

Prague School of Creative
Communication
Visual Arts

Bachelor's Thesis

AI-assisted Photographic Production and
Post-Production

2025

Ivan Riapov



Prague School of Creative
Communication
Visual Arts
Photography and Audiovisual
Arts

**AI-assisted Photographic Production and
Post-Production**

Theoretical Part: Theoretical Foundations of Photography, CGI,
and Artificial Intelligence

Practical Part: Photographic Production and Post-Production
Workflow Optimization Using AI Tools

Author: Ivan Riapov

Supervisor: Jan Macuch

2025

Declaration

I declare that I have elaborated my bachelor's thesis independently and that I have listed all the sources and literature from which I have drawn. I confirm that all submitted copies of my bachelor's thesis correspond to the electronic version in the UCC information system and I agree that the thesis will be made available to the public for study and research purposes.

In Prague on

Author's Signature:

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Jan Macuch, for his incredible support, effort, constructive feedback, and the time he devoted to guiding me throughout this journey. Since the very first semester of my studies, he has consistently helped me unlock my potential, providing mentorship, encouragement, and practical advice.

His guidance introduced me to fascinating intersections between visual art and latest technologies. The inspiration he sparked in me to explore the development of artificial intelligence in the creative field became the foundation for this thesis.

I am truly thankful for his continued support and invaluable help.

Abstract

This bachelor's thesis explores the integration of Artificial Intelligence (AI) tools in photography, with a specific focus on interior and architecture fields. It examines how technologies such as AI-powered generative models and automated editing tools can streamline photographic workflows, enhance image quality, and reduce time spent on post-production. The theoretical part presents the historical and technological evolution of photography, CGI, and AI, while the practical part demonstrates a custom multi-tool workflow developed and tested in real-world scenarios. Emphasis is placed on HDR merging, object replacement, noise reduction, and style transfer using tools like Luminar Neo, Adobe Photoshop powered by Adobe Firefly, Stable Diffusion, and FLUX AI. The thesis concludes that AI, as a part of CGI (Computer-Generated Imagery), is not a replacement for traditional photographic techniques but a powerful collaborative tool that increases both efficiency and creative potential.

Keywords: photography, artificial intelligence, workflow optimization, post-production

Table of Contents

Theoretical Part	1
1 Introduction and Key Theory	1
1.1 Foundational understanding of photography	3
1.2 The Development of Early Photographic Processes	4
1.3 Industrial and Technological Advancements	7
1.4 The Digital Revolution	7
1.5 Foundational understanding of CGI and AI	8
1.6 A Brief History of CGI	9
1.7 CGI, AI, and Photography	11
1.8 Machine learning in CGI	13
1.9 Contributions of OpenAI to AI-Assisted Photography and Post-Production	16
1.10 Stable Diffusion and Advances in Generative AI	17
Practical Part	26
2 Methodology	26
2.1 Tools Used in the Practical Work	28
Luminar Neo	28
Adobe Lightroom	28
Adobe Photoshop with Adobe Firefly Generative AI	29
Stable Diffusion with LoRa Models	29
3 Editing Process and Use Cases	30
3.1 Airbnb Workflow	30
3.2 Lightroom vs Luminar Neo (HDR Merge)	32
3.3 Comparison of Luminar Neo and Lightroom HDR Merging	40
Luminar Neo Advantages	40
Lightroom Limitations	40
3.4 AI-Assisted Cleanup and General Corrections	42
4 Generative AI Tools	46
4.1 Brief Overview of the Photoshop Generative AI	46
4.2 Demonstration of Basic Capabilities	46
4.3 Brief Overview of Luminar NEO AI-Powered Tools	49
4.4 FLUX AI by Black Forest Labs	52
4.5 Stable Diffusion and Its Application in My Workflow	61
Why LoRA?	61
LoRA Training: From Idea to Results	61
Initial Attempts and Lessons Learned	61
5 Step-by-Step LoRA Training Process	62
5.1 Tagging and Configuration for LoRA Training	65
5.2 Training Configuration and Key Parameters	65

5.3 Training Insights	67
6 AI Upscaling with Freepik Magnific	70
7 Practical Conclusion: Efficiency Gains through AI Tools in Photography Workflow	72
8 Final Statement	75
9 Personal Applications of AI-Assisted Post-Production in Interior Photography	77
Glossary of Terminology	83
References	85
List of Figures	88
List of Tables	91

Theoretical Part

1 Introduction and Key Theory

As a photographer, I always look for the most efficient ways to optimize and automate my workflow toward time, energy, and resource economy. The client doesn't usually see the real amount of post-production behind the image and believes that the photographer's work is nothing else than the actual photoshoot and extremely fast software editing.

Most people think that a photographer's job is not that significant nowadays because of artificial intelligence and CGI, which "does everything" and takes zero effort. But the truth is that qualitative image processing is also a part of a professional photographer's job and currently, it consumes way more time and energy than actual photoshoots. It is still demanding and requires a decent amount of manual work. Even the most efficient and powerful AI tools for graphic work still have to be "guided" and programmed to perform better and complete tasks according to expectations. It would be fair to mention that besides the practical part of a photography project, in particular, photoshoot and post-production there is a huge pre-production step that includes an organization that sets the whole direction of the project.

The thesis aims to find an optimal workflow combining traditional photography techniques with generative AI tools and to define what constitutes an optimal workflow and the criteria for its evaluation. I would like to compare a certain number of AI tools, their "main focus" and how they work with different types of photography, like portraits, products, and interior photography in particular. Also, I challenge myself to automate the post-production process as much as possible. Recently, I decided to do research where I measured the time I spent editing one single photograph step by step. In this case, I have chosen an interior photo because it's related to my work - interior and architecture photography services for house management companies, real estate agents, and private clients, mostly for platforms such as Booking.com, Airbnb, websites and social media.

I compared the results of my research with statistics provided by Perplexity AI (date of AI research 25.10.24). The time I spend editing one photograph matches the time the average interior photographer spends on the same type of editing. Surprisingly, it was accurate to

the minute. The results are presented in *Table 1*. According to this information, my workflow meets the average interior photography editing. If we take a look at the time the average photographer spends on editing, we might be frustrated because it ten times exceeds the actual photoshoot. For example, 40 hours of pure editing against only four hours of shooting.

	Action	Approximate time
1	Selection	2 minutes
2	Blending images into HDR and lens distortion correction	26 seconds
3	Basic camera RAW adjustments, perspective correction	2 minutes
4	Clean-up (unwanted objects, sensor dust, cables, etc) • Remove Tool • Generative Fill	7 minutes
5	Lighting balance setting by using luminance masks	1 min 10 sec
6	Second step of removal (reflections, glares, more complicated objects)	4 min 30 sec
7	Surface Blur to remove stains and for smoother surfaces	5 minutes
8	Final Touch	2 min 40 sec
	Total	~ 22 min 42 sec

Table 1: - Approximate time of editing one interior photograph using the manual approach with minimal AI assistance.

1.1 Foundational understanding of photography

Photography is a form of visual art, science, and practice of creating images by recording light using either an image sensor or light-sensitive materials such as photographic paper or film. The term “photography” comes from the Greek words *phos* (light) and *graphê* (drawing), meaning “drawing with light”. Photography has approximately 200 years of history, but the optical effect behind photography, called *camera obscura*¹, was known much earlier. In the 4th-5th centuries BC, ancient Chinese philosophers, and Greek scholars were already observing how light passes through a small and narrow hole and projects an upside-down image on the surface opposite that hole. This ancient discovery laid the groundwork for photography in the early 19th century when Joseph Nicéphore Niepce captured the first permanent photograph, “View from the Window at Le Gras”. It was taken in 1826 or 1827, and these years might be considered the birth of photography.

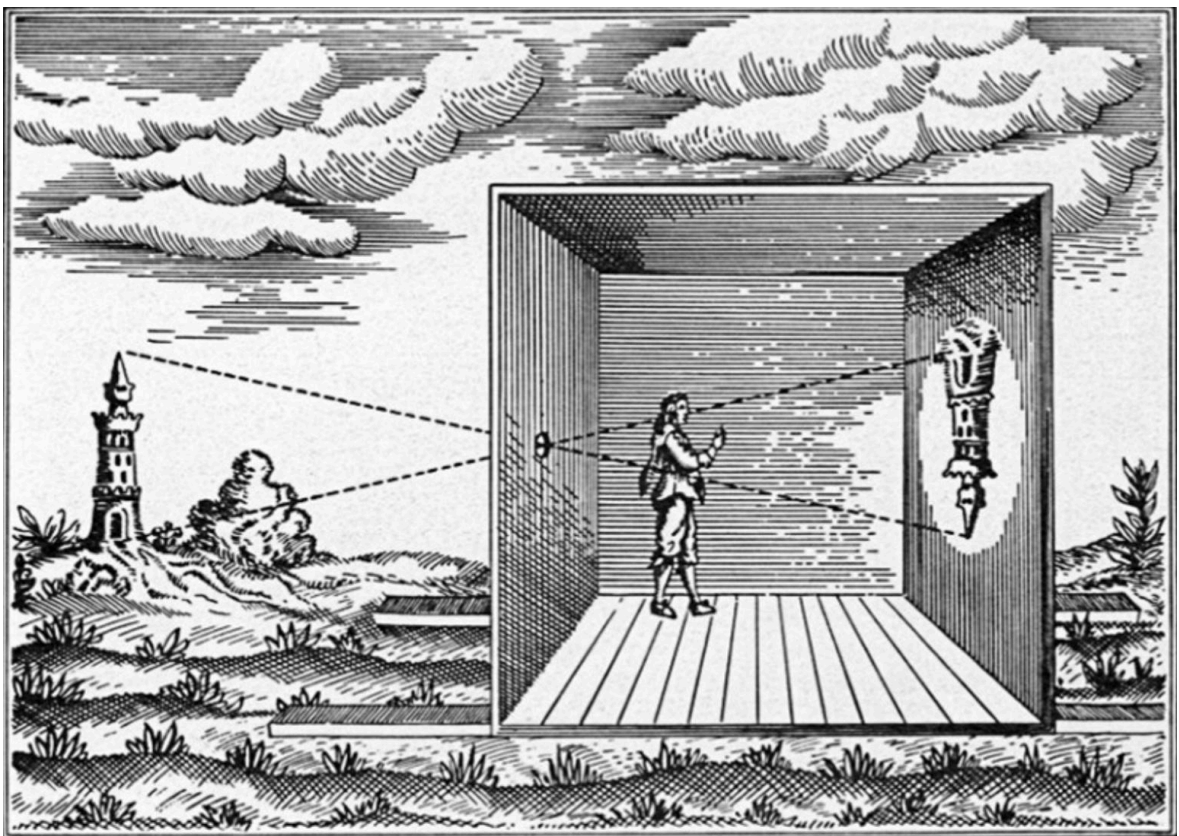


Fig. 1: Camera Obscura

¹Lumas. (n.d.). *History of photography*. Retrieved February 2, 2025, from <https://www.lumas.com/history-photography/>

1.2 The Development of Early Photographic Processes

Following Niépce's breakthrough, significant advancements were made that shaped the evolution of photography as both a science and art:

Daguerreotype Process (1839)

Louis Daguerre refined Niépce's work and started creating highly detailed images on copper plates with a thin layer of silver. This process was named the Daguerreotype², and it was the first publicly announced and commercially efficient photographic method. But it required a long exposure time initially, and gradual improvements reduced it to mere seconds, making photography available to a broader audience worldwide.

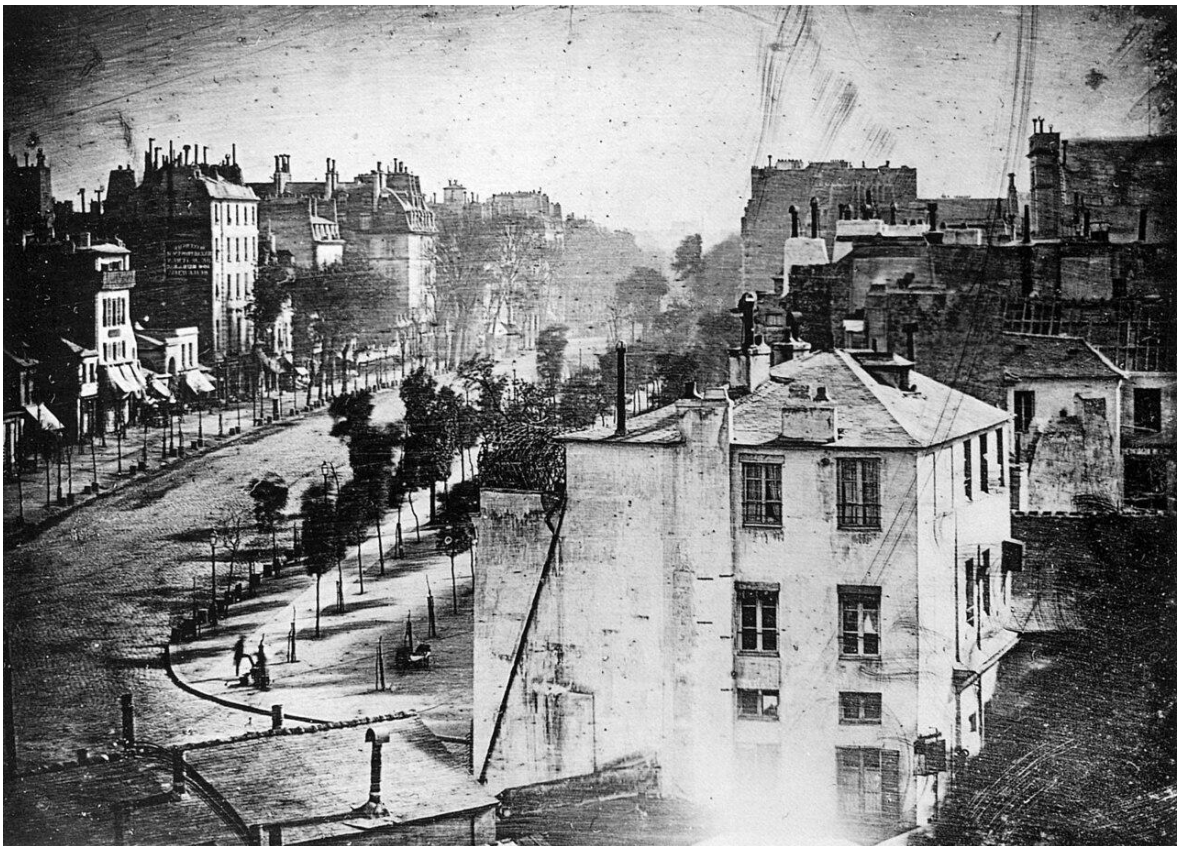


Fig. 2: Boulevard du Temple, daguerreotype by Louis Daguerre, 1838.

² The daguerreotype medium. *Library of Congress*. Retrieved from <https://www.loc.gov/collections/daguerreotypes/articles-and-essays/the-daguerreotype-medium#:~:text=The%20daguerreotype%20is%20a%20direct,surface%20looked%20like%20a%20mirror> [Accessed 25.06.2025]

Direct Positive Process (1839)

Hippolyte Bayard developed the direct positive process, which created unique photographic prints directly on paper without a negative. His method involved exposing chloride silver chloride paper to light, producing atmospheric images valued for their artistic quality. Bayard also staged conceptual photographs, including his notable *Self-Portrait as a Drowned Man* (1840)³, showcasing photography's potential for storytelling.



Fig. 3. *Self-Portrait as a Drowned Man*, photograph by Hippolyte Bayard, 1840.

Calotype Process (1841)

William Henry Fox Talbot introduced the calotype process, which utilized paper negatives to produce multiple positive prints⁴. This innovation laid the groundwork for modern photography by making image reproduction possible, unlike the daguerreotypes that existed only in single copies.

³ Open Culture. (2019, October 22). *The First Faked Photograph (1840)*. Retrieved February 3, 2025, from <https://www.openculture.com/2019/10/the-first-faked-photograph-1840.html>

⁴ Britannica, *Calotype*, Accessed March 18, 2025, <https://www.britannica.com/technology/calotype>.

Scientific Contributions by John Herschel (1839-1842)

John Herschel made significant contributions by discovering that sodium thiosulfate could be used as a fixer to stabilize photographs and by inventing the cyanotype process, which produced blue-toned images that later became famous for copying documents and architectural blueprints. It involves coating paper or fabric with a light-sensitive emulsion made from ferric ammonium citrate and potassium ferricyanide. When exposed to UV light, chemicals react, creating a pigment called Prussian Blue. The unexposed areas are then washed away with water, leaving behind a permanent cyan-colored image.

British botanist Anna Atkins used this process to create photograms (a technique where objects are placed directly on the coated paper before exposure) and was credited with creating the first photographic book using cyanotype prints of plants⁵ in 1840. John Herschel also coined essential terms such as “photography”, “negative,” and “positive”.

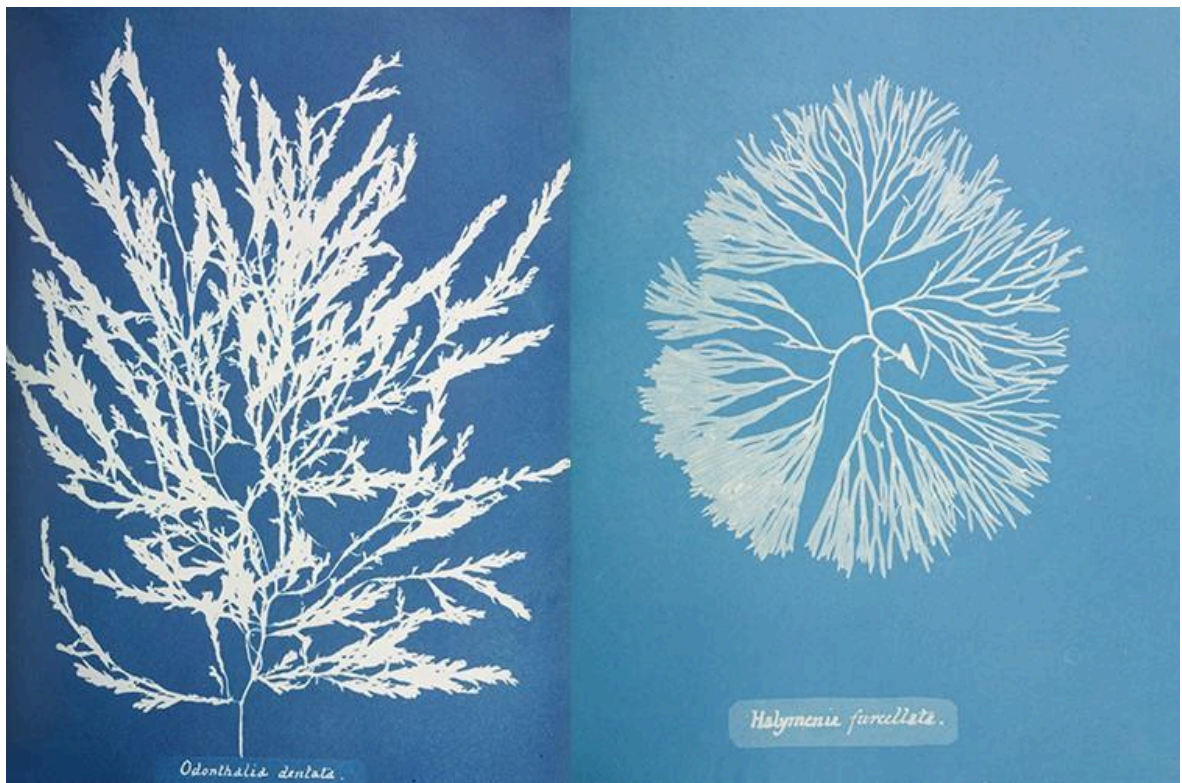


Fig. 4. Cyanotype photograph of British algae from Anna Atkins's book *Photographs of British Algae: Cyanotype Impressions* (1843).

⁵ Natural History Museum. (n.d.). *Anna Atkins's cyanotypes: the first book of photographs*. Retrieved March 18, 2025, from <https://www.nhm.ac.uk/discover/anna-atkins-cyanotypes-the-first-book-of-photographs.html>

Collodion Wet Plate (1851)

Frederick Scott Archer developed the Collodion Process⁶, which combined the sharpness of daguerreotypes with the reproducibility of calotypes. It was widely used for decades due to its ability to produce high-quality images on glass plates.

1.3 Industrial and Technological Advancements

The late 19th century saw rapid industrialization and technological progress that transformed photography⁷ into a practical tool for amateurs and professionals.

Introduction of Roll Film (1888)

George Eastman revolutionized photography with roll film and portable cameras like Kodak's box camera. This introduction basically eliminated inconvenient glass plates and made photography accessible to everyday users.

Color Photography (1936)

Kodachrome film introduced high-quality color photography, enabling photographers to capture vibrant images and expand creative possibilities.

Instant Photography (1948)

Polaroid cameras allowed users to develop photos instantly, marking another milestone in convenience and accessibility.

1.4 The Digital Revolution

The transition to digital photography in the late 20th century marked a crucial moment in its history.

Digital Sensors

In 1986, Kodak scientists developed the world's first megapixel sensor⁸ capable of producing high-quality images. This innovation started to replace film as a primary medium for capturing photographs.

⁶ National Media Museum. (n.d.). *Wet collodion negative (1851–1885)*. Retrieved March 18, 2025, from <https://www.nms.ac.uk/discover-catalogue/inventors-of-photography-exploring-victorian-photographic-techniques>

⁷ Cooke, A. (2025, April 25). *The Birth of the Digital Camera: From Film to Filmless Revolution*. Fstoppers. Retrieved March 18, 2025, from <https://fstoppers.com/historical/birth-digital-camera-film-filmless-revolution-699530>

⁸ Digital Camera Museum. (n.d.). *Videk Megaplust (1986)*. Retrieved March 19, 2025, from <https://www.digitalkameramuseum.de/en/cameras/item/videk-megaplust>

Consumer Digital Cameras

The introduction of consumer-friendly digital cameras like Kodak DC40⁹ in 1995 made digital photography mainstream. By 2000, mobile phones with inbuilt cameras democratized photography even more, making it universal in daily life.

Post-Processing Tools

Software like Adobe Photoshop (released in 1990) revolutionized post-production workflows by allowing photographers to manipulate images digitally with precision.

