IMAGE ENHANCEMENT & SUPER-RESOLUTION UPSCALING

Applied Linear Algebra For Machine Learning

GROUP 5

PRATHAM HARSHVARDHAN DAVE - 20BCS102

SOMISETTY SAI PRANEETH - 20BCS125

SPARSH JAIN - 20BCS127

SRI HARI L - 20BCS129

EXECUTIVE SUMMARY

PROBLEM STATEMENT

The current image upscaling solutions are either purely algorithmic or only apply neural networks. These cause the resultant upscaled images to lack resolution or appear distorted, due to the addition of details from the neural network model. Moreover, neural networks can only be trained to upscale images to a certain factor, despite its ability to take inputs of varying resolution. To mitigate this issue, we need to come up with a method that takes into consideration both the formula based and neural network approach and merges them into a single solution.

PROPOSED SOLUTION

In our approach, we combine the benefits of image interpolation and neural network-based image upscaling to tackle the issue of having fixed upscaling factor as well as reduce the degree of distortion caused if a standalone neural network model is used. More specifically, instead of using basic interpolation methods such as Bi-Cubic, Nearest Neighbor etc. we apply Lanczos interpolation to retain as much information as possible from the original image, after which a Convolutional Neural Network is employed to upscale the image by a fixed factor of 4. This Convolutional Neural Network is trained as a component of a Generative Adverserial Network, wherein the Generator CNN is employed as the upscaler, while the Discriminator is discarded. We devise a custom loss function wherein we also take into consideration the “content loss” which is measured using a pretrained VGG-19 model and is weighted with a specific value.

VALUE

Our proposed solution can provide better results than the novel methods. While it is not possible to mathematically determine the “superiority” or the level of detail from one image to the other (attempts have been made in the past using factors like PSNR), we have noticed a significant improvement in the perceived level of detail in the images.

Upscaling plays a huge role in the image restoration and video upscaling technologies wherein large technology companies employ such methods to ‘restore’ classical images and videos from before early 2000s for viewership on media streaming platforms. High performance computing has enabled real-time upscaling, which is now widely used in the gaming industry to boost the performance of videogames.

FINAL THOUGHTS & NEXT STEPS

Image Upscaling is a field in which a lot of work can still be done as to create a generalized model that can work in every scenario, including but not limited to animated videos, recorded videos and Space Exploration Videos captured from Satellites and Space Imaging devices.

INTRODUCTION

Image upscaling is the process of raising the resolution of an input image to generate a higher-quality output image. In recent years, machine learning (ML) approaches have been used for image upscaling instead of traditional linear methods such as nearest neighbor, bilinear, bicubic, Lanczos etc. interpolation techniques. This is because ML approaches produce more accurate and visually appealing results than classic interpolation methods.

Although traditional interpolation methods are still useful, their outcomes can be improved by feeding them as inputs to ML methods for image upscaling. This hybrid technique enables the ML model to learn from the outputs of the interpolation methods, which serve as a starting point for generating higher-quality images.

We propose a novel approach to image upscaling in this paper that combines the advantages of feed-forward convolutional neural networks and perceptual loss functions. Our method trains the networks with perceptual loss functions, which measure picture similarities more reliably than per-pixel losses and produce high-quality images. We demonstrate that our approach outperforms standard interpolation approaches and produces state-of-the-art results in real-time.

LINEAR-UPSCALING METHODS

2.1 NEAREST NEIGHBOR INTERPOLATION

Nearest neighbor interpolation is a basic and fast method for image upscaling. It assigns the value of the closest pixel in the input image to each pixel in the output image. Mathematically, it can be expressed as

$$\boxed{f(x, y) = f(\mathrm{floor}(x + 0.5), \mathrm{floor}(y + 0.5))}$$

2.2 BILINEAR INTERPOLATION

Bilinear interpolation considers the values of the four nearest neighboring pixels to compute the new pixel value. It interpolates linearly between these values to produce a smoother image. The mathematical expression for bilinear interpolation is

$$\boxed{f(x, y) = (1 - u)(1 - v)f(i, j) + u(1 - v)f(i + 1, j) + (1 - u)vf(i, j + 1) + uvf(i + 1, j + 1)}$$

where u and v are the fractional parts of x and y, respectively, and i and j are the integer parts.

2.3 BICUBIC INTERPOLATION

Bicubic interpolation is a higher-order method that uses a weighted average of 16 nearest neighboring pixels to compute the output pixel value. It produces sharper images than bilinear interpolation. The mathematical expression for bicubic interpolation is more complex, involving a set of coefficients that depend on the distance of the pixel from the center.

2.4 LANCZOS INTERPOLATION

Lanczos interpolation is a more sophisticated method that uses a windowed sinc function to compute the output pixel value. The sinc function is a mathematical function that is commonly used in signal processing. Lanczos interpolation produces sharper images than bicubic interpolation, but it is also slower. The mathematical expression for Lanczos interpolation is

$$\boxed{f(x, y) = \Sigma i=0,1,...,k-1\ \Sigma j=0,1,...,k-1\ w(i/k)w(j/k)f(x + i - (k-1)/2, y + j - (k-1)/2)}$$

where w(t) is the Lanczos windowed sinc function.

