Image Super Resolution Project

IPD Report

60004190003 - Adithya Sanyal
60004190008 - Aditya Thaker
60004190011 - Akshath Mahajan

Guide: Sudhir Bagul

Introduction:

In super resolution, we aim at transforming a low resolution (L.R) image into a high resolution (H.R) image without storing any excess information about the media.

The domain of this project is a subset of artificial intelligence known as deep neural networks. Our goal is to develop an algorithm, or design a model that has higher speed and accuracy as compared to the current super resolution techniques. By doing so, we expect the product to contribute towards the research dedicated to real-time video super resolution which may one day be used across all devices to access media on the web faster.

Literature Survey:

“Super resolution is the problem of artificially enlarging a low resolution photograph to recover a plausible high resolution version. In the regime of high magnification factors, the problem is dramatically underspecified and many plausible, high resolution images may match a given low resolution image. In particular, traditional super resolution techniques fail in this regime due to the multimodality of the problem and strong prior information that must be imposed on image synthesis to produce plausible high resolution images” [1]

[2, 3, 4, 5, 6, 7, 8, 9, 10]Deals with various techniques for solving this problem mainly using GANs (Generative Adversarial Networks). The approach has produced remarkable results in most papers, and has a lot of potential as a feasible solution that could be made better.

[2] provides us insights about the potential of GANs in the field of Single Image Super-Resolution. It states that existing methods are not able to directly optimize perceptual metrics. To solve this issue, a new method is proposed which is given by “Super-Resolution Generative Adversarial Networks with Ranker(RankSRGAN) to optimize generators in the direction of perceptual metrics. Specifically, we first train a Ranker which can learn the behavior of perceptual metrics and then introduce
a novel rank-content loss to optimize the perceptual quality. The most appealing part is that the proposed method can combine the strengths of different SR methods to generate better results.”[2]

[3] gives us another way to counter this. It states “To further enhance
the visual quality, we thoroughly study three key components of SRGAN
– network architecture, adversarial loss and perceptual loss, and
improve each of them to derive an Enhanced SRGAN (ESRGAN).Moreover,
we borrow the idea 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. The proposed ESRGAN achieves consistently better visual quality with more realistic and natural textures than SRGAN ” [3]

[4] proposes another method to improve the quality of the upscaled image. This is done by “we propose a method (ProSR) that is progressive both in architecture and training: the network upsamples an image in intermediate steps, while the learning process is organized from easy to hard, as is done in curriculum learning. To obtain more photorealistic results, we design a generative adversarial network (GAN), named ProGanSR, that follows the same progressive

multi-scale design principle. This not only allows to scale well to high upsampling factors (e.g., 8×) but constitutes a principled multi-scale approach that increases the reconstruction quality for all upsampling factors simultaneously."[4]

[5] provides an alternative approach to tackle super resolution problems. Here, instead of supervised learning, unsupervised learning algorithms are used. This is achieved in the following way "a new SR approach to mitigate
such an issue using unsupervised learning in Generative Adversarial Network (GAN) framework - USISResNet. In an attempt to provide high quality SR image for perceptual inspection, we also introduce a new loss function based on the Mean Opinion Score (MOS). We demonstrate the generalizable nature of proposed network by evaluating real-world images as against other state-of-the-art methods which employ synthetically downsampled LR images."[5]

[6] provides another method to achieve the objective. It states that "Recognizing that a wavelet transform provides a "coarse" as well as "detail" separation of image content, we design a deep CNN to predict the "missing details" of wavelet coefficients of the low-resolution images to obtain the Super-Resolution (SR) results, which we name Deep Wavelet Super-Resolution (DWSR). Our network is trained in the wavelet domain with four input and output channels respectively. The input comprises 4 sub-bands of the low-resolution wavelet coefficients and outputs are residuals (missing details) of 4 sub-bands of high-resolution wavelet coefficients. Wavelet coefficients and wavelet residuals are used as input and outputs of our network to further enhance the sparsity of activation maps. A key benefit of such a design is that it greatly reduces the training burden of learning the network that reconstructs low frequency details. The output prediction is added to the input to form the final SR wavelet coefficients." [6]

[7] proposes a GAN based edge-enhancement network for robust satellite imagery. This is possible by "EEGAN consists of two main subnetworks: an ultradense subnetwork (UDSN) and an edge-enhancement subnetwork (EESN). In UDSN, a group of 2-D dense blocks is assembled for feature extraction and to obtain an intermediate high-resolution result that looks sharp but is eroded with artifacts and noises as previous GAN-based methods do. Then, EESN is constructed to extract and enhance the image contours by purifying the noise-contaminated components with mask processing. The recovered intermediate image and enhanced edges can be combined to generate the result that enjoys high credibility and clear contents."[7]