1.5 Foundational understanding of CGI and AI

Computer-Generated Imagery¹⁰ (CGI) refers to the visual content creation using computer software. It is a subfield of computer graphics that focuses on creating realistic or stylized images, animation, or visual effects for use in film, photography, television, video games, advertising, and other industries. CGI covers both static images and dynamic sequences, making it perfect for modern visual storytelling and design. By simulating real-world elements (or fictional objects with a high degree of realism), such as lighting, textures, particles, physics, and objects, CGI creates visuals that would otherwise be impossible or impractical to achieve using traditional methods.

CGI is often associated with 3D imaging, where objects are modeled in three dimensions and rendered into photorealistic and artistic visuals. It also includes 2D graphics for applications like digital painting and compositing. The methodology behind CGI draws heavily from disciplines such as geometry, optics, physics, and computer science.

⁹ Digital Camera Museum. (n.d.). *Kodak DC40 (1995)*. Retrieved March 24, 2025, from <https://www.digitalkameramuseum.de/en/cameras/item/kodak-dc40>

¹⁰ TechTarget. (n.d.). *What is CGI (Computer-Generated Imagery)?* Retrieved March 26, 2025, from <https://www.techtarget.com/whatis/definition/CGI-computer-generated-imagery>

1.6 A Brief History of CGI

Early Beginnings (1950s-1970s)

The origins of CGI¹¹ can be traced back to the 1950s when it was primarily used as a visualization tool for scientists and engineers. Early development took place in research centers like Boeing and Bell Labs, where computers were used to create simple graphical representations for technical purposes. Universities such as MIT and the University of Utah became main hubs for computer graphics research in the 1960s and 1970s, leading to breakthroughs such as hidden surface determination (essential for rendering 3D objects) and shading models.

Key Milestones:

- **1968:** Arthur Appel introduced ray casting, a precursor of ray tracing, which simulates light paths for realistic rendering.
- **1973:** The ACM SIGGRAPH conference was established as a leading forum for computer graphics research.
- **1970s:** Henri Gouraud and Bui Tuong Phong developed shading techniques that added depth and realism to computer-generated objects.

Rise of Cinematic CGI (1980s - 1990s)

The 1980s marked the appearance of CGI in cinematography¹², especially in mainstream cinema. Movies like *Tron* (1982), *The Last Starfighter* (1984), and *Young Sherlock Holmes* (1985), showcased an early use of CGI for creating wireframe models, digital environments, and fully animated characters. The Pixar Studio developed RenderMan¹³ software in the late 1980s and introduced shaders - programs designed to simulate complex lighting effects, preparing the ground for photorealistic rendering.

By the 1990s advancements in hardware and software, allowed for more complex applications:

¹¹ The Ohio State University. (n.d.). *CG Historical Timeline – Computer Graphics and Computer Animation*. Retrieved March 26, 2025, from <https://ohiostate.pressbooks.pub/graphicshistory/back-matter/cg-historical-timeline/>

¹² Pixflow. (2025, March 10). *How Toy Story Changed Animation History: Pixar's First CGI Animated Movie*. Retrieved March 26, 2025, from <https://pixflow.net/blog/how-toy-story-changed-animation-history-pixars-first-cgi-animated-movie/>

¹³ Levy, S. (1995, December 1). *The Movie That Changed Everything: Toy Story*. *Wired*. Retrieved March 26, 2025, from <https://www.wired.com/1995/12/toy-story/>

- Jurassic Park (1993) set a new standard for integrating CGI into live-action footage
- Toy Story (1995) became the first fully computer-animated feature film
- Global illumination techniques were introduced to simulate realistic lighting interactions between objects

Modern Era (2000s - Present)

In the 21st century, CGI has become widely adopted across various industries. Real-time rendering engines like Unreal Engine¹⁴ enable interactive experiences in video games and virtual reality. Films heavily rely on CGI to create expansive worlds, environments (Avatar, Interstellar), and objects. Architectural visualizations use CGI to produce lifelike renders of buildings before the construction begins.

AI Tools in Image Generation

Integrating artificial intelligence (AI) into image generation has revolutionized how visuals are created and manipulated. AI-powered tools leverage deep learning techniques to automate complex tasks and expand creative possibilities.

Key AI Technologies

1. Generative Adversarial Networks (GANs)¹⁵

GANs consist of two neural networks - a generator that creates synthetic images and a discriminator that evaluates their realism. Introduced in 2014, GANs have enabled the creation of high-quality images with realistic textures. They are widely used in applications, such as image synthesis, style transfer, and data augmentation.

2. Diffusion Models

Diffusion Models like Stable Diffusion generate images by gradually refining random noise into coherent visuals based on textual prompts, or reference image. These models are good at producing photorealistic results while offering precise control over styles and compositions.

3. AI-assisted Editing

¹⁴ Epic Games. (n.d.). *Unreal Engine and real-time rendering*. Retrieved March 27, 2025, from <https://www.unrealengine.com/en-US/real-time-rendering>

¹⁵ Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). *Generative adversarial nets*. Advances in neural information processing systems, 27. Retrieved March 28, 2025, from <https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>

AI tools such as Adobe Sensei automate tasks such as object removal, background replacement, and color-grading in post-production workflows.

1.7 CGI, AI, and Photography

CGI and Artificial Intelligence are reshaping photography¹⁶ by bringing full creative control, efficiency, and innovation. CGI integrates computer-generated elements into product, studio, architectural, and conceptual photography. Product photographers can use it to create realistic 3D models, refine lighting, or design realistic virtual scenes without physical limitations. In interior and architecture photography CGI helps visualize interiors and exteriors before construction or virtually stage existing interiors, simulating natural lighting and spatial depth. It also enables photographers to merge real objects with CGI-generated backgrounds.

AI tools enhance photography through automation, precision, creative potential, and generative capabilities. Automated editing optimizes control over exposure, color-grading, resizing, and noise-reduction. AI-powered retouching tools refine skin, and textures, remove blemishes, and upscale low-resolution images without compromising detail. Additionally, AI identifies faces, objects, and specific scenes in images, assisting in organization, and tagging for social media or client projects.

The integration of CGI and AI in Photography has transformed the medium into a dynamic mix of technical precision and creative freedom. CGI helps photographers visualize concepts beyond physical limitations, such as futuristic environments, interiors, and pre-production product designs, keeping them photorealistic. AI accelerates workflows by automating edits, increasing image quality, and saving tremendous amounts of time.

Together, these tools redefine photographic areas such as advertising, interior and architecture, and fine art. For instance, product photographers use CGI to create flawless, customizable backgrounds, while portrait artists leverage AI to refine lighting or skin tone and texture. As technology advances, the line between captured and generated will blur further, offering photographers unparalleled flexibility to merge reality with imagination. The future of photographic pre-production and post-production lies in using CGI and AI

¹⁶ Wonderful Machine. (2023). *Specialty: AI/CGI Photography*. Retrieved March 28, 2025, from <https://wonderfulmachine.com/article/specialty-ai-cgi-photography/>

not as a replacement for traditional techniques, but as collaborative tools that can elevate efficiency and creative expression.

According to the key theory, I'm providing the most significant areas of research:

- **Machine Learning and Computer Vision:** Research in this area focuses on developing algorithms that enable computers to interpret and understand visual data from the world. This is crucial for tasks like image recognition, object detection, and scene understanding in photography and CGI.
- **Image Processing and Enhancement:** This area involves developing techniques to improve the quality of images, such as noise reduction, color correction, and super-resolution. AI can automate and enhance these processes, making it easier for photographers and CGI artists to achieve high-quality results.
- **Automated Editing and Post-Production:** Research in this field focuses on using AI to automate various editing tasks, such as cropping, color grading, and adding special effects. This can significantly speed up the post-production process and allow photographers to focus on creative aspects.
- **Ethics and Legal Issues:** As AI becomes more integrated into photography and CGI, ethical and legal questions arise. Research in this area explores issues like copyright, authenticity, and the potential for manipulation, as well as how to address these concerns.

To understand the impact, we need to consider the underlying history of computer-generated imagery in photography, its path, potential vector, and the influence of artificial intelligence on visual art. First of all, CGI has revolutionized photography and helped artists create, manipulate, or modify images that would be difficult or impossible to capture using traditional methods. CGI includes many different techniques, such as 3D modeling, texturing, and rendering. Using these techniques is mostly based on creative purposes and client needs in various industries, but mostly in commercial photography. In recent years, the continuous development of artificial intelligence and its integration into visual art tremendously elevated and shifted CGI, allowing artists to spend less time working on post-production but still maintain quality, realism, and artistic balance, which is also important. Generative tools and AI-powered platforms, such as Open AI's Midjourney, DALL-E, Stable Diffusion, and Adobe Firefly, significantly automate complicated tasks, speeding up workflows and enhancing creative techniques. These tools are available to create, modify, and seamlessly edit already existing visual material and

generate high-quality backgrounds and images from simple text prompts. First of all, it would be fair to define the difference between program editing software used in post-production and AI-powered platforms and tools. The difference is that software is programmed to follow certain algorithms that may not apply to specific kinds of scenarios and may not fit certain purposes. Programs still need constant human supervision and a lot of manual control to perform the best results, unlike AI-based tools that can work independently and make decisions due to machine learning technologies that provide advanced mechanisms of pattern recognition, predictive analysis, and adaptive learning. It can analyze extensive amounts of data, identify trends and make adjustments in real-time. Artificial intelligence as a part of editing instrumentary and a separate autonomous component of CGI is a powerful mechanism that accelerates workflows and helps to stay focused on creative aspects. Delegation of repetitive and complex processes efficiently speeds up the project duration, opens new possibilities, and completely shifts an approach to post-production for artists and photographers in particular.

1.8 Machine learning in CGI

The term " machine learning " was first applied by an American computer scientist focused on artificial intelligence, Arthur Samuel¹⁷. Machine learning is an application of artificial intelligence evolution that focuses on developing algorithms and statistical models that allow one to perform specific tasks without being additionally programmed. This helps computers to learn from input data, recognize patterns, and make predictions and decisions.

Machine learning¹⁸ is divided into three groups:

- **Supervised learning** - when a machine has sample input and output data that is labeled or tagged. The main condition here is that data must be tagged correctly, and sample data should remain unchanged so that the machine can match them, compare them, and make predictions about the possible output relying on the trained dataset. It is also able to predict, identify, and fix errors while learning.
- **Unsupervised learning** - when a machine is trained using input data only and output samples are not provided. It looks for structures or repetitive patterns and

¹⁷ Dataversity. (n.d.). *A Brief History of Machine Learning*. Retrieved March 28, 2025, from <https://www.dataversity.net/a-brief-history-of-machine-learning/>

¹⁸ Data Science Dojo. (2022, October 4). *Machine Learning 101: The types of ML explained*. Retrieved March 28, 2025, from <https://datasciencedojo.com/blog/machine-learning-101/>

predicts an output that may not be that accurate or may be different from expected results.

- **Reinforcement learning** is a feedback-based technique where models are required to interact with an environment, performing specific actions to receive feedback in the form of rewards or penalties.

According to this information, it is necessary to mention the main contributors to artificial intelligence development and its integration into digital image processing-related tools and platforms. These companies were at the forefront of visual artificial intelligence development, creating powerful tools such as generative AI, and image recognition algorithms which quickly became the main pillars of AI-assisted digital post-production. One of the pioneers in this area was NVIDIA, which developed an Instant NeRF¹⁹ technology that can transform 2D images into entirely 3D scenes using neural radiance fields (NeRFs). Unlike other 3D technologies, Instant NeRF doesn't use a set of grids with placed and tracked objects but uses density and radiation function in 3D space, where density is responsible for the opacity and transparency of points in 3D space and radiance defines the color and radiance direction of the points. When the NeRF algorithm is being trained, it basically learns to predict the color and density of the light at every point of the space. That allows the creation of realistic 3D scenes considering the number of shadows and reflections. It uses an optimized learning method that is based on multi-resolution hash-grid encoding²⁰. This method is not the case for the groups of machine learning mentioned before (supervised learning, unsupervised learning, and reinforcement learning) it is different and might be described as a self-supervised learning method. It is similar to supervised learning because it has input data. In this case, input data is many similar photographs of the same scene or an object, taken from different angles. It is good if the photographs overlap each other, it makes the learning process much smoother. This technology might even be applied to certain areas of photography, for example, interior and architecture photography, providing professional 3D photoshoots without having complex scanners or it makes it possible to visually recreate an original and authentic look of historical buildings or entire buildings themselves. It can also be used in other areas of

¹⁹ Stephens, J. (2022, May 12). *Getting Started with NVIDIA Instant NeRFs*. NVIDIA Technical Blog.

Retrieved March 28, 2025, from <https://developer.nvidia.com/blog/getting-started-with-nvidia-instant-nerfs/>

²⁰ Müller, T., Evans, A., Schied, C., & Keller, A. (2022, January 16). *Instant Neural Graphics Primitives with a Multiresolution Hash Encoding* [PDF]. arXiv. Retrieved March 28, 2025, from <https://arxiv.org/pdf/2201.05989>

photography, such as fashion, portrait, or destination photography. In fashion and portrait photography, these algorithms can help photographers adjust an angle or change the scene around the subject. In destination or travel photography, this technology can help create 3D versions of sights and landmarks or put the viewer right into the scene on Google Maps. NVIDIA's and Getty Images game-changing collaboration²¹ started in 2023 and focused on developing and customizing generative AI tools and text-to-image and text-to-video models. The idea of this collaboration revolves around the creation of fully licensed and customized visual content based on text concepts. Advanced generative AI models developed by Getty Images were trained on their own libraries of content and powered by NVIDIA's Edify/Picasso, which is presented as multimodal architecture for developing visual generative AI models for image, 3D, and 360 HDRi. As a result, users can generate high-quality images up to 4K, add or remove objects from existing images (inpainting and outpainting), and integrate images into larger scenes with various aspect ratios, maintaining seamless transition. It is a way for creators to receive fully custom visuals that meet legal standards and avoid copyright infringement. Additionally, in 2021, NVIDIA presented 57 research papers at the Computer Vision and Pattern Recognition (CVPR)²² conference dedicated to visual-related AI innovations. NVIDIA presented a technique called JeDi (**J**oint-**i**mage-**D**iffusion **M**odels)²³ for custom personalized text-to-image generation without finetuning using reference images. JeDi can be merged with Retrieval-Augmented-Generation as known as RAG to generate images within a specific profile, for example, it might keep the same visual style or appearance as the content that might be useful for brands or companies. During the CVPR conference, NVIDIA also presented FoundationPose, a new foundation model for object pose estimation and tracking that can be applied to other objects without fine-tuning. This model can use a small set of references or 3D presentation to understand the shape of the object, its movement, and its rotation despite poor lighting conditions or lack of visibility. This particular model can find its application in assisting robots and helping them identify and track moving objects or in augmented reality applications for educational or scientific purposes. Potentially, it also might be used in certain areas of visual art, for example in CGI. FoundationPose can

²¹ NVIDIA. (2023). *NVIDIA and Getty Images collaborate to build AI-powered creative tools*. Retrieved March 29, 2025, from <https://blogs.nvidia.com/blog/generative-ai-getty-images/>

²² NVIDIA Research. (2023). *Visual generative AI research highlights from CVPR 2021*. Retrieved March 29, 2025, from <https://blogs.nvidia.com/blog/visual-generative-ai-cvpr-research/>

²³ Müller, T., Evans, A., Schied, C., & Keller, A. (2022). *Joint Image Diffusion for Generative AI* [PDF]. Retrieved March 29, 2025, from <https://research.nvidia.com/labs/dir/jedi/assets/joint-image-diffusion-combine.pdf>

precisely place 3D objects into the scene according to the conditions and orientation. This helps to improve compositing, maintaining a balance between seamless blending, realism, and the artist's idea and creative vision. It might be useful in animation, as it requires complex and precise interaction between moving objects.

1.9 Contributions of OpenAI to AI-Assisted Photography and Post-Production

Another significant contributor to the advancement of AI-based tools is OpenAI, which is best known for developing AI language models like GPT²⁴ and the text-to-image generative AI model, DALL-E. According to the variety of tasks that are usually included in photographic production and post-production, language models like GPT-4 may provide advanced assistance in certain steps, such as:

- **Organization and Planning.** - Helping photographers to create detailed shooting plans and checklists, selecting locations, and providing production schedules.
- **Communication** - Outlining texts ensuring clear communication between photographers, clients, and project leaders by drafting emails or contracts.
- **Concept Writing** - Assisting in the creation of captivating visual or text concepts and mood boards by generating themes or descriptions that guide the direction of the shoot.
- **Reference Research** - Bringing relevant references or inspiration by selecting similar works, projects, styles, or trends. Research on specific techniques and equipment.
- **Brainstorming** - Generating suggestions or alternative approaches to problems that appear during a photoshoot or post-production stage.
- **Post-Production** - Analyzing the user's images, and providing feedback for image-enhancing by offering edits and possible adjustments, retouching styles, and color grading solutions.
- **SEO and Captions** - Generating SEO-optimized captions, metadata, and tags that can help improve visual material performance through search engines or social media platforms.

²⁴ OpenAI. (2023). *GPT-4 Technical Report* [PDF]. Retrieved March 30, 2025, from <https://arxiv.org/pdf/2303.08774>

DALL-E²⁵ is a generative AI model that creates images from text prompts (text-to-image model). It can generate detailed, creative, and realistic visuals based on the text concept, which might be an extremely useful tool for photographers. For example, it allows for the generation of custom references and sketches tailored to the specific requirements of a photoshoot. By uploading reference images, DALL-E can guide the creation of visual content, making it closely resemble the reference or partially align with it by adopting its color palette, structure, or conceptual elements. DALL-E can also examine existing images and provide insights on their composition, color grading, and even the studio lighting techniques likely used to achieve specific lighting effects. The accuracy of the result is strongly based on the quality of the input data and the number of details provided. The whole process still needs to be supervised and controlled because every small error or inaccuracy can significantly affect the entire result.

The idea of optimization and automation of the manual processes during pre-production and post-production work in photography opens the door for creativity and the application of more advanced and sophisticated tools, such as Stable Diffusion.

1.10 Stable Diffusion and Advances in Generative AI

Stable Diffusion - is a deep-learning, text-to-image, and image-to-image generative AI software developed by Stability AI and released in 2022²⁶. It represents a “latent-diffusion” model²⁷ class that is capable of quickly transferring text prompts into photo-realistic images. This allows developers to build their own tools using the Stable Diffusion base while users can run the model on their computing hardware - even on under 10GB of VRAM.

First of all, it is necessary to clarify how Stable Diffusion works and what exactly are diffusion and latent-diffusion models used in generative AI. Stable Diffusion uses CLIP (Contrastive Language-Image Pretraining), an AI model developed by OpenAI, to understand text descriptions and match them to relevant visual elements. After that, the algorithm learns to identify relationships between pixels and their spatial positions and

²⁵ Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical Text-Conditional Image Generation with CLIP Latents*. Retrieved March 30, 2025, from <https://arxiv.org/pdf/2204.06125>

²⁶ Stability AI. (2022). *Stable Diffusion: The Latent Diffusion Model*. Retrieved March 30, 2025, from <https://stability.ai/blog/stable-diffusion-announcement>

²⁷ Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). *High-Resolution Image Synthesis with Latent Diffusion Models*. Retrieved March 30, 2025, from <https://arxiv.org/pdf/2112.10752>

generate random noise that will gradually be “diffused” revealing and shaping the final result step-by-step. During the entire generation process, an algorithm keeps comparing the current step with both the previous step and the text prompt, ensuring accuracy and relevance to the given input. This describes the basic operational mechanism of the diffusion model’s work. The latent-diffusion model compresses the space and reduces its resolution using variational autoencoder (VAE), instead of operating in the regular pixel space (which requires enormous computing capabilities), which significantly increases the speed, and accuracy of generation and takes fewer resources, making it able to operate on a regular user’s hardware. After the main set of generations is finished, right before the final generation, Stable Diffusion decodes the images back to high resolution.

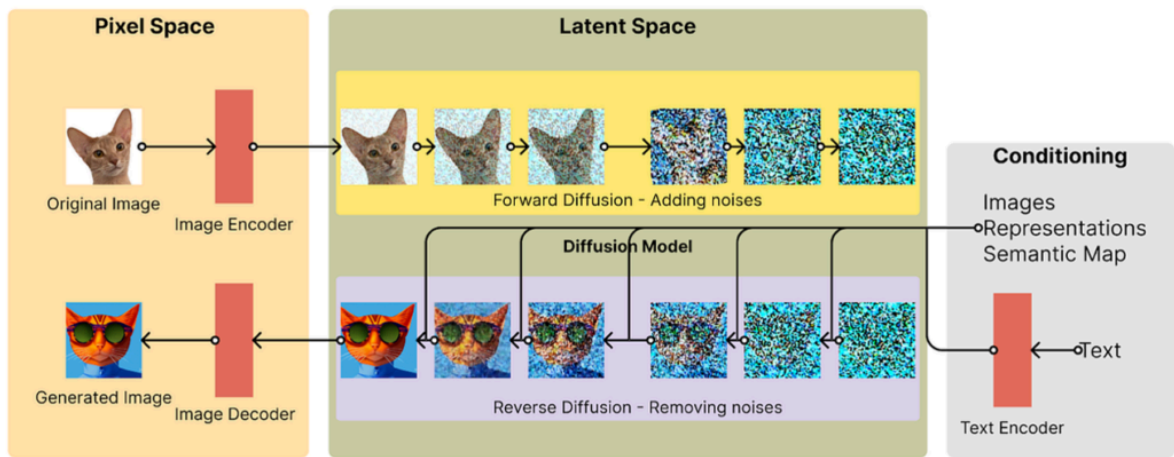


Fig. 5. Latent space noise reduction guided by textual conditioning to generate the final image.



Fig. 6. Image decoded from a latent diffusion model.

The collective effort of a Stable Diffusion community²⁸ led to the appearance of a large number of specialized models for photorealism, architecture, CGI, portraiture, animation, and more. Stable Diffusion supports such functions as inpainting and outpainting.

