Image Super Resolution

Team Kota

Mentor - Gowri Lekshmy

Project Details

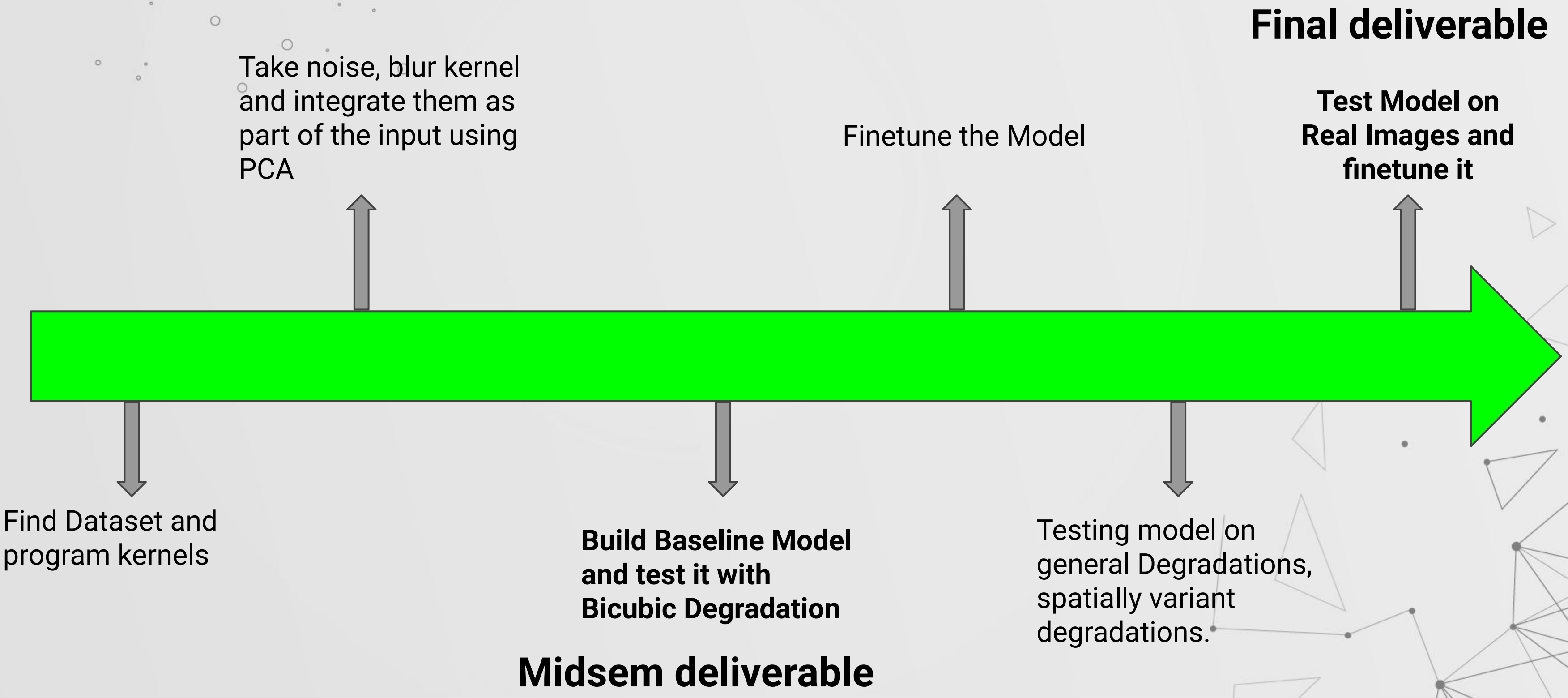
Team Name - Kota

Topic ID - 7

Members

1. Manvith Reddy - 2018101057
2. Sai Srinivasa Likhit Kota - 2018101043
3. Ainsly Dsouza - 2018101060
4. Anshul Padhi - 2018114013

Progress


Find Dataset and program kernels
Take noise, blur kernel and integrate them as part of the input using PCA
Build Baseline Model and test it with Bicubic Degradation
Midsem deliverable
Finetune the Model
Testing model on general Degradations, spatially variant degradations.
Final deliverable
Test Model on Real Images and finetune it

