AI Applied to Upscaling

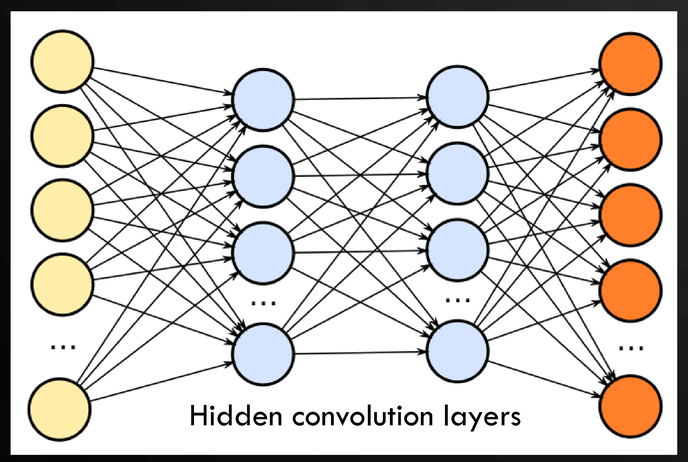
Omar Iqbal
Jacob Turner
Mohammed Alqarni

Introduction

Media upscaling is the act of taking low-resolution media and increasing the resolution while also trying to preserve as much image quality as possible. Media upscaling is important because it preserves older media such as 20th century films, old photographs, and retro video games. Without upscaling, old media played on modern-day TVs and monitors will have low clarity. Artificial intelligence (AI) can drastically increase the efficiency of upscaling and preserve the media in a fraction of the time that manual upscaling would take. AI upscaling is not perfect and often needs human intervention when desired outputs are not met. Different types of media require different algorithms, so there is no universal AI that can accurately upscale everything. We seek to understand the varying kinds of upscaling that utilizes AI.

How AI Learns

Convolutional neural networks (CNN) are used to detect patterns in imagery. Images are inserted into the input layer, which are then sent into hidden convolution layers that perform pattern detection using filters. A filter is a matrix of random initialized values that pass over every pixel of the image. These filters can discern shapes, textures, edges, and more. As the image is passed through more and more convolutional layers, the filters become more advanced to the point where they can recognize more complex structures like cars or animals. The images are then passed to the output layer once the AI decides how to classify the images.



Common AI Upscaling Methods

SRGAN

Also known as a super-resolution generative adversarial network, SRGAN is a single-image upscaling method that addresses the main problem of CNN, which is the loss of finer details as resolution is increased. SRGAN uses a generative adversarial network (GAN) which is composed of two neural networks, the generative network and the discriminative network. The generative network creates samples for the discriminative network to analyze. The objective for the discriminative network is to accurately discern whether they are upscaled images (fake) or naturally photo-realistic images (real). Training in this manner allows the AI to more accurately recover finer details during the upscaling process.

ESRGAN

ESRGAN is an enhanced version of SRGAN. It aims to eliminate upscaling artifacts. ESRGAN was created by improving on the perceptual loss function of SRGAN. The perceptual loss function has two main parts: adversarial loss and content loss. The former is responsible for ensuring images look natural, while the latter ensures the upscaled image still has features similar to the original image. ESRGAN also shifts the discriminator network's focus towards discerning which image is "more real" than the other, rather than deciding if an image is real or fake as with SRGAN. Images using ESRGAN have consistently higher visual fidelity than images upscaled with SRGAN.

Before

After

A before and after comparison of a screenshot that's been upscaled to 4x the original resolution by using an AI software called Gigapixel. The screenshot is from a video game called Final Fantasy VII which was released in 1997.

Traditional methods of upscaling such as nearest neighbor interpolation involve duplicating or replacing pixels as the image is expanded. This causes artifacts like jagged lines and blurriness. AI upscaling can avoid these issues.

Conclusion & Future of Upscaling

While not perfect, AI upscaling drastically outperforms traditional upscaling methods. ESRGAN, for example, picks out fine details that would otherwise be lost when expanding the image size. As AI is exposed to greater variations of training samples, it becomes faster and more accurate. Companies like Samsung and Sony are already beginning to incorporate AI in television sets to improve picture quality and resolution dynamically. Fans of old video games have used AI to improve visual quality to levels that have never been seen before. As older technology becomes exceedingly obsolete, AI will likely play an essential role in media preservation.

References:

[deeplizard]. (2017, December 9). Convolutional Neural Networks (CNNs) explained [Video File]. Retrieved from https://youtu.be/YRhxdVk_sIs
CaptRobau. (2019, May 22). Remako Mod Version 1.0 Released! Retrieved March 25, 2020, from https://captrobau.blogspot.com/2019/05/remako-mod-version-10-released.html
Wang, Yu, Ke, Wu, Gu, Liu, … Xiaoou. (2018, September 17). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. Retrieved March 28, 2020, from https://arxiv.org/abs/1809.00219
Deshpande, A. (n.d.). Deep Learning Research Review Week 1: Generative Adversarial Nets. Retrieved March 28, 2020, from https://adeshpande3.github.io/Deep-Learning-Research-Review-Week-1-Generative-Adversarial-Nets
Sony Computer Entertainment. (1997). Final Fantasy VII [PlayStation video game]. Japan: Square.
Ledig, Theis, Lucas, Huszar, Ferenc, Caballero, … Wenzhe. (2017, May 25). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. Retrieved March 28, 2020, from https://arxiv.org/abs/1609.04802
Goodfellow, I. J., Pouget-Abadie†, J., Mirza, M., Xu, B., Warde-Farley, D., Orzair, S., … Bengio, Y. (n.d.). Generative Adversarial Nets. Retrieved from https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf
Hicks, M. (2019, August 26). Here's the secret behind 8K AI upscaling technology. Retrieved March 28, 2020, from https://www.techradar.com/news/heres-the-secret-behind-8k-ai-upscaling-technology