Ayush Majumdar

Individualized Real Time Automation of Color Grading and Editing Using CNNs (Convolutional Neural Networks), and/or GANs (Generative Adversarial Networks) WITHIN Mirrorless Cameras based on user history and stylistic choices, while maintaining photographic and artistic integrity.

Abstract

Our research explores the integration of Artificial Intelligence (AI) within mirrorless cameras to automate color grading and editing processes, utilizing Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs). The project aims to harness AI's potential to analyze and learn from a photographer's historical stylistic choices, enabling personalized image enhancement directly in-camera. Current DSLR and mirrorless cameras already leverage AI for various functions such as autofocus, noise reduction, and image stabilization. Building on these applications, my project seeks to add a layer of creative automation by training models on a photographer's edited images, thereby recognizing and applying their unique stylistic traits to new photos. Our approach aims to streamline the editing workflow, offer inspiration, and maintain artistic individuality, addressing concerns over the potential homogenization of creative expression in photography. By embedding this technology directly into camera hardware, it provides an innovative tool for photographers to instantly visualize and apply their preferred aesthetics, challenging the traditional post-processing workflow. Our research not only advances the technical capabilities of camera systems but also explores the balance between technological innovation and the preservation of artistic integrity in this age of photography.

Ayush Majumdar

Technical Implementation of CNNs and GANs in Automated Color Grading for Mirrorless Cameras

Overview

The technical foundation for embedding real-time AI-driven color grading and editing directly in mirrorless cameras centers on the effective use of Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs). These AI methodologies enable the system to analyze photographic elements, understand stylistic nuances, and apply personalized enhancements based on historical data of a photographer's previous edits. Implementing these technologies within the confined hardware of a camera presents unique challenges that involve resource management, model efficiency, and real-time processing capabilities.

Convolutional Neural Networks (CNNs)

CNNs are pivotal for feature detection within images, making them ideal for understanding compositional elements and style characteristics in photography. In our proposed system, CNNs will serve the primary role of analyzing incoming images to detect patterns, textures, and other elements that define a photographer's style. Key considerations for CNN implementation in cameras include:

Model Size and Efficiency

Given the limited processing power and memory in cameras, the CNN architectures need to be highly optimized for speed and size. Techniques such as pruning, quantization, and knowledge distillation could be employed to compress the model without significant loss in accuracy.

Edge Computing

Processing must be done locally (on-device) to ensure that style applications are real-time and do not require cloud dependency. This involves leveraging the camera's GPU (if available) or dedicated image processing chips.

Generative Adversarial Networks

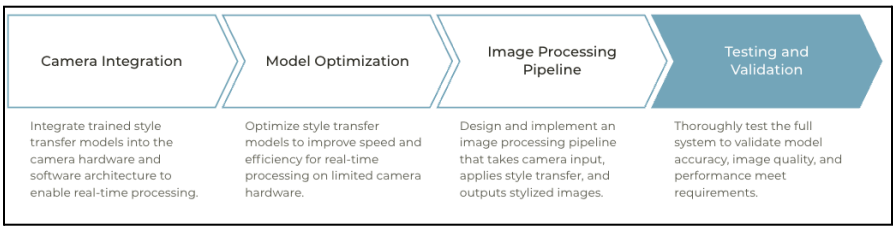
GANs in this context are used to generate or modify images in a way that reflects a photographer's unique style based on learned characteristics from their image editing history. The GAN architecture would typically consist of two main components:

Generator: Creates images or image adjustments that aim to mimic the photographer's style.
Discriminator: Assesses whether the generated images fit the style of the original edits, providing feedback to the generator.

For in-camera integration:

Real-Time Processing: The GANs must be optimized for low latency to function in real time. This could involve simplifying the generator and discriminator networks and employing faster convergence techniques.

Style Application: The discriminator's feedback is crucial for adjusting the generator's output, ensuring that the final image respects the artistic integrity and uniqueness of each photographer's style.



Hardware Considerations

GPU and Image Processing Chips: Modern mirrorless cameras are equipped with powerful GPUs and dedicated image processing chips that can handle sophisticated algorithms. The AI models would need to be tailored to exploit these hardware capabilities fully.