Image 1.1 Original VS Interpolated from our code. (50% Lanczos + 50% Bicubic)

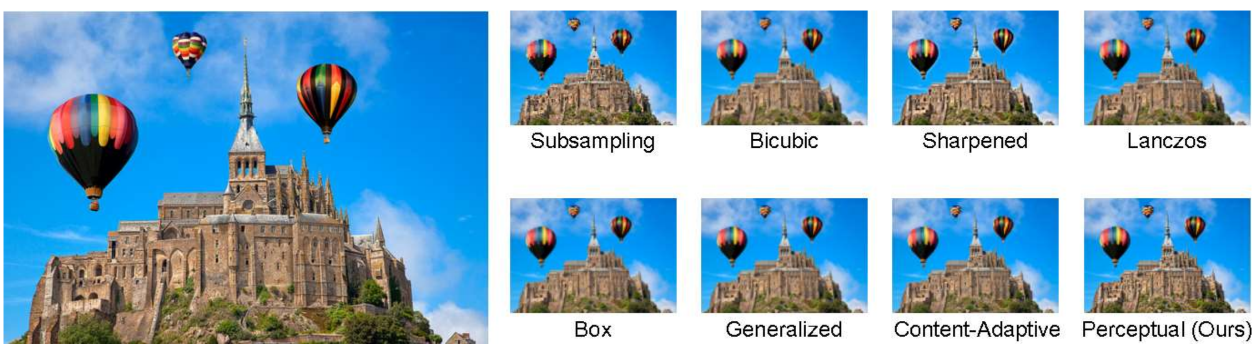

Image 1.2 – A Comparison of Different types of Algorithmic Upsampling methods.

[Figure 1 from Content-Adaptive Perceptual Bicubic Sharpened Lanczos Subsampling Box Generalized | Semantic Scholar]

MACHINE-LEARNING UPSCALING METHODS

3.1 EDSR

EDSR, which stands for Enhanced Deep Residual Networks, is a deep learning-based image upscaling method that has gained popularity in recent years for its ability to produce high-quality images with fine details and textures.

EDSR uses a deep neural network architecture with residual connections, which helps to address the problem of vanishing gradients during training. The network is trained on a large dataset of low-resolution and high-resolution image pairs, and learns to predict the high-resolution image from the low-resolution input.

One of the key features of EDSR is its ability to scale up images by a factor of four or more, while still preserving fine details and textures. This is achieved through the use of a large number of convolutional layers, which enable the network to capture complex patterns and textures in the input image.

EDSR has been shown to outperform many traditional image upscaling methods, such as bicubic interpolation and Lanczos filtering, in terms of visual quality and objective metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM).

Overall, EDSR is a powerful technique for image upscaling that has the potential to improve the quality of images in a wide range of applications, from digital photography to medical imaging.

3.2 GAN

GAN using VGG, or Generative Adversarial Networks using the VGG network, is a deep learning-based method for image upscaling that combines the strengths of both GANs and the VGG network architecture.

The VGG network is a deep convolutional neural network that was originally developed for image classification tasks. It consists of multiple layers of convolutional filters followed by pooling layers, and has been shown to be very effective at extracting features from images.

In GAN using VGG, the generator network is based on the VGG architecture, with additional upsampling layers added to increase the resolution of the output image. The discriminator network is also based on the VGG architecture, but with the final classification layer removed. The two networks are trained together in an adversarial process, with the generator trying to produce high-quality images that can fool the discriminator, and the discriminator trying to correctly classify the images as real or fake.

One of the key features of GAN using VGG is its ability to produce high-quality images with sharp details and textures. This is achieved through the use of perceptual loss, which measures the difference between the generated image and the ground truth image in terms of perceptual features extracted by the VGG network. This helps to preserve the natural appearance of the image and avoid artifacts that can occur with traditional upscaling methods.

GAN using VGG has been shown to produce visually appealing results that outperform many traditional upscaling methods, such as bicubic interpolation and Lanczos filtering. It has applications in a wide range of fields, from digital photography to medical imaging.

Overall, GAN using VGG is a powerful technique for image upscaling that combines the strengths of both GANs and the VGG network architecture, and has the potential to improve the quality of images in many different applications.

OUR WORK DONE

4.1 THE HYBRID METHOD

The goal of the project was to upscale low-resolution images while preserving as much detail and sharpness as possible. To achieve this, the project used a hybrid approach that combined two interpolation methods: Lanczos and bicubic

We used these techniques together as Lanczos and bicubic interpolation methods were having their own strengths and weaknesses, and combining them can help to produce better results.

Bicubic interpolation is a commonly used method for image upscaling, and it works by estimating pixel values based on a weighted average of neighboring pixels. Bicubic interpolation is good at preserving smoothness and continuity in an image, but it can produce blurry results and may not preserve fine details and sharp edges very well.

