Image Enlargement using Generative Adversarial Networks

Dikshita Poojari[1], Trupti Raut[2], Faiz Sayyed[3], Rupali Pashte[4]

[1,2,3]*B.E Student, Shree L.R Tiwari College of Engineering, Mumbai University, Mumbai, Maharashtra, India*
[4]*Assistant Professor, Shree L.R Tiwari College of Engineering, Mumbai University, Mumbai, Maharashtra, India*

Abstract - *Low resolution image enhancement is a classical computer vision problem. Selecting the best method to reconstruct an image to a higher resolution with the limited data available in the low-resolution image is quite a challenge. A major drawback of the existing enlargement techniques is the introduction of color bleeding while interpolating pixels over the edges that separate distinct colors in an image. The color bleeding causes to accentuate the edges with new colors as a result of blending multiple colors over adjacent regions. This paper proposes a novel approach to mitigate the color bleeding by segmenting the homogeneous color regions of the image using Enhanced SRGAN (ESRGAN). In particular, we introduce the Residual-in-Residual Dense Block (RRDB) without batch normalization as the basic network building unit. Moreover, we borrow the thought from relativistic GAN to let the discriminator predict relative realness instead of the absolute value. Finally, we improve the perceptual loss by using the features before activation, which could provide stronger supervision for brightness consistency and texture recovery.*

***Key Words*: Residual-in-Residual Dense Block (RRDB), SR- GAN –network architecture, adversarial loss, perceptual loss.**

1. INTRODUCTION

Here goes the introduction the digital images are affordable and easy to preserve, transmit, and modify; they are widely used across different fields. Since digital image sampling is non sequential and incomplete, digital image resolution is often limited. Therefore, it is necessary to employ image enlargement technologies when viewing portions of an image in detail. The quality of an image is commonly determined by factors in the natural environment. Factors present in the natural environment usually relate to light. If the light distribution is extreme, the target object in the picture becomes hard to identify. This study proposes an image enlargement method that combines image enhancement with image enlargement.

Image enlargement is a technique to generate a higher resolution image from a lower resolution one. Some applications of image enlargement are video conference, medical imaging, and digital photographs. Image enlargement has applications in many industries: investigation department, medical department, traffic department, and others. Image enlargement has made it possible to regenerate small pixel images into vivid and clear images. Traffic cameras capture blur and distorted images with the help of image enlargement, we can solve this problem too. There are many different algorithms and it may vary depending upon the situation. Earlier there was the problem of rough and bleeding edges and images were torn while enlarging that's the reason for the image enlargement system. From the viewpoint of digital signal processing, image enlargement can be considered as a re-sampling problem. To re-sample a digital signal, an interpolator is required. Popular interpolation schemes for image enlargement are the nearest neighbor interpolation (NNI), the bilinear interpolation (BLI), and the bicubic interpolation (BCI). These interpolators are 2- D polynomials of zero-, first-, and third-order, respectively. In general, a polynomial interpolator with higher order has better performance. Since those interpolators result in blur and zigzag edge problems. Deep learning provides better solution to get optimized images.

To recover or restore high resolution image from low resolution image. There are many forms of image enhancement which includes noise-reduction, up-scaling image and color adjustments. This post will discuss enhancing low resolution images by applying deep network with adversarial network (Generative Adversarial Networks) to produce high resolutions images. Our main target is to reconstruct super resolution image or high resolution image by up-scaling low resolution image such that texture detail in the reconstructed SR images is not lost. In modern days, the amount of data available to train machine learning algorithms is getting larger, however, the cost of labeling is expensive. Many of the machine learning approaches require annotated data for training and testing, therefore, the need for labeled data is growing. One of the solutions for this problem is to use generated data for training instead of real data. One of the prominent approaches for image generation using the Generative Adversarial Networks. Image Enlargement uses various techniques like deep learning and AI to provide the better user interface and to reach the expectation of the user. This system will provide better resolution images using low-quality images with fast processing and in an efficient way. Based on the small pixels and real life information using various algorithms. The system will provide a high-resolution image.

