Research Design II

1 SECTION

1.1 TITLE AND AUTHOR INFORMATION

Enhancing Realism in Photo Editing: A Comparative Study of GAN-based Techniques versus Traditional Methods

Author: Kledi Azzopardi

Supervisor: Daniele Gravina

Date: October 2024

1.2 CHOSEN RESEARCH

1.2.1 Description of Theme and Topic Rationale

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 regards to computational requirements, ethical issues, and simplicity in dealing with realism.

1.2.2 Positioning and Research Onion

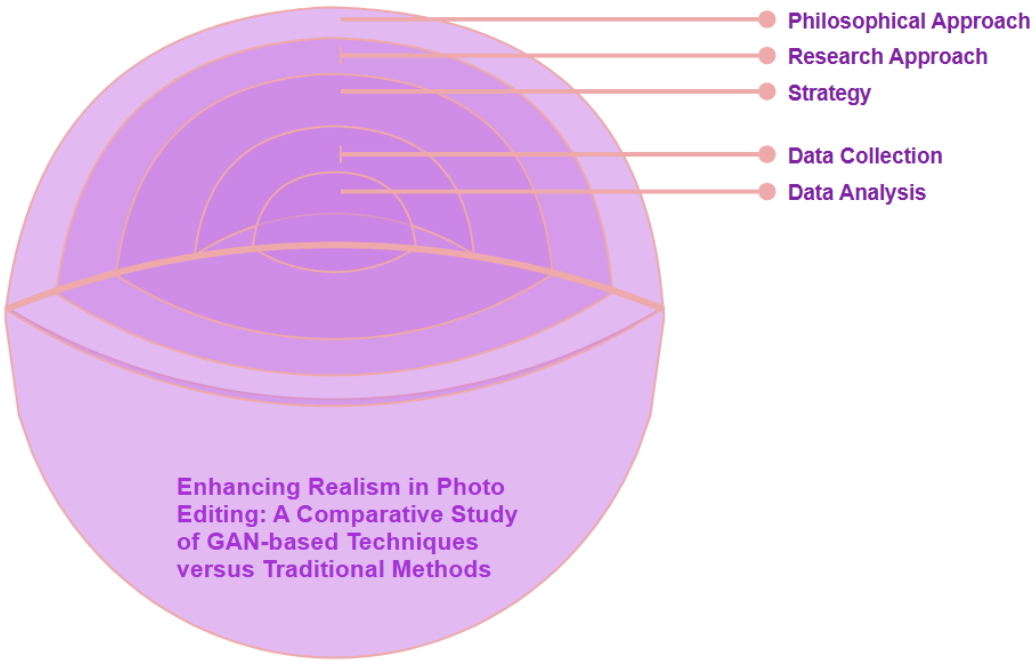
1. **Philosophical Approach:** The ideal approach is adopted in this research, whereas the objective is to explore and compare quantitative findings based on AI-generated images.

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.

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.



1.2.3 Background to This Research Theme

Image editing is a critical part of the creative industry, digital marketing, and forensic analysis. AI has revolutionised the field with GANs and VAEs to facilitate high-quality automatic transformations and hence super-resolution, inpainting, and style transfer. Photoshop is a premier software tool, but is time-consuming and skill-dependent, and therefore AI-based automation can be a game-changer. GANs work with two parts: a discriminator and a generator. The generator repeatedly alters the image to make it look more realistic. VAEs work differently, however. They alter image features in a clear way using a concept called latent space. Both techniques have their advantages:

i. GANs have the capability of generating high-resolution images that look aesthetically pleasing but are unstable.

ii. VAEs provide interpretable and structured image transformation but can yield blurry outcomes. The study evaluates AI-editing versus Photoshop based on accuracy, ease, and overall realism.

1.2.4 Hypothesis

The research is built on the following hypothesis:

1. Image editing using GAN is more realistic than with standard Photoshop techniques.
2. VAE-based methods give controlled image alterations in a well-structured form but are less sharp compared to GANs.
3. Both GANs and VAEs can be merged to get a trade-off between realism and organised latent space editing.
4. AI-driven photo editing is significantly quicker than manual Photoshop techniques regarding usability and time.

Previous research, including Karras et al. (2019) on StyleGAN, has demonstrated that AI models can achieve realistic editing with identity preservation. Previous research does not provide a comparison of AI-based editing and standard manual editing techniques, which this study seeks to address.

1.2.5 Research Aim and Purpose Statement

The aim of this research is to:

1. **Compare** how well GAN-based and VAE-based image editing techniques fare compared to normal Photoshop editing.
2. **Analyse** the realism, structure preservation, and efficiency of the AI-driven approaches with respect to their weaknesses and benefits.
3. **Provide** insight into the viability and utility of AI-driven image editing for common use, including digital marketing, fashion, or forensic imaging.

2 SECTION

2.1 REVIEW OF RESEARCH METHODOLOGY

2.1.1 Undertake a short literature review about methodologies used in other studies

Multiple research investigations explore the utilization 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 photorealistic images. The work of Karras et al. represents an excellent example of StyleGAN [1]. 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 organized results although they introduce image blur [2].

2.1.2 Distinguish between academic and non-academic material

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.

2.1.3 Recommended 5 articles from peer reviewed journals

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.

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.