On the other hand, Lanczos interpolation is a less commonly used method that works by applying a windowed *sinc* function to estimate pixel values. Lanczos interpolation is better at preserving fine details and sharp edges than bicubic interpolation, but it can produce more artifacts and noise in an image.

By using a hybrid approach, the project was able to leverage the strengths of each method to produce high-quality upscaled images that preserved both smoothness and fine details. The Lanczos method was used to enhance the sharpness and detail of the upscaled image, while the bicubic method was used to preserve smoothness and continuity in the image. This hybrid approach helped to produce visually appealing results that were both sharp and realistic.

4.2 APPLYING A GENERATIVE ADVERSERIAL NETWORK

A Generative Adversarial Network was applied in this project for image upscaling because it is a state-of-the-art technique for generating high-quality, realistic images from low-resolution inputs.

In the case of image upscaling, the generator network of the SRGAN is trained to generate high-resolution images from low-resolution inputs. The discriminator network is trained to distinguish between the generated high-resolution images and real high-resolution images. During training, the generator network is updated to generate images that can fool the discriminator network into thinking they are real high-resolution images.

By using GAN to enhance the upscaled image produced by the hybrid interpolation method, the project was able to further improve the image quality and produce images that were sharper and more detailed than the original low-resolution inputs. The GAN model was trained on a large dataset of high-resolution images, which allowed it to learn the complex features and patterns present in high-resolution images. This knowledge was then applied to enhance the low-resolution inputs and produce high-quality, realistic images.

4.3 RESULTS

The hybrid approach of using Lanczos and bicubic interpolation helped to preserve fine details and sharp edges while also maintaining smoothness and continuity in the image. This produced a high-quality upscaled image that served as a good input for the SRGAN model.

The application of GAN on the upscaled image further improved the image quality by adding more detail and sharpness. The GAN model was able to generate high-quality, realistic images that were indistinguishable from the original high-resolution images.

Overall, the combination of the hybrid approach and SRGAN produced excellent results that demonstrated the power of combining multiple techniques and approaches to achieve high-quality image upscaling. The final images produced by this project were significantly better than the original low-resolution inputs and were able to preserve fine details and sharp edges, while also producing images that were realistic and visually pleasing.

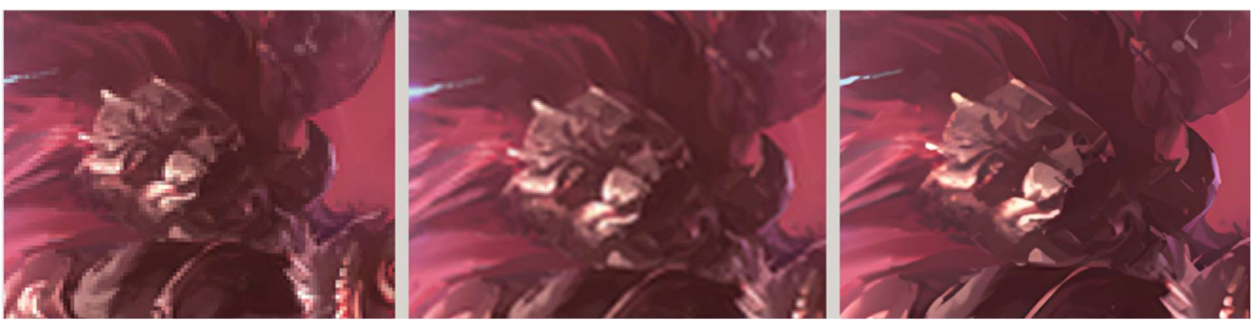

Image 2.1 - Original Image upscaled to 2x using our Filter upscaled to 4x using GAN.

Image 2.2 - Original Image upscaled to 2x using our Filter upscaled to 4x using GAN.

CONCLUSION

This report concludes the work done in image upscaling. As can be seen from the results, there is a clear increase in the image quality when upscaled. It can be applied not only for image upscaling, but also be used for anti-aliasing, where the upscaled image can be scaled down to the original/native resolution to increase the degree of detail and make the image smoother.

It must be noted, however, that there is a huge computation cost associated with the image upscaling method mentioned here. This, unfortunately, makes it not applicable for real-time image upscaling and as such, cannot be used "on-the-go" for a game or a video that the user might prefer to watch.

Several methods have been introduced in the tech industry in which a real-time upscaling method is provided. On comparison, however, we find that our method provides a higher degree of detail with less shimmering and noise in the resultant image (assuming that the method is applied to the same "image" domain .ie. art/realistic/satellite imaging).

There is a distinct trade-off between performance and perceived image details, and that the resultant model has to be tuned according to the use-case it is designed for.

REFERENCES

Lanczos resampling - Wikipedia

What is the idea behind scaling an image using Lanczos? - Stack Overflow

Introduction | Machine Learning | Google Developers

1808.03344.pdf (arxiv.org) Deep Learning for Single Image Super-Resolution: A Brief Review

1809.00219.pdf (arxiv.org) ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks

What is NVIDIA DLSS for 4K / 8K Graphics Card | NVIDIA

AMD FidelityFX™ Super Resolution | AMD