Method Overview

- Pose SISR using MAP framework to be a function of noise, blur kernel and the low-resolution image
- Perform dimensionality stretching on blur kernel using PCA
- Create degradation maps, which are concatenated to the original image
- This modified input is sent through a plain CNN, to show effectiveness of dimensionality stretching

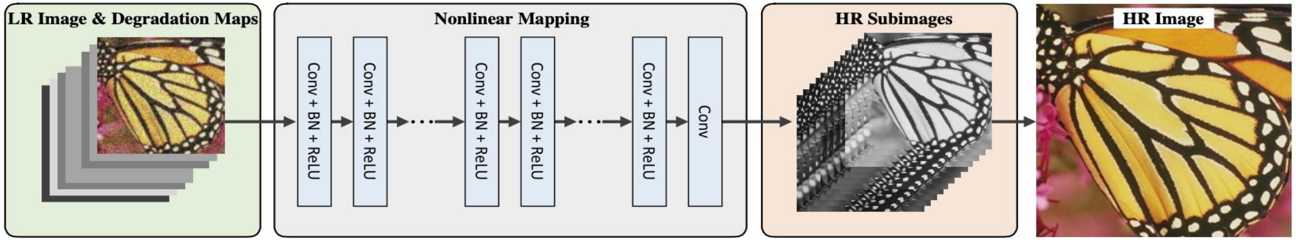


Figure 3. The architecture of the proposed convolutional super-resolution network. In contrast to other CNN-based SISR methods which only take the LR image as input and lack scalability to handle other degradations, the proposed network takes the concatenated LR image and degradation maps as input, thus allowing a single model to manipulate multiple and even spatially variant degradations.

Pipeline

Data → **Kernel** → **Train**

Data

Load Images from our Dataset to send to kernel

Kernel

Using blur kernels and different types of downsampling, we convert our hi-resolution images into low-resolution images to feed to model

Train

Pass the generated LR images through our model and optimise model based on training loss

Data

- Make a compositional dataset using 3 datasets: DIV2K, Waterloo Exploration Dataset and BSDS300
- DIV2K dataset of RGB images (only this for mid-eval)
  - Contains 800 2K resolution images
- Waterloo Exploration Dataset
  - Contains 4744 natural pristine images
- Berkeley Segmentation Dataset
  - Contains 300 pristine images

Data

- Preparing data
  - Random crop (behaves like data augmentation, improve generalisation of model)
  - Blur (applying blur kernel on random crop)
  - Downsampling happens along with random crop. We use linear and bicubic downsampling approaches.
  - Concatenate degradation map to image
    - Degradation maps are obtained using PCA
  - Now we have (x, y) = (degraded image, high res image)
- Test-Train split
  - 80/20

Kernels

To model the degradation between high resolution and low resolution images, there are two major steps: blurring and downsampling. For the blurring phase, we use gaussian blur kernels of varying standard deviations. The deviations lie anywhere in the range [0.2,0.3,0.4,...,2].

Before we begin training, the model, we create all the possible kernels (i.e, one kernel corresponding to each possible standard deviation) of size 15x15.

At train time, we use a random kernel to degrade each high resolution image and also pass on information regarding the kernel as input to the model. But this is not easy as there is a dimensionality mismatch between the image and the kernel itself.

Hence we reshape the kernels to a 125x1 vector and apply **PCA** to get a low dimensional 15x1 vector approximation of the blur kernel.

