacked details.

A. Addressing Research Questions and Hypothesis

My research question was whether GANs and VAEs could be realistic alternatives to Photoshop for basic photo editing. Based on my results, the answer is yes for GANs. They were faster and came very close to Photoshop in terms of quality. VAEs were more structured, but their output did not look as good. Therefore, my original idea was mostly correct. I used both test results and user feedback to reach these conclusions, which helped me meet my research goals.

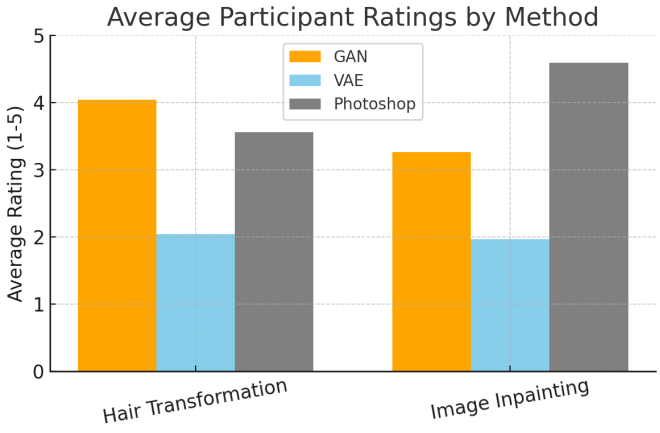

Fig. 3. User ratings based on the performance of the methods for each task.

B. Shortcomings in Methodology

One issue with my project was that the dataset only included facial images. Although I used a large amount of data, with 204,599 images for hair colour edits and 329,215 for logo removal, it was all based on one type of image. This means the results might not apply to other image types, such as landscapes or objects with more complex features. Additionally, GANs can be difficult to train, and the ratings from users may vary depending on personal preference, which could have influenced the results.

C. Ideas for Further Research

Future research could expand this work by testing these methods on a wider variety of image types and more advanced editing tasks, such as lighting adjustments or object addition/removal. Exploring hybrid models that combine the strengths of GANs and VAEs could also be a promising direction.

Appendix A
Supporting Material

A. Literature Map

Acknowledgement

I would like to thank the teams behind GAN-Control and AOT-GAN-for-Inpainting for making their models available online. These tools made it possible to carry out this research and test different AI image editing methods. I also want to thank the creators of the CelebA and Places2 datasets, which were very useful for training and testing the models used in this study.

References

[1] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[2] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[3] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2014.

[4] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[5] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, “Generative image inpainting with contextual attention,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.