[8] aims at one central problem which is to recover the finer texture details when we super-resolve at large upscaling factors.The approach used is a super-resolution generative adversarial network (SRGAN) for a deep residual network (ResNet) with skip-connection is employed and diverge from MSE as the sole optimization target.
"In this paper, we present SRGAN, 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 4× upscaling factors. To achieve this, we propose a perceptual loss function which consists of an adversarial loss and a content loss. The adversarial loss pushes our 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, we use a content loss motivated by perceptual similarity instead of similarity in pixel space. Our deep residual network is able to recover photo-realistic textures from heavily downsampled 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. ”[8]

[9] this paper proposes a solution to license plate number checking at speeds faster than humans using progressive vehicle search and Domain Priori GAN (DP-GAN),
“Particularly, we design a null space based progressive vehicle search approach to retrieve the relevant images captured by different cameras given one vehicle with a low-resolution license plate. To handle the extremely varied license plate images caused by different sensors, times, depths, and viewpoints, we also propose a DP-GAN framework to generate multiple spatial correspondences and high-resolution plate images. In the generator network of DP-GAN, a license plate synthesis pipeline is exploited to generate the nearly canonical license plates. In the discriminator network, a spatial split layer is designed to simultaneously preserve the global and local manufacture standards of the license plate. Finally, a multiple images super-resolution GAN is exploited to combine all the synthetic license plates into one high-resolution image.”[9]

[10] this paper proposes a deep generative adversarial network for super-resolution considering the trade-off between perception and distortion, “Based on good performance of a recently developed model for super-resolution, i.e., deep residual network using enhanced upscale modules (EUSR) [20], the proposed model is trained to improve perceptual performance with only slight increase of distortion. For this purpose, together with the conventional content loss, i.e., reconstruction loss such as L1 or L2, we consider additional losses in the training phase, which are the discrete cosine transform coefficients loss and differential content loss. These consider perceptual part in the content loss, i.e., consideration of proper high frequency components is helpful for the trade-off problem in super-resolution.”[10]

The papers [11, 12, 13, 14, 15, 16, 17, 18] propose various techniques for image super resolution using deep CNNs. This approach has proven to be quite accurate in generating a higher resolution image from a lower resolution image.

The paper [11] proposes a model called “Super-Resolution Convolutional Neural Network (SRCNN)” [11]. It claims to be equivalent to a sparse-coding based method.

“The proposed SRCNN has several appealing properties. First, its structure is intentionally designed with simplicity in mind, and yet provides superior accuracy compared with state-of-the-art example-based methods” [11]

[12] has another CNN solution for the problem. It gives us “a highly efficient and faster Single Image Super-Resolution (SISR) model with Deep Convolutional neural networks (Deep CNN).” [12]. The goal of this paper was to achieve SISR while utilizing (comparatively) lesser computational resources. It uses a feature extraction network and a reconstruction network in its model.
“This paper proposed a fast and accurate Image Super Resolution method based on CNN with skip connection and network in network. In the feature extraction network of our method, the structure is optimized and both local and global features are sent to the reconstruction network by skip connection. In the reconstruction network, network in network architecture is used to obtain a better reconstruction performance with less computation. In addition, the model is designed to be capable of processing original size images. Using these devices, our model can achieve state-of-the-art performance with less computation resources.” [12]

[13] Proposes a highly accurate solution using deep convolutional networks. The authors call the approach “Very Deep Convolutional Networks”
“We find increasing our network depth shows a significant improvement in accuracy. Our final model uses 20 weight layers. By cascading small filters many times in a deep network structure, contextual information over large image regions is exploited in an efficient way” [13] They point out speed being a critical issue during training, but also propose an effective solution for it.

[14] proposes another highly accurate solution. It uses a “deeply-recursive convolutional network (DRCN)” to achieve SISR. “Increasing recursion depth can improve performance without introducing new parameters for additional convolutions” [14]
“For image super-resolution (SR), receptive field of a convolutional network determines the amount of contextual information that can be exploited to infer missing high-frequency components” [14]

[15] uses deep residual learning to enhance the performance of deep convolutional networks, ”In this paper, we develop an enhanced deep super-resolution network (EDSR) with performance exceeding those of current state-of-the-art SR methods. The significant performance improvement of our model is due to optimization by removing unnecessary modules in conventional residual networks. The performance is further improved by expanding the model size while we stabilize the training procedure. We also propose a new multi-scale deep super-resolution system (MDSR) and training method,which can reconstruct high-resolution images of different upscaling factors in a single model.”[15]

