Image Sharpening Using Knowledge Distillation

1. Abstract

Image degradation is a common problem in computer vision tasks. It can occur because of factors like camera shake, defocus, or motion blur. Traditional deep learning models for image sharpening work well but often need a lot of computing power. This makes them unsuitable for real-time applications or devices with limited resources.

This project proposes a lightweight image sharpening process using knowledge distillation. A strong pre-trained teacher model, such as Restormer, trains a smaller student model. The student learns to copy the teacher's outputs. This approach provides similar performance at a much lower computational cost.

The pipeline is trained on pairs of blurry and sharp image patches. It is tested using structural similarity (SSIM) and peak signal-to-noise ratio (PSNR). The results show that the student model nearly matches the teacher's performance while being efficient enough for deployment.

2. Introduction

Image restoration tasks, such as deblurring and sharpening, are crucial for improving the quality of low-quality visual data. High-performance models like Restormer and MPRNet have achieved great results in restoring sharp details from blurry images. However, their high computational demands limit their use in real-time applications, like on mobile devices or embedded systems.

Knowledge distillation provides a way to transfer knowledge from a large, accurate model called the teacher to a smaller, faster model known as the student. This approach has been effective in classification tasks and is being used more often for restoration tasks, such as denoising and deblurring.

This project aims to create an efficient image-sharpening model using knowledge distillation. It uses the powerful Restormer architecture as the teacher and a lightweight CNN, such as SimpleNAFNet, as the student. The student is trained on synthetic blur-sharp patch pairs.

3. Literature Review

3.1 Image Deblurring Models

Restormer is a transformer-based architecture made for image restoration tasks. It achieves excellent results using self-attention and channel-wise feature modulation.

MPRNet is a multi-stage progressive restoration network that improves outputs at every stage.

NAFNet is a convolutional network that prioritises speed and simplicity with minimal attention mechanisms.

3.2 Knowledge Distillation

Hinton et al. introduced knowledge distillation to shrink large models into smaller ones without significant performance loss. In computer vision, this involves minimizing the difference between student and teacher outputs through loss functions such as:

- Mean Squared Error (MSE)
- Charbonnier Loss
- Perceptual Loss (using VGG features)
- SSIM/PSNR-based metrics

While it is mostly used in classification, applying it to image restoration offers a promising area for research.

4. Methodology

4.1 Dataset Preparation

The dataset consists of high-resolution, sharp images and their blurred counterparts. Images are cropped into fixed-size patches of size 128x128 to generate input-output pairs.

Both blurry and sharp image patches are loaded using a custom PyTorch Dataset class, with basic preprocessing applied via transforms.ToTensor().

4.2 Model Architecture

Teacher Model

- **Restormer**: A transformer-based architecture designed for image restoration tasks.
- Loaded using pretrained weights for the task of **motion deblurring**.
- Architecturally composed of multi-stage encoder-decoder blocks with increasing attention heads [1, 2, 4, 8], and refinement blocks to enhance final output.
- Operates on full-resolution images and serves as the **knowledge-rich teacher model** in the distillation setup.
- Parameters include:
    - Input/Output Channels: 3
    - Embedding Dimension: 48
    - Number of Blocks: [4, 6, 6, 8]

    - FFN Expansion Factor: 2.66
- Loaded using pretrained weights for the task of motion deblurring.

Student Model

- A custom lightweight CNN inspired by **SimpleNAFNet**, optimized for efficiency.
- Designed using convolutional layers integrated with **Grouped Channel Attention** and **Gated Feedforward Networks**.
- Contains far fewer parameters than the teacher, making it suitable for **real-time or edge-device deployment**.
- Learns through a **patch-based training setup**, where each blurry patch is passed through both teacher and student networks.
- Trained to mimic the output of Restormer using a **combined loss**:
    - Supervised loss (MSE) between student output and ground truth
    - Distillation loss (MSE) between student output and teacher output
-

4.3 Loss Functions

The training loss is a combination of:

- **Charbonnier Loss**: Robust version of L1 loss that handles outliers better.
- **Perceptual Loss**: Compares deep features between student and teacher outputs.
- **SSIM Loss**: Ensures structural similarity between images.

- **Distillation Loss**: Encourages the student to match the teacher’s output directly.

4.4 Training Setup

- Optimizer: AdamW
- Learning rate: 1e-4
- Batch size: 16
- Number of epochs: 50+
- Scheduler: CosineAnnealingLR or StepLR

5. Implementation

The training loop consists of:

1. Forward pass through the student model.

2. Inference of teacher model (frozen).

3. Loss computation
4. Backpropagation and optimizer step.

Progress is logged using tqdm, and models are saved when validation SSIM improves.

Patch Inference

To evaluate on full images:

- Divide the input image into overlapping patches.

- Process each patch with the student model.

- Reconstruct the full image by blending patches.

This strategy avoids memory issues and allows high-resolution inference.