- **Inpainting** - is an image editing technique where an AI-powered tool fills selected areas, maintaining the visual coherence of the image. It works as follows: First, the user uploads the original image and selects the areas for generative fill with a selection brush. After that, the AI algorithm analyzes an image's context and surrounding pixels and then matches it with the text prompt to generate a seamless and natural-looking replacement. The main advantage of this technique is that generative content applies only to selected areas, while the rest of the image is untouched. As a result, inpainting is an advanced tool for post-production and can be used to assist in object removal, content-aware editing, and photo restoration.
- **Outpainting** - a technique that extends the image beyond its original frame by generating visual content that matches the original's style, context, lighting,

²⁸ Stable Diffusion community contributions and advanced image editing functions. (2025, March 29). In *Hyperstack*. Retrieved March 29, 2025, from <https://www.hyperstack.cloud/blog/case-study/everything-you-need-to-know-about-stable-diffusion>

composition, and quality within a given input, such as a text prompt or reference image.

StabilityAI gathered a dynamic community of developers and artists who actively work on Stable Diffusion's enhancement and refinement. A collaborative ecosystem is characterized by developing custom plugins such as ControlNet²⁹, which is a significant Stable Diffusion advancement that is created to improve user's control over the generation process. Depending on the idea, specialization, and desired result, there are many specific conditioning methods³⁰ of the ControlNet plugin available:

- **Edge Detection** - edge detection methods like Canny Edge and Holistically-Nested Edge Detection (HED) are used to extract outlines and contours from the original image. These outlines serve as a structural guide during the image generation. The edges of reference images are detected and used as input, ensuring that the generated image is aligned with the reference edges. It might be used to: transform simple line drawings into full artworks, maintain structural consistency while experimenting with colors and styles, and create stylized interpretations of real-world scenes.
- **Sketch-to-image** - This allows artists to use rough sketches to create clean, detailed, and polished images. This method is also good for quickly prototyping concept art or storyboarding.
- **Segmentation Maps** divide images into distinct regions based on object categories (sky, person, building, etc). This method provides control over the object's placement and composition. It might be used to compose complex scenes maintaining visual coherence, quality, and clarity or to generate consistent variations of scenes with individual elements changing.
- **Normal Maps** are used to add intricate surface details, such as textures and lighting effects. It might be useful to add realism to the generated images. Normal maps are extracted from reference images (which are possible to upload right into ControlNet)

²⁹ Yan, X., Lei, Z., Wu, L., Li, S., & Zeng, X. (2023). ControlNet: Adding Conditional Control to Text-to-Image Diffusion Models. *arXiv*. Retrieved March 29, 2025, from <https://arxiv.org/pdf/2302.05543.pdf>

³⁰ Illyasviel. (2023). ControlNet GitHub Repository. Retrieved March 30, 2025, from <https://github.com/Illyasviel/ControlNet>

and used to guide the diffusion process, ensuring surface details are accurately represented in the final image. It is a powerful tool for CGI that can create photorealistic images for advertising and elevate the overall realism of the generated images.

It is possible to combine several ControlNet conditionings simultaneously. This approach gives unparalleled control over various aspects of image generation, ensuring all of the elements can work together. This flexibility makes ControlNet an essential tool for generating complex compositions with many layered details. In addition to ControlNet's conditioning methods, Stability AI expanded creative possibilities through LoRA³¹ (low-rank adaptation) models. LoRAs allow users to train Stable Diffusion for specific purposes by tuning it on custom datasets with minimum computing power. For example, photographers can train LoRAs on particular directions or styles of photography, such as high-contrast black and white, fashion, portraits, interior, and architecture. As a result, it is possible to achieve a separate and fully personalized processing model that is tailored to special artistic needs. When LoRA is trained on specific photographic styles, it might be used for automatic skin retouching, color-grading, or lighting adjustments with minimum manual intervention. Stable Diffusion can significantly enhance image quality through upscaling. An upscaling function can increase the resolution of low-quality images while preserving fine details and textures. It is perfect for preparing images for large-scale printing or high-definition displays. As the main focus of the thesis revolves around Interior and Architecture photography and the application of CGI and AI elements in this field, it would be necessary to examine the proper setting of tools such as LoRA and discover all the benefits of its use. To fully prepare LoRA for work in the field of interior and architecture photography, photographers need to train the model on specific datasets that reflect the characteristics of interior and architecture photography. The training process involves only modifying the subset of the model's parameters (e.g., cross-attention layers), making it computationally lightweight while retaining high-quality outputs. Training LoRA to work with Interior and Architecture images includes the following steps:

³¹ Hu, E. J., et al. (2021). *LoRA: Low-Rank Adaptation of Large Language Models*. arXiv preprint. Retrieved March 29, 2025, from <https://arxiv.org/pdf/2106.09685.pdf>

Step 1: Define the Objective

Before the beginning of the training³² process, it is essential to define the specific purposes for the LoRA model. Interior and architectural photography includes various styles and applications, so defining the intended objective is essential for guiding dataset preparation and training.

Examples of Objectives:

- Create LoRA that modifies specific lighting conditions of the photo, such as natural daylight for interiors and evening dusk lighting for exteriors.
- Train LoRA to enhance textures and surfaces like polished marble, wooden floors, glass facades, and iron-steamed bedsheets.
- Develop LoRA that specializes in specific architectural styles like modern minimalism, industrial design, or classic architecture.
- Automate tasks like perspective correction, color-grading, or virtual staging for real estate photography.

Step 2: Curate a High-Quality Dataset

The dataset is the most crucial component of training LoRA. It should accurately represent the needed characteristics of interior and architecture photography to ensure the model learns effectively.

Dataset Characteristics

1. **Diversity:** Include images that cover different lighting conditions and angles (wide shots vs close-ups) and environments (residential interiors, commercial spaces, office or urban exteriors)
2. **Resolution:** Use high-resolution images that capture fine details, such as textures and patterns.
3. **Consistency:** Ensure that the dataset aligns with the needed style or objective.

For example:

- For modern interiors: images with clean lines, neutral tones, and minimalist furniture.
- For urban architecture: images of classic facades or cityscapes at dusk, or skyscrapers.

Sources for Datasets

³² ScottBaker.ca. (2023). *LoRA training (Stable Diffusion 1.5)*. Retrieved April 2, 2025, from <https://www.scottbaker.ca/AI/LoRA-Training>

- Personal photography portfolios
- Stock photo libraries (e.g., Getty Images, Adobe Stock, Unsplash)
- Synthetic datasets that are generated using existing AI tools.

Dataset Size

LoRA requires relatively small datasets compared to full model fine-tuning. A set of 10-50 high-quality images is enough for training specialized LoRA. But it is recommended to use larger sets of images.

Preprocessing

Before training starts, it is necessary to complete the following steps:

- Resize images to match an input resolution of Stable Diffusion (512x512 or 768x768 pixels)
- Normalize colors to ensure uniformity across the dataset.
- Annotate the images if necessary (labeling lighting types, textures, or architectural features)

Step 3: Setting Up Training Environment

Hardware Requirements

- A GPU with at least 8GB of VRAM is recommended
- More powerful GPUs will speed up the training but are not strictly necessary due to LoRA's lightweight nature.

Software Tools

1. Training Frameworks:

- **DreamBooth:** Ideal for personalizing models based on small datasets
- **Kohya Trainer:** A convenient tool designed specifically for fine-tuning Stable Diffusion with advanced customization options.
- **Automatic1111 WebUI Extensions:** Provides a convenient and accessible interface for integrating LoRA's into workflows after training.

Dependencies:

- Python libraries such as PyTorch, Hugging Face Transformers, and diffusers.
- Pretrained StableDiffusion weights as the base model

Configuration Settings

During Setup:

- Select cross-attention layers as the focus of fine-tuning to reduce memory usage

- Adjust learning rates and batch sizes based on dataset size and hardware capabilities
- Use early-stopping techniques to avoid overfitting

Step 4: Train the LoRA model

The actual training involves “feeding” curated dataset into the base Stable Diffusion model, while modifying only specific parameters using LoRA techniques.

Steps in Training

1. Loading pre-trained Stable Diffusion weights as the base model.
2. Feeding curated dataset into training pipeline.
3. Fine-tuning only selected layers (e.g., cross-attention layers) while freezing others to save computational resources.
4. Monitoring loss functions during training to ensure optimization without overfitting.

Training Duration

The process duration typically takes from 10 minutes to several hours depending on dataset size, GPU performance, and desired level of details in the final LoRA.

Step 5: Evaluate and Refine

After training is completed, it is necessary to evaluate the LoRA by generating the test outputs based on prompts related to interior and architecture photography. The evaluation should focus on identifying any artifacts or unwanted details and assessing the quality of surfaces, lighting, and placement of the objects in generated images.

Evaluation Criteria

1. Artifacts and Unwanted Details

- Are there any noticeable artifacts such as noise, pixelation, or patterns?
- Are there any unwanted objects or details in the generated images (e.g., ghosting, reflections, unrealistic structures, or non-existing objects not intended by the prompt)?

2. Surface Quality

- Are textures realistic? (e.g., patterns, fabric textures)?
- Are surfaces smooth and detailed if it is necessary (e.g., glass surfaces)

3. Lighting Accuracy

- Is the lighting consistent with the intended style or prompt? (e.g., natural daylight, artificial lights)
- Are shadows and highlights realistic and well-balanced?

4. Placement and Composition

- Are the objects and elements correctly positioned in the scene?
- Is the overall composition visually appealing?

If the evaluation reveals issues with these criteria:

- Refine the dataset by adding more examples or improving preprocessing steps
- Adjust hyperparameters such as learning rate or number of epochs in subsequent training runs.

Practical Part

Building upon the theoretical framework and research presented in the previous chapters, this practical part demonstrates the application of CGI and AI tools to optimize a photographer's workflow. The primary objective is to explore and evaluate how these technologies can be integrated into photographic practice to enhance both creative expression and technical efficiency.

In this section, I outline the step-by-step process undertaken to implement CGI and AI-based solutions. This includes the careful selection of appropriate software and tools, the preparation of visual materials, and the development of specific workflows tailored to the goals of the project. Throughout the practical work, particular attention is given to the challenges encountered, the criteria used to evaluate results, and how the theoretical concepts discussed earlier are realized in practice.

The focus of this practical work is primarily on interior and architectural photography. This specialization aligns closely with my professional experience, providing opportunities to apply and refine these techniques across a diverse range of real-world projects. Working in this field allows for practical experimentation with lighting, spatial composition, texture enhancement, and virtual staging - areas where CGI and AI tools can have a significant impact. The variety of projects I engage with offers a broad perspective on how these technologies can be adapted to different styles, environments, and client needs.

By systematically documenting each stage of the process, this part aims to provide a comprehensive understanding of the practical implications and creative potential of CGI and AI in photography. The following methodology section details the specific tools, processes, and criteria used to implement and assess the integration of CGI and AI in interior and architectural photography projects.

2 Methodology

This section outlines the methods and tools employed to integrate CGI and AI technologies into the workflow of interior and architecture photography. The choice of software and techniques was guided by the need to balance technical precision, creative flexibility, and practical applicability to real-world projects.

The practical work begins with the acquisition of photographic material from diverse interior and architectural environments, reflecting the variety of projects encountered in my

professional practice. These images serve as the foundation for subsequent CGI enhancements and AI-assisted processing.

CGI techniques are applied to generate or augment visual elements such as object removal or replacement, lighting simulation, and texture enhancement. The software selected for this purpose includes industry-standard 3D modeling and rendering tools, chosen for their compatibility with photographic workflows and ability to produce photorealistic results.

AI tools are employed to automate and refine post-processing tasks, including noise reduction, color correction, and detail enhancement. Additionally, generative AI models are explored to expand creative possibilities by generating objects, filling missing areas, or stylistic variations based on textual prompts.

Throughout the process, each stage is documented and evaluated against criteria such as visual coherence, realism, and workflow efficiency. Challenges encountered and solutions developed are discussed to provide insight into the practical integration of these technologies.

The practical part of this thesis relies on a selection of software tools that support the integration of AI and digital editing techniques into the photographic workflow, specifically tailored to interior and architectural photography. These tools were chosen for their ability to enhance image quality, streamline editing processes, and enable creative experimentation.

The core photo editing and enhancement software used includes Luminar Neo, Adobe Photoshop, and Adobe Lightroom. Photoshop provides extensive capabilities for detailed image manipulation, compositing, and retouching, while Lightroom facilitates efficient color correction and organization of photographic material. Additionally, Luminar Neo is employed for AI-powered enhancements such as automated batch processing, structure adjustments, and noise reduction, which contribute to improving workflow efficiency.

For generative AI applications, Stable Diffusion is utilized to explore creative possibilities such as background extension, style variation, and image synthesis based on textual prompts. This tool allows for innovative experimentation beyond traditional photographic techniques.

This selection of software represents the current toolkit applied in this thesis, but the workflow and methodology are designed to accommodate additional tools or updated versions in the future without requiring fundamental changes to the process. This

flexibility ensures that the practical approach remains adaptable to evolving technologies and software capabilities.

2.1 Tools Used in the Practical Work

Luminar Neo

Recently, I have adopted Luminar Neo as the central hub of my photographic workflow due to its powerful AI-driven editing capabilities and integrated organizational features. Luminar Neo combines batch processing, catalog management, and advanced AI tools such as structure adjustment, noise reduction, and the AI-assisted Relight feature, which allows selective relighting of image areas to enhance depth and balance. While sky replacement is available and occasionally used, the software's strengths lie in streamlining complex corrections and accelerating the editing of large image sets. Its user-friendly interface and all-in-one approach enable me to efficiently manage and enhance my interior and architectural photography projects from start to finish, reducing the need to switch between multiple programs.

Adobe Lightroom

Adobe Lightroom is used more selectively in my workflow, primarily for projects requiring specific file formats such as DNG, often requested by clients like Airbnb. Lightroom excels in RAW file management, offering robust cataloging, keyword tagging, and metadata handling. It also supports initial color correction and exposure adjustments to establish a consistent baseline before further editing. While Lightroom's organizational features are more comprehensive than Luminar Neo's, I rely on it mainly when client requirements or workflow demands necessitate its use.

FLUX by Black Forest Labs

Flux is a recent addition to my workflow, tested for object replacement and reconstruction - such as entire beds. It shows promise in automating complex edits with minimal manual input. I use it experimentally, focusing on large-object retouching.

Adobe Photoshop with Adobe Firefly Generative AI

Adobe Photoshop, enhanced by Adobe Firefly's generative AI capabilities, is a highly versatile and powerful tool central to my photographic workflow. Firefly's integration introduces advanced generative features such as Generative Fill, which allows for intuitive object removal, replacement, and the generation of missing image areas using natural language prompts. These AI-driven tools significantly accelerate traditionally time-consuming tasks like cleanup, texture refinement, compositing, and extending or altering image content. Photoshop's generative layers enable non-destructive editing, providing flexibility to experiment and iterate rapidly. While results can occasionally vary, the combination of Photoshop's manual precision and Firefly's AI creativity offers a highly efficient and effective environment for detailed retouching and creative manipulation across interior and architectural photography projects.

Stable Diffusion with LoRa Models

Stable Diffusion, fine-tuned with LoRa (Low-Rank Adaptation) models specifically trained for my purposes, is a powerful generative AI tool designed for large-scale scene regeneration and creative experimentation. Training LoRa models is a complex and time-intensive process, often requiring significant computational resources and careful parameter tuning. Despite fine-tuning, the model can sometimes produce unrealistic scenes or fail to preserve critical details as requested. In such cases, I selectively use only the well-generated portions of the AI output and blend them with the original photographs in Photoshop to maintain visual coherence and quality.

After generation, I apply an upscaling tool to enhance resolution and detail, ensuring the final images meet professional standards required in interior and architectural photography. This approach allows me to efficiently recreate or modify entire scenes, such as replacing a bed with a new one featuring soft, smooth textures, much faster than manually editing each element piece by piece. Throughout the practical part of this thesis, I will gradually describe how I develop and refine this multi-tool workflow to balance automation, precision, and creative flexibility.

Having outlined the core tools and technologies that form the foundation of my workflow, I now turn to their practical application within real interior and architectural photography projects.

3 Editing Process and Use Cases

The following section presents a step-by-step description of how these resources are integrated and utilized in practice. By moving from a discussion of capabilities to actual implementation, I aim to demonstrate how each tool contributes to an efficient, creative, and adaptable process. This practical workflow illustrates not only the technical sequence of actions but also the decision-making and problem-solving involved at each stage.

With this context established, I will now detail the workflow from initial image acquisition through to final output, highlighting the role and impact of each tool along the way.

When working on interior photography projects, I generally follow a consistent workflow, though the exact steps and choice of software can vary depending on the client's needs and the intended use of the images. My clients include apartment owners, real estate agents, residential management companies, and Airbnb, with whom I work on a contract basis. Each type of client shapes how I approach both the photo shoot and the post-production process.

3.1 Airbnb Workflow

For Airbnb assignments, the workflow is adapted to fit their specific requirements. Since Airbnb has its own editing team, my main responsibility is to deliver high-quality images that are ready for their editors. After receiving a request from Airbnb, I get in touch with the apartment owner to arrange the shoot. During the session, I use HDR (High Dynamic Range) photography - typically capturing 3 to 5 photos of each scene at different exposures, two stops apart, a technique known as exposure bracketing. This approach is especially valuable in interior photography, where lighting conditions can be challenging and there are often strong contrasts between bright windows and darker interior spaces. By exposure bracketing, I can capture a wider dynamic range, preserving important details in both highlights and shadows, and later merge these images into a balanced, natural-looking

final photo³³. This method is widely used in real estate and property photography to achieve high-quality, visually appealing results even in difficult lighting situations.

Once the shoot is complete, I merge the exposures into single HDR images, save them in DNG format as required by Airbnb, and upload the files to Airbnb's Media Box. Their editing team then handles the rest of the post-processing.

In this workflow, I use Adobe Lightroom for two primary reasons. First, it helps me maintain organized files, structuring each project separately and allowing easy retrieval or re-editing of images later if needed. Lightroom's cataloging features are essential for managing RAW files efficiently and ensuring that I can quickly locate specific images even months or years after the initial shoot.

Second, Lightroom provides a very convenient and effective tool for merging bracketed exposures into high-quality HDR images. The HDR merge feature, combined with the editing tools in the Develop section, allows me to fine-tune the merged images and save them as DNG files - the required format for Airbnb. This process is intuitive, can be applied step-by-step to individual images, and supports the processing of multiple images.

This streamlined approach allows me to focus on capturing strong images from the start, while meeting Airbnb's standards. In the following sections, I will describe how my workflow changes when working with other clients, where I am responsible for both the photography and the entire editing process. After scheduling the photoshoot, I focus on understanding the client's specific needs and priorities, identifying where to place the most attention during the shoot. I discuss important details such as project duration, pricing, editing style, and the number of images required. At this stage, I use ChatGPT for several purposes: to ensure clear and professional communication, to help create moodboards, and to find reference images.

Once the photoshoot is scheduled, I move into preparation and pre-production. Here, ChatGPT assists me in creating detailed checklists of rooms and the number of shots needed, helping me stay organized and avoid missing any important images.

On the day of the shoot, I arrive at the location and spend the first 10 to 15 minutes assessing the space, lighting conditions, and any details that need special attention. After setting up my gear, I follow my checklist closely while photographing the space.

³³ Expert Photography. (n.d.). *Exposure bracketing and HDR in interior photography*. Retrieved April 18, 2025, from <https://expertphotography.com/bracketing-photography>

3.2 Lightroom vs Luminar Neo (HDR Merge)

After carefully reviewing the material, I upload it into Lightroom or Luminar Neo - the latter being a powerful tool I recently discovered and now use as the main starting point for my post-production process. Luminar Neo offers an integrated catalog system, allowing me to import RAW files and immediately begin reviewing and processing images within a single interface.

One of Luminar Neo's standout features is its automatic batch HDR merge capability. The software intelligently groups bracketed exposures by capture time, sorting them automatically. This is especially helpful when working with dozens of similar brackets that may have slight angle differences not always visible before merging. Luminar reads the capture timestamps and pauses between each bracket series to group images, presenting them as manageable portions. This lets me remove completely dark or unwanted photos - for example, in series without windows, I often exclude the darkest brackets since there are no strong highlights to recover.

To merge, I select and drag all relevant files into the HDR Merge section and configure preferences such as auto-alignment and distortion correction. Luminar Neo then aligns all the selected images - merging any number of exposures at the same time - into high-quality HDR photos. The merged images are saved as 16-bit TIFF files, which contain more image data than PSD or DNG formats, allowing for greater flexibility and detail in further editing. This streamlined, intelligent workflow saves significant time and ensures consistent, detailed results across large image sets, making it ideal for interior and architecture photography projects.

Here, I compare the HDR merging process in Lightroom and Luminar Neo, focusing on how each software handles merging bracketed exposures. I highlight differences in workflow and the quality of the merged HDR images, showing how the programs produce different results in terms of dynamic range and image data. This comparison helps illustrate the strengths and limitations of each tool in creating high-quality HDR photos.

Adobe Lightroom

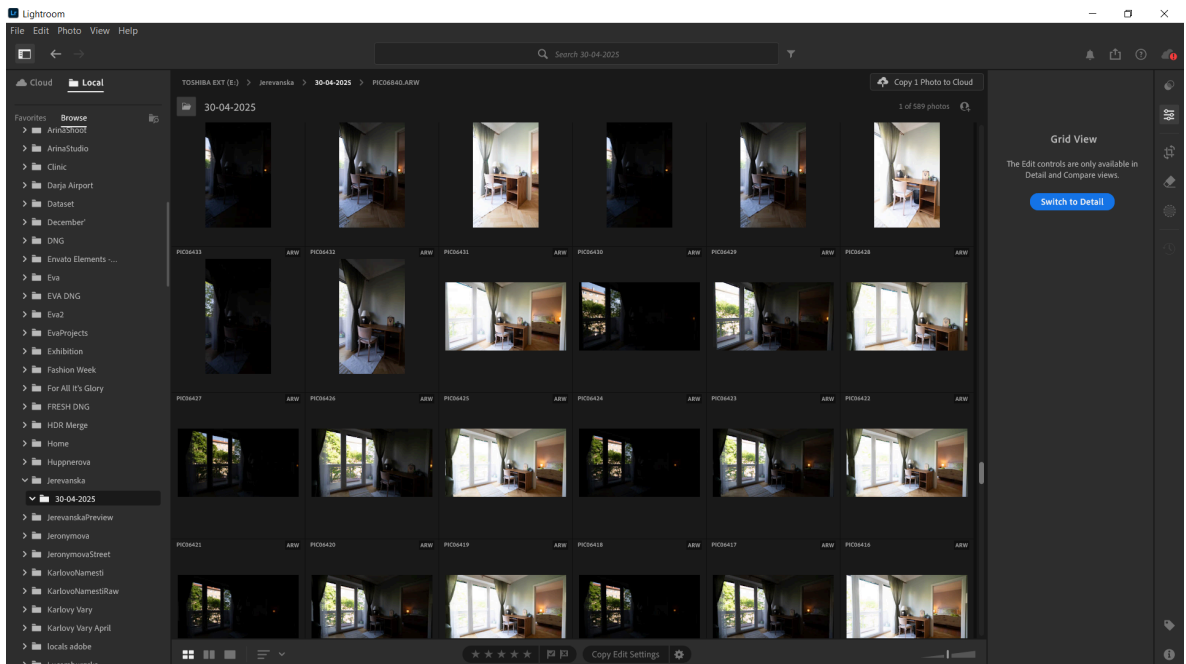


Fig. 7. Adobe Lightroom interface showing catalog and image organization.