[16] differs fundamentally from existing external example-based approaches, in that ours does not explicitly learn the dictionaries or manifolds for modeling the patch space. These are implicitly achieved via hidden layers. The paper directly considers a convolutional neural network which is an end-to-end mapping between low and high-resolution images."Our method directly learns an end-to-end mapping between
the low/high-resolution images. The mapping is represented as a deep convolutional neural network (CNN) [15] that takes the low resolution image as the input and outputs the high-resolution one. We further show that traditional sparse-coding-based SR methods can also be viewed as a deep convolutional network. But unlike traditional methods that handle each component separately, our method jointly optimizes all layers."[16]

[17] Light-field photography (also known as plenoptic photography) is an imaging technology that makes it possible to adjust the focus in an existing picture. Unlike conventional images, any area in an image taken with a light-field camera can be brought into focus to make a particular feature sharper or make details of a particular area more visible. In this paper, the authors present a novel method for Light-Field image super-resolution via a deep convolutional neural network. " Rather than the conventional optimization framework, we adopt a data-driven learning method to simultaneously up-sample the angular resolution as well as the spatial resolution of a Light-Field image. We first augment the spatial resolution of each sub-aperture image to enhance details by a spatial SR network. Then, novel views between the sub-aperture images are generated by an angular super-resolution network. These networks are trained independently but finally fine-tuned via end-to-end training. "[17]

[18] In recent successful methods, the low resolution (LR) input image is upscaled to the high resolution (HR) space using a single filter, commonly bicubic interpolation, before reconstruction. The super-resolution (SR) operation is performed in HR space. "We demonstrate that this is sub-optimal and adds computational complexity.
we present the first convolutional neural network (CNN) capable of real-time SR of 1080p videos on a single K2 GPU. To achieve this, we propose a novel CNN architecture where the feature maps are extracted in the LR space. In addition, we introduce an efficient sub-pixel convolution layer which learns an array of upscaling filters to upscale the final LR feature maps into the HR output. By doing so, we effectively replace the handcrafted bicubic filter in the SR pipeline with more complex upscaling filters specifically trained for each feature map, whilst also reducing the
computational complexity of the overall SR operation. We evaluate the proposed approach using images and videos from publicly available datasets and show that it performs significantly better (+0.15dB on Images and +0.39dB on Videos) and is an order of magnitude faster than previous CNN-based methods. "[18]

[19] Deals with taking super-resolution to the real world. It states - “Most of the existing learning-based single image super-resolution (SISR) methods are trained and evaluated on simulated datasets, where the low-resolution (LR) images are generated by applying a simple and uniform degradation (i.e., bicubic downsampling) to their high-resolution (HR) counterparts. However, the degradations in realworld LR images are far more complicated” [19]

[21]Deals with an algorithm for super resolution of footage of security cameras. It states - “Security cameras have infrared imaging modes for low-light conditions. However, infrared imaging sensitivity is low, and the quality of images recorded in low-light conditions is often poor as they do not always possess sufficient contrast and resolution; thus, infrared imaging devices produce blurry monochrome images and videos. A real-time nonlinear signal processing technique that improves the contrast and resolution of low-contrast infrared images and video is proposed. The proposed algorithm can be installed in a field programmable array.” [21]

Need Of The Product

In today's world there are multiple applications which require a massive use of high resolution images. The need for high resolution is common in computer vision applications for better performance in pattern recognition and analysis of images. High resolution is of importance in medical imaging for diagnosis. Many applications require zooming of a specific area of interest in the image wherein high resolution becomes essential, e.g. surveillance, forensic,space and satellite imaging applications.

However, high resolution images are not always available. This is since the setup for high resolution imaging proves expensive and also it may not always be feasible due to the inherent limitations of the sensors, optics manufacturing technology. These problems can be overcome through the use of image processing algorithms, which are relatively inexpensive, giving rise to the concept of super-resolution. It provides an advantage as it may cost less and the existing low resolution imaging systems can still be utilized.

Super-resolution is based on the idea that a combination of low resolution (noisy) sequence of images of a scene can be used to generate a high resolution image or image sequence. Thus it attempts to reconstruct the original scene image with high resolution given a set of observed images at lower resolution. The central aim of Super-Resolution (SR) is to generate a higher resolution image from lower resolution images. High resolution image offers a high pixel density and thereby more details about the original scene. The general approach considers the low resolution images as resulting from resampling of a high resolution image. The goal is then to recover the high resolution image which when resampled based on the input images and the imaging model, will produce the low resolution observed images. Thus the accuracy of the imaging model is vital for super-resolution and an incorrect modeling, say of motion, can actually degrade the image further.