Memory and Storage: Efficient memory management is crucial, as storing and processing large neural networks can be resource-intensive. Techniques such as model partitioning could be used, where only parts of the model needed for immediate tasks are loaded into memory.

Connectivity and Cloud Integration

While the primary processing would be done on-device, optional cloud connectivity could be offered for more intensive processing tasks or for updates and improvements to the AI models:

Cloud-Assisted Processing: For exceptionally complex edits or when new styles are being learned, the camera could offload tasks to the cloud, provided there is connectivity available.

Continuous Learning and Updates: The models can be periodically updated via the cloud with new data or improvements in AI algorithms to enhance performance and adapt to evolving photographic styles.

Integrating CNNs and GANs into mirrorless cameras for automated, real-time photo editing is a promising advancement that blends technical innovation with artistic expression. By carefully balancing the demands of real-time processing with the creative nuances of photography, this technology can offer photographers a powerful tool to enhance their workflow while preserving the individuality and artistic integrity that define their work. This approach not only pushes the boundaries of what's technically feasible in photography but also respects the deeply personal and expressive nature of the medium.

Ethics and Market

To edit my own images, I use Adobe Lightroom. Within Adobe Lightroom, there are amazing cc I recognize that this is an INCREDIBLE feat of technology, but not one without key flaws.

1) **Destruction of Creative Aspect in an Art-Based Interest/Industry.** It is also incredibly useful for already experienced photographers with a creative blockage and need for inspiration, or at time constraint. This type of tool seems useful for beginners who do not have a great idea about how to work with every individual slider available in Lightroom, Photoshop, or etc, but in reality provides shortcuts for newcomers to photography to bypass the learning curve and more importantly, FAIL to find their own style. Photography, at the end of the day, is an art because of its INDIVIDUALITY.

By no means do I believe that photography is not also an influential and historically impacted field (if anything, it is all that), but I am a strong believer that the EXACT work of other photographers should not be imprinted on top of a photographer's art. It not only kills the individuality and expression of the piece, but also does not allow the photograph to be its best version, as no two pictures use identical slider statistics to best self-present.

Let me move on from my rant about AI misusage in photography.

I do believe there is a huge potential for AI in the industry without killing it.

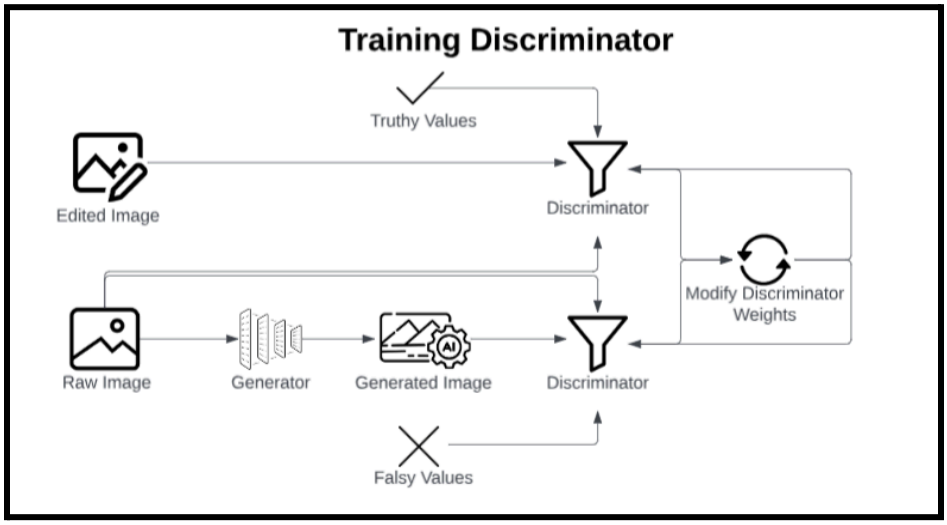
Similarly to what I said before regarding Adobe Lightroom's AI tools, I believe there are similar methods of AI usage that still protect the novelty of photography. Image data in RAW or JPEG format not only have their original code written into them, but also any modifications made to the image made upon user performed edits. Essentially, the files contain both the original image and the impacted picture as well. On top of this, it is not common practice for a photographer to store the impacted picture in the SD card that is stored in the camera.