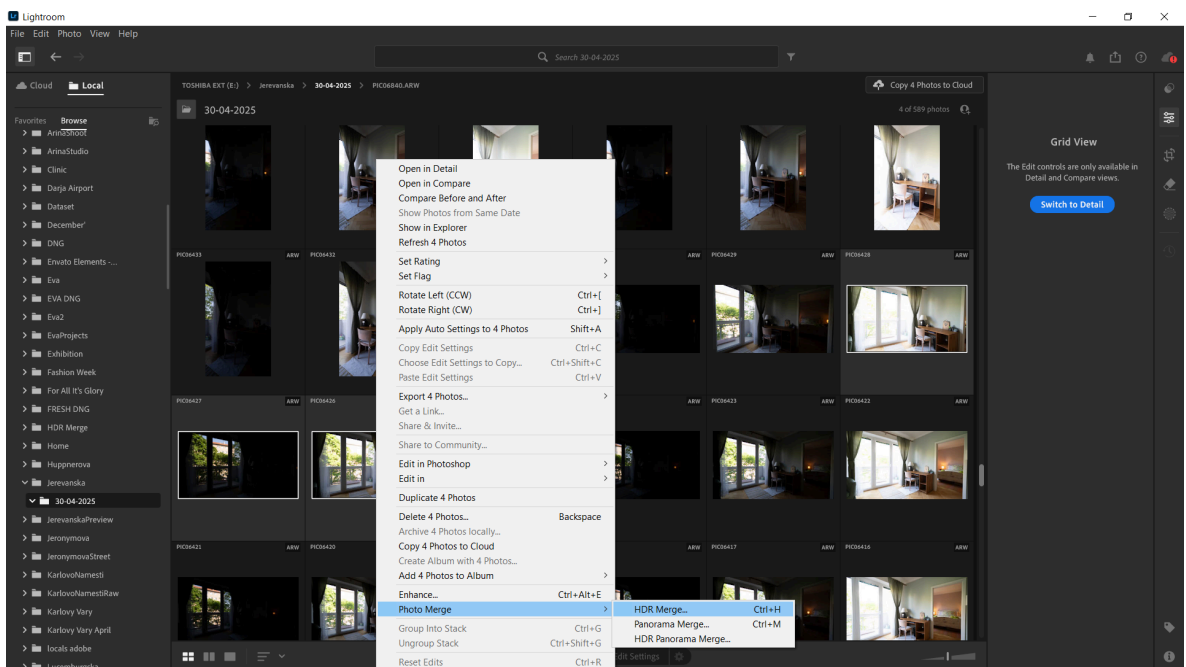


Fig. 8. HDR Merge interface in Adobe Lightroom, showing exposure alignment and merge options.

Steps: Selecting brackets holding CTRL/Cmd - RMB click - Photo Merge - HDR Merge.

Note: No Batch HDR Merge.

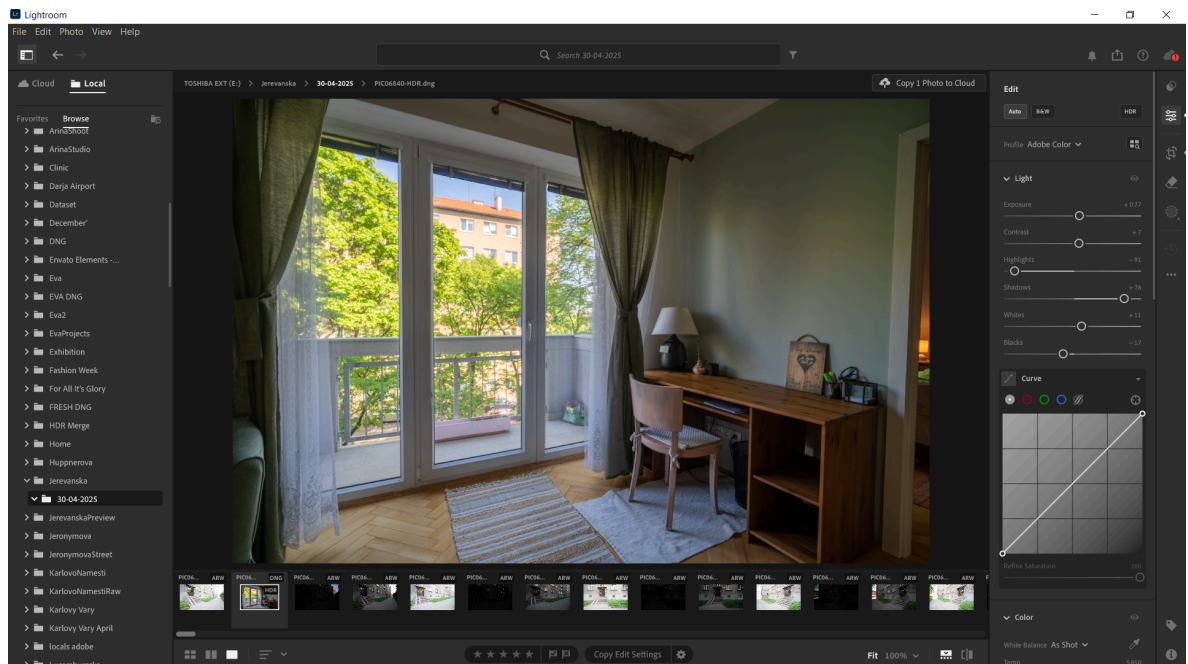


Fig. 9. HDR image after merging in Lightroom, illustrating tone correction and editing limits.

Steps: Basic exposure adjustment, lowering highlights, and increasing shadows to bring back details.

Note: No exposure adjustments are made at this stage to avoid losing important image data when exporting the file to Photoshop.

Luminar Neo

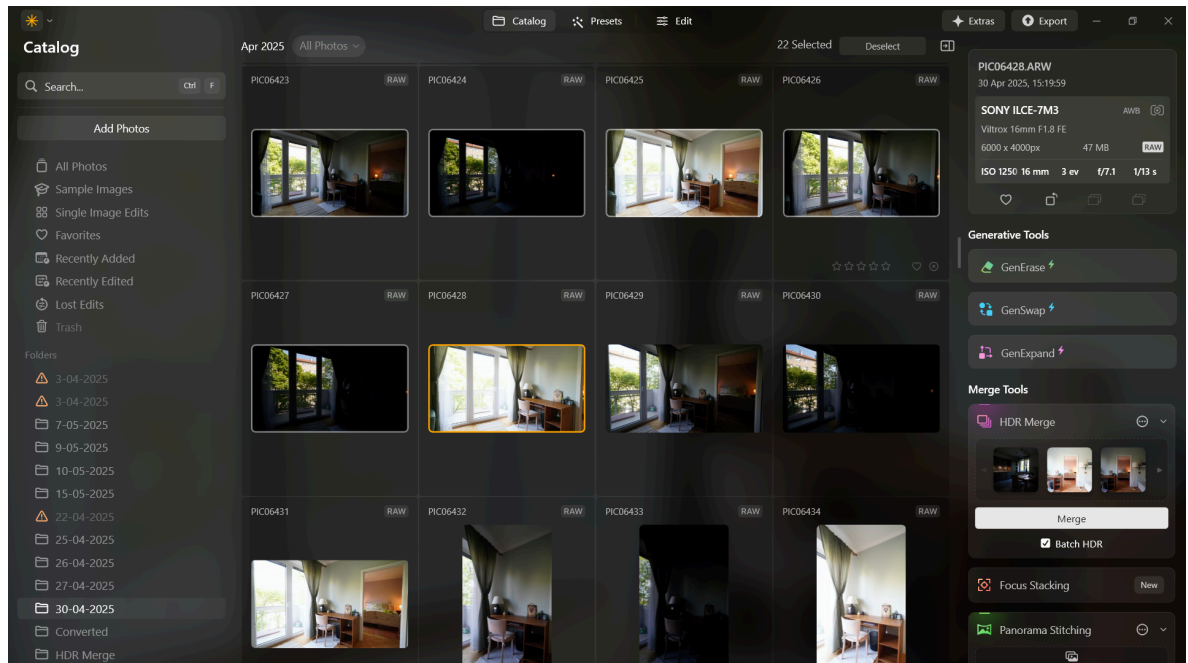


Fig. 10. Luminar Neo interface and image catalog view.

Steps: Selecting multiple photos from the Catalog (up to 1000 images) - Dragging them to the HDR Merge menu on the right panel - Turning on Batch HDR

Note: Similar to Lightroom, but easier to navigate

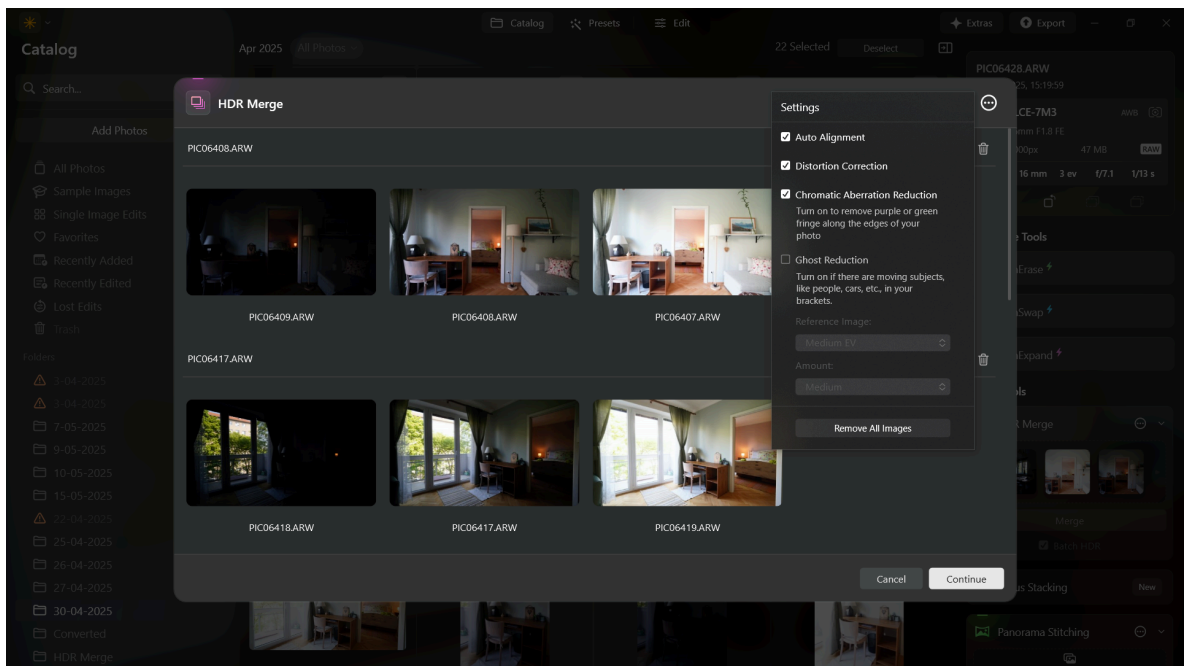


Fig. 11. Luminar Neo Batch HDR Merge feature in the main workspace.

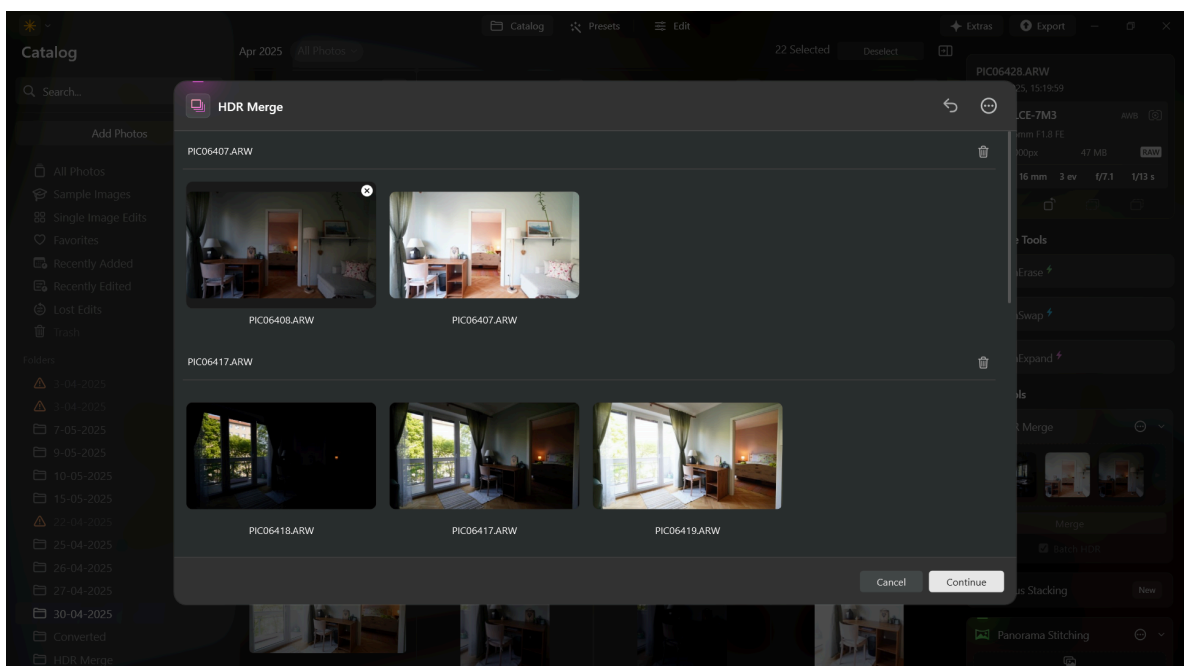


Fig. 12. Luminar Neo Batch HDR settings showing image exclusion options.

Steps: Set batch options using the menu at the top right corner of the HDR Merge window. Exclude unwanted images - for example, the darkest exposures in bracket sets without windows - since these don't contain highlights to recover and can make the merged image darker, which may increase noise during later editing.

Note: More steps, but more control and ability to automate the entire merging process

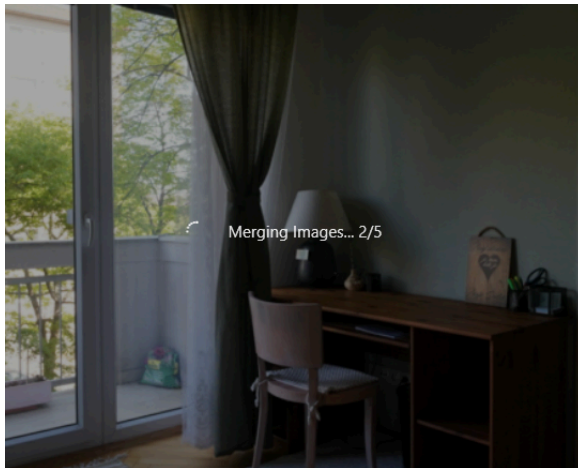


Fig. 13. Preview window of Batch HDR Merge process in Luminar Neo.

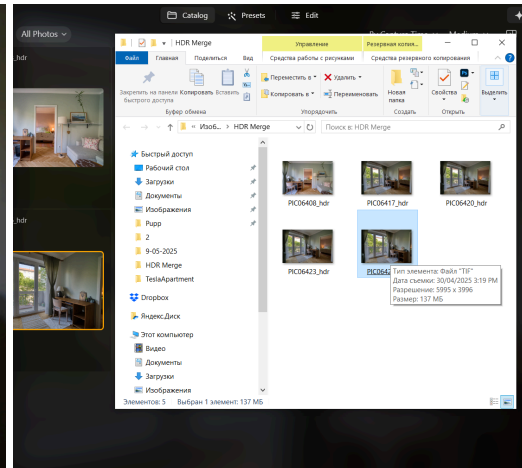


Fig. 14. Luminar Neo saving merged HDR image as TIFF in a separate folder.

Luminar merges selected brackets and saves the merged HDR photos as TIFFs into a separate folder.

In my experience, Luminar Neo's merged TIFF files consistently retain more detail in both the brightest and darkest areas of the image, making it significantly easier to recover information during further editing. This advantage is significant when working with interior spaces that often present challenging lighting conditions, such as strong contrasts between sunlit windows and shadowed corners.

To clearly demonstrate this difference in data retention and recovery, the following screenshots will compare Luminar Neo and Lightroom by adjusting the merged HDR images in Photoshop's Camera Raw. By pushing highlights down and shadows up, I will show how much detail can be recovered from each file, illustrating the practical impact of each program's HDR merging capabilities. This comparison will help highlight the strengths of Luminar Neo in preserving image data for advanced post-production work. First of all, I would like to compare the merged results from Luminar Neo and Lightroom to highlight differences in visual balance and detail. Luminar Neo's merges generally offer smoother highlight-to-shadow transitions and better detail retention, while Lightroom's results may show less recovery in extreme areas, affecting overall image quality.



Fig. 15. Luminar Neo – HDR Merge preview without adjustments.

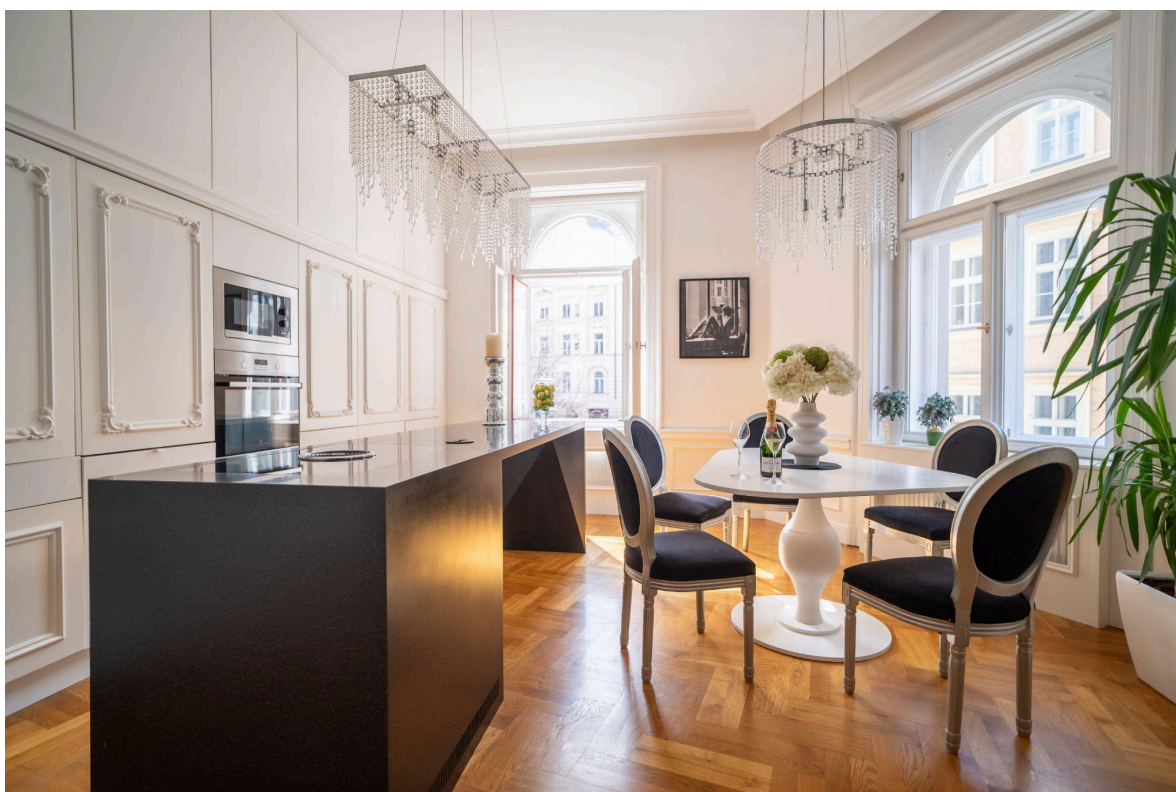


Fig. 16. Adobe Lightroom – HDR Merge preview without adjustments.

