TRIBHUVAN UNIVERSITY
INSTITUTE OF ENGINEERING
PURWANCHAL CAMPUS

COMPARISON OF DIFFERENT IMAGE SUPER RESOLUTION TECHNIQUES

BY:
ABHINAV JAISHI (PUR076BEI001)
MEGHA THAKUR (PUR076BEI021)
SANGHARSHA DAHAL (PUR076BEI029)
KABIN POUDEL (PUR076BEI047)

A MID TERM PROGRESS REPORT TO THE DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE BACHELOR'S DEGREE IN ELECTRONICS, COMMUNICATION AND INFORMATION ENGINEERING

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING
PURWANCHAL CAMPUS
DHARAN, NEPAL

January 7, 2024

ABSTRACT

The instructors are responsible to ensure the smoothness of the classroom activities alongside with the monitoring the student's attendance, attention and activities. Manual observation is a tedious job and affects the whole learning process. With the incorporation of IOT devices and computational algorithms

The collected preliminary data-set from area around Kathmandu valley and the country's major lines are able to map some interesting features and environmental proxies that are visualised and the patterns and variations in it are explored using various models namely such as Autoregressive integrated moving average (ARIMA), Recurrent neural network (RNN), which worked best for time series database.

Keywords: a, b, c, d

TABLE OF CONTENTS

List of Figures

List of Tables

1. INTRODUCTION

Image super resolution is a technique used to enhance the resolution and quality of image beyond its original size. It is useful technique in the field of image processing. It can be used to enhance the quality of image captured in night, shaky image etc. There are various techniques and algorithm used for image super resolution, including both traditional methods and deep learning-based approaches. Traditional super image resolution techniques are based on mathematics techniques. These techniques have been used for many years. Some of the traditional super resolution techniques include: Interpolation, Edge-Directed interpolation, Bayesian methods etc. Traditional methods have certain limitations compared to deep learning approaches. They may struggle to capture complex patterns and textures, and their performance is often constrained by the hand-crafted features and assumptions used in the algorithms. Deep learning based super resolution typically employs Convolution Neural Network (CNN) to learn the mapping between low-resolution and high-resolution images. The network is trained on large dataset of paired low-resolution and high- resolution images. During training, the networks learn to identify patterns and features that enables it to generate high-resolution details from low-resolution inputs.

1.1. Statement of problem

The problem of super image resolution arises when we encounter low-resolution images that lack the desired level of detail and clarity. These low-resolution images can be the result of various factors such as limitations in capturing hardware, or resizing operations. The inadequate resolution in these images hinders their effective use in various applications, where higher resolution and finer details are essential for accurate analysis, interpretation, and visualization.

1.2. Objectives

- Compare different Super Resolution algorithms.
- Recommend

1.3. Scope and Applications

Solving the super image resolution problem has numerous practical applications in field like photography, surveillance, and digital art where enhanced image quality is crucial for accurate decision-making, analysis, and interpretation. The development of efficient and accurate super-resolution method involves exploring traditional signal processing techniques, as well as cutting-edge deep learning approaches, to achieve optimal results while considering computational efficiency and resource constraints.

Super image resolution can be used for following things:

1. Super resolution can be used in digital cameras and smartphones to provide better zoom capabilities without significant loss of image quality.

2. Super resolution can enhance the resolution of individual frames in videos, leading to improved video quality and better extraction of information.

3. Super resolution can be used to improve the resolution of old and degraded images, preserving historical photographs, artworks, and documents.

4. In surveillance systems, low resolution camera feeds can be upscaled to improve the ability to identify faces, license plates, or other critical details, helping in investigations.

2. LITERATURE REVIEW

In the timeline of image scaling process, the first used methods were different interpolation techniques and sparse representation-based methods. But with the advent of deep-learning based image scaling techniques interpolation methods were less commonly used in image enhancement.