With the proposed project, we will be attempting to increase the speed and accuracy of the existing super resolution techniques to improve the quality of the image we receive.

Problem Formulation

The proposed project focuses on improving existing image/video super resolution technology, by working with new techniques and existing models. The proposed project aims at increasing the speed and accuracy of image super resolution network architectures using novel techniques and by extension improving video super resolution as well.

A few of the existing solutions are as follows:

1. Downgrading the video quality
2. Buying more storage
3. Formating the database server frequently.

The above mentioned solutions are practical but at the same time restricting, downgrading the video quality is something that can be done to only a certain extent as it requires losing a lot of valuable information. Whereas buying more storage for the server is the best of all solutions but isn't very cost effective and only so much storage can be bought by a small organization or standalone users. Deletion of older videos is an inevitable part of the software , as the hard drive reaches its limit but storing as much data as possible for the least cost and highest resolution would be desirable.

A reliable software, which can somehow retrieve high resolution video from a small-sized low resolution is required to increase the effectiveness of storage devices, as the video captured is in lower resolution it not only saves storage space but also decreases latency when streaming live footage on other devices and reduces the time required for encoding the video to higher resolution.

Product Objectives

The main objectives of the project are:

- Improve on the speed of existing video super resolution techniques.
- Improve on the accuracy of existing video super resolution techniques.
- Reduce video encoding time.
- Reduce the storage requirements for any video storing or streaming system.
- Increase quality of videos.
- Reduced latency for live streaming videos.

Applications

Super-resolution imaging (SR) is a class of techniques that enhance (increase) the resolution of an imaging system.Aside from regular video quality enhancements in our mobile phones and TVs for better picture quality, the few applications of SR in today's world and the future are:

Surveillance

Nowadays, digital video recorder (DVR) devices are everywhere, and they play a significant role in applications such as traffic surveillance and security monitoring. It is, however, impossible for the moment to equip large-scale HR devices. Thus, it is necessary to study image SR techniques.Although the techniques have developed progressively, the practical use of video SR is still a challenge. Firstly,outdoor video devices are vulnerable to the impact of weather conditions. Moreover, video data usually feature a huge amount of data and complex motion. Some algorithms can deal with the motion outliers, but the computational efficiency limits their application.

Biometric information identification

SR is also important in biometric recognition, including resolution enhancement for faces ,fingerprints ,and iris images . The resolution of biometric images is pivotal in the recognition and detection process. To deal with the LR observations, a common approach is the development of high-quality images from multiple LR images. Based on the redundancy and similarity in the structured features of biometric images, example based single-frame SR with an external database is an effective way of resolution enhancement .Using SR, the details of the shapes and structural texture are clearly enhanced, while the global structure is effectively preserved,which can improve the recognition ability in the relevant applications.

Astronomical observation

The physical resolution of astronomical imaging devices limited by system parameters also provides a chance for SR techniques to play a role. Astronomical systems can typically collect a series of images for SR. By improving the resolution of astronomical images,SR can help astronomers with the exploration of outer space.

Medical diagnosis

Various medical imaging modalities can provide both anatomical information about the human body structure and functional information. However, resolution limitations always degrade the value of medical images in the diagnosis. SR technologies can be used with the key medical imaging modalities, including magnetic resonance imaging (MRI), functional MRI (fMRI), and positron emission tomography (PET) . The goal is to increase the resolution of medical images while preserving the true isotropic 3-D imaging. Medical imaging systems can be operated under highly controlled environments, and thus continuous and multi-view images can be easily acquired. Example-based SR for single frames has also been applied in the medical imaging field, by collecting similar images to establish a database.

Novelty

In the proposed project, we aim at creating a super resolution technique with improved speed and accuracy. In this technique we would try to devise an algorithm to improvise the quality of the image or video received from a video camera. We will also focus on increasing the accuracy of the model we are making by integrating different approaches we have encountered during the making of the project. In our model, the stored image files will be upscaled to a better resolution. If necessary, we might also deviate from the current approaches which we have studied and work our way up from ground level.

Scope of the project

Our goal is to design an algorithm which can efficiently and effectively solve our selected problem statement. The scope of this project is to build a system which can be used by image/video management systems to increase their effectiveness. The proposed project aims at using SR to improve the quality of videos on demand. This product will provide a better model for performing image and video super resolution.