I made a short data limit test. I set Highlights on -100 and Shadows on +100 to see how much data is recovered from overexposed and underexposed areas, and here are the results:

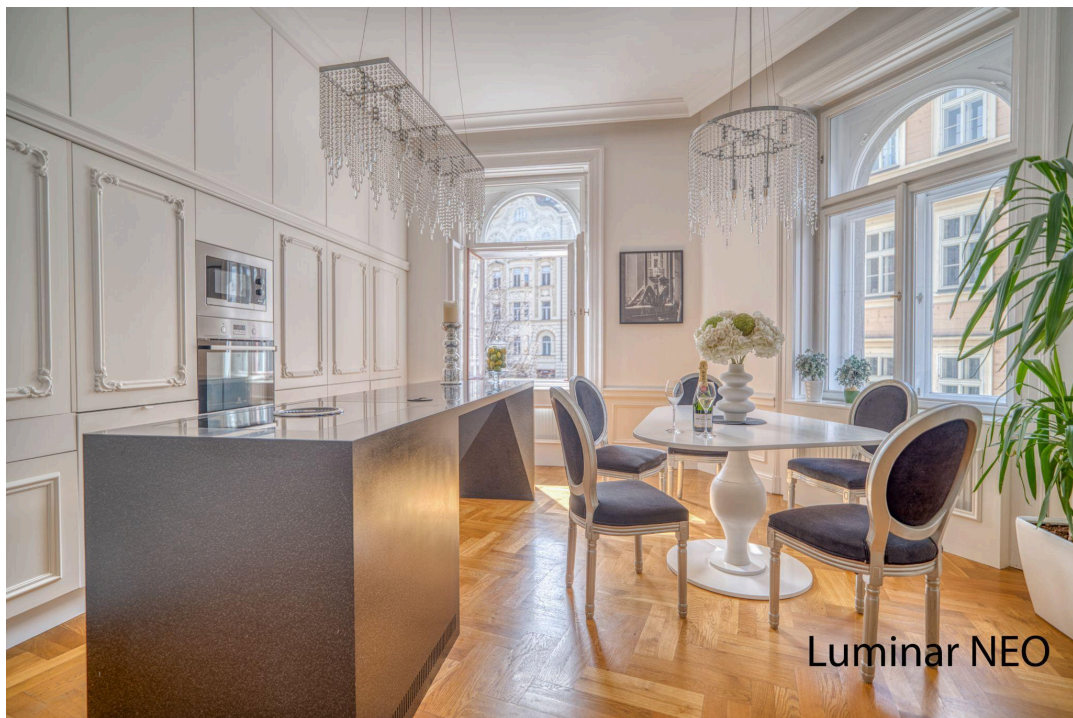


Fig. 17. Luminar Neo – Data limit test with -100 highlights and +100 shadows.

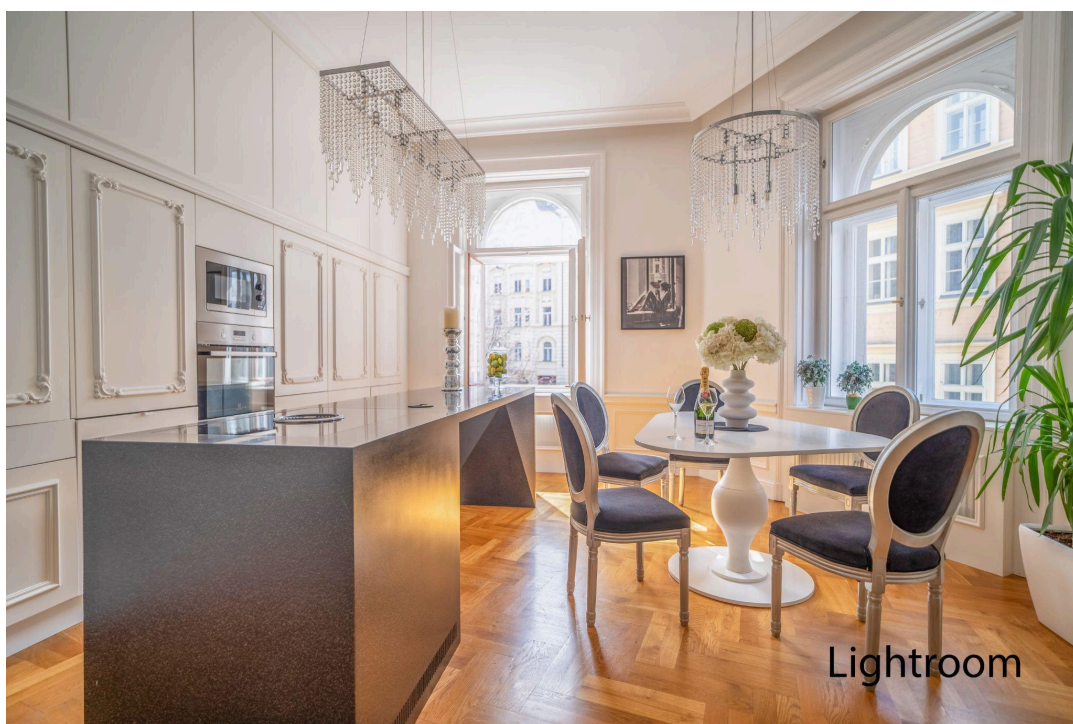


Fig. 18. Adobe Lightroom – Data limit test with -100 highlights and +100 shadows.

3.3 Comparison of Luminar Neo and Lightroom HDR Merging

Luminar Neo Advantages

Luminar Neo leverages cutting-edge AI technology to provide a highly efficient and high-quality HDR merging experience. Its AI-assisted exposure blending evenly mixes data from all bracketed exposures, resulting in a well-balanced final image with smooth shadows and rich highlight details. The merged files are saved as 16-bit TIFFs, which retain more image data, including greater color depth and tonal range, compared to Lightroom's DNG output. This richer data retention is crucial for interior photography, where preserving subtle lighting nuances and details in bright windows or shadowed areas is essential.

Luminar Neo's HDR merge also produces images with less digital noise, even when shadows are aggressively lifted, allowing for cleaner edits and less reliance on noise reduction. This means post-production is faster and less complicated, as highlight and shadow recovery can be achieved with fewer adjustment layers or complex masking techniques.

Lightroom Limitations

Lightroom's HDR merge workflow differs significantly. It applies auto-settings to each bracket before merging and consistently uses the darkest bracket as a reference image. This approach results in merged files that contain less overall data because not all bracket exposures are fully incorporated. Consequently, overexposed areas are harder to recover, often appearing bleached and lacking detail. The reliance on darker brackets also introduces more digital noise, which requires additional noise reduction steps during editing.

Moreover, Lightroom tends to base the merged image on middle or darker exposures. If these exposures are darker than usual, the final image will also be darker, leading to challenges in shadow recovery. When shadows are pushed beyond a certain limit, Lightroom begins to bleach highlights and reduce contrast, further degrading image quality and increasing noise. This workflow demands more manual intervention and prolongs the editing process, especially for interior photography where balanced lighting is critical.

Key Differences Luminar NEO vs Lightroom

HDR Merge Feature	Luminar Neo	Lightroom
Output File Format	16-bit TIFF with high data retention	DNG with less data than TIFF
Exposure Blending	AI-assisted, balanced mixing of all brackets	Auto-settings on each bracket; darkest bracket as reference
Batch Processing	Automatic grouping by capture time	Manual grouping often required
Highlight and Shadow Recovery	Superior detail retention; smooth shadows; less noise	Limited highlight recovery; increased noise; darker overall results
Data Incorporation	Uses full data from all brackets	Limited data, biased toward middle/dark exposures
Noise Handling	Reduced noise even with aggressive shadow lift	Noise from darker brackets requires extra reduction steps
Workflow Efficiency	Streamlined with fewer manual corrections	More manual adjustments and masking needed

Table 2. Key differences between Luminar Neo and Adobe Lightroom.

Luminar Neo's HDR merging process offers clear advantages in preserving image data, balancing exposures, and reducing noise, which directly benefits the quality and flexibility of interior photography edits. Due to these strengths, I will continue to use Luminar Neo for HDR merging in my workflow.

3.4 AI-Assisted Cleanup and General Corrections

A thorough cleanup step is crucial for professional, non-destructive interior photo editing. The process usually begins with standard adjustments like exposure correction, light and color balancing, and geometry or distortion fixes. However, achieving a polished final image also requires careful cleanup - removing dust spots, stains, unwanted objects, distractions, and reflections that can reduce image quality. This cleanup is often time-consuming and draining, which is where AI tools prove especially helpful.

To demonstrate AI features, I will use different images suited to various scenes. To start, we'll focus on a bright interior with mostly white tones, where dust spots are visible.

Dust spot removal is an important part of the cleanup process and can be done in a couple of ways using AI tools. Luminar Neo's Erase feature finds and removes dust spots automatically, making the process much faster. While it works well for most spots, some tricky areas might still need a little manual fixing. On the other hand, Photoshop's generative fill requires to select the dust spots yourself, which takes more time and effort.

To make dust spots easier to see, I make a copy of the image layer and add a Curves adjustment layer on top. By lowering the midtones and moving the highlights closer to the middle, the dust spots stand out more clearly. This helps find and remove even small dust spots and blemishes more easily.

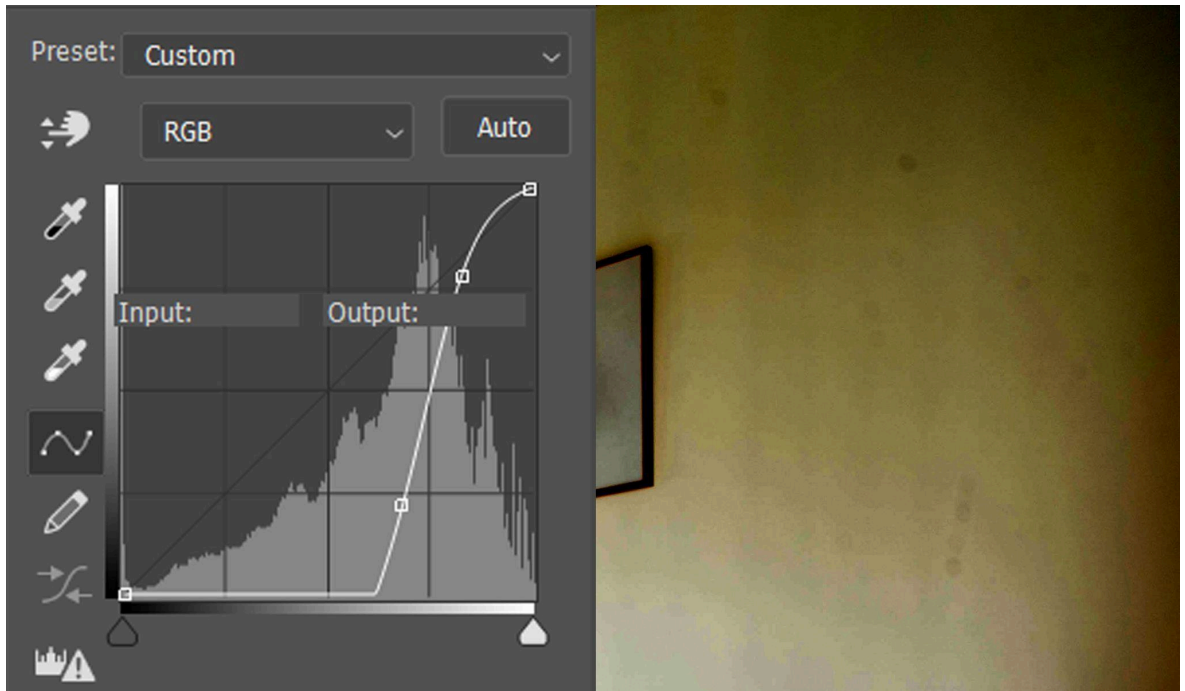


Fig. 19. Curves adjustment layer as dust spot revealing overlay.



Fig. 20. Curves adjustment layer application.

Let's test both tools and compare their convenience, time required, and results.

First, I'm using Luminar Neo's Erase feature, accessed directly through Photoshop since it works as a plugin. After opening the tool and selecting "Erase – Dust Spots," the software automatically finds and removes the dust spots. The whole process takes about 10–15 seconds, making it extremely fast and efficient.

Next, I'll remove dust spots using Photoshop's generative fill. I start by creating an empty layer as my working base, then add a Curves adjustment layer above to make the dust spots stand out. Using a selection brush, I carefully select each visible spot - adjusting the brush size slightly larger than the spots themselves. Once all spots are selected, I click "Generate," and Photoshop fills the selected areas by matching the surrounding textures, tones, and colors. This method gives full control and can deliver technically flawless results, but it is much more time-consuming (around 3-5 minutes) since every spot must be selected manually. Here, I'm providing interfaces comparison and results.

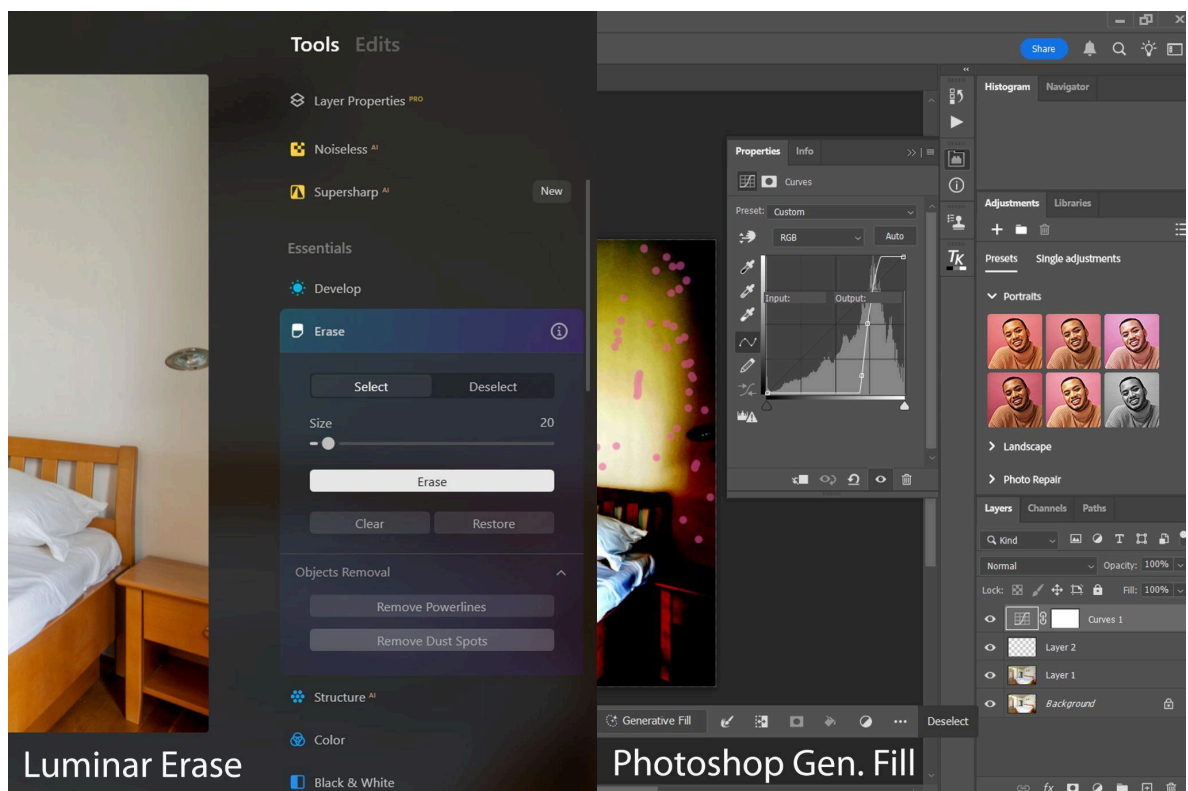


Fig. 21. Demonstration of dust spots removal techniques in Luminar Neo and Photoshop.

To clearly see the differences, I'm using the same Curves adjustment layer above each version, which helps highlight any remaining dust spots.



Fig. 22. Demonstration of dust spots removal results using Luminar Neo and Photoshop's Generative Fill.

As we can see, Luminar provides a faster and decent solution, but some spots remain. On a regular scale without adjustment layers, these spots aren't very visible, but it's still better to retouch them using a surface blur with a local mask. Generative fill takes much longer since it requires manually selecting all spots and then waiting for the algorithm to remove them, but the result is much cleaner from the first try. My choice mainly depends on the amount of time available for a project. When speed is more important than perfection, I choose Luminar Neo because it delivers good results quickly. However, if technical quality and a flawless finish are priorities, I prefer using Photoshop's generative fill, as it offers greater control and cleaner results, though it requires more time and effort.

4 Generative AI Tools

4.1 Brief Overview of the Photoshop Generative AI

Generative Fill is an AI-powered feature in Adobe Photoshop, based on the Adobe Firefly generative model. It allows users to add, remove, or modify image content by making selections and optionally typing text prompts. Unlike traditional tools such as the clone stamp or content-aware fill, Generative Fill analyzes the surrounding textures, lighting, perspective, and colors to create realistic and seamless results that blend naturally with the original image.

4.2 Demonstration of Basic Capabilities

- **Object Removal**

I select unwanted objects or distractions in the image and apply Generative Fill. The AI analyzes the surrounding area and fills the selection with matching textures and lighting, creating a seamless removal.

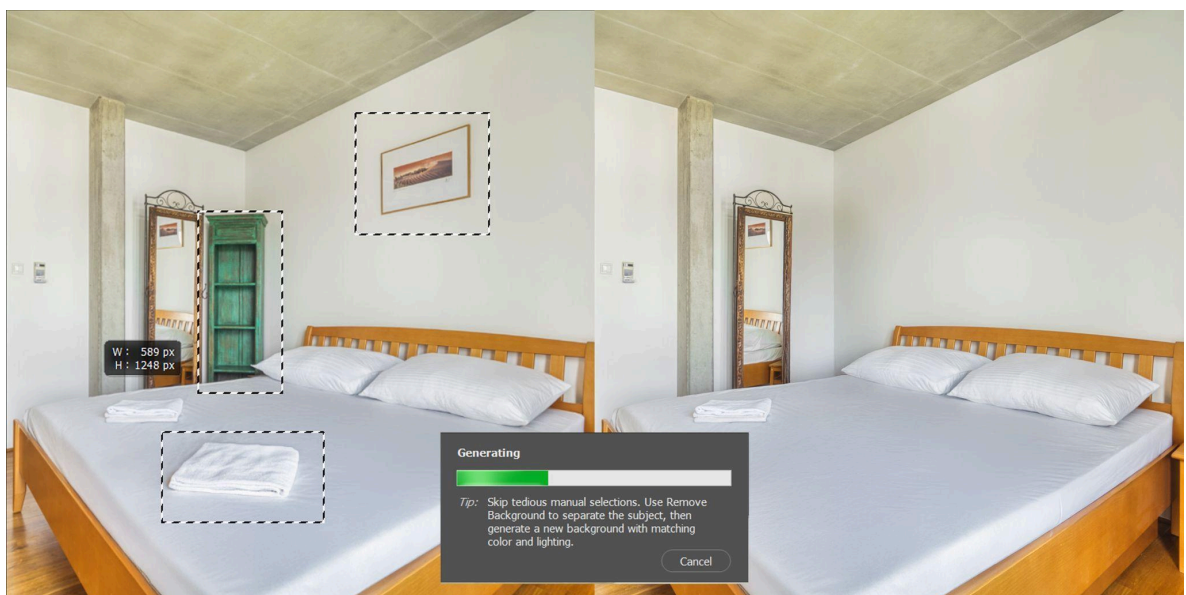


Fig. 23. Object removal process using Generative Fill in Photoshop.

- **Wires, Cables, and People Removal**

Using Photoshop's Remove Tool with the new "Find Distractions" feature, I can remove wires, cables, and even unwanted people very easy. I select the layer, click "Find

Distractions,” then choose either “Wires and Cables” or “People.” Photoshop automatically detects and removes these elements by generating new background content that blends naturally. For example, I’ve used this to instantly erase power lines or distracting cables from photos without manual brushing.

People Removal

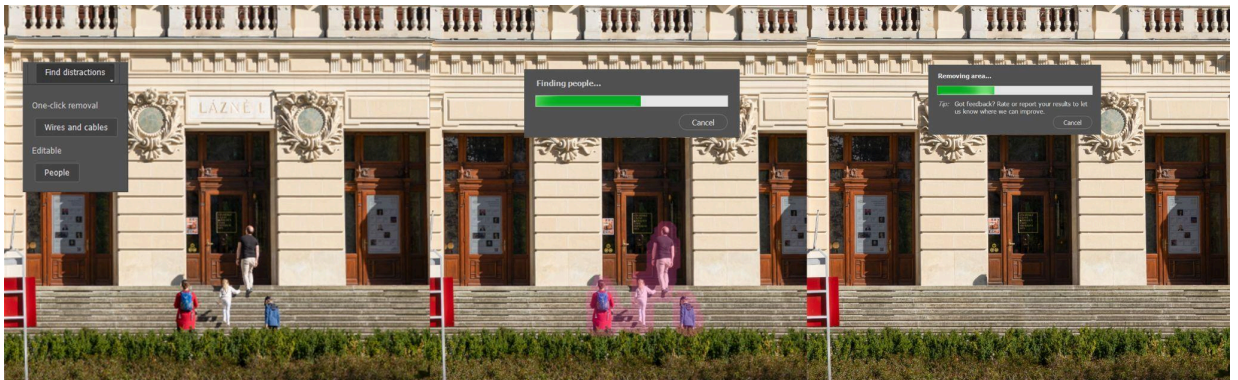


Fig. 24. Automatic People Removal feature demonstration with automatic detection and results.

Wires and Cables Removal

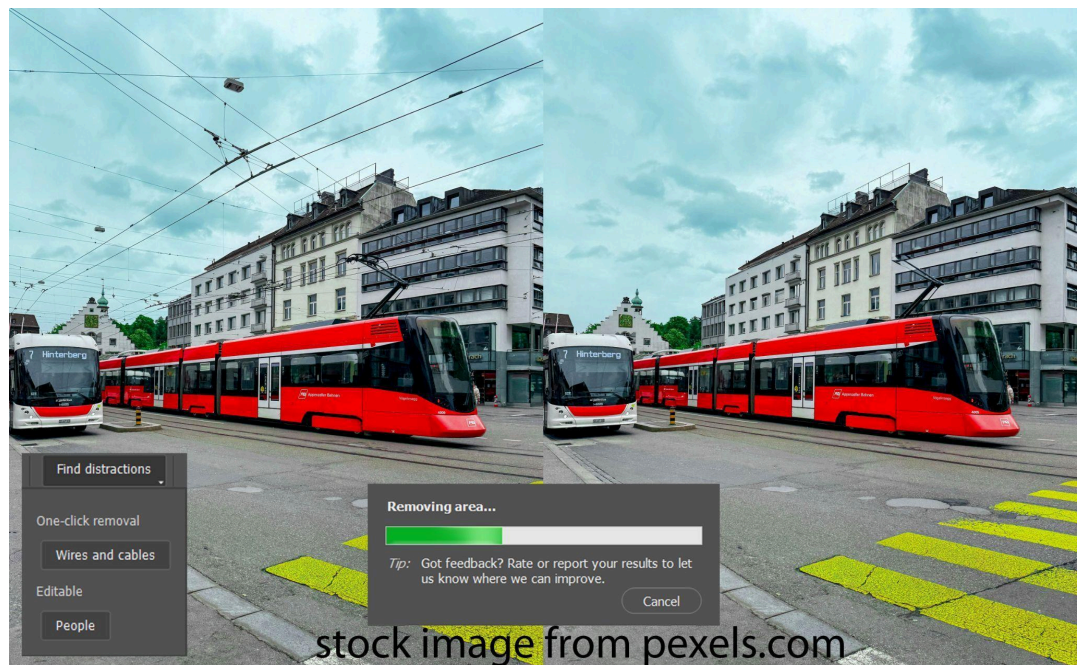


Fig. 25. Automatic wires and cables removal feature in Photoshop with auto detection and clean results demonstrated on stock image from pexels.com

- **Background Expansion**

After increasing the canvas size, I select the new empty areas and use Generative Fill to generate background content that fits naturally with the existing scene's perspective and colors.