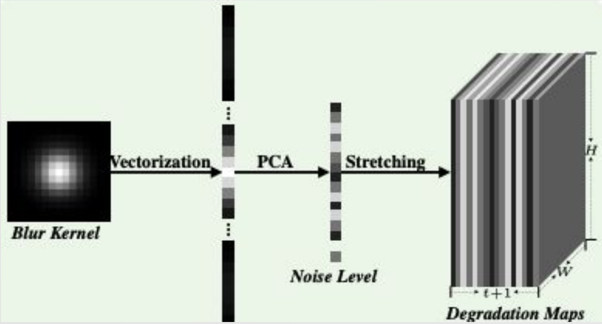
Dimensionality stretching is applied on each value of the new vector to obtain 15 degradation maps of size HxW. These Degradation maps are concatenated with the low resolution image and passed as input to the model

Kernels

Before the mid evals we had only trained our model on a downsampler of scale 2. The model performed better as we had to super-resolve our lr image by 2x.
Later we also trained our model on a scale of 3 to see if our model is powerful enough to produce accurate super resolutions at such scaled

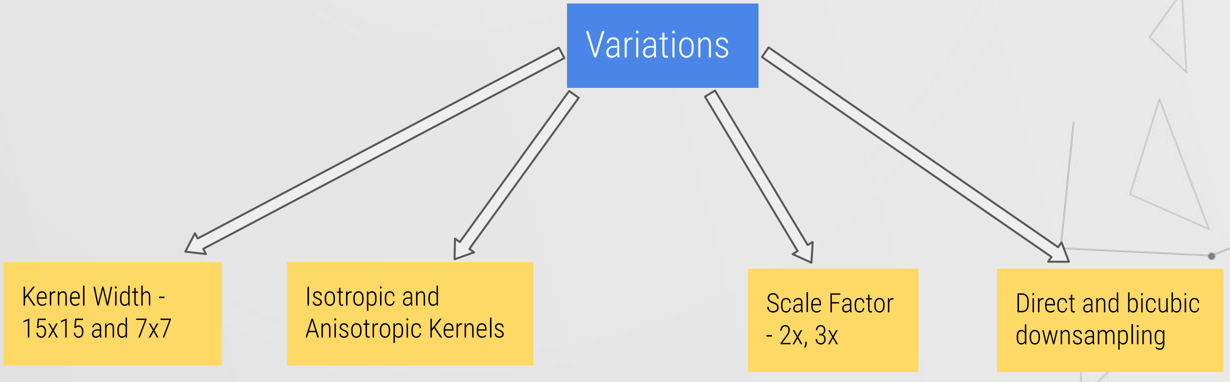
Firstly, we had trained our model on isotropic kernels which only accounted for situations in which the degradation was uniform in all directions (i.e equal std deviations along x and y axis).

In reality, the degradations can be better explained using more complex assumptions (i.e unequal standard deviations along x and y axis). To account for the same, we use anisotropic kernels which should theoretically allow our models to generalise better to real world images and provide better results.



Training

- Once LR image is created it is passed through the SRMD model
- SRMD Model is a CNN with 12 2D-Conv layers
- SRMD Model returns a hi-res version of the input
- ReLU is used as activation function
- Adam is used to optimise model over MSE loss
- Final MSE loss of our model for the dataset is 0.0011



Testing

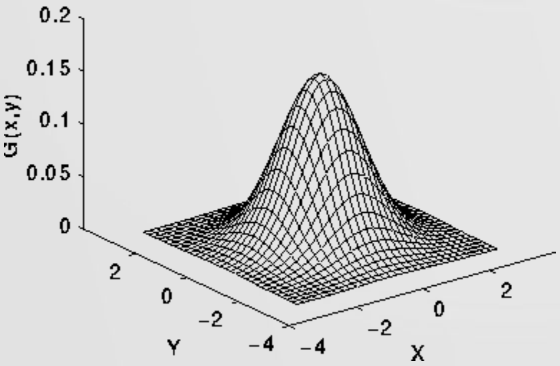
Before the mid evaluation our model was evaluated primarily on bicubic degradations. Bicubic degradations are primitive and limits the applicability for the model as real life degradations are much more complex and random. So most models before were basically trained to work on images which had bicubic degradation.

