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

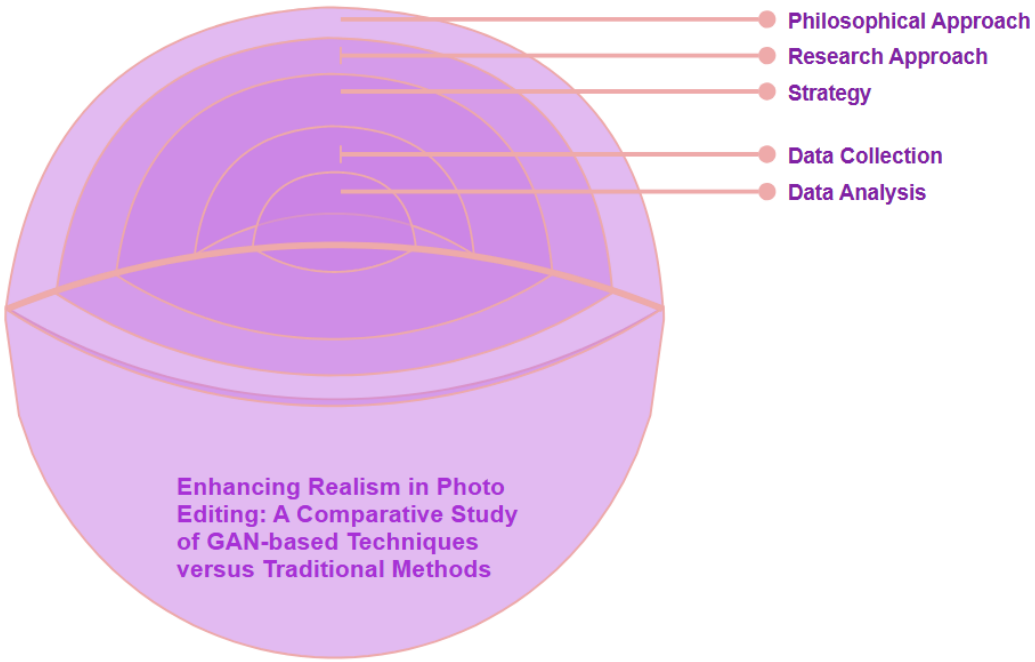
1. **Philosophical Approach:** The ideal approach is adopted in this research, whereas the objective is to explore and compare quantitative findings based on AI-generated images.

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.

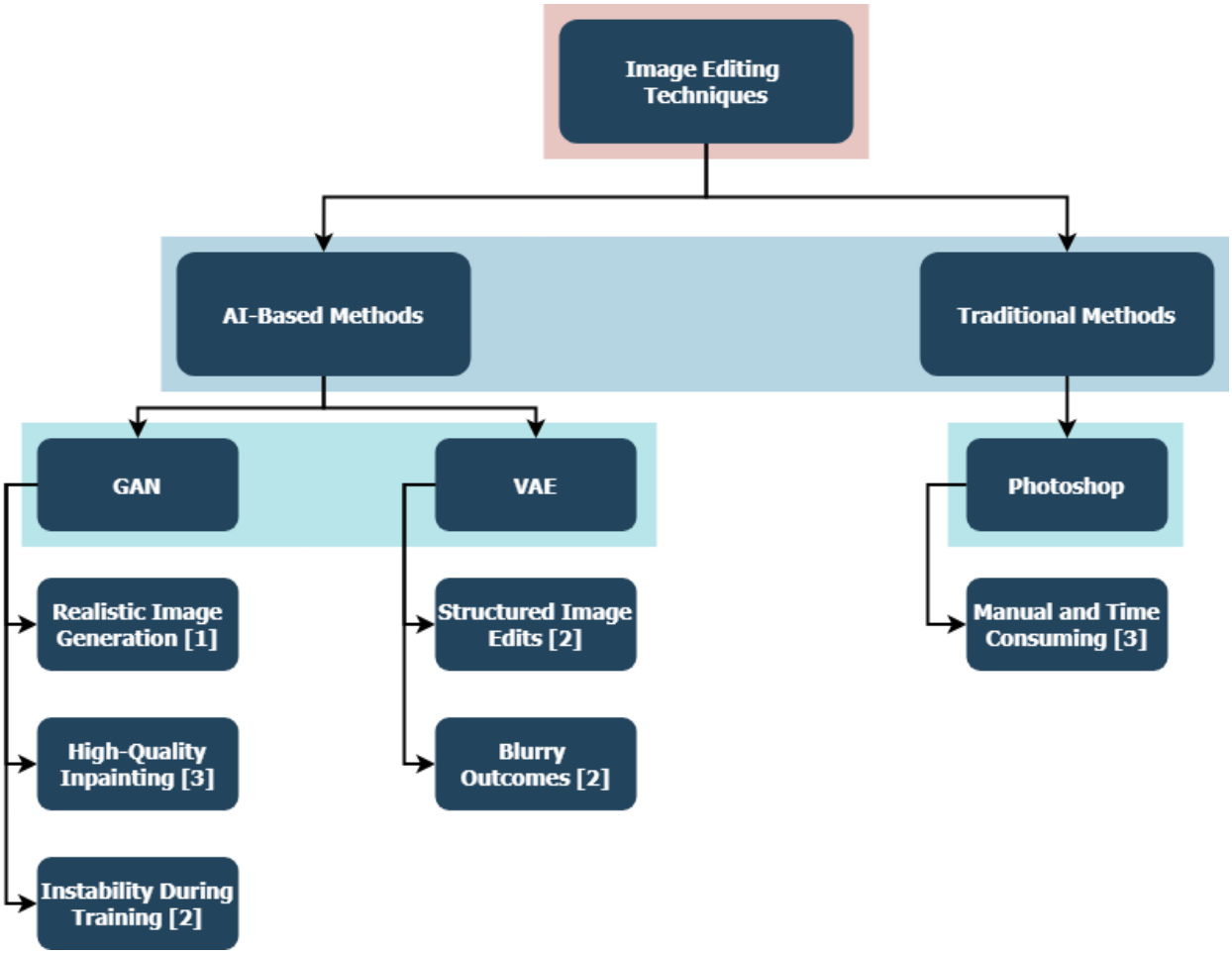
2.1.4 Contextualised literature and research material

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



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.

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 generalizability 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.