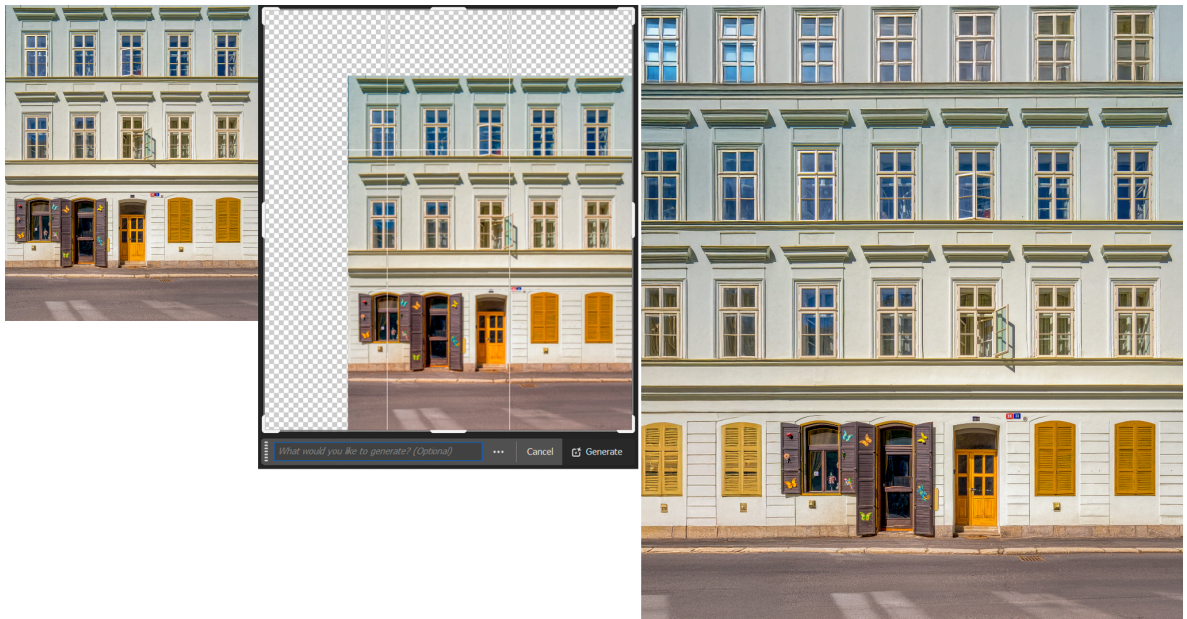


Fig. 26. Collage that demonstrates background expansion feature step-by-step.

- **Replacing Details**

I select elements, then type a descriptive prompt to replace them and AI generates new content that blends smoothly with the rest of the image.

Prompt: White Kettle



Fig. 27. Demonstration of the object replacement process using Generative Fill in Photoshop.

4.3 Brief Overview of Luminar NEO AI-Powered Tools

Luminar Neo offers several AI-powered tools that can significantly improve interior photos, making the editing process faster and more intuitive.

How AI in Luminar Works

Luminar Neo's AI features are powered by the Skylum AI Engine, an in-house technology developed by Skylum³⁴. This engine uses advanced algorithms to analyze the content, structure, and depth of each photo, automatically identifying objects, lighting, and details. By doing so, it applies targeted adjustments - like exposure, noise reduction, and relighting - based on what each image needs.

Enhance AI

Enhance AI is often the first tool I use when editing in Luminar. With just one slider, it automatically analyzes the image and adjusts exposure, contrast, shadows, and highlights for a well-balanced look. It also includes Accent AI for overall scene improvement and Sky Enhancer AI for refining sky details if visible in interior shots. This tool helps achieve a polished baseline quickly, saving time on basic corrections.



Fig. 28. Before-and-after demonstration of the Enhance AI tool in Luminar Neo.

³⁴ <https://skylum.com/luminar>

Sky Replacement

Luminar's Sky Replacement tool uses AI to detect and replace the sky in photo seamlessly. It matches the new sky's lighting and colors to image for a natural, realistic look. The AI also suggests sky options that best fit photo's mood and lighting, helping quickly find the perfect match. It is possible to fine-tune the sky's position, brightness, and blending to ensure a smooth transition with the rest of the scene, including realistic reflections on water and other reflective surfaces. I mostly use Luminar's Sky Replacement for exterior shots to replace dull, gray skies with cleaner, brighter ones. This helps create a more polished, advertising-friendly look by making the scene appear clearer and more inviting.



Fig. 29. Before-and-after demonstration of the Sky Replacement AI tool in Luminar Neo.

Noiseless AI

Noiseless AI analyzes the image to detect and separate noise from real details. It then applies targeted noise reduction, preserving textures and sharpness while removing grain, which is especially valuable for interior photos taken in low light. Unlike traditional noise reduction, it keeps important details clear.



Fig. 30. Before-and-after demonstration of the Noiseless AI tool in Luminar Neo.

Structure AI

Structure AI enhances details and contrast in the image, especially in backgrounds and textures, without overprocessing people or soft areas. This tool is content-aware, so it brings out the richness of surfaces like wood, fabric, or architectural features, making interiors pop while maintaining a natural look.



Fig. 31. Before-and-after demonstration of the Structure AI tool in Luminar Neo.

Relight AI

Relight AI creates a 3D depth map of the scene, letting me adjust exposure and color temperature separately for foreground and background. This is especially convenient for interiors, where lighting can be uneven. With a few sliders, I can rebalance the light, brighten dark corners, or add warmth, all while keeping the result photorealistic.



Fig. 32. Before-and-after demonstration of the Relight AI tool in Luminar Neo.

How AI in Luminar Works

Luminar Neo's AI features are powered by the Skylum AI Engine, an in-house technology developed by Skylum. This engine uses advanced algorithms to analyze the content, structure, and depth of each photo, automatically identifying objects, lighting, and details. By doing so, it applies targeted adjustments - like exposure, noise reduction, and relighting - based on what each image needs.

4.4 FLUX AI by Black Forest Labs

FLUX AI is a text-to-image generative model developed by Black Forest Labs, a startup founded in 2024 by former Stability AI researchers who played key roles in creating Stable Diffusion. Black Forest Labs has quickly gained recognition for pushing the boundaries of generative AI with its FLUX.1 model suite, which is designed to deliver exceptional image detail, prompt adherence, and editing precision.

The FLUX.1 suite includes several specialized tools such as FLUX.1 Fill for advanced inpainting and outpainting, FLUX.1 Depth for depth-aware control, and other models that enable highly controllable and context-aware image editing. These models are available in multiple variants, including open-weight versions for research and fast, efficient versions for personal use.

I discovered FLUX AI after I had already developed my main AI workflow. As a photographer working in interior and architectural photography, I often need to replace complex objects - like beds with messy linens or wrinkled pillows - without disturbing the rest of the scene. Traditionally, I use inpainting techniques that require careful manual selection of areas bit by bit, using tools like Photoshop's Generative Fill or Stable Diffusion. Manual selection is usually the trickiest and most time-consuming part of this process.

Soft objects and fabrics are particularly challenging because the AI fills areas based on the prompt and surrounding textures. If the selected area has low contrast or multiple materials and patterns, the AI may generate a blended, surreal mess. To achieve a good result, I must separately select and work on each surface or texture, often modifying selections or prompts multiple times. This can take 40 minutes to an hour for a single object, even though it's still faster than fully manual retouching.

FLUX AI offers a more advanced and efficient approach. Instead of requiring detailed manual masking, I can simply describe the change I want using a text prompt. The AI recognizes the exact object and applies edits precisely where needed, without affecting the rest of the image. This object-level understanding and modification significantly speed up complex replacements and retouching, making FLUX a powerful addition to my editing toolkit.



Fig. 33. Flux.1 Kontext model announcement banner from Black Forest Lab official website.

So far, FLUX stands out as a game-changer among the many generative AI tools available. Its capabilities are especially valuable for advertising photographers, including those working in interior and architectural photography - fields that often overlap with commercial work like rental listings on platforms such as Booking.com and Airbnb. In these cases, speed and efficiency can matter more than absolute perfection, making FLUX an ideal solution thanks to its rapid, targeted editing.

In my workflow, FLUX has proven particularly effective for tasks like “reassembling” beds and bed covers, making them look inviting and tidy without altering any other part of the frame. While FLUX offers a wide range of advanced applications, for the purposes of

this thesis - which focuses on interiors and architecture - I will concentrate on using it for this specific scenario.

Next, I will showcase the process, highlight the main advantages and disadvantages, and explain how FLUX has helped me streamline my editing and save significant time.

Tools like Flux AI and Stable Diffusion typically require working with images that are not too large - usually under about 5 MB or within certain pixel dimensions (commonly up to around 2 megapixels or roughly 1400×1400 pixels). This limitation exists because processing very high-resolution images demands significant computational power and memory, which can slow down generation or cause errors.

To handle this, the common workflow is to save and upload a lower-resolution version of the image, sacrificing some quality to fit within the tool's limits. After the AI processing is complete, the image is then upscaled back to the desired resolution using specialized upscaling tools like Stable Diffusion's upscalers or dedicated software such as Upscayl.

By following this approach - resize down, process, then upscale - ensures maintaining a good balance between manageable file sizes and image quality, enabling efficient and effective use of generative AI tools for complex edits.

I'm going to test FLUX AI to see if it meets my specific needs for precise and efficient editing in interior photography. This will help determine how well it integrates into my workflow and whether it can speed up complex object replacements without compromising quality.

I have an interior photo of a bedroom where I have already applied some general corrections. I briefly balanced the lighting, geometry, and exposure, erased visible dust spots using Luminar, and applied Accent AI and Relight AI tools. The goal at this stage was to fix the messy bed, making it look properly prepared with soft white pillows and smooth, wrinkle-free bedsheets with straight edges. I saved the image as a JPEG weighing 3.94 MB.

It is important to note that I perform these fast, general corrections early in the editing process to balance exposure and geometry without overworking the file. Each correction adds to the file size, so I intentionally keep the resolution lower to fit Flux's upload limits (usually under 5 MB) while sacrificing minimal quality. This way, the file remains balanced and nearly lossless, ready for AI processing.

When saving the JPEG in Photoshop, I used quality level 11 on the 0-to-12 scale, which is very close to the maximum quality. This setting applies minimal compression, preserving about 96% of the image details with very little visible difference compared to the absolute maximum (quality 12), but results in a somewhat smaller file size. The resolution of the saved image is 5997×3998 pixels, ensuring that fine details and textures are well preserved for further editing and upscaling if needed.



Fig. 34. Merged and resized HDR image prepared for test run in FLUX.

Next, I upload the bedroom photo to FLUX AI. Among the available options - Generate, Edit, Fill, and Expand - I choose Edit because I want to modify only a specific object without affecting the rest of the image. I upload the file and enter the text prompt: *“Smooth bedsheets, clean bed cover, fluffy white pillows neatly arranged, no wrinkles.”* This instructs FLUX to focus precisely on improving the bed’s appearance while preserving the rest of the scene.



Smooth bedsheets, clean bed cover, fluffy white pillows neatly arranged, no wrinkles.



Fig. 35. FLUX Playground interface displaying test image and text prompt.

Let's run the algorithm and see the results. I set the output format to PNG or JPEG and the batch size to 4, which means each generation will produce four different variations. This allows me to choose the one that best matches my vision.

Here is the result:



Fig. 36. First image result generated using FLUX.1 Kontext that demonstrates targeted object rearrangement.



Fig. 37. Before-and-after comparison of an image processed with FLUX.1 Kontext.

From the four variations, I chose the one where the bed is arranged best - pillows and bedsheets are almost completely wrinkle-free, with all shapes looking natural and authentic. This image remains true to the original, and the entire process took me about a minute. Next, I will upscale this photo to restore its quality and prepare it for the final editing and fine-tuning steps.

There are many options to upscale the image and restore its original quality, but I will focus mainly on two tools. The first is Upscayl, a free and open-source AI image upscaler available as an application for Windows. It offers several AI models for different needs and processes images locally on my computer, making it a versatile and accessible choice.

The second tool is Stable Diffusion, which also supports upscaling with AI-driven models. For now, let's see what Upscayl can do. I am going to use the High Fidelity model in Upscayl because it focuses on preserving realistic details and smooth textures, making it ideal for enhancing interior photography.



Fig. 38. Interface of the Upscayl AI upscaling tool.



Fig. 39. First result after upscaling using Upscayl AI.

After 3 minutes of processing, here is the result I received from Upscayl. The image looks good when viewed at this size, suitable for display on a website without zooming in. However, it still doesn't fully meet my quality expectations. When I zoom in and compare the before and after, I notice that the upscaled image appears segmented, with object edges outlined unnaturally. This suggests some loss of smoothness and fine detail, which affects the overall realism.

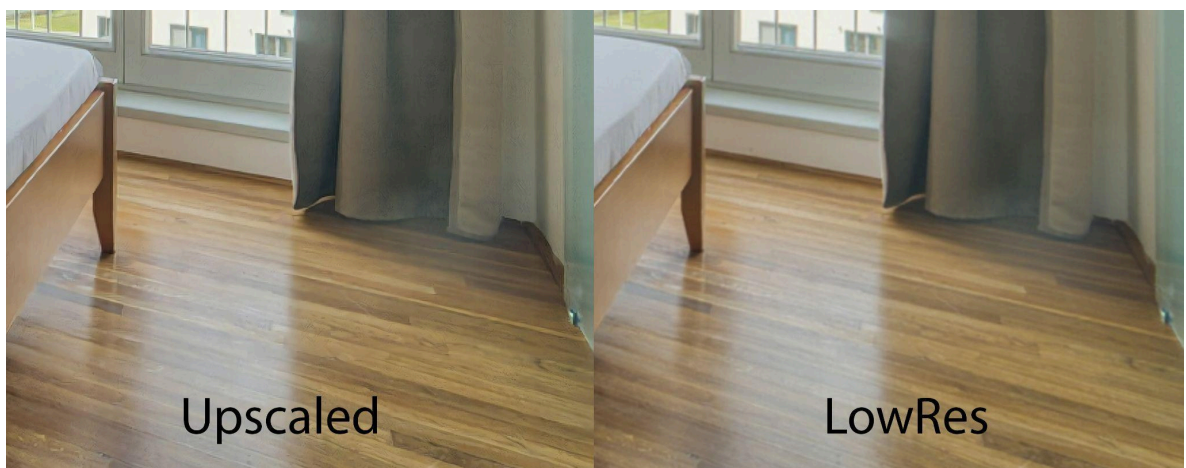


Fig. 40. Comparison demonstrating before-and-after results in detail.

A practical compromise and interesting solution I found is to use FLUX only for quickly rearranging the bed, which works well. However, downscaling the image and running it through various tools causes some quality loss affecting the whole image.

To avoid this, I place two images as layers in Photoshop - the original at the bottom and the FLUX-edited version with the new bed on top. Then, I add a layer mask to the FLUX layer and mask out everything except the bed, effectively blending the improved bed seamlessly over the original image without compromising overall quality.



Fig. 41. Original image with a layer mask revealing the bed from the upscaled version.

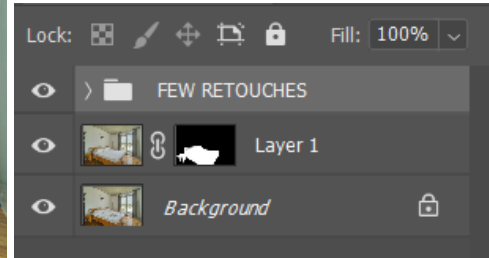


Fig. 42. Layer management and mask overlay.

The result is excellent because the FLUX and original images are nearly identical, making it easy to select the bed precisely. I refined the selection, applied a 10px feather, and used an inverted mask to blend the edited bed seamlessly with the original photo. Shadows, edges, and contours align perfectly, creating a natural look. I'm satisfied because the most complex object is now edited, while the rest of the image retains its original full data, allowing me to continue working comfortably.

The entire process - including general corrections, preparation for FLUX, generation, and blending with the original - took only about 7 minutes, compared to the 40 minutes to 1 hour typically required when retouching manually or using generative fill. Using Photoshop layer masks in this way ensures nondestructive, precise blending by hiding and revealing parts of the FLUX layer with soft edges, preserving a natural transition between layers.

4.5 Stable Diffusion and Its Application in My Workflow

In my workflow, Stable Diffusion is primarily used for targeted object replacement - specifically, rearranging and fixing beds in interior photos - while preserving the rest of the image untouched. This selective editing is possible through LoRA (Low-Rank Adaptation) training, a technique that fine-tunes Stable Diffusion on custom datasets to learn specific editing styles and tasks without retraining the entire model.

Why LoRA?

LoRA modifies only small parts of the model's weights, focusing on cross-attention layers, which control how text prompts influence image generation. This approach allows efficient, fast training with relatively small datasets and computational resources, making it ideal for adapting Stable Diffusion to precise, domain-specific edits like bed reassembly in interior photography.

LoRA Training: From Idea to Results

My thesis supervisor and I hypothesized that training a LoRA model on my own images could teach Stable Diffusion to rearrange beds according to my editing style. The goal was to automate complex retouching tasks, reducing manual work while maintaining quality and consistency.

Initial Attempts and Lessons Learned

Our first training attempt used only 86 images, but the results were inaccurate. The dataset included images with folded towels and various accessories that confused the algorithm, producing distorted beds and unusual artifacts. This highlighted the need for a larger, more focused dataset and longer training time to capture the variability and nuances of interior scenes.



Fig. 43,44. LoRA training first results at epochs 15–20.

5 Step-by-Step LoRA Training Process

Dataset Collection

- Selected 240 unedited bedroom images from completed projects, covering diverse angles, exposures, and color tones.
- Gathered 56 edited images as regularization data, showing the desired “neat bed” outcome for the model to learn.
- This larger dataset improved accuracy and reduced artifacts significantly compared to the initial run.

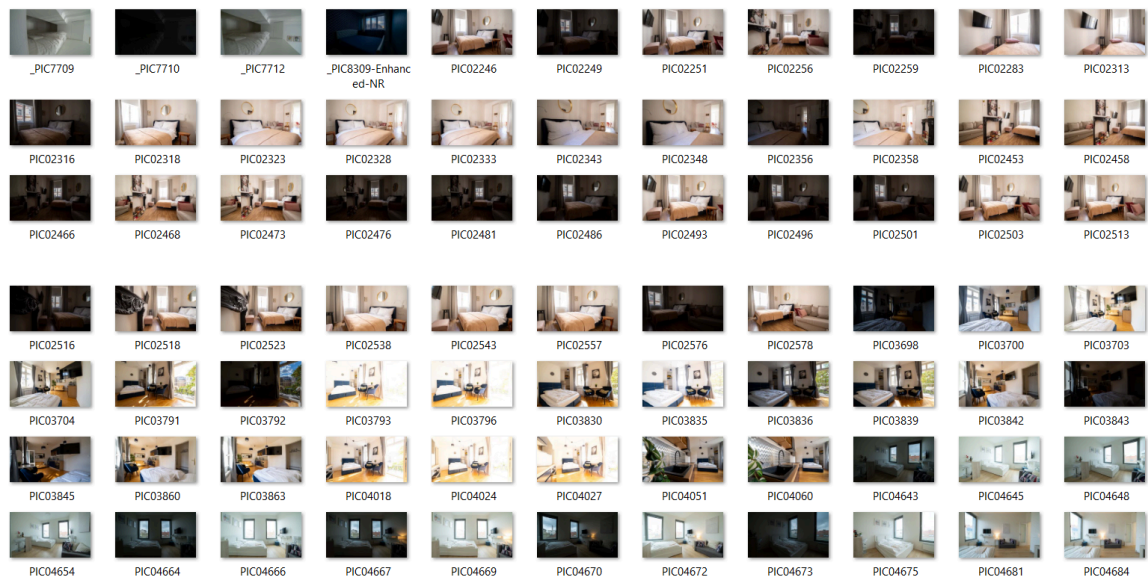


Fig. 45. Dataset prepared for LoRA training.

Image Preprocessing

- Resized all images to 768×512 pixels using a Photoshop action for uniformity and efficient GPU processing.
- Organized images into /train (inputs) and /reg (regularization) folders.

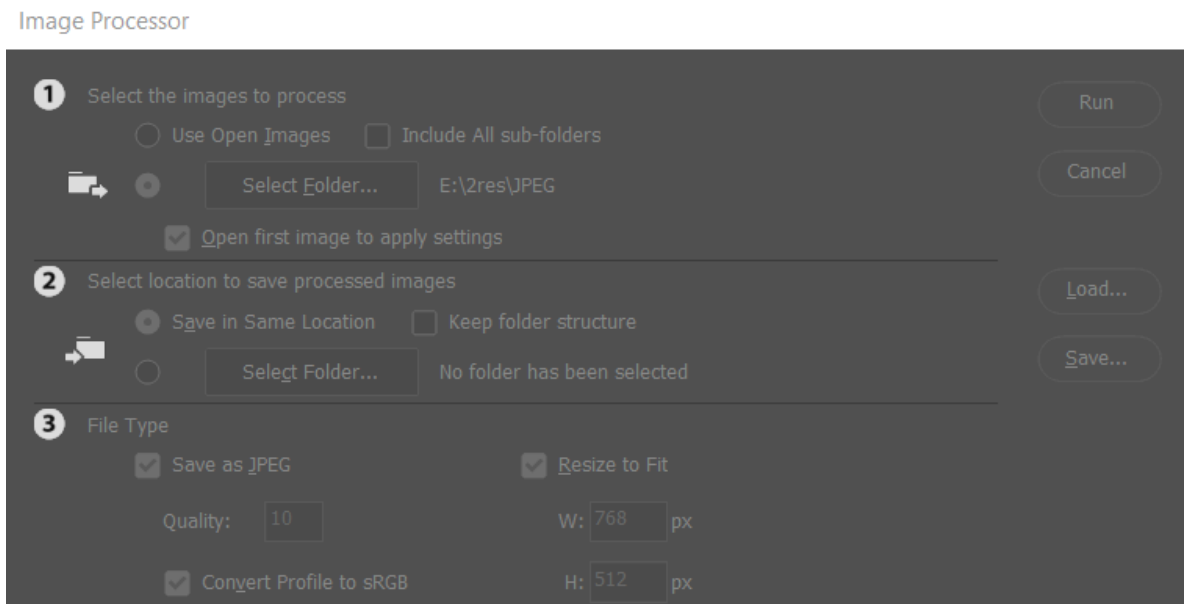


Fig. 46. Image Processor interface in Photoshop with settings for automated batch resizing.