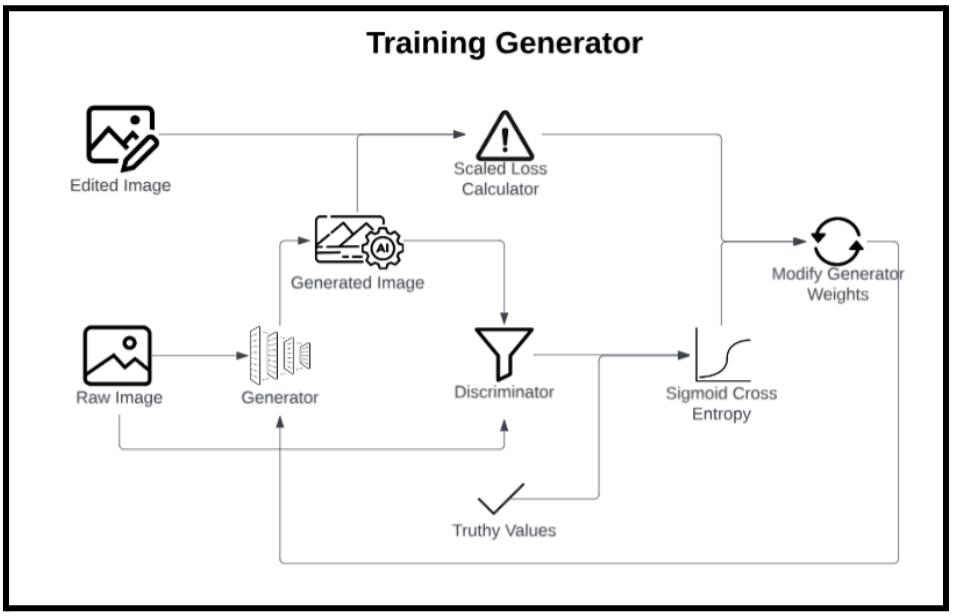
My idea is to train all of a photographer's impacted pictures (or those that the photographer chooses) using a CNN and GAN model to identify and recognize the photographer's stylistic traits. The trained model with all of the stored impacted images will be helped within the camera, and this can be used as a tool WITHIN the camera itself.

When taking an image, the photographer has the option to modify the image based on his/her past stylistic decisions or not to. The photographer ideally would have the ability to view the potential impacted image on the LCD screen prior to capture, in order to see if the style should be maintained or not for the specific situation. The model would only intake pictures similar enough to the current situation to identify the style for the image that the photographer would be most interested in.

One KEY factor about this tool, is that this tool is not only available on Lightroom and other mobile/computer softwares, but an in-built mirrorless camera feature.