GANs are class of AI algorithms used in Unsupervised Ma- chine learning. GANs are deep neural network architectures comprised of two networks (Generator and Discriminator) pitting one against the other (thus the "adversarial"). GANs is about creating, like drawing a portrait or composing a symphony. The main focus for GANs is to generate data from scratch. To understand GANs, first we need to understand what a generative model is. In machine learning, the two main classes of models are generative and discriminative. A discriminative model is one that discriminates between two (or more) different classes of data — for example a convolutional neural network that is trained to output 1 given an image of a car and 0 otherwise. A generative model on the other hand doesn't know anything about classes of data. Instead, its purpose is to generate new data which fits the distribution of the training data. GANs consist of a Generator and Discriminator. Think it like a game where Generator tries to produce some data from probability distribution and Discriminator acts like a judge. Discriminator decides whether input is coming from true training data set of fake generated data. Generator tries to optimize data so that it can match true training data. Or we can say discriminator is guiding generator to produce realistic data. They just work like encoder and decoder. This has numerous applications like satellite and aerial image analysis, medical image processing, compressed image/video enhancement etc.

2. LITERATURE REVIEW

In[1],The paper explores Bilinear Interpolation applied to image enlargement after a fuzzification pre-processing. Interpolation methods are basically used to rescale the image i.e. whether to shrink or enlarge the original image. Methods used in interpolation are Nearest Neighbor Interpolation, Bilinear Interpolation, Bicubic Interpolation and Lanczos Interpolation."Bilinear Interpolation" is being widely used as it brings great balance between computation speed and precision of the interpolation. The fuzzification is based on the idea that the gray intensity assigned to one pixel has an inherent uncertainty. Specifically, the gray level perceived in one pixel depends on the pixel in its neighborhood. The goal is to substitute the intensity value of each pixel by a fuzzy number, which represents the relationship between the intensity of pixels and the intensities in their surroundings.

In [2], we studied an algorithm is proposed that enhances the pixel number in high-speed camera image acquisition. High speed cameras are capable of capturing images more than one hundred frames per second (fps), but with a decrease of pixels when fps increases. Reason, time for swipe out is proportional to the number of image pixels while the increase of fps number suppresses the time for swipe out. In this, a novel approach based on the combination of L1 norms within each frame and between adjacent frames is being proposed. By Douglas-Rachford splitting algorithm, the proposed method outperforms conventional Methods with small computational cost. Steps as follows:-1) We formulate the image reconstruction problem and define a criterion (cost function) for the image reconstruction, By convex Optimization. 2) We propose a fast reconstruction algorithm based on the Douglas-Rachford splitting method, Reconstruction algorithms.3) We show the effectiveness of the proposed method by computer simulations. In this paper we studied, optical setup that randomly selects some percent of pixels in an image and developed an algorithm that reconstructs the entire image from the selected partial pixels. In this Algorithm, sparsity not only within each frame but also induced from the similarity between adjacent frames was exploited.

In [3], in this paper, by applying simple enhancement equations, we studied a novel super-resolution method which can produce a sharp enlarged image with low calculation cost. Methods are as follows:-1) TOTAL VARIATION REG- ULARIZATION FILTER: Separate image into two structural parts consist of low frequency component and edge and textural part consist of high-frequency vibration signals. Each component is enlarged by linear interpolation methods. Blur is sharpened with non-linear structural enhancement filtering. Chambolle's projection algorithm is used for solving each component.2) SHOCK FILTER: A Shock Filter can construct sharp edges by repeating the following calculation: 3) EDGE ENHANCEMENT FILTER: In this, a new value is generated by multiplying the high-frequency component derived from the Unsharp Mask and the gradient strength value derived from the Shock Filter.

We aim to generate high-quality edges. In the previous Shock Filter, the center point of the edge was strongly stimulated, resulting in jaggy noise. This supplementary processing is expected to remove ringing noise.