Automated Tagging with Kohya_ss

- Uploaded images to Kohya_ss trainer.
- Used its automatic tagging feature (CLIP Captioning) to generate descriptive text files for each image, identifying objects and styles.
- Manually refined tags to ensure accuracy and relevance.

Model

Pretrained model name or path

runwayml/stable-diffusion-v1-5

Trained Model output name

ivan6

Image folder (containing training images subfolders)

C:/Kohya/kohya_ss/dataset/img

Dataset config file (Optional. Select the toml configuration file to use for the dataset)

Training comment

(Optional) Add training comment to be included in metadata

Save trained model as

☐ ckpt
☒ safetensors

Save precision

☐ float
☒ fp16
☐ bf16

Folders

Output directory for trained model

C:/Kohya/kohya_ss/dataset/model

Regularisation directory (Optional. containing regularisation images)

C:/Kohya/kohya_ss/dataset/reg

Logging directory (Optional. to enable logging and output Tensorboard log)

C:/Kohya/kohya_ss/dataset/log

Fig. 47. Kohya_SS trainer interface.

Dreambooth

LoRA

Textual Inversion

Finetuning

Utilities

About

Captioning

Convert model

Group Images

LoRA

Basic Captioning

BLIP Captioning

BLIP2 Captioning

GIT Captioning

WD14 Captioning

Manual Captioning

This utility allows you to create simple caption files for each image in a folder.

Image folder to caption (containing the images to caption)

C:/Users/PC/Desktop/Dataset/InputData/img

Caption file extension

.txt

☐ Overwrite existing captions in folder

Prefix to add to caption

(Optional)

Caption text

A professionally edited bedroom interior with enhanced materials and lighting.

- Surfaces: Polished, smooth finishes (marble-like or lacquered), no visible texture.
- Fabrics: Crisp, neatly arranged bedsheets (satin or linen), no wrinkles.
- Lighting: Balanced volumetric lighting, soft shadows, cinematic contrast.
- Geometry: Symmetrical composition, aligned perspective lines.
- Colors: Adjusted white balance, harmonious tones (cool/warm palette).

Postfix to add to caption

(Optional)

Find text

e.g., "by some artist". Leave empty if you only want to add a prefix or postfix.

Replacement text

e.g., "by some artist". Leave empty if you want to replace with nothing.

Caption images

v25.0.3

Fig. 48. Kohya SS caption text for automated image tagging.

Training Configuration and Execution

- Model: juggernaut_reborn.safetensors [338b85bc4f] as base.
- Parameters:
 - Resolution: 768×512
 - Batch size: 4
 - Epochs: 500 (longer training for better style capture)
 - Network rank: 128
 - Learning rate: 1e-4

Training took approximately 38 hours on an NVIDIA RTX 4070 Ti Super, 16GB RAM

5.1 Tagging and Configuration for LoRA Training

LoRA training is a highly specific process that demands careful setup of numerous parameters to achieve precision and avoid artifacts. One crucial step is tagging - each training image must have an associated descriptive text file that guides the model in learning the correct features and styles.

- Automated Tagging: Using Kohya_ss, I employed the built-in CLIP Captioning feature to generate initial tags for each image, identifying objects, textures, and styles automatically.
- Manual Refinement: After auto-tagging, I reviewed and refined tags to ensure accuracy, adding or removing descriptors to focus the model on relevant features (e.g., “messy bed,” “wrinkled sheets” for inputs; “neat bed,” “smooth white sheets” for regularization images).

This tagging process helps the model understand the difference between input and desired output states, essential for effective LoRA training.

5.2 Training Configuration and Key Parameters

The training setup involves tuning many parameters in Kohya_ss to balance learning speed, precision, and generalization. Below is a summary of the key parameters used for my project:

Parameter	Value	Purpose/Effect
Base Model	juggernaut_reborn.safetensors [338b85bc4f]	Starting checkpoint for fine-tuning
Resolution	768×512	Matches preprocessed image size
Batch Size	4	Number of images processed simultaneously
Epochs	500	Number of full dataset passes; longer for style capture
Learning Rate	1e-4	Controls update speed; balances stability and speed
Network Rank (Dim)	128	Controls LoRA matrix size; higher captures more detail
Network Alpha	64	Scales LoRA weights; balances influence
LR Scheduler	Constant with warmup	Gradually warms learning rate for stable training
Clip Skip	2	Improves generation quality for SD 1.5 models
Mixed Precision	Enabled (fp16)	Reduces VRAM usage and speeds up training
Cache Latents	Enabled	Speeds up training by caching intermediate data
Save Frequency	Every 10 epochs	Saves checkpoints regularly for monitoring progress

Table 3. LoRA training configuration and key parameters.

```

15:39:57-290076 INFO bucket_no_upscale = true
bucket_reso_steps = 64
cache_latents = true
caption_extension = ".txt"
clip_skip = 1
dynamo_backend = "no"
enable_bucket = true
epoch = 500
flip_aug = true
gradient_accumulation_steps = 1
huber_c = 0.1
huber_scale = 1
huber_schedule = "snr"
logging_dir = "C:/Kohya/kohya_ss/dataset\\log"
loss_type = "l2"
lowvram = true
lr_scheduler = "cosine"
lr_scheduler_args = []
lr_scheduler_num_cycles = 1
lr_scheduler_power = 1
lr_warmup_steps = 0.1
max_bucket_reso = 2048
max_data_loader_n_workers = 0
max_grad_norm = 1
max_timestep = 1000
max_token_length = 75
max_train_steps = 220000
mem_eff_attn = true
min_bucket_reso = 256
mixed_precision = "fp16"
network_alpha = 1
network_args = []
network_dim = 8
network_module = "networks.lora"
network_train_unet_only = true
noise_offset_type = "Original"
optimizer_args = []
optimizer_type = "AdamW8bit"
output_dir = "C:/Kohya/kohya_ss/dataset\\model"
output_name = "ivan6"
pretrained_model_name_or_path = "runwayml/stable-diffusion-v1-5"
prior_loss_weight = 1
reg_data_dir = "C:/Kohya/kohya_ss/dataset\\reg"
resolution = "512,512"
sample_every_n_epochs = 1
sample_prompts = "C:/Kohya/kohya_ss/dataset\\model\\sample\\prompt.txt"
sample_sampler = "euler-a"
save_every_n_epochs = 20
save_model_as = "safetensors"
save_precision = "fp16"
text_encoder_lr = []
train_batch_size = 1
train_data_dir = "C:/Kohya/kohya_ss/dataset\\img"
unet_lr = 0.0001
wandb_run_name = "ivan6"
xformers = true

```

Fig. 49. LoRA training log showcasing training status and key parameters.

Additional advanced parameters such as denoising strength and ControlNet assistance can be configured to improve training precision by guiding the model on noise reduction and structural consistency.

5.3 Training Insights

LoRA training is an iterative and sometimes unpredictable process. During the early epochs (1–10), the model begins to grasp basic features, but outputs often remain rough and inconsistent. Between 10 and 20 epochs, some progress becomes visible: generated images start to resemble interiors influenced by the training data, though still lacking realism and coherence.



Fig. 50, 51. LoRA training results at 200–300 epochs.

As training continues into the 200 - 300 epoch range, the model occasionally produces realistic beds with smoother shapes and fewer wrinkles. However, a persistent challenge is that the model frequently introduces non-existent structures and artifacts, especially in fine details such as fabric textures and edges. Instead of refining details consistently, it can generate distorted patterns or phantom objects that were not present in the original images. This behavior reflects the complex nature of LoRA training, where the model struggles to perfectly replicate intricate textures and patterns while balancing generalization and overfitting. Despite these imperfections in raw training outputs, these images primarily serve as visual confirmation that the model is learning.



Fig. 52. LoRA training results after first run completed.

The true measure of success comes from applying the trained LoRA in practical workflows - using it in Stable Diffusion's image-to-image mode with the LoRA enabled to edit specific images, especially when I use it in inpaint mode. Since LoRA training is a long-term oriented project that demands significant computational resources and precise fine-tuning, it makes sense to approach training and refinement with a long-term perspective. When performing inpainting, I enable two ControlNets - depth and tiled blur - to guide the model in maintaining the original scene's structure and textures. The depth ControlNet provides spatial information about object distances and shapes, helping the model preserve three-dimensional consistency, while the tiled blur ControlNet applies localized blurring to smooth transitions and protect areas outside the masked object from unwanted changes. Together, these ControlNets ensure that only the targeted object (the bed) is altered, while the rest of the image remains intact and natural.

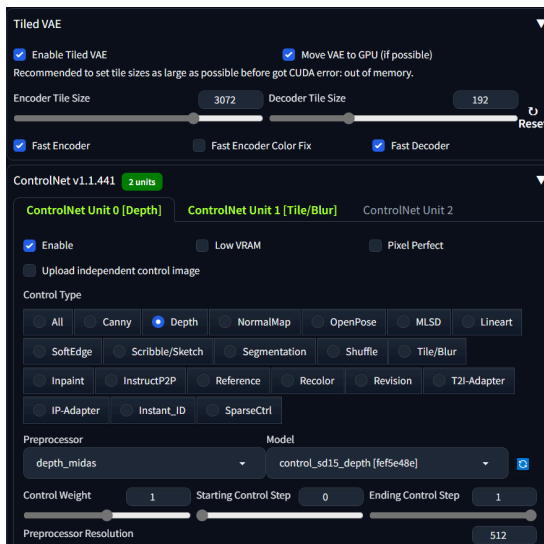


Fig. 53. Stable Diffusion's ControlNet 0 interface with settings and Depth condition enabled.

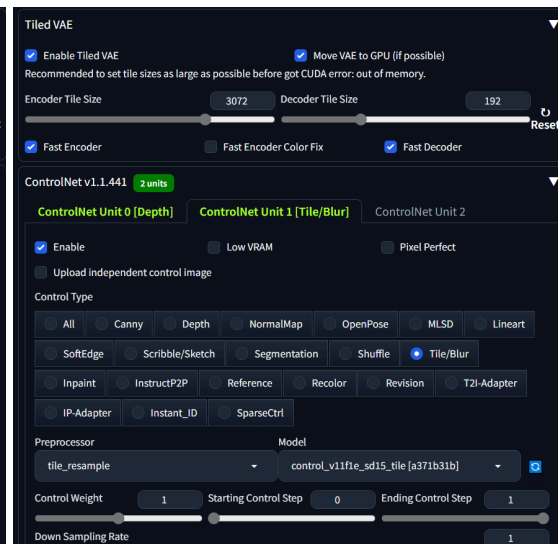


Fig. 54. Stable Diffusion's ControlNet 1 interface with settings and Tile/Blur condition enabled.

Applying image-to-image (img2img) with LoRA and inpainting at a denoising strength of around 0.35 has proven to be a good balance for my workflow so far. This approach allows targeted, controlled edits - like rearranging beds - while preserving the rest of the image. However, it is important to note that this process is not perfect and requires significant time and effort. Preparing a high-quality dataset, training the LoRA model, fine-tuning parameters, and setting up the workflow all demand patience and computational resources. Multiple training attempts and adjustments are often necessary to achieve satisfying results.

Despite these challenges, from a broader perspective, this method holds great promise for optimizing photographic workflows. It can drastically reduce manual retouching time while maintaining stylistic consistency and image integrity. With continued refinement and more efficient tooling, applying img2img combined with LoRA and inpainting could become a powerful standard approach in professional photo editing.

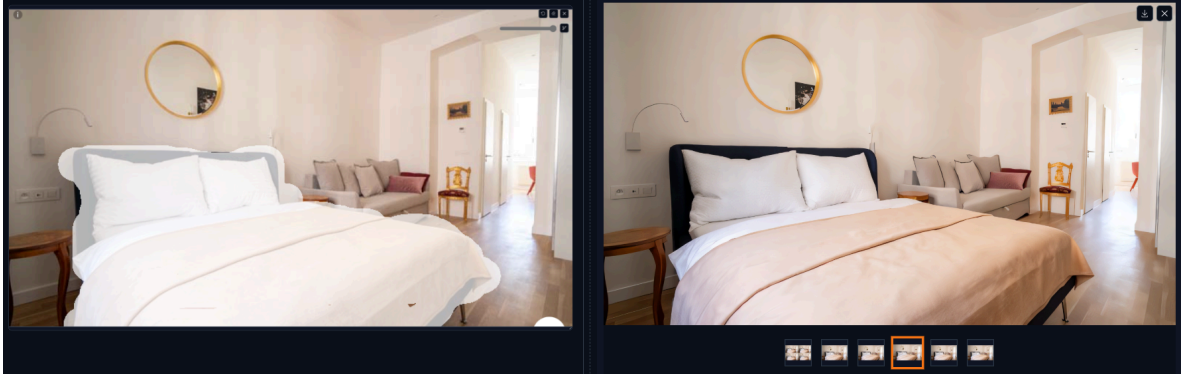


Fig. 55. Stable Diffusion Image-to-Image interface using trained LoRA with inpaint mask and before-and-after results.

For now, I have decided to continue developing my own LoRA model but lean towards using Flux for my workflow. Flux operates on a similar principle of targeted image editing but comes pre-trained on a much larger and diverse dataset, with its models already fine-tuned and optimized. This makes Flux significantly easier to use, as it requires far fewer preparation steps, less manual setup, and minimal training on my part.

Flux's user-friendly interface and powerful Flux.1 Kontext technology allow precise, instruction-based edits using natural language, enabling me to quickly and reliably make changes - such as rearranging beds - while keeping the rest of the image untouched. Its fast generation speed and support for multiple editing modes (text-to-image, image-to-image) further streamline the process.

Overall, Flux offers a practical, efficient solution that fits well with my current needs, saving time and effort compared to the more involved LoRA training and fine-tuning process, without sacrificing quality.

6 AI Upscaling with Freepik Magnific

Recently, I tested the Freepik Upscaler using its Magnific mode, which is one of the more effective tools I've tested so far, although it requires a subscription. This upscaler uses

deep learning algorithms to enhance image resolution by reconstructing details and refining textures rather than simply enlarging pixels.

Visually, Magnific produces images with sharp edges and soft, natural textures, avoiding common upscaling problems such as segmentation artifacts or unwanted texture overlays. The output maintains clarity and coherence, making it suitable for detailed photographic and illustrative content. One minor limitation I noticed is that the tool sometimes “continues” or generates very low-contrast structures - subtle patterns or slight generative artifacts that were not present in the original image. However, these imperfections can usually be handled by using the upscaled image as an overlay and masking out the unwanted areas during post-processing, similar to workflows I have applied with other upscaling tools.

Overall, Freepik’s Magnific mode provides a reliable option for improving image quality and resolution, complementing other AI-based editing techniques within photographic workflows.

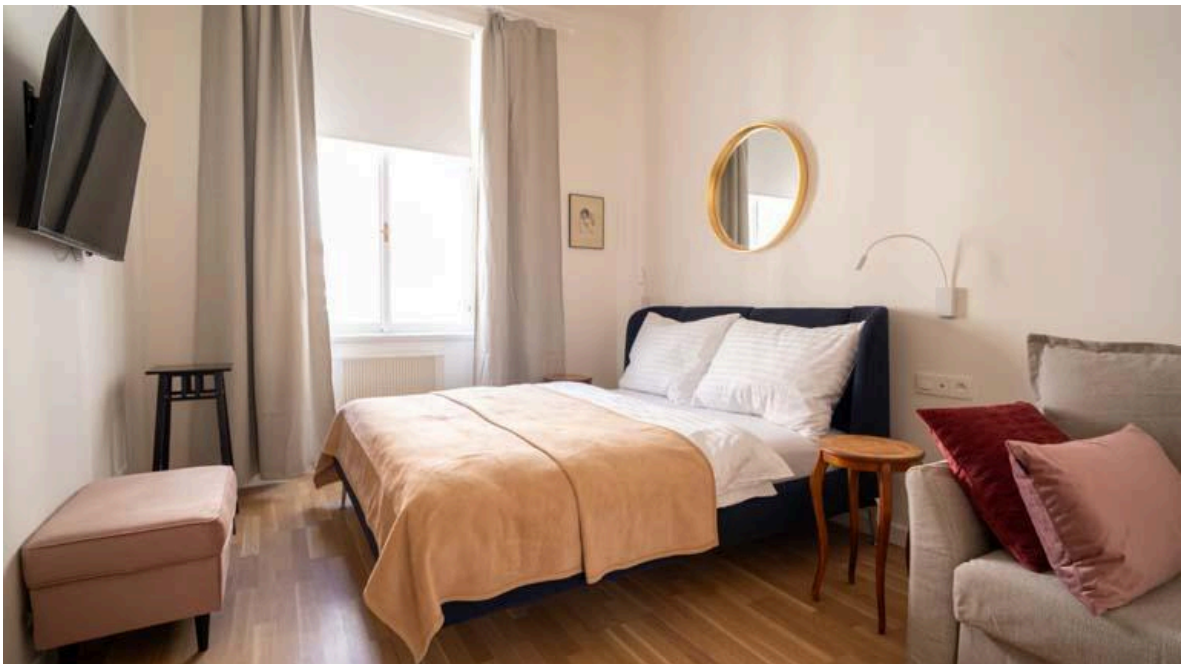


Fig. 56. Interior photo prepared for upscaling using Magnific AI by Freepik.



Fig. 57. Upscaling results after Magnific AI by Freepik.

7 Practical Conclusion: Efficiency Gains through AI Tools in Photography Workflow

In the course of this project, I explored various AI-based tools and their integration into the professional workflow of an interior photographer. The transition from purely manual processes to AI-assisted methods led to a significant reduction in time spent on repetitive or technical tasks, while maintaining - or even enhancing - image quality.

The following table summarizes key stages of the workflow, comparing traditional manual methods with their AI-enhanced alternatives. It highlights both the time savings achieved and the qualitative improvements gained through automation and intelligent processing:

See *Table 4 : Time efficiency comparison: manual vs. AI-assisted photo editing workflows.*

This comparison demonstrates that AI tools are not simply shortcuts, but reliable collaborators that improve consistency, streamline routine work, and enable photographers to focus more on creative and narrative decisions.

Editing Stage	Manual Process	AI / Tool Used	Manual Time	AI Time	Comment
Dust spot removal	Healing Brush in Photoshop	Luminar Neo / PS Gen Fill	3-5 min	10-15 sec / 3-5 min	Luminar faster; Firefly cleaner
Object removal (cables, stains)	Clone Stamp / manual masking	PS Remove Tool	~10 min	~2 min	Clean results, faster workflow
Background expansion / fill	Manual compositing	Firefly / Stable Diffusion	15-30 min	~3-5 min	Preserves perspective and tones
Complex object replacement	Masking + retouching	FLUX AI	40-60 min	~7 min	Fast replacement of beds, fabric
Texture/detail enhancement	Clarity / Sharpen manually	Luminar Structure AI	2-3 min	15-30 sec	Texture boost without artifacts
Light correction	Manual curves / gradients	Luminar Relight AI	3-5 min	~1 min	Foreground/background separation
Color correction	Manual Lightroom / PS	Luminar Enhance AI	5-7 min	~30 sec	One-slider global correction
HDR merging (30 images)	Lightroom manual merge	Luminar Neo Batch Merge	~2 hrs	~30 min	Major time reduction for batch merging

Editing Stage	Manual Process	AI / Tool Used	Manual Time	AI Time	Comment
Background generation	Manual compositing from libraries	Stable Diffusion / LoRA	30-60 min	5-10 min	Fast, stylistically accurate
AI personalization (style)	Not available manually	LoRA (Stable Diffusion)	—	15-60 min (training + tuning)	LoRA requires training + manual fine-tuning
Image upscaling	Photoshop interpolation	Upscayl / SD upscalers	~5 min	~2-3 min	Detail preserved, higher print resolution
Planning & pre-production	Manual checklists and moodboards	ChatGPT	30-60 min	5-10 min	Speeds up conceptual prep and structure

Table 4. Time efficiency comparison: manual vs. AI-assisted photo editing workflows.

The comparative analysis presented in Table 1 clearly demonstrates the substantial impact of AI integration on the interior photography workflow. By systematically replacing or augmenting traditional manual methods with advanced AI-powered tools - such as Luminar Neo, Photoshop's Generative Fill, Stable Diffusion, FLUX AI, and automated planning assistants - I achieved significant reductions in the time required for nearly every stage of post-production. Tasks that previously demanded extensive manual effort, such as object removal, background expansion, complex object replacement, and HDR merging, can now be completed in a fraction of the time, often with improved consistency and quality.

In summary, by integrating AI tools into my interior photography workflow, I save approximately four hours per project (for a typical set of tasks including a 30-image HDR batch) compared to traditional manual methods. This represents a reduction of more than 80% in editing and preparation time for many key tasks, while also improving consistency and creative flexibility.

Beyond accelerating routine operations, these AI tools contribute to a more streamlined and creative workflow. The automation of technical adjustments and repetitive edits allows the photographer to focus more on artistic decisions and other projects. Overall, the integration of AI has transformed the editing process from a series of labor-intensive steps into a much more efficient and dynamic system, highlighting the practical benefits and future potential of intelligent automation in professional photography.

