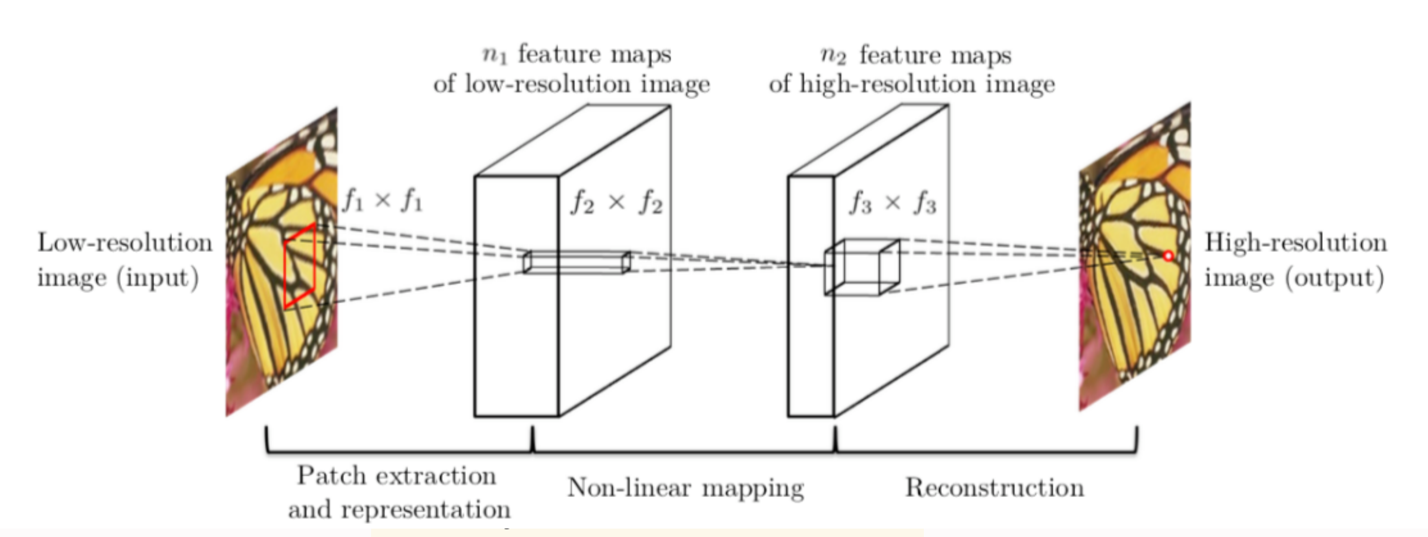
Image Super-Resolution Via a Convolutional Neural Network

Machine Learning Project

Mahendra Kausik V – PES1UG23AM163, Purandar Puneet - PES1UG23AM906

Motivation and Overview

Image super-resolution attempts to counteract the effects of faulty imaging hardware or software degradation by interpolating between pixels. A super-resolution convolutional neural network (SRCNN) uses a pair of convolutional layers--a feature extraction layer and a feature reconstruction layer--to relate patches of low-resolution pixels to higher-resolution improvements.



Dataset: T91 and Set5

•91-image is collection of 91 high-quality images, preprocessed into smaller patches to train the SRCNN model used for training.

•Each image is downscaled by a specific scale factor (×2, ×3, ×4).

•Then bicubic interpolation is applied to upscale it back to the original size producing a low-resolution (LR) input.

•The high-resolution (HR) counterpart is the original image.

•Small patches (typically 33×33 HR / 21×21 LR) are extracted and stored in .h5 format for efficient batch loading during training.

•Set5 is a small dataset specifically used for evaluating and testing the performance of trained super-resolution models

•Contains 5 high-quality natural images:

baby, bird, butterfly, head, woman

•Similar to the training set:

