CSE445.03
Group #4

Md. Misbah Khan
ID: 2132089642

Md Abdula Al Shyed
ID: 2212592042

Rakibul Islam
ID: 2212058642

Raju Ahamed Rabby
ID: 2212592042

I. Introduction

Super-Resolution is a method in machine learning that involves increasing the resolution of an image by generating missing pixels from low-resolution input. The goal is to produce an output image with a higher resolution than the input image. And our goal is to preserve the original content and structure [1]. For this project we selected 3 machine learning model and 1 deep learning model. They are XgBoost [2], CNN [3], Random Forest [4] and SRGAN [5]. We split each model to each one of the members of the group. We used the same dataset for all model.

II. Dataset Preparation

We Downloaded 100 high-resulation images from pexel [6]. After that we saved them to `data/highRes` folder with sequential filenames (HR001.jpg, HR002.jpg, etc.). Then we under-sampling these data by the factor of 10 using high-quality Lanczos resampling filter. We preprocesses high-resolution images by reducing their dimensions using the Pillow library. The processed images are then saved in a separate folder `data/lowRes` with sequential filenames (LR001.jpg, LR002.jpg, etc.). So the pair for each high and low resulation image is differentiate by HR and LR.

III. The Problem

Now how can we upscale an image? The simplest way to do so is to spread out all the pixels and then fill the blank pixels value with the closest pixel value. KNN [7] model work like this. But that just makes a image with bigger pixel rather than a high resulation image. We can solve this problem a bit but not taking exact values of the nearest neighbors, rather taking their weighted average of neighbor pixels. This method is known as Bilinear Interpolation [8].

But according to information theory Processing data cannot add information content. This is known as Data Processing inequality [10].

$$X \rightarrow Y \rightarrow Z$$

$$I(X;Y) \geq I(X;Z)$$

So is super-resulation model is an impossible task? No. We need some additional source of information to recreate these information. Or we should say make hesitation of the recreation of information. This is where models like CNN comes into play. CNN trains itself with a large amount of data. Using this knowledge-base this model can learn how to hallucinate these detail. Dong *et al.* proposed an approach that simply train a convolutional Neural Network(CNN) that can fill information in low resulation image in this way. After that they take the min squared error and minimize it [11]. But this also does not solve the problem properly.Because mean squared error only cares about pixel wise intensity differences, but not the structural differences of the content of an image. For measuring the output quality of an image, we use metrics like structural Similarity Index Measurements(SSIM) [12] and PSNR [13]. There are some other pretrained model like SRGAN [5] and ESRGAN [14]. But they all are deep learning based.

Fig. 1. KNN approach [9].

Fig. 2. Bilinear Interpolation approach [9].

IV. Our Approaches

In this section we will talk about how we intended to solve the problem. As we discussed before, all work in this field was done with different deep learning model. So it was quite hard to find a machine learning model that can do our job, as not all of our teammate has idea of deep learning. We choose 4 models initially. For the final implementation, we will choose the model with best performance. We also used some AI tools like ChatGPT [15] and DeepSeek [16] for our research and coding help.

A. *XgBoost*

XgBoost, short for eXtreme Gradient Boosting, is an open-source machine learning library. It is widely used for supervised learning tasks such as regression and classification.

- **Why we Chose This Model:**
  - XgGBoost is fast, efficient, and works well with GPU.
  - It captures non-linear relationships between low and high-resolution image patches.
- **Problems:**
  - High memory usage with large patches.
  - Risk of overfitting with limited data (100 images).

B. *Random Forest*

Random forest is a widely-used machine learning algorithm that operates by constructing an ensemble of decision trees to improve predictive performance.

- **Why we Chose This Model**
  - Computationally efficient for smaller tasks.
  - Easy to interpret model predictions.
- **Problems of This Model**
  - Less effective for high-quality image super-resolution.
  - Takes too much time.

C. *CNN*

CNNs are specialized deep learning models designed for processing structured datalikes images, by utilizing convolutional layers to automatically learn spatial hierarchies of features.

- **Why I Chose This Model:**
  - To enhance image quality by improving the resolution of low-quality images.
  - To leverage advancements in deep learning for better image reconstruction.
- **Problems:**
  - The model requires high computational resources for training and inference.
  - Our dataset is too small, only 100 images.

D. *SRGAN*

Super-Resolution Generative Adversarial Networks (SRGANs) are deep learning models designed to enhance image resolution by generating high-quality images from low-resolution inputs.

- **Why I Chose This Model:**
  - High-Quality Generation: Enhances images with fine details.
  - GAN Framework: Uses generator and discriminator for realism.
- **Problems:**
  - High Computational Cost and Long Training Time.
  - Dependence on Data Quality.

References

[1] P. with Code, "Super-resolution," 2025, accessed: 2025-02-20. [Online]. Available: https://paperswithcode.com/task/super-resolution

[2] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 785–794. [Online]. Available: https://arxiv.org/abs/1603.02754

[3] Y. LeCun, "Generalization and network design strategies," in *Connectionism in Perspective*, R. Pfeifer, Z. Schreter, F. Fogelman, and L. Steels, Eds. Zurich, Switzerland: Elsevier, 1989, pp. 1–14, an extended version was published as a technical report of the University of Toronto. [Online]. Available: http://yann.lecun.com/exdb/publis/pdf/lecun-89e.pdf

[4] T. K. Ho, "Random decision forests," in *Proceedings of the 3rd International Conference on Document Analysis and Recognition*. IEEE, 1995, pp. 278–282. [Online]. Available: https://ieeexplore.ieee.org/document/598994

[5] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 105–114. [Online]. Available: https://arxiv.org/abs/1609.04802

[6] Pexels, "Pexels api - free stock photos and videos api," 2025, accessed: 2025-02-24. [Online]. Available: https://www.pexels.com/api/

[7] E. Fix and J. L. H. Jr., "Discriminatory analysis. nonparametric discrimination: Consistency properties," Randolph Field, Texas, 1951, report Number 4, Project Number 21-49-004. [Online]. Available: https://www.jstor.org/stable/1403797

[8] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it?" *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009. [Online]. Available: https://ece.uwaterloo.ca/~z70wang/publications/SPM09.pdf

[9] Computerphile, "How super resolution works," 2018, accessed: 2025-02-24. [Online]. Available: https://www.youtube.com/watch?v=KULkSwLk62I

[10] Z. Goldfeld, "Lecture 7: Data processing inequality," 2020, scribed by Aaron Lou, Net ID: al968; Assistant Editor: Kia Khezeli. [Online]. Available: https://people.ece.cornell.edu/zivg/ECE_5630_Lectures7.pdf

[11] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *arXiv preprint*, vol. arXiv:1501.00092, 2015. [Online]. Available: https://arxiv.org/abs/1501.00092

[12] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. [Online]. Available: https://ece.uwaterloo.ca/~z70wang/publications/ssim.pdf

[13] D. Sethi, S. Bharti, and C. Prakash, "A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work," *Artificial Intelligence in Medicine*, vol. 129, p. 102314, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365722000793

[14] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, "ESRGAN: Enhanced super-resolution generative adversarial networks," *arXiv preprint arXiv:1809.00219*, 2018. [Online]. Available: https://arxiv.org/abs/1809.00219

[15] OpenAI, "Chatgpt," 2023, accessed: 2025-02-24. [Online]. Available: https://chat.openai.com/

[16] DeepSeek, "Chat with deepseek ai," 2025, accessed: 2025-02-24. [Online]. Available: https://chat.deepseek.com/