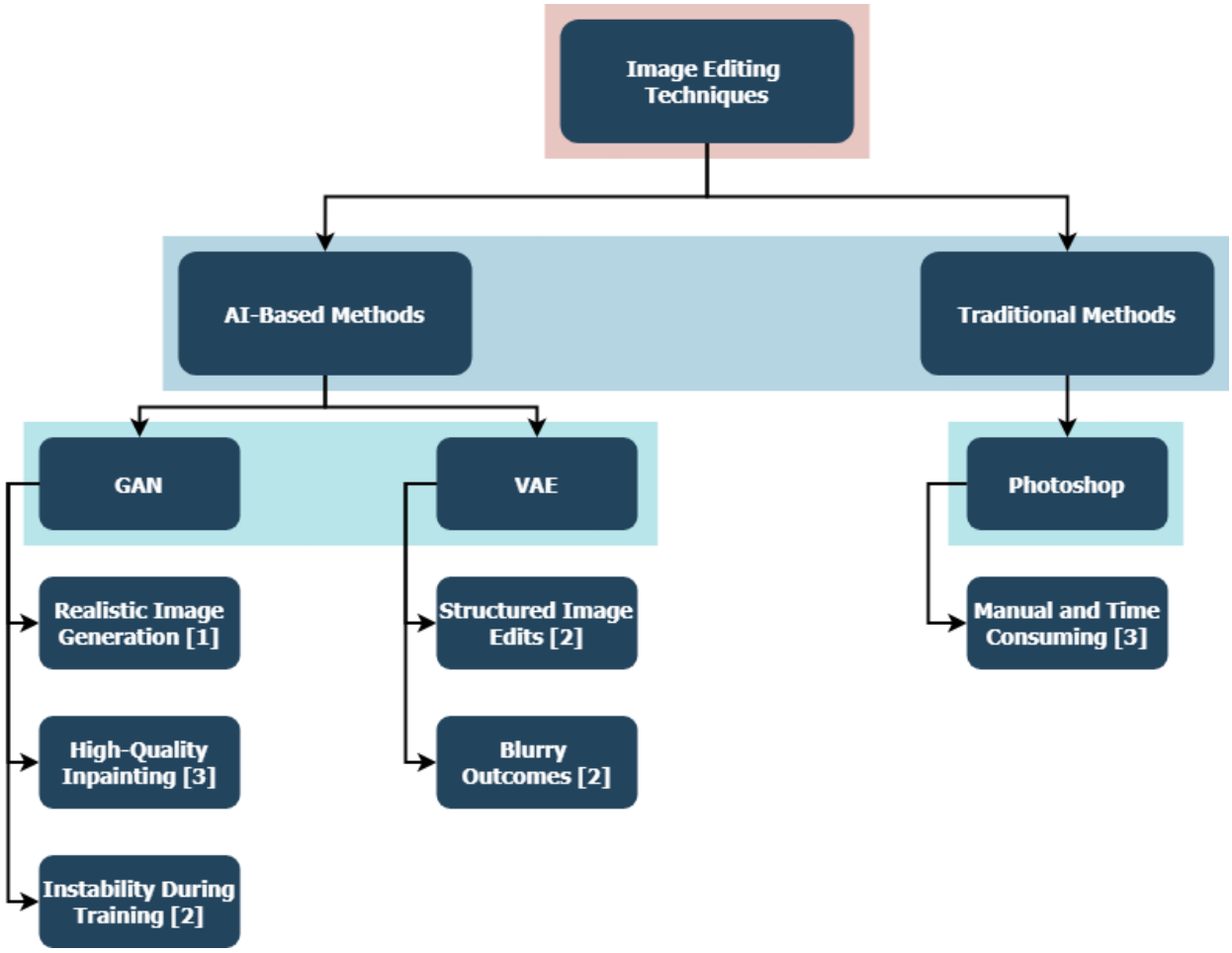
2.1.4 Contextualised literature and research material

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

2.1.5 Add a good element of critical literature arguments (compare, contrast and identify knowledge gaps)

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 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.

2.2 Literature Map



3 Section

3.1 Reflection on the Chosen Methodology

3.1.1 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?

3.1.2 Research Objectives

To address my research question, I have set the following 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.

3.1.3 Understanding 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 then validate them through empirically collected data. The research framework follows primarily a positivist approach because it concentrates on quantitative assessments of observable details.

3.1.4 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.

3.1.5 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 color 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.

3.1.6 Reflections on Validity, Reliability, Generalisability/Transferability

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.

3.1.7 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.

4 SECTION

4.1 RESULTS, ANALYSIS & DISCUSSION

4.1.1 Presentation of Results

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.

4.1.2 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

4.1.3 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

These ratings were gathered from a group of 9 participants who evaluated the output based on how realistic the images looked, how well the edits blended into the image, and how much original detail was preserved.

4.1.4 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 lack of detail and poor blending.

4.1.5 Comparative Criticisim

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.

4.1.6 Discussion in Relation to Hypothese 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 Karras et al. [1] and Yu et al. [3] 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.

5 SECTION

5.1 CONCLUSION

5.1.1 Main Conclusions

In this study, I found that GANs were able to edit images almost as well as Photoshop, especially for tasks like changing hair color 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.

5.1.2 Addressing Research Questions and Hypotheses

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.

5.1.3 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.

5.1.4 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.

References

[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.

[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.