INTEL UNNATI INDUSTRIAL TRAINING INTERNSHIP

Problem statement 2

Image Sharpening using knowledge distillation

Objective-

Develop a model to enhance image sharpness during video conferencing, addressing issues like reduced clarity due to low bandwidth or poor internet connections.

By-

Akshat Singh

Enrolement-230905250

Batch of 23-27

Manipal Institute of Technology

Acknowledgement

This project was carried out by me under the guidance of my mentor, Dr. Roopashree Shetty as part of my industrial training internship program at Intel. I would like to express my sincere gratitude to Dr. Shetty for her continuous guidance, encouragement and technical support throughout the course of this project.

I also extend my heartfelt thanks to Intel for providing me with the opportunity and resources to work on a real-world research problem. This project served as a significant learning experience, helping me deepen my understanding of knowledge distillation, computer vision, and model optimization.

It also gave me valuable exposure to industry-grade practices in research, model evaluation, and deployment.

Akshat Singh
230905250
MIT manipal
Batch of 23-27

Image Sharpening using Knowledge Distillation

This project implements an image sharpening system using Knowledge Distillation, where a powerful pretrained Restormer model acts as a teacher and a lightweight Residual UNet model is trained as a student to sharpen blurred images efficiently. The aim is to produce sharper images with fewer computational resources while maintaining high quality.

Models

Teacher Model — Restormer

Restormer is a transformer-based architecture specifically designed for image restoration tasks such as deblurring and denoising. It features a multi-stage encoder–decoder structure and long-range feature modeling through self-attention.

- Source: https://github.com/swz30/Restormer
- Task: Motion Deblurring
- Checkpoint: motion_deblurring.pth
- Used for inference only (teacher guidance)

Student Model — Residual UNet

The student model is a compact Residual UNet architecture designed to balance efficiency and performance. Key components include:

• Encoder and decoder paths with downsampling and upsampling

• Residual blocks inserted into each stage to preserve fine details

• Skip connections between encoder and decoder layers to retain spatial context

• Final output layer with 3 channels for RGB image restoration

Trained using L1 + KD Loss

Efficient, lightweight & fast to train

Loss Function Used:

combined two types of loss:

- **L1 Loss between student output and ground truth**
- **L1 Loss between student output and teacher (Restormer) output**

Formula:

TotalLoss = $\alpha \times$ L1(Student, GroundTruth) + $\beta \times$ L1(Student, TeacherOutput)

Where:

- $\alpha = 0.8$
- $\beta = 0.2$

Dataset

The dataset used in this project consists of 1800 blurred–sharp image pairs. The sharp images are high-quality samples, while the blurred counterparts are synthetically generated using Gaussian blur with σ = 0.5. The dataset is divided into: • 900 original blurred–sharp pairs • 900 additional pairs obtained by cropping sharp images and generating corresponding blurred versions All images are resized to 256×256 pixels to ensure compatibility with model input requirements and to reduce memory usage during training

- Source: Custom dataset on Kaggle
- Total: **1800 image pairs**
  - 900 original (sharp + blurred)
  - 900 additional cropped & blurred images
- Blur Type: **Gaussian Blur**, σ = 0.5
- Format: .png, paired image names (e.g. 001.png in both folders)

The code and notebooks used have all been linked below

GITHUB REPO- https://github.com/akshat0817/Knowledge-distillation

TRAINING

- Framework: PyTorch
- Platform: Kaggle Notebook
- Batch size: 4
- Image size: 256×256
- Optimizer: Adam (learning rate = 1e-4)
- Epochs: 5
- Evaluation metrics: SSIM (Structural Similarity Index)

RESULTS

Evaluated the performance of the student model trained via Knowledge Distillation from the Restormer teacher model on a dataset of 1800 sharp–blur image pairs. The student model was trained for 5 epochs using a combination of L1 loss and distillation loss. The training loss steadily decreased as follows:

Epoch	AvgLoss
1	0.0447
2	0.0142
3	0.0121
4	0.0109
5	0.0100

After training evaluated the student model on the test set using two widely accepted image quality metrics:

- Structural Similarity Index (SSIM): 93.72%

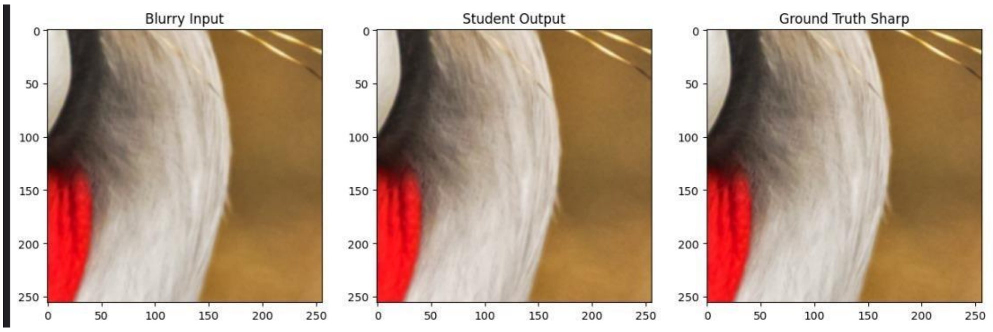

Fig3- Comparison of blurry input, student output, and ground truth.

REFERENCES

Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M.-H., & Shao, L. (2022). Restormer: Efficient Transformer for High-Resolution Image Restoration. In CVPR 2022. GitHub: https://github.com/swz30/Restormer

scikit-image: Image processing in Python. Documentation: https://scikit-image.org/

https://arxiv.org/abs/2111.09881