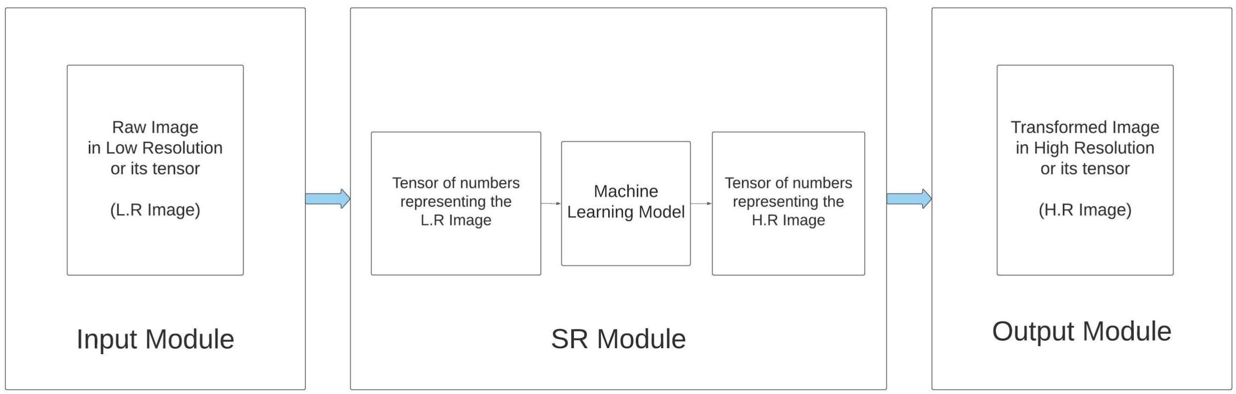
We aim at building a software which can upscale the video quality using SR techniques, the video will be stored at lower resolution to save storage space on the database server. We also seek to make this technology possible for real time video, i.e. for streaming to mobile phones and other computers for surveillance purposes.

Proposed Design

There are 3 main steps involved when approaching the super resolution problem:

- Taking input
- Feeding the input to selected model
- Getting output

We will design our application based on this simple architecture. The following diagram is a block representation of the generic project design (it will be refined as per use-case):



The first module (called the input module) is designed to take the input. The implementation of the input module can vary depending on where our solution is to be used. For example, if it has to be used by the common public to upscale images on demand, the input module will reside on either a website or an app so that it is user-friendly and accessible. On the other hand, in case this project is meant to be integrated into a larger software (in some computer vision application), we might have to accept multiple images simultaneously, so we need to be capable of directly receiving tensors of numbers as the input instead of an actual image.

The input module passes the image to the SR Module. This module first transforms the input raw image into a tensor (assuming the input given is a raw image and not a tensor) that the model can work on. Then this tensor is passed on to the trained model which generates a new tensor of numbers. This resulting tensor is then passed on to the output module. The purpose of this module is to transform the input low resolution (L.R) image into a high resolution (H.R) image, hence the name SR Module.

The output module again depends on the use-case of our tool. In case of a public use, we have to transform the tensor of the high resolution image back to an actual high resolution image and display it to the user. Whereas, in case our project has to be integrated in a larger software, it's probably better to give the tensor itself as the output.

Use Cases

The use cases of our project can be seen in different real world applications some of them which are used in our day to day activities.

PDF Scanner

Most of the documents which we make involve scanning pictures and converting them into pdf documents. In such cases, there might be a chance that the images taken on the phone are blurred or if the camera used for scanning is not a good one. This will result in making the pdf difficult to interpret and understand. Super resolution will enable us to convert the low resolution pictures to high resolution improving the quality and readability of the files.

Medical Imaging

Many medical images may be unclear due to reasons like light diffraction in X rays. This results in very unclear images and so diagnosis will be difficult. It will not be possible to check the affected part properly if we have unclear images. With super resolution, the quality of these images can be considerably improved making diagnosis and detection easier and better.