D. Han (2013) **"Comparison of commonly used image interpolation methods"** discusses multiple interpolation methods like nearest neighbor, bilinear, bicubic, etc. Although interpolation is fast but have some disadvantages. But using these methods was not optimal for our use case as they may lead to the loss of fine details and sharpness in image and can amplify noise present in image which may not be accurate in real-world images.[1]

C. Dong, et. al, **"Image super-resolution using deep convolutional networks"** is a seminal paper on use of Deep Learning for the use of for the task of SR. It only consists of three layers and requires the LR image to be up-sampled using bicubic interpolation prior to being processed by the network, but it was shown to outperform the state-ofthe-art methods of that time.[2]

C. Ledig et al., **"Photo-realistic single image super-resolution using a generative adversarial network"** discusses about a Generative Adversarial Network (GAN) for image super-resolution (SR). To our knowledge, it is the first framework capable of inferring photo-realistic natural images for 4x upscaling factors. To achieve this, it uses a perceptual loss function which consists of an adversarial loss and a content loss. The adversarial loss pushes the solution to the natural image manifold using a discriminator network that is trained to differentiate between the super-resolved images and original photo-realistic images. In addition, it uses a content loss motivated by perceptual similarity in VGG space instead of similarity in pixel RGB space. This deep residual network is able to recover photo-realistic textures from heavily down sampled images on public benchmarks. An extensive mean-opinion-score (MOS) test shows hugely significant gains in perceptual quality using SRGAN. The MOS scores obtained with SRGAN are closer to those of the original high-resolution images than to those obtained with any state-of-the-art method at the time.[3]

3. HARDWARE AND SOFTWARE REQUIREMENTS

3.1. Hardware Requirements

We used the system with following specifications for training:

- PC with Windows 11
- 16 GB RAM
- Nvidia Geforce GTX 1050(4GB)
- 256 GB storage
- Processor: Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz

3.2. Software Requirements

The software requirements for developing, training, evaluating, and deploying the model effectively are listed below:

1. **Integrated Development Environment (IDE):** We extensively used Visual Studio Code (VS Code) as IDE for training, testing and documentation also.
2. **Version Control System:** We used Git and Github.
3. **Web framework:** We simply used Streamlit framework.
4. **Deep-Learning Frameworks:** We used PyTorch.

4. METHODOLOGY

Image Super Resolution is a branch of Artificial Intelligence that deals with upscaling a Low-Resolution Image to High Resolution Image, filling in the missing pixels with the help of learning from the environment using machine learning. There are other methods of upscale images like Linear or Bi-cubic interpolation but they do not generate any new information based on the environment and hence are not super useful to upscale an LR image. The retrieval of High-Resolution images based on underlying image is not new, there were techniques such as sparse representation-based methods. However, it was the advent of deep learning and convolutional neural networks that arguably brought about the most significant leaps forward, with the seminal work being the Super-Resolution Convolutional Neural Network (SRCNN) proposed by Dong et al. in 2014. Much work has been done since then, not only on the design and structure of the neural networks but also on the data used to train and evaluate these networks. Deep Learning based methods require huge amounts of data so that the model is not overfitted, the dataset needs to contain an HR and LR version of the same image. Images must be perfectly aligned to each other. To do so, the LR image is obtained by synthetically degrading image using a degradation model.

The 'classical' degradation model is the most commonly used, which considers down-sampling, blurring, and noise:

$$I^{LR} = (I^{HR} \circledast k) \downarrow_s +n$$

where $\circledast$ represents a convolution operation, k is a kernel (typically a Gaussian blurring kernel, but it can also represent other functions such as the Point Spread Function (PSF)), n represents additive noise, and $\downarrow$s is a downscaling operation that is typically assumed to be bicubic down-sampling with scale factor s. The aim of SR is to then reverse whichever degradation process is considered, to retrieve the original underlying high-fidelity image. Some of popular methods to perform this task are:

4.1. Non-Blind SR Methods

1. **SRCNN:**The Super-Resolution Convolutional Neural Network (SRCNN) is consid-

ered to be the pioneering work in using deep learning and convolutional neural networks for the task of SR. It only consists of three layers and requires the LR image to be up-sampled using bicubic interpolation prior to being processed by the network, but it was shown to outperform the state-of-the-art methods of the time such as A+ and the sparse representation-based method. It was also shown that sparse-coding-based methods are equivalent to convolutional neural networks, which influenced SRCNN's hyperparameter settings. While SRCNN has been used as a benchmark by numerous researchers, it is now comprehensively outclassed and is no longer used so frequently.