8 Final Statement

While this thesis clearly shows the efficiency and creative benefits of using artificial intelligence in interior photography, it also offers a step-by-step example of how manual tasks can be gradually handed over to AI in a professional workflow. Many of the tools and methods described here - such as generative fill, object removal, and AI upscaling - are not limited to interiors. They are already being used in other areas of photography. For example, background expansion with Adobe Firefly is now common in advertising to create striking visuals; automated skin retouching is widely used in fashion and portrait photography; object removal tools help product photographers clean up images for e-commerce; and AI color grading is used in editorial shoots to quickly set the mood. These examples show that AI is becoming a practical tool across the entire industry.

Of course, artificial intelligence is not perfect. Sometimes it creates visual errors, struggles with complex textures, or misunderstands what we want. AI still cannot replace the eye and judgment of an experienced photographer. But the speed of improvement is remarkable, and right now, we are at the front line of this change. This means it is important to start using AI tools, learning their strengths and weaknesses, and setting good standards for their use.

There are also important ethical questions. In documentary, reportage, and photojournalism, the main goal is to show the truth and record real events. Using AI to change or invent content in these fields can damage trust and mislead viewers. Still, some AI tools - like those for restoring old photos, colorizing, or upscaling - can be helpful if used with care and honesty. The key is to fully understand what these tools do and use them responsibly.

In my view, professional photographers should treat artificial intelligence as a helpful assistant. AI can take over repetitive and technical tasks, letting us spend more time on creative work, other projects and even improve our work-life balance. As long as we stay in control, know the tools, and respect ethical limits, there is no reason to be afraid of this new technology. We cannot prevent everyone from misusing AI - like the recent case where an AI-generated image was entered into a wildlife photo contest - but we can set a good example. The more of us who use AI responsibly, the more likely others will follow.

In conclusion, we are living through a time of fast and dramatic change in visual production. Artificial intelligence is growing at a speed we have never seen before. Those who start using it now will have a real impact on how it is used in photography in the future. By adding AI to our workflows in a careful and thoughtful way, we do not just save time and improve quality - we also open up new creative possibilities. The work in this thesis shows how powerful AI can be, and I believe that with ongoing learning and a clear sense of responsibility, AI will become a valuable tool for photographers for many years ahead.

9 Personal Applications of AI-Assisted Post-Production in Interior Photography

Here, I share the examples from my interior and architecture photography projects to show how AI tools fit into my editing workflow. It highlights how AI-powered tools helped streamline the process, solve creative challenges, and produce polished, and visually appealing results that match the goals of each project.



Fig. 58.

Interior presentation for a short-term rental. Edited for online advertising purposes (Airbnb/Booking). Client: House Management Company, Prague, 2025.

Tools used: Luminar Neo - EraseAI (dust spot removal) , RelightAI (lighting balance), EnhanceAI (depth accentuation). Adobe Photoshop - automatic wires and cables removal, Generative Fill for object removal and rearrangement of bed and curtains.



Fig. 59.

Exterior presentation for a short-term rental listing. Client: House Management Company, Prague, 2024.

The goal was to emphasize the building's clean modernist aesthetic - aiming for a look reminiscent of architectural 3D renders.

Tools used: Adobe Photoshop - automatic wires and cables removal, people removal, Generative Fill for reconstructing hidden areas, street cleanup, adding trees and figures. Sky replacement and automatic lens distortion correction.



Fig. 60.

Interior photography for online listing platforms (Booking, Airbnb) and website. Private client, Prague, 2025.

The aim was to capture the atmosphere - a sense of soul, warmth, and spaciousness within a rustic and loft-style space.

Tools used: Luminar Neo - Erase AI for dust spot removal, Enhance AI for depth and visual emphasis, Relight AI for balanced lighting. Adobe Photoshop Generative Fill was used to clean the window view from construction cranes, refine soft object contours (pillows, curtains), and remove distracting electronic devices.



Fig. 61.

Editing project for an architectural firm from New Jersey, 2024. The task was to process a low-resolution Google Maps screenshot showing the building after renovation by the firm. One of the most challenging projects to date, due to low resolution and the presence of people, cars, trash, and trucks in the image. The photo was meticulously reconstructed using existing details. Tools used include wires and cables removal, people removal, sky replacement, and 48 layers of generative fill for complete scene reconstruction (removing a truck blocking part of the building, restoring windows, removing reflections and unwanted objects, cleaning the street from debris, adding trees, adjusting perspective and geometry), automatic perspective correction, remove tool, and luminance masks in Photoshop.



Fig. 62.

Exterior photograph of a residential complex for a House Management Company, Prague 2023 (re-edited in 2025). The objective was to create a universal cover image for online platforms such as Booking, Airbnb, the company website, and social media - usable across listings of various apartments within the complex. Tools used: Luminar Neo's Erase Tool to remove dust spots, Enhance AI to create the desired visual focus. In Photoshop: wires and cables removal, remove tool for eliminating blemishes. Luminance masks of varying intensity were applied to emphasize the monumental structure of the building.



Fig. 63

Interior photograph for the House Management Company, 2024.

The aim was to capture a clean, minimalist interior with understated tones and elegant textures in the spirit of contemporary minimalism. Intended for use across online platforms - Booking, Airbnb, the company's website, and social media. Tools used: Photoshop's Generative Fill to remove dust and unwanted elements; Remove Tool and Wires and Cables Removal to clean up distractions. Generative Fill was also used to enhance aesthetic appeal by subtly replacing unappealing items with visually cohesive ones, and to introduce context-appropriate objects - such as a laptop, books, or shelf decor - that suggest a professional atmosphere. All generative edits were created using Adobe Firefly.

Glossary of Terminology

Term	Definition
Artificial Intelligence (AI)	The simulation of human intelligence by machines, enabling learning, reasoning, and decision-making.
Batch Size	The number of training examples processed simultaneously in one iteration during AI model training.
Diffusion Model	A generative AI model that creates images by iteratively transforming random noise into structured visuals.
Generative Adversarial Network (GAN)	A machine learning framework with two neural networks competing to generate realistic synthetic images.
Latent Space	A compressed, abstract representation of data used by AI models to efficiently learn and generate new content.
Latent Diffusion Model (LDM)	A diffusion model operating in latent space to generate high-quality images efficiently.
Low-Rank Adaptation (LoRA)	A fine-tuning method that updates a small subset of parameters in large AI models to reduce computational cost.
Overfitting	When a model learns noise or details from training data too well, reducing its ability to generalize.
Prompt	Natural language input used to guide AI models to generate desired outputs, often

	refined iteratively.
Style Transfer	AI technique that applies the visual style of one image to another.
Text-to-Image Generation	AI models generating images based on textual descriptions.
Training Data	The labeled or unlabeled data used to teach AI models to recognize patterns or generate content.
Variational Autoencoder (VAE)	A generative AI model that learns compressed latent representations and can generate new data samples.
Fine-Tuning	Further training of a pretrained AI model on a specific dataset to improve performance on a task.
Generative Fill	AI-powered image editing that fills selected areas with contextually appropriate content.
Training Epoch	One complete pass through the entire training data during the training process of an AI model.
Unsupervised Learning	A type of machine learning where models learn patterns from unlabeled data without explicit instructions.

References

- Britannica. (n.d.). *Calotype*. Retrieved March 18, 2025, from <https://www.britannica.com/technology/calotype> [Accessed 25.06.2025]
- Cooke, A. (2025, April 25). The birth of the digital camera: From film to filmless revolution. *Fstoppers*. Retrieved March 18, 2025, from <https://fstoppers.com/historical/birth-digital-camera-film-filmless-revolution-699530> [Accessed 25.06.2025]
- Daguerreotype. (n.d.). The daguerreotype medium. *Library of Congress*. Retrieved from <https://www.loc.gov/collections/daguerreotypes/articles-and-essays/the-daguerreotype-medium#:~:text=The%20daguerreotype%20is%20a%20direct,surface%20looked%20like%20a%20mirror> [Accessed 25.06.2025]
- Dataversity. (n.d.). A brief history of machine learning. Retrieved March 28, 2025, from <https://www.dataversity.net/a-brief-history-of-machine-learning/> [Accessed 25.06.2025]
- Data Science Dojo. (2022, October 4). Machine learning 101: The types of ML explained. Retrieved March 28, 2025, from <https://datasciencedojo.com/blog/machine-learning-101/> [Accessed 25.06.2025]
- Digital Camera Museum. (n.d.). Kodak DC40 (1995). Retrieved March 24, 2025, from <https://www.digitalkameramuseum.de/en/cameras/item/kodak-dc40> [Accessed 25.06.2025]
- Digital Camera Museum. (n.d.). Videk Megaplug (1986). Retrieved March 19, 2025, from <https://www.digitalkameramuseum.de/en/cameras/item/videk-megaplug> [Accessed 25.06.2025]
- Epic Games. (n.d.). Unreal Engine and real-time rendering. Retrieved March 27, 2025, from <https://www.unrealengine.com/en-US/real-time-rendering> [Accessed 25.06.2025]
- Expert Photography. (n.d.). Exposure bracketing and HDR in interior photography. Retrieved April 18, 2025, from <https://expertphotography.com/bracketing-photography> [Accessed 25.06.2025]
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27. Retrieved March 28, 2025, from <https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf> [Accessed 25.06.2025]
- Hu, E. J., et al. (2021). LoRA: Low-rank adaptation of large language models. *arXiv preprint*. Retrieved March 29, 2025, from <https://arxiv.org/pdf/2106.09685.pdf> [Accessed 25.06.2025]

Levy, S. (1995, December 1). The movie that changed everything: Toy Story. *Wired*. Retrieved March 26, 2025, from <https://www.wired.com/1995/12/toy-story/> [Accessed 25.06.2025]

Lumas. (n.d.). History of photography. Retrieved February 2, 2025, from <https://www.lumas.com/history-photography/> [Accessed 25.06.2025]

Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding [PDF]. *arXiv*. Retrieved March 28, 2025, from <https://arxiv.org/pdf/2201.05989> [Accessed 25.06.2025]

Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Joint image diffusion for generative AI [PDF]. Retrieved March 29, 2025, from <https://research.nvidia.com/labs/dir/jedi/assets/joint-image-diffusion-combine.pdf> [Accessed 25.06.2025]

National Media Museum. (n.d.). Wet collodion negative (1851–1885). Retrieved March 18, 2025, from <https://www.nms.ac.uk/discover-catalogue/inventors-of-photography-exploring-victorian-photographic-techniques> [Accessed 25.06.2025]

Natural History Museum. (n.d.). Anna Atkins's cyanotypes: The first book of photographs. Retrieved March 18, 2025, from <https://www.nhm.ac.uk/discover/anna-atkins-cyanotypes-the-first-book-of-photographs.html> [Accessed 25.06.2025]

NVIDIA. (2023). NVIDIA and Getty Images collaborate to build AI-powered creative tools. Retrieved March 29, 2025, from <https://blogs.nvidia.com/blog/generative-ai-getty-images/> [Accessed 25.06.2025]

NVIDIA Research. (2023). Visual generative AI research highlights from CVPR 2021. Retrieved March 29, 2025, from <https://blogs.nvidia.com/blog/visual-generative-ai-cvpr-research/> [Accessed 25.06.2025]

OpenAI. (2023). GPT-4 technical report [PDF]. Retrieved March 30, 2025, from <https://arxiv.org/pdf/2303.08774> [Accessed 25.06.2025]

Open Culture. (2019, October 22). The first faked photograph (1840). Retrieved February 3, 2025, from <https://www.openculture.com/2019/10/the-first-faked-photograph-1840.html> [Accessed 25.06.2025]

Pixflow. (2025, March 10). How Toy Story changed animation history: Pixar's first CGI animated movie. Retrieved March 26, 2025, from

<https://pixflow.net/blog/how-toy-story-changed-animation-history-pixars-first-cgi-animated-movie/> [Accessed 25.06.2025]

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with CLIP latents. Retrieved March 30, 2025, from <https://arxiv.org/pdf/2204.06125> [Accessed 25.06.2025]

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. Retrieved March 30, 2025, from <https://arxiv.org/pdf/2112.10752> [Accessed 25.06.2025]

ScottBaker.ca. (2023). LoRA training (Stable Diffusion 1.5). Retrieved April 2, 2025, from <https://www.scottbaker.ca/AI/LoRA-Training> [Accessed 25.06.2025]

Stable Diffusion community contributions and advanced image editing functions. (2025, March 29). In *Hyperstack*. Retrieved March 29, 2025, from <https://www.hyperstack.cloud/blog/case-study/everything-you-need-to-know-about-stable-diffusion> [Accessed 25.06.2025]

Stephens, J. (2022, May 12). Getting started with NVIDIA Instant NeRFs. *NVIDIA Technical Blog*. Retrieved March 28, 2025, from <https://developer.nvidia.com/blog/getting-started-with-nvidia-instant-nerfs/> [Accessed 25.06.2025]

TechTarget. (n.d.). What is CGI (Computer-Generated Imagery)? Retrieved March 26, 2025, from <https://www.techtarget.com/whatis/definition/CGI-computer-generated-imagery> [Accessed 25.06.2025]

The Ohio State University. (n.d.). CG historical timeline – Computer graphics and computer animation. Retrieved March 26, 2025, from <https://ohiostate.pressbooks.pub/graphicshistory/back-matter/cg-historical-timeline/> [Accessed 25.06.2025]

Wonderful Machine. (2023). Specialty: AI/CGI photography. Retrieved March 28, 2025, from <https://wonderfulmachine.com/article/specialty-ai-cgi-photography/> [Accessed 25.06.2025]

Yan, X., Lei, Z., Wu, L., Li, S., & Zeng, X. (2023). ControlNet: Adding conditional control to text-to-image diffusion models. *arXiv*. Retrieved March 29, 2025, from <https://arxiv.org/pdf/2302.05543.pdf> [Accessed 25.06.2025]

Illyasviel. (2023). ControlNet GitHub repository. Retrieved March 30, 2025, from <https://github.com/Illyasviel/ControlNet> [Accessed 25.06.2025]

List of Figures

Fig. 1: Camera Obscura. Retrieved February 2, 2025, from

<https://www.matrise.no/2018/07/camera-obscura-and-the-world-of-illusions/>

Fig. 2: Boulevard du Temple, daguerreotype by Louis Daguerre, 1838. Retrieved February 2, 2025, from

https://en.wikipedia.org/wiki/Boulevard_du_Temple_%28photograph%29#/media/File:Boulevard_du_Temple_by_Daguerre.jpg

Fig. 3. *Self-Portrait as a Drowned Man*, photograph by Hippolyte Bayard, 1840. Retrieved February 3, 2025, from

https://en.wikipedia.org/wiki/Hippolyte_Bayard#/media/File:Hippolyte_Bayard_-_Drownedman_1840.jpg

Fig. 4. Cyanotype photograph of British algae from Anna Atkins's book *Photographs of British Algae: Cyanotype Impressions* (1843). Retrieved March 18, 2025, from

<https://www.nhm.ac.uk/discover/anna-atkins-cyanotypes-the-first-book-of-photographs.html>

Fig. 5. Latent space noise reduction guided by textual conditioning to generate the final image. Retrieved June 25, 2025, from

https://www.researchgate.net/figure/An-illustration-of-stable-diffusion-working-with-an-example_fig2_383326974

Fig. 6. Image decoded from a latent diffusion model. Retrieved June 25, 2025, from

<https://victordibia.com/blog/stable-diffusion-denoising/>

Fig. 7. Adobe Lightroom interface showing catalog and image organization. (screenshot by author)

Fig. 8. HDR Merge interface in Adobe Lightroom, showing exposure alignment and merge options. (screenshot by author)

Fig. 9. HDR image after merging in Lightroom, illustrating tone correction and editing limits. (screenshot by author)

Fig. 10. Luminar Neo interface and image catalog view. (screenshot by author)

Fig. 11. Luminar Neo Batch HDR Merge feature in the main workspace. (screenshot by author)

Fig. 12. Luminar Neo Batch HDR settings showing image exclusion options. (screenshot by author)

Fig. 13. Preview window of Batch HDR Merge process in Luminar Neo. (screenshot by author)

Fig. 14. Luminar Neo saving merged HDR image as TIFF in a separate folder. (screenshot by author)

Fig. 15. Luminar Neo – HDR Merge preview without adjustments. (screenshot by author)

Fig. 16. Adobe Lightroom – HDR Merge preview without adjustments. (screenshot by author)

Fig. 17. Luminar Neo – Data limit test with -100 highlights and +100 shadows. (screenshot by author)

Fig. 18. Adobe Lightroom – Data limit test with -100 highlights and +100 shadows. (screenshot by author)

Fig. 19. Curves adjustment layer as dust spot revealing overlay. (screenshot by author)

Fig. 20. Curves adjustment layer application. (screenshot by author)

Fig. 21. Demonstration of dust spots removal techniques in Luminar Neo and Photoshop. (screenshot by author)

Fig. 22. Demonstration of dust spots removal results using Luminar Neo and Photoshop's Generative Fill. (screenshot by author)

Fig. 23. Object removal process using Generative Fill in Photoshop. (screenshot by author)

Fig. 24. Automatic People Removal feature demonstration with automatic detection and results. (screenshot by author)

Fig. 25. Automatic wires and cables removal feature in Photoshop with auto detection and clean results demonstrated on stock image from pexels.com (screenshot by author)

Fig. 26. Collage that demonstrates background expansion feature step-by-step. (collage by author)

Fig. 27. Demonstration of the object replacement process using Generative Fill in Photoshop. (screenshot by author)

Fig. 28. Before-and-after demonstration of the Enhance AI tool in Luminar Neo. (collage by author)

Fig. 29. Before-and-after demonstration of the Sky Replacement AI tool in Luminar Neo. (collage by author)

Fig. 30. Before-and-after demonstration of the Noiseless AI tool in Luminar Neo. (collage by author)

Fig. 31. Before-and-after demonstration of the Structure AI tool in Luminar Neo. (collage by author)

Fig. 32. Before-and-after demonstration of the Relight AI tool in Luminar Neo. (collage by author)

Fig. 33. Flux.1 Kontext model announcement banner from Black Forest Lab official website. (screenshot by author using Black Forest Labs Playground (Flux))

Fig. 34. Merged and resized HDR image prepared for test run in FLUX. (screenshot by author using Black Forest Labs Playground (Flux))

Fig. 35. FLUX Playground interface displaying test image and text prompt. (screenshot by author using Black Forest Labs Playground (Flux))

Fig. 36. First image result generated using FLUX.1 Kontext that demonstrates targeted object rearrangement. (screenshot by author)

Fig. 37. Before-and-after comparison of an image processed with FLUX.1 Kontext. (collage by author)

Fig. 38. Interface of the Upscayl AI upscaling tool. (screenshot by author using Upscayl AI software)

Fig. 39. First result after upscaling using Upscayl AI. (photo by author)

Fig. 40. Comparison demonstrating before-and-after results in detail. (collage by author)

Fig. 41. Original image with a layer mask revealing the bed from the upscaled version. (screenshot by author)

Fig. 42. Layer management and mask overlay. (screenshot by author)

Fig. 43,44. LoRA training first results at epochs 15–20. (screenshot by author)

Fig. 45. Dataset prepared for LoRA training. (screenshot by author)

Fig. 46. Image Processor interface in Photoshop with settings for automated batch resizing. (screenshot by author using Adobe Photoshop)

Fig. 47. Kohya_SS trainer interface. (screenshot by author using Kohya_ss trainer)

Fig. 48. Kohya SS caption text for automated image tagging. (screenshot by author using Kohya_ss trainer)

Fig. 49. LoRA training log showcasing training status and key parameters. (screenshot by author)

Fig. 50, 51. LoRA training results at 200–300 epochs. (screenshot by author using Stable Diffusion)

Fig. 52. LoRA training results after first run completed. (screenshot by author using Stable Diffusion)

Fig. 53. Stable Diffusion's ControlNet 0 interface with settings and Depth condition enabled. (screenshot by author using Stable Diffusion)

Fig. 54. Stable Diffusion's ControlNet 1 interface with settings and Tile/Blur condition enabled. (screenshot by author using Stable Diffusion)

Fig. 55. Stable Diffusion Image-to-Image interface using trained LoRA with inpaint mask and before-and-after results. (screenshot by author using Stable Diffusion)

Fig. 56. Interior photo prepared for upscaling using Magnific AI by Freepik. (screenshot by author)

Fig. 57. Upscaling results after Magnific AI by Freepik. (photo by author)

Fig. 58. Interior presentation for a short-term rental. (photo by author)

Fig. 59. Exterior presentation for a short-term rental listing. Client: House Management Company, Prague, 2024. (photo by author)

Fig. 60. Interior photography for online listing platforms (Booking, Airbnb) and website. Private client, Prague, 2025. (photo by author)

Fig. 61. Editing project for an architectural firm from New Jersey, 2024. (photo provided by client via freelance platform; edited by author)

Fig. 62. Exterior photograph of a residential complex for a House Management Company, Prague 2023 (re-edited in 2025). (photo by author)

Fig. 63 Interior photograph for the House Management Company, 2024. (photo by author)

List of Tables

Table 1: - Approximate time of editing one interior photograph using the manual approach with minimal AI assistance.

Table 2. Key differences between Luminar Neo and Adobe Lightroom.

Table 3. LoRA training configuration and key parameters.

Table 4. Time efficiency comparison: manual vs. AI-assisted photo editing workflows.

Attachment no 1

Use of technical means in the preparation of the final thesis (AI)

Use	Description	Tool
Research - searching for sources	Used exclusively for locating relevant, publicly available sources	perplexity.ai (deep research mode)
Resource processing	Used AI for summarizing technical articles and simplifying complex English-language sources. The processed content helped me develop a framework for writing the technically focused sections of the thesis.	perplexity.ai ChatGPT
Work on the text done by AI	Was used for creating a personalized writing plan based on requirements, rephrasing complex sentences for clarity, verifying sources that are relevant, correcting lexical and grammatical errors, and suggesting more precise definitions where needed	perplexity.ai Grammarly
Work with data done by AI	AI was used for assisting in calculating the efficiency achieved through AI-assisted post-production compared to traditional editing methods. The analysis was based on my personal workflow data	ChatGPT
Visual output	Used in personal photography projects for post-production. Results from this work are included in the thesis, and the tools I also used to demonstrate their capabilities within the context of the thesis itself. Applications included object replacement, background extension, and upscaling.	Stable Diffusion, Adobe Photoshop (powered by Firefly), Flux AI, Luminar Neo, Upscayl AI