Photography

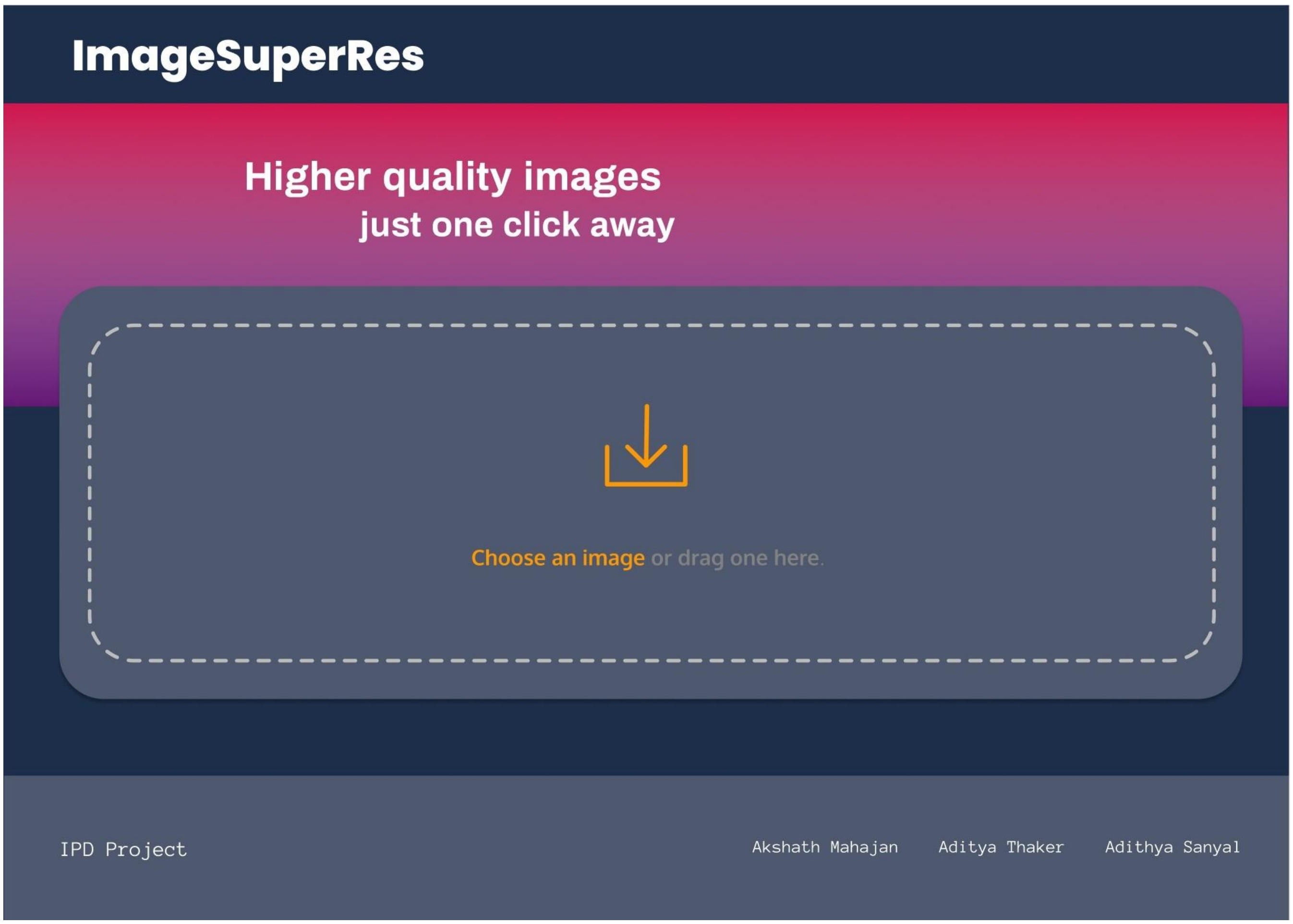
To give proper images with higher resolution, super resolution can be used to upscale the quality of images captured by the camera. Sometimes it may take place that the image may be unclear because the camera gets shaken or some other kind of disturbance. To avoid such issues which might occur we can use super resolution to up the quality.

Astronomical observations

It is necessary to obtain images of faraway planets and heavenly bodies. Hubble telescope and satellites will not always be successful in getting crystal clear images in order to facilitate and carry on with quality research. In such cases, super resolution will up the quality of the image and help us in making better progress in astronomical research.

GUI Design

The current implementation plan is to develop a stand alone project as a web service which processes images and gives the desired image file as output. Since we haven't decided yet whether to use the project as a service for other applications or to make a standalone project we have made design for a website. Below is our design draft for the website:



Module Implementation

We studied about GANs(Generative Adversarial Networks) this semester and we managed to implement a GAN network. We decided to perform super resolution using GAN techniques. To learn more about this kind of technology, we implemented a basic GAN model. This model is not the one which will be used in super resolution, but instead is being used to simply study how to implement GANs. The dataset used for this implementation is the mnist dataset in the tensorflow library of Python.

Generator model

```
def make_generator_model():
    model = tf.keras.Sequential()
    model.add(layers.Dense(7*7*256, use_bias=False,
input_shape=(100,)))
    model.add(layers.BatchNormalization())
    model.add(layers.LeakyReLU())

    model.add(layers.Reshape((7, 7, 256)))
    assert model.output_shape == (None, 7, 7, 256)  # Note: None
is the batch size

    model.add(layers.Conv2DTranspose(128, (5, 5), strides=(1,
1), padding='same', use_bias=False))
    assert model.output_shape == (None, 7, 7, 128)
    model.add(layers.BatchNormalization())
    model.add(layers.LeakyReLU())

    model.add(layers.Conv2DTranspose(64, (5, 5), strides=(2, 2),
padding='same', use_bias=False))
    assert model.output_shape == (None, 14, 14, 64)
    model.add(layers.BatchNormalization())
    model.add(layers.LeakyReLU())

    model.add(layers.Conv2DTranspose(1, (5, 5), strides=(2, 2),
padding='same', use_bias=False, activation='tanh'))
    assert model.output_shape == (None, 28, 28, 1)

    return model
```