In [4], This paper proposes a hybrid method that combines bicubic interpolation method and undecimated wavelet trans- formation that interpolates an image in a way that preserves the edge and detail region of the input. Interpolation is a process of estimating unknown values that fall between known values. There are two kinds of image interpolation, namely, downscaling and upscaling. Interpolation from a higher resolution to a lower resolution is referred to as down-scaling or down sampling. On the other hand, interpolation from a lower resolution to a higher resolution is called up-scaling or up sampling. Interpolation algorithms can be grouped into two categories, namely, Adaptive and Non-adaptive techniques. Adaptive methods change depending on what they are interpolating (sharp edges vs. smooth texture), whereas non-adaptive methods treat all pixels equally. Non adaptive or linear interpolation is a fixed pattern of computation that is applied in every pixel location to recover the missing components. The proposed interpolation method consists of 3 modules, (i) A Region Classification Module (ii) Edge Enhancement Module and (iii) Detail Enhancement Module. The region analyzer that separates the edge and non-edge regions. After separating the edge region from the input mammogram image, the edge enhancement algorithm interpolates the edge region. The detail enhancement module can be considered as a prediction of detailed pixel procedure.

In [5], This paper proposes a novel approach to mitigate the color bleeding by segmenting the homogeneous color regions of the image using Simple Linear Iterative Clustering (SLIC) and applying a higher order interpolation technique separately on the isolated segments. The interpolation at the boundaries of each of the isolated segments is handled by using a morphological operation. The approach is evaluated by comparing against several frequently used image enlargement methods such as bilinear and bicubic interpolation by means of Peak Signal-to- Noise-Ratio (PSNR) value. The results obtained exhibit that the proposed method outperforms the baseline methods by means of PSNR and also mitigates the color bleeding at the edges which improves the overall appearance.

3. PROPOSED APPROACH

The proposed system is a trained network i.e. a neural net- work by which the function is trained by some sort of example ie. Neural network upscaling which takes low resolution photo and turns it in high resolution photo by comparing with the pair of data available in network of low resolution and high resolution which is a GAN structure. Our algorithm consists of a GAN structure consisting of generator and discriminator. Generator: The network will make a best guess with the low resolution image and then the neural network guessed image pixels will be compared with the real image pixel , the pre- pixel subtraction is carried out i.e. the loss function , the loss function is trained to minimize the total pre-pixel subtraction between two image.

Discriminator :It act as a police where, generator image is analyze whether the image is fake real image by guess of neural network or actual real world image, it gives feedback to generator to make correction and iteration is carried out until the accuracy of the discriminator is minimized. The proposed network interpolation enjoys two merits. First, the interpolated model is able to produce meaningful results for any feasible without introducing artifacts. Second, we can continuously balance perceptual quality and fidelity without re-training the model.

To remove unpleasant noise in GAN-based methods while maintain a good perceptual quality, we propose a flexible and effective strategy – network interpolation. Specifically, we first train a PSNR-oriented network GPSNR and then obtain a GAN-based network GGAN by fine-tuning. We interpolate all the corresponding parameters of these two networks to derive an interpolated model GINTERP. The proposed network interpolation enjoys two merits. First, the interpolated model is able to produce meaningful results for any feasible without introducing artifacts. Second, we can continuously balance perceptual quality and fidelity without re-training the model.

A. User Interfaces

The product will exist on a real life system. The Interface will be a simple user interface where the system will take any image. Once the image is fed into the dataset it will take some time to process it and give a better quality image as a result. It will have a common and user friendly UI so it can be useful for all the sectors.

B. Software Interfaces

The following program will take the raw input and convert it into grayscale first for better edge detection and further with the help of Generative Adversarial Network (GAN). Generative modeling is an unsupervised learning task in machine learning that involves automatically discovering and learning the regularities or patterns in input data in such a way that the model can be used to generate or output new examples that plausibly could have been drawn from the original dataset.

C. Communications Interfaces

The proposed framework will analyze each pixel of the image and will generate adjacent pixels using the GAN algorithm. It will give a clear and much bigger image with smooth edges.