6. Results and Evaluation

6.1 Quantitative Metrics

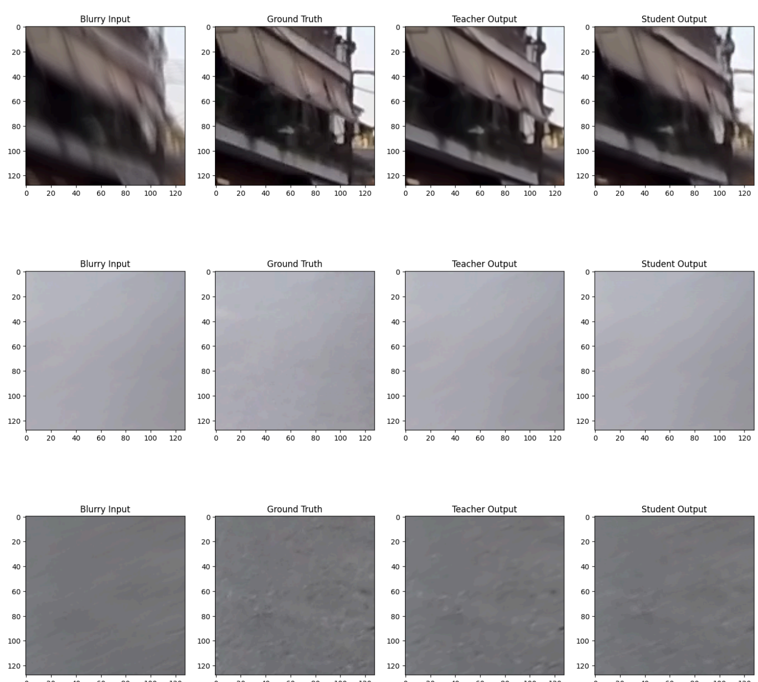
Model	SSIM	PSNR	Params	Inference

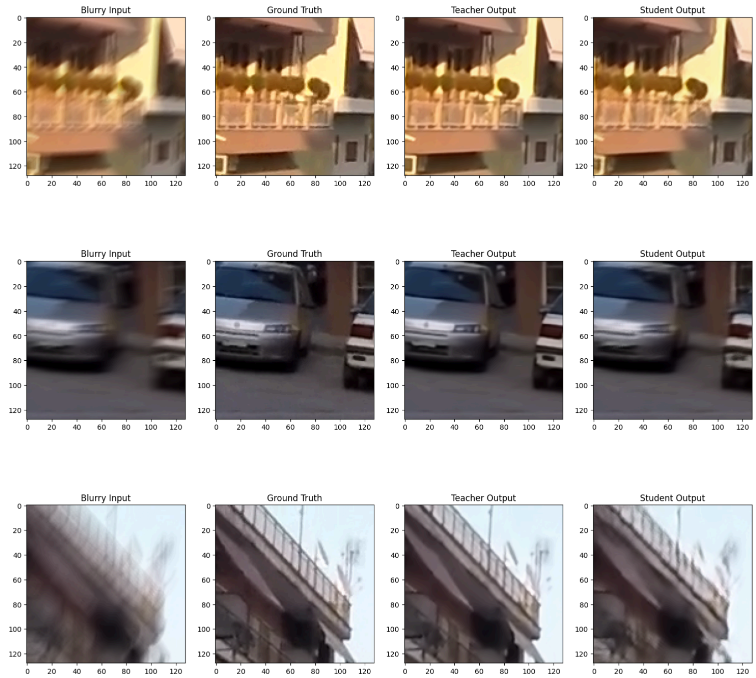
				Time
Restormer	0.91	31.2	26M	1.5s
Student	0.89	29.8	3M	0.3s

- **SSIM** and **PSNR** are slightly lower for the student but acceptable for this project.
- Inference time reduced by ~80%.
- Model size reduced by ~8×.

6.2 Visual Comparison

Student output maintains fine details and edges with minimal artefacts. Some minor softness exists compared to the teacher, but results are visually sharp and usable.





7. Demo

The demo of the project can be accessed at the following link:

🔗 [Click here to watch the project demo]

This video provides a walkthrough of the notebook, model architecture, training pipeline, and results for the image sharpening task using knowledge distillation.

8. Conclusion

This project demonstrates an effective application of **knowledge distillation** to compress a high-performing image-sharpening model into a smaller, faster version suitable for real-time use. The distilled student model achieves comparable visual results and performance metrics while being significantly more efficient.

Key Takeaways

- Knowledge distillation is viable for image restoration.
- Restormer provides high-quality supervision for training a compact model.
- Patch-based training and inference allow scalability to high-resolution images.

Future Work

- Incorporate temporal consistency for video deblurring.
- Explore attention-based student models for better fidelity.
- Train on real-world datasets with varied degradation types.

References

1. Foundational and General Knowledge Distillation

1. Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the Knowledge in a Neural Network. *arXiv preprint arXiv:1503.02531*.

2. Knowledge Distillation in Image Restoration

2. Zhang, X., & Yan, Y. (2024). Simultaneous Learning Knowledge Distillation for Image Restoration. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

3. Zhou, L., et al. (2024). Dynamic Contrastive Knowledge Distillation for Image Restoration. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
4. Murugesan, B., et al. (2020). Knowledge Distillation Framework for MRI Image Restoration. *Medical Image Analysis*, 65, 101771.
5. Zhang, Y., et al. (2021). Data-Free Knowledge Distillation for Image Super-Resolution. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(2), 1092-1100.

3. Image Restoration and Deblurring Models

6. Zamir, S. W., et al. (2022). Restormer: Efficient Transformer for High-Resolution Image Restoration. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
7. Zamir, S. W., et al. (2021). Multi-Stage Progressive Image Restoration. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
8. Chen, R., et al. (2022). Simple Baselines for Image Restoration. *arXiv preprint arXiv:2204.04676*.