Discriminator model

```
#The discriminator is a CNN-based image classifier.
def make_discriminator_model():
    model = tf.keras.Sequential()
    model.add(layers.Conv2D(64, (5, 5), strides=(2, 2),
padding='same',
                                     input_shape=[28, 28, 1]))
    model.add(layers.LeakyReLU())
    model.add(layers.Dropout(0.3))

    model.add(layers.Conv2D(128, (5, 5), strides=(2, 2),
padding='same'))
    model.add(layers.LeakyReLU())
    model.add(layers.Dropout(0.3))

    model.add(layers.Flatten())
    model.add(layers.Dense(1))

    return model
```

Discriminator loss

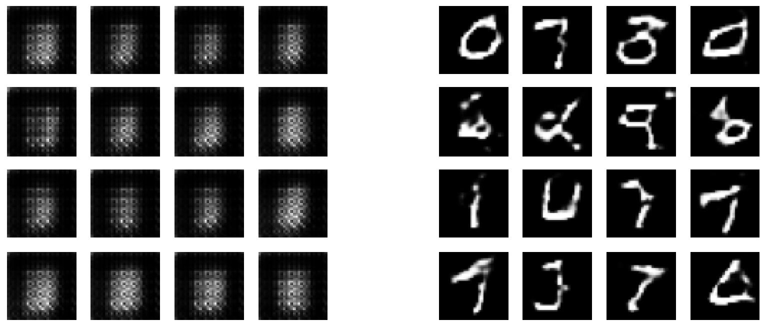
```
def discriminator_loss(real_output, fake_output):
    real_loss = cross_entropy(tf.ones_like(real_output),
real_output)
    fake_loss = cross_entropy(tf.zeros_like(fake_output),
fake_output)
    total_loss = real_loss + fake_loss
    return total_loss
```

Generator loss

```
def generator_loss(fake_output):
    return cross_entropy(tf.ones_like(fake_output), fake_output)
```

Output

Result at the first iteration (left) and after 50 iterations (right)



Datasets

After deciding to go with the GAN based approach, the first step was acquiring data to train our networks.

One of the most popular datasets for image superresolution is the Diverse 2k or DIV2K dataset. It offers about 800 high resolution images along with their low resolution counterparts for training, and another 100 for validation. Many variants of low resolution images were available and for the time being we downloaded the following data:

- DIV2k_train_HR (800 images)
- DIV2K_valid_HR (100 images)
- DIV2K_train_LR_Bicubicx2 (800 images)
- DIV2K_valid_LR_Bicubicx2 (100 images)
- DIV2K_train_LR_Bicubicx4 (800 images)
- DIV2K_valid_LR_Bicubicx4 (100 images)

Sample HR image (left) and it’s LR downgraded bicubicx4 counterpart (left)

This data was stored on the disk with the following directory structure:

- `DIV2K/Train/HR`
- `DIV2K/Train/X2`
- `DIV2K/Train/X4`
- `DIV2K/Valid/HR`
- `DIV2K/Valid/X2`
- `DIV2K/Valid/X4`

Test cases

After downloading all data, it is important to be able to get the data correctly. Meaning, it is important to be able to read the high resolution and low resolution variants of the same image at once. This will definitely be needed when training any model.
To verify that we can get different variants of the same image at once from our dataset, we use the following code

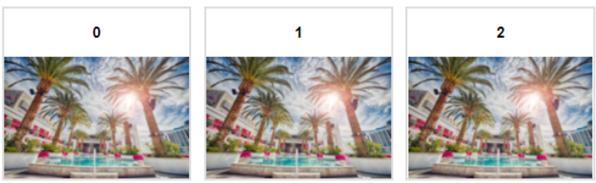
```
from PIL import Image
import os
import ipyplot

path_train_hr = "Train/HR"
path_train_x2 = "Train/X2"
path_train_x4 = "Train/X4"
train_hr = os.listdir(path_train_hr)
train_x2 = os.listdir(path_train_x2)
train_x4 = os.listdir(path_train_x4)
for i in range(3):
    arr = [
        Image.open(os.path.join(path_train_hr, train_hr[i])),
        Image.open(os.path.join(path_train_x2, train_x2[i])),
        Image.open(os.path.join(path_train_x4, train_x4[i]))
    ]
    ipyplot.plot_images(arr, max_images=3, img_width=150)
```

On running this piece of code, we obtained the following output show html

show html

show html



Results

From the above test, it is inferred that we can get the data in the desired manner. We do not require to rearrange the data as we read it, it is already present in a proper manner. When the same code is executed with more number of iterations in the for loop, we see that all HR images map perfectly to their corresponding LR variants. As we can see, all three lists that generate on calling os.listdir(filename) have the same ordering of images, the only difference being whether they are high resolution or low.

