

Assignment 1

Applied Deep Learning

Helena Cooper 12505729

October 21, 2025

1 Choice of topic and project type

I will be doing a computer vision project specifically on image generation/reconstruction of images. The project type will be 'bring your own method' where I will use an existing dataset and then build a neural network architecture that takes the dataset as input.

2 Project idea

For this project I will experiment with reconstructing astronomical images. The idea is to take an existing dataset and reduce the resolution of the images by blurring them. The low-resolution or blurred input images will be processed by a U-Net autoencoder to produce high-resolution reconstructions. If time permits, I would like to experiment with Vision Transformers by combining this with the U-Net autoencoder.

The dataset I intend to use is "ESA Hubble Images - 3 classes" from Kaggle

(link: https://www.kaggle.com/datasets/subhamshome/esa-hubble-images-3-classes?select=galaxies_heic1110a.jpg).

The dataset contains three classes (galaxies, nebula, solarsystem) with approximately 200-300 images per class.

The approach is fully end-to-end: the model should learn to map bad quality images to their original high quality. Evaluation will include appropriate metrics and manual comparison of the images.

3 Workplan

The following is a work-breakdown schedule for the project. I've based the schedule on an expected workload of 8 hours per week.

- **Dataset collection & preprocessing:** Download dataset, dataloader, and preprocessing (blurring, creating more sample etc.) **12 hours**
- **Model design & implementation:** Implement U-Net autoencoder, experiment whether ViT is feasible/realistic **18 hours**
- **Training & fine-tuning:** Train baseline U-Net, evaluate on validation set, hyperparameter tuning **25 hours**
- **Further experiments experiments:** Either continue experimenting with just U-Net or do further experiments with ViT **15 hours**
- **Evaluation & visualization:** Find relevant ways to evaluate the output **10 hours**
- **Application / demo:** Build interactive demo, not sure how yet **8 hours**
- **Report writing** **8 hours**
- **Presentation preparation** **5 hours**

Total estimated time: 101 hours ($\tilde{8}$ hours/week \times 12–13 weeks)

4 References

- Dosovitskiy, A., et al., “*An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*,” ICLR 2021.
- Zhang, K., et al., “*Image Super-Resolution Using Deep Convolutional Networks*,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.
- Lim, B., et al., “*Enhanced Deep Residual Networks for Single Image Super-Resolution*,” CVPR Workshops, 2017.