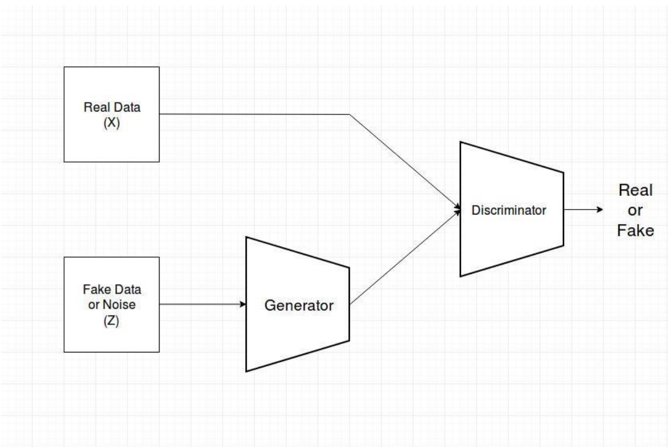


Fig 1- Flow Model

4. IMPLEMENTATION

The algorithm consists of the following steps:

1) Data: We mainly used the DIV2K dataset which has 800 images, which is a high-quality(2K resolution) dataset for image restoration tasks. We also seek for other datasets with rich and diverse datasets for our training.

2) Generator: We used two setting for our generator:

- One of them contain 16 Residual blocks
- And other is deeper model with 23 RRDB blocks

We implemented our models with the PyTorch framework and trained them using NVIDIA Titan Xp GPUs.

3) Relativistic Discriminator : Estimate the probability that one input x is real and natural, a relativistic discriminator tries to predict the probability that a real image xr is relatively more realistic than a fake one xf .

4) Training model: We performed all experiments with a scaling factor of x4 between LR and HR images. We obtain LR images by down-sampling HR images using the MATLAB bicubic kernel function. The mini-batch size is set to 16. The spatial size of the cropped HR patch is 128 x 128. The training process is divided into two stages. First, we train a PSNR- oriented model with the L1 loss. The learning rate is initialized as 2 104 and decayed by a factor of 2 every 2 105 of mini- batch updates. We then employ the trained PSNR- oriented model as an initialization for the generator. The generator is trained using the loss function. The learning rate is set to 1 104 and halved at [50k, 100k, 200k, 300k] iterations. Pre- training with pixel-wise loss helps GAN- based methods to get more visually pleasing results. The reasons are that 1) it can avoid undesired local optima for the generator; 2) after pre-training, the discriminator receives relatively good super- resolved images rather than extreme fake ones (black or noisy images) at the very beginning, which helps it to focus more on texture discrimination. For optimization, we use Adam with 1 = 0.9, 2 = 0.999. We alternately update the generator and discriminator network until the model converges.

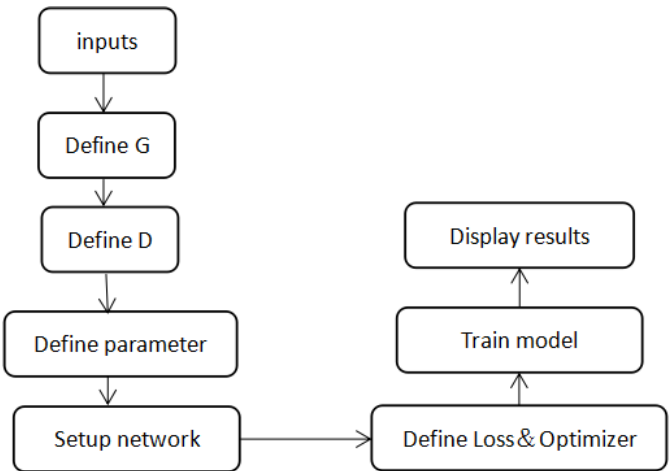
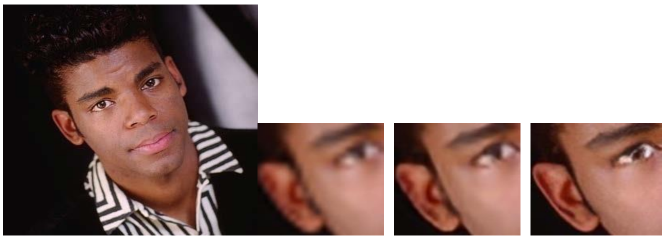