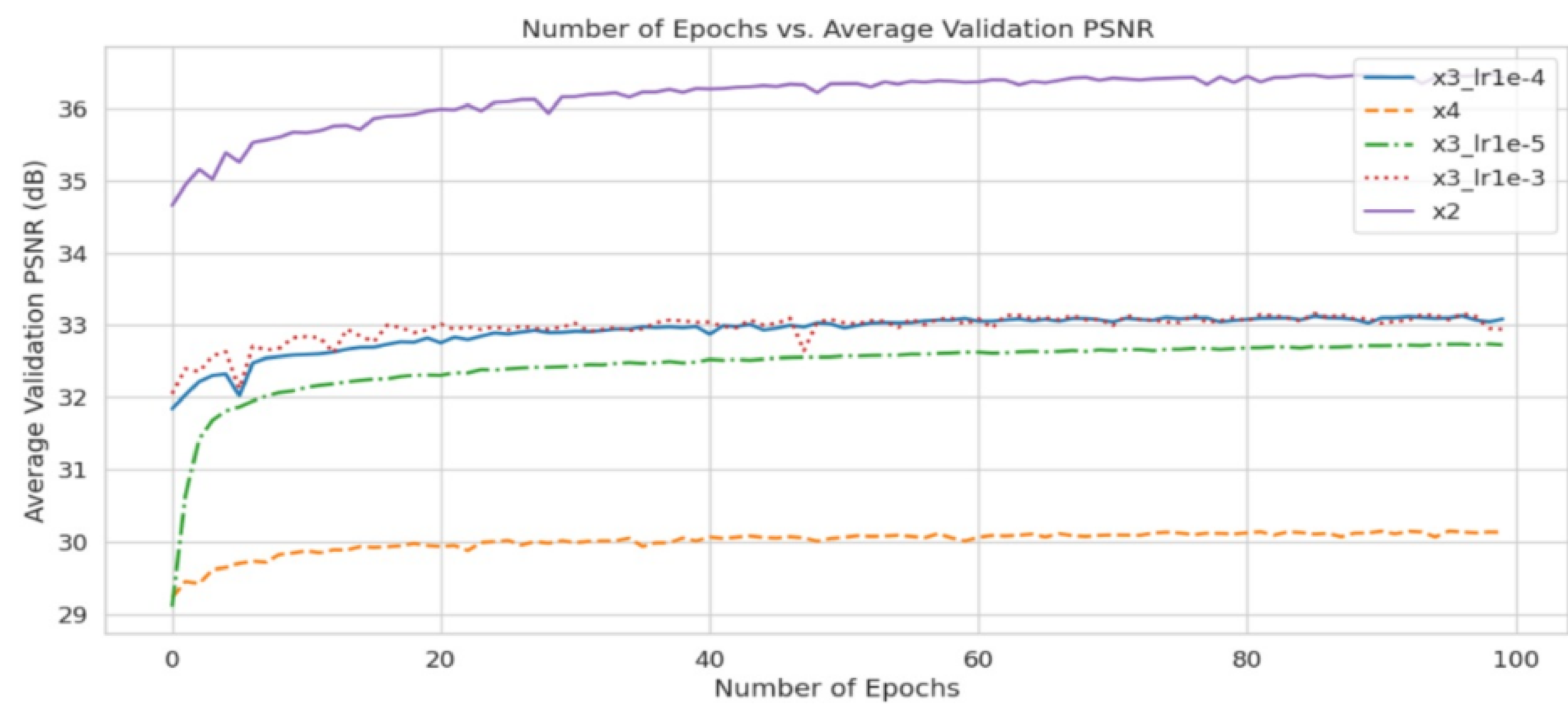
Each HR image is downscaled and then bicubically upscaled.

Stored as .h5 for quick evaluation.

Experiments

- Metric: peak signal-to-noise ratio $PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right)$
- Learning rate of 0.0001 was most effective
  - highest test PSNR: 33
- Learning Rate Shows Classic Behavior:
  - The learning rate of 0.001 is too high. Its performance is erratic and unstable, jumping up and down without converging properly.
  - The learning rate of 0.00001 is too low. The PSNR increases rapidly in the start but continues to stay below 0.001
  - The learning rate of 0.001 provides the best balance, achieving the highest PSNR quickly and maintaining stability. This confirms it's the optimal choice among the options tested.
- Lower the Upscaling factor, higher the PSNR
- For a 3x upscale, the model trained with a learning rate of 1e−3 (x3_lr1e-3) is technically the winner, but the 1e-4 model is a very close and reliable alternative.

Number of Epochs vs. PSNR for Hyperparameters



Model Summary

Model	Best PSNR	Best Epoch	Final Loss
x2	36.4782	98	0.000543483
x3_lr1e-3	33.1602	85	0.00113819
x3_lr1e-4	33.1287	96	0.00116326
x3_lr1e-5	32.7378	98	0.00126744
x4	30.1533	95	0.00183

Results

Bicubic LR

SRCNN Output

Future Work

- Test SRCNN with larger datasets like DIV2K which has 800 training examples.
- Test response with noise added to input images
- With GPU availability, test U-Net and ResNet model ability to add more detail

References

Garber, B., Grossman, A., & Johnson-Yu, S. (2020). *Image Super-Resolution Via a Convolutional Neural Network*. Stanford University.