References:

[1] Pixel Recursive Super Resolution
- Ryan Dahl, Mohammad Norouzi, Jonathon Shlens

[2] RankSRGAN: Generative Adversarial Networks with Ranker for Image Super-Resolution
- Wenlong Zhang, Yihao Liu, Chao Dong, Yu Qiao - ShenZhen Key Lab of Computer Vision and Pattern Recognition, SIAT-SenseTime Joint Lab, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, University of Chinese Academy of Sciences.

[3] ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks
- Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy - CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, The Chinese University of Hong Kong, Shenzhen University of Chinese Academy of Sciences, Nanyang Technological University, Singapore

[4] A Fully Progressive Approach to Single-Image Super-Resolution
- Yifan Wang , Federico Perazzi, Brian McWilliams, Alexander Sorkine-Hornung, Olga Sorkine-Hornung, Christopher Schroers - ETH Zurich ,Disney Research

[5] Unsupervised Single Image Super-Resolution Network (USISResNet) for Real-World Data Using Generative Adversarial Network
- Kalpesh Prajapati, Vishal Chudasama, Heena Patel, Kishor Upla, Raghavendra Ramachandra, Kiran Raja, Christoph Busch - Sardar Vallabhbhai National Institute of Technology (SVNIT), Surat, India. ,Norwegian University of Science and Technology (NTNU), Gjøvik, Norway.

[6] Deep Wavelet Prediction for Image Super-resolution
- Tiantong Guo, Hojjat Seyed Mousavi, Tiep Huu Vu, Vishal Monga - School of Electrical Engineering and Computer Science, The Pennsylvania State University, State College, PA, 16803.

[7] Edge-Enhanced GAN for Remote Sensing Image Super Resolution
- Kui Jiang , Zhongyuan Wang , Member, IEEE, Peng Yi , Guangcheng Wang, Tao Lu , and Junjun Jiang - Wuhan University, Shanghai University.

[8] Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network
- Christian Ledig, Lucas Theis, Ferenc Husz´ar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi.

[9] Beyond Human-level License Plate Super-resolution with Progressive Vehicle Search and Domain Priori GAN
- Wu Liu, Xinchen Liu, Huadong Ma, Peng Cheng - Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications

[10] Generative Adversarial Network-based Image Super-Resolution using Perceptual Content Losses
- Manri Cheon, Jun-Hyuk Kim, Jun-Ho Choi, and Jong-Seok Lee - School of Integrated Technology, Yonsei University

[11] Image Super-Resolution Using Deep Convolutional Networks
- Chao Dong, Chen Change Loy, Member, IEEE, Kaiming He, Member, IEEE, and Xiaoou Tang, Fellow, IEEE

[12] Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network
- Jin Yamanaka , Shigesumi Kuwashima and Takio Kurita

[13] Accurate Image Super-Resolution Using Very Deep Convolutional Networks
- Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee

[14] Deeply-Recursive Convolutional Network for Image Super-Resolution
- Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee

[15] Enhanced Deep Residual Networks for Single Image Super-Resolution
- Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee

[16] Learning a Deep Convolutional Network for Image Super-Resolution
- Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang

[17] Learning a Deep Convolutional Network for Light-Field Image Super-Resolution
- Youngjin Yoon, Hae-Gon Jeon, Donggeun Yoo, Joon-Young Lee, In So Kweon

[18]Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network

- Wenzhe Shi, Jose Caballero, Ferenc Huszar, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, Zehan Wang, Magic Pony Technology Imperial College London

[19] Toward Real-World Single Image Super-Resolution: A New Benchmark and A New Model
- Jianrui Cai , Hui Zeng, Hongwei Yong, Zisheng Cao, Lei Zhang - The Hong Kong Polytechnic University, 2DJI Co.,Ltd, 3DAMO Academy, Alibaba Group

[20] Deep residual network with enhanced upscaling module for super-resolution. In: Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (2018)
- Kim, J.H., Lee, J.S.

[21]Real-time Super Resolution Algorithm for Security Cameras
- Seiichi Gohshi - Kogakuin University, 1-24-2 Nishi-Shinjuku, Shinjuku-ku, Tokyo, 163-8677, Japan