Fig 2- Implementation/training Model

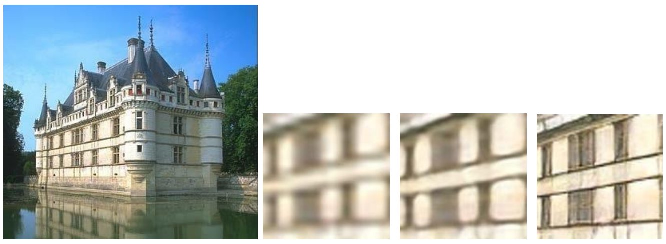
5. RESULT DISCUSSION

We have given some train data and test data to the generator model. The fundamental aim of the generator is, The network will make a best guess with the low resolution image and then the neural network guessed image pixels will be compared with the real image pixel , the pre- pixel subtraction is carried out i.e. the loss function , the loss function is trained to minimize the total pre-pixel subtraction between two images. While Discriminator model, It acts as a police where, generator image is analyze whether the image is fake real image by guess of neural network or actual real world image, it gives feedback to generator to make correction and iteration is carried out until the accuracy of the discriminator is minimized. We compared our final models on several public benchmark datasets state- of- the-art PSRN-oriented methods including bicubic, SRCNN, EDSR and also with perceptual-driven approaches including SRGAN. Qualitative result produce by ours Network produce more natural texture e.g., animal fur, building structure.

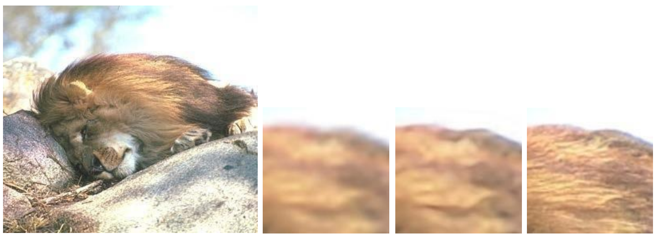
Examples:



(PSNR/Perceptual Index) Bicubic SRCNN ESRGAN



(PSNR/Perceptual Index) Bicubic SRCNN ESRGAN



(PSNR/Perceptual Index) Bicubic SRCNN ESRGAN

Fig 3- Results-Pictures

6. CONCLUSION

The Image Enlargement proposed in this paper overcomes problems that we encountered in the Image Resolution. We use ESRGAN Network that achieve consistently better perceptual quality. We have formulated a novel architecture containing several RDDB blocks without BN layers. In addition, useful techniques including residual scaling and smaller initialization are employed to facilitate the training of the proposed deep model. We have also introduced the utilization of relativistic GAN because of the discriminator, which learns to judge whether one image is more realistic than another, guiding the generator to recover more detailed textures. Moreover, we have enhanced the perceptual loss by using the features before activation, which offer stronger supervision and thus restore more accurate brightness and realistic textures. Thus this is a step forward in producing images more pleasing to the eye at the least possible storage space required to store the image.

7. FUTURE SCOPE

1) The network can produce texture detail, and realistic content, which will be an advantage to many applications, such as texture synthesis, super-resolution, image inpainting, etc.
2) To improve the architecture of the network to make ES- RGAN capable of tackling images with strong repetitive structure.
3) Real images can be identified by sketch image.
4) It is hoped that the future work will bring a step closer in our attempt to embed human qualities such as creativity in the intelligence we build.

8. REFERENCES

[1] Petr Hrutik, Nicolas Madrid, "Bilinear Interpolation over Fuzzified image: enlargement",Institute of Research and Applications of Fuzzy Modeling University of Ostrava Czech Republic,2015.
[2] A. Hirabayashi, N. Nagami, J. White, L. Condat, "Pixel enlargement in high speed camera image acquisition based on 3D sparse representation", University College of Information Science and Engineering Kusatsu, Shinga 525-8577, Japan, 2015.
[3] Tsubasa Munezawa, Tomio Goto, Santoshi Hirano "Super-resolution method for enlarged images utilizing non-linear filtering", 2017 IEEE 6th Gobal Conference on Consumer Electronics (GCCE), 2017, pp. 1-2, doi: 10.1109/GCCE.2017.8229293.
[4]K. Mirunalini, Dr. (mrs.)S.N Geethalashmi, Dr. (mrs.)Vasantha Kalyani David, "Enhanced interpolation method for enlargement of mammogram Images", International conference on Advance and communication system, 2017.
[5]M.Z.F. Amara, R.Bandara, Thushari Silva, "SLIC Based Digital Image Enlargement", International conference on Advance in ICT For Emerging Regions (ICTer): 113-118, 2018.