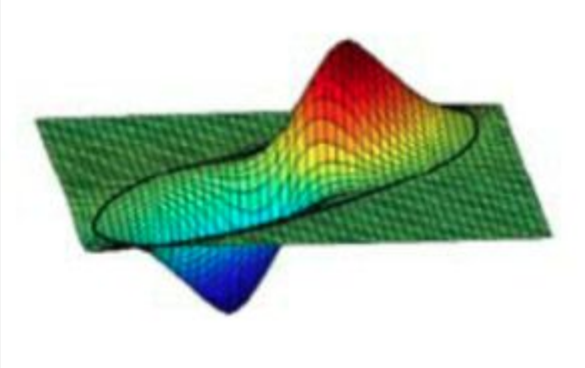
To show that our model can overcome this issue we also test our model on more general degradations. For ex. We test our model on a 7×7 Gaussian kernels and direct downsampler with scale factor 3. From the paper we saw that the older models like NCSR and IRCNN had a huge dip in performance but ours did not.

Testing

After this we also evaluated our models on spatially variant degradations. To do this we simply use spatially variant blur kernels and add Gaussian noise like before. We use a mixture of isotropic and anisotropic kernels to make it spatially variant. We see that our scores still remain fairly consistent heavily showing that the proposed model is highly generalised and this applicable for real world images.

Finally we test our model on real images. This is the best test for the applicability of our model as it doesn’t use predetermined kernels and noise levels. There is no ground truth for these images, but our predicted results seem to be quality super resolutions of the low resolution image.





Scoring Metrics

Peak signal to noise ratio (PSNR) and structural index similarity (SSIM) are two measuring tools that are widely used in image quality assessment.

Peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.

SSIM is a perception-based model that considers image degradation as perceived change in structural information, while also incorporating important perceptual phenomena

$$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right)$$

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{+}c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Results - 2x Isotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :35.67
SSIM Score :0.96

PSNR Score :35.22
SSIM Score :0.93

Results - 2x Isotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :34.86
SSIM Score :0.93

PSNR Score :35.78
SSIM Score :0.94

Results - 2x Anisotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :34.21
SSIM Score :0.911

PSNR Score :33.55
SSIM Score :0.92

Results - 2x Anisotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

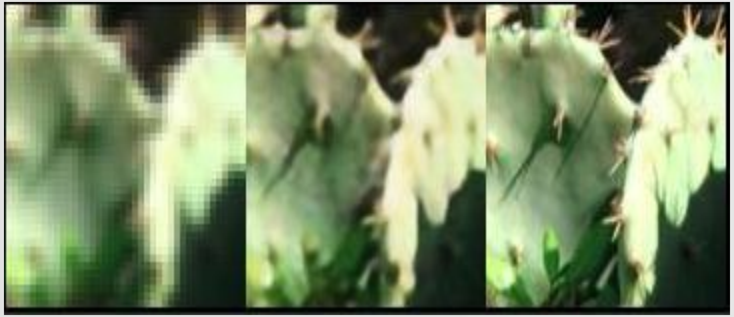
PSNR Score :33.43
SSIM Score :0.901

PSNR Score :33.19
SSIM Score :0.905

Results - 3x Isotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :33.05
SSIM Score :0.905



PSNR Score :32.6
SSIM Score :0.89

Results - 3x Isotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :32.01
SSIM Score :0.907

PSNR Score :31.61
SSIM Score :0.881

Results - 3x Anisotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :31.11
SSIM Score :0.886

PSNR Score :30.7
SSIM Score :0.891

Results - 3x Anisotropic

Left: LR image (input) ; Middle:HR image from model (prediction); Right:Actual HR image (ground truth)

PSNR Score :29.15
SSIM Score :0.886

PSNR Score :29.08
SSIM Score :0.87

Results on Real Images

Scores

Scale	Degradation	PSNR	SSIM	PSNR Paper Results Best Score	SSIM Paper Results Best Score
2x	Bicubic	34.57	0.92	37.53	0.959
2x	Spatially Variant	31.22	0.89	N/A	N/A
3x	Bicubic	30.91	0.88	33.86	N/A
3x	Spatially Variant	28.76	0.86	N/A	N/A
2x	General	32.14	0.904	33.74	N/A
2x	7x7 Kernel	32.07	0.908	32.59	N/A