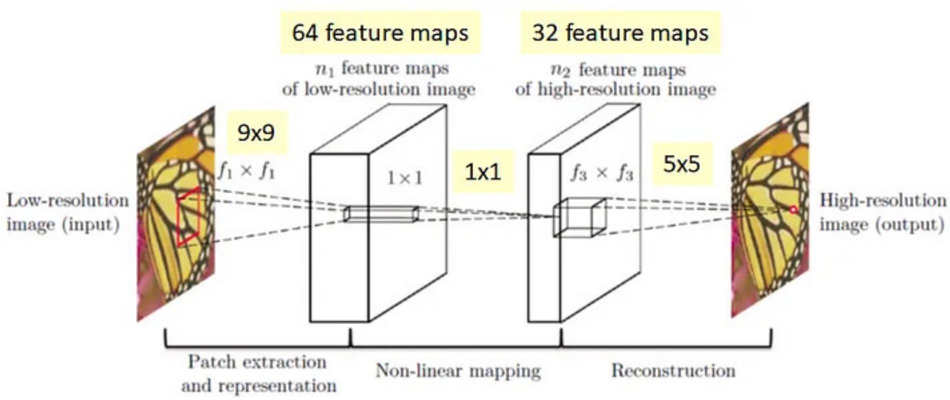

Figure 1: SRCNN Representation

2. **ResNet:** The Residual Network (ResNet) architecture was primarily designed to ease the training of networks as the number of layers increases. Indeed, while many works in literature have indicated that deeper networks provide superior performance, there comes a point when accuracy gets saturated and then quickly degrades. It is shown that this is not due to overfitting, but due to the difficulty in optimizing and training very deep networks. In particular, it was noted that deep networks may find it hard to learn identity functions.

   The Residual Network (ResNet) architecture was primarily designed to ease the training of networks as the number of layers increases. Indeed, while many works in literature have indicated that deeper networks provide superior performance, there comes a point when accuracy gets saturated and then quickly degrades. It is shown that this is not due to overfitting, but due to the difficulty in optimizing and training very deep networks. In particular, it was noted that deep networks may find it hard to learn identity functions.

To counter this problem, a number of skip/shortcut connections are utilized to directly feed feature maps at any given level of the network to higher layers, thus corresponding to the identity function. In this way, the network would only need to learn a function that suppresses the responses (drive the outputs to zero) of intermediary layers.

ResNet was applied to the SR domain to create SRResNet in, where it was also used as the basis of a Generative Adversarial Network (GAN)-based network termed SRGAN.

5. FEASIBLITY ANALYSIS

We are planning to use deep learning to develop a system that can increase the resolution of image. The system is trained on dataset of high-resolution images and low-resolution images generated using degradation model to generate high resolution images from low resolution images.

Deep learning is shown to be very effective for ISR. There are many models developed for this task we can take inspiration from. The technical feasibility of this project is therefore high. The cost of developing the system will depend on the size of the dataset used to train the model, and the complexity of the model. As we are planning to train the model on server and deploy the system on cloud-based service. So, server costs will be our main cost for this project.

The technical and economic feasibility analysis show that we can complete this project successfully.

6. RESULT AND DISCUSSION

6.1. Work Completed

6.2. Work Remaining

6.3. Limitations

6.4. Problem Faced

6.5. Budget Analysis

6.6. Work Schedule

References

[1] Dianyuan Han. Comparison of commonly used image interpolation methods. In *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*. Atlantis Press, 2013/03.

[2] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks, 2015.

[3] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew P. Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. *CoRR*, abs/1609.04802, 2016.