Discriminator and Generator

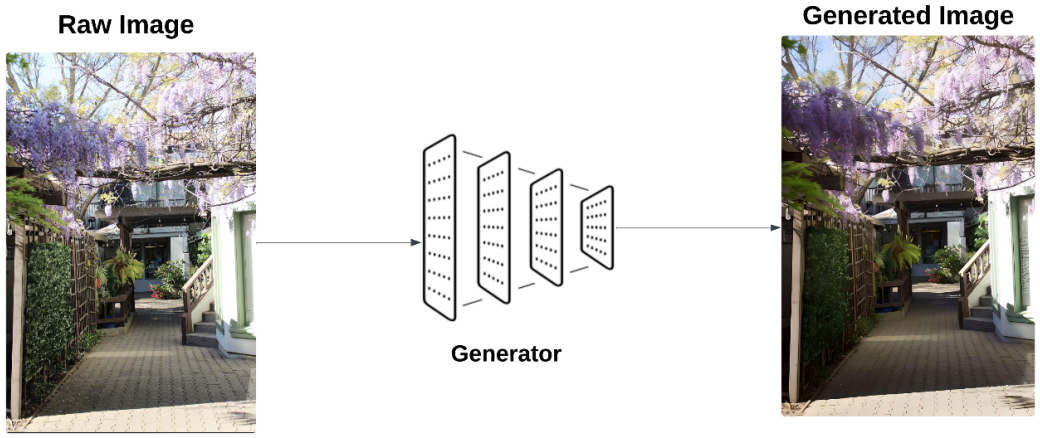


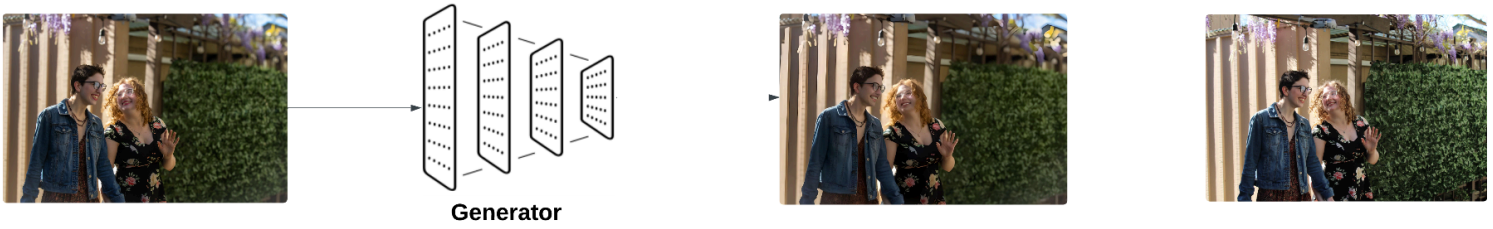


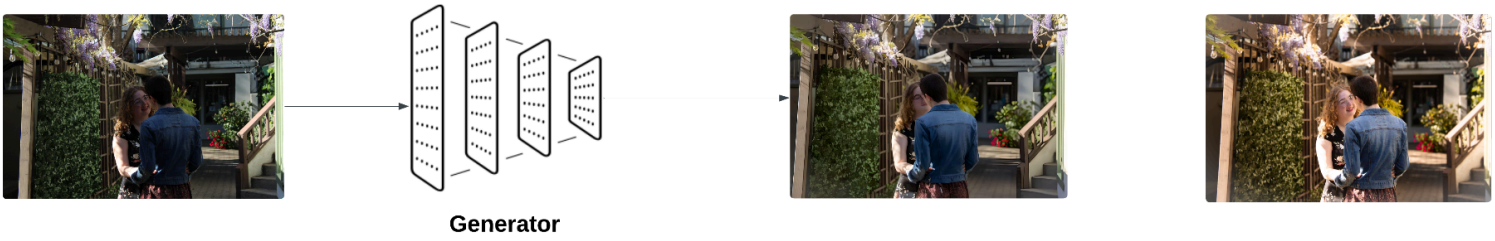
AI and Deep Learning Tools Used in Current DSLR and Mirrorless Cameras

- **HARDWARE**
  - Image Sensors
    - Image Processing and Enhancement
    - Noise Reduction
      - In newer mirrorless cameras, AI is used to even reduce the noise in low light situations.
    - Improve AutoFocus and even selectivity in Manual Focus
      - Recognizes the subject at extremely high speeds without user input, but still with user maneuverability to CHANGE the subject of focus involved by the user.
    - HDR Processing
      - Scanning environmental lighting and choosing the perfect exposure using base level artificial intelligence - aligns with shutter priority and ISO manufacturing, also in auto modes.
    - Computational Photography
      - Digital Zoom, bokeh backgrounds for blurred backgrounds, fill flash.
    - Video
      - Auto adjustments in footage stabilization, frame rate adjustments given a situation, and resolution enhancement.
    - Image Sensor Optimization
      - Dynamically adjust the sensor's sensitivity, light capture speed, capture light methods.

Examples


Raw Image
Generated Image
Ground Truth
Generator


Generator


Generator

References:

Majumdar, A. (2024). *Ayush Majumdar Photography*. Retrieved April 7, 2024, from https://ayushmajumdarphotography.mypixieset.com/

TensorFlow. (n.d.). *Pix2Pix*. Retrieved April 7, 2024, from https://www.tensorflow.org/tutorials/generative/pix2pix