

---

# POVa Project Report

---

**Luís Alves (xalvesl00)**  
Brno University of Technology  
xalvesl00@stud.fit.vutbr.cz

**Tiago Portugal**  
Brno University of Technology  
tiagodiasportugal@tecnico.ulisboa.pt

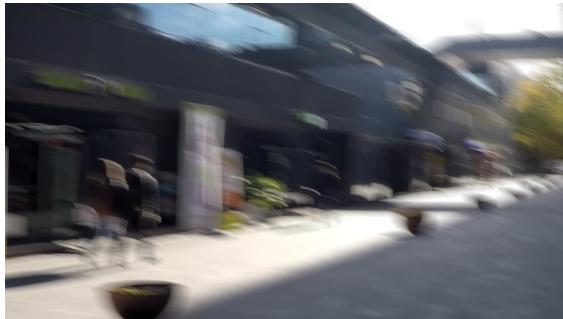
**Přemek Janda (xjanda28)**  
Brno University of Technology  
xjanda28@stud.fit.vutbr.cz

## Image Deblurring

Task definition: *Implement an existing algorithm which can sharpen images with motion blur. It may have information about the motion as an input.*

## 1 Introduction

Image deblurring is a fundamental computer vision task aimed at recovering sharp images from blurred observations (Fig. 1). In dynamic scenes, such as those featured in the **GoPro dataset** [2], blur is often non-uniform and spatially variant, making the restoration an ill-posed problem.



(a) Blurred Observation ( $y$ )



(b) Sharp Ground Truth ( $x$ )

Figure 1: GoPro dataset sample showing dynamic motion blur.

This project evaluates the transition from classical signal processing to modern deep learning architectures. We compare deterministic, non-blind filters against data-driven models, focusing on their ability to generalize across complex motion patterns.

The report is organized as follows: Section 2 provides the *Preliminary Work*; Section 3 details the *Dataset and Preprocessing*; Section 4 describes the *Implementation*; Section 5 presents the *Evaluation and Results*; and Section 6 ends with the *Conclusions*.

## 2 Preliminary Work

The image degradation process is modeled as a convolution operation:

$$y = (x * k) + n \quad (1)$$

where  $x$  is the latent sharp image,  $k$  is the Point Spread Function (PSF),  $n$  represents additive noise, and  $y$  is the blurry observation. Recovery of  $x$  is an **ill-posed inverse problem** due to the loss of high-frequency information during blurring.

Restoration approaches are categorized into two paradigms:

- **Traditional Methods:** Rely on mathematical inversion. *Non-blind deconvolution* (e.g., Wiener filtering, Richardson-Lucy [1]) assumes a known kernel  $k$ .
- **Deep Learning:** Architectures like **U-Net** utilize an encoder-decoder structure with **skip connections** to learn direct blurry-to-sharp mappings without explicit kernel estimation.

Restoration quality is quantified via standard metrics:

1. **PSNR (Peak Signal-to-Noise Ratio):** Measures pixel-level fidelity based on Mean Squared Error (MSE).
2. **SSIM (Structural Similarity Index):** Evaluates structural preservation, luminance, and contrast to better align with human perception.

### 3 Dataset and Preprocessing

The **GoPro Deblur Dataset** provides high-resolution RGB pairs consisting of blurred observations and sharp ground truths.

#### 3.1 Data Partitioning

The training pool (2,103 pairs) was subdivided into a 90/10 ratio for training and validation. The independent testing pool (1,111 pairs) was reserved for final evaluation.

Split	Pool Source	Percentage	Samples
Training	Training Set	≈59.0%	1,893
Validation	Training Set	≈6.5%	210
Testing	Test Set	≈34.5%	1,111
<b>Total</b>		100%	3,214

Table 1: Dataset distribution across GoPro pools.

#### 3.2 Deep Learning Pipeline

The models utilize a stochastic Torchvision pipeline:

- **Configuration:** Training over 50 epochs with a batch size of 16 and a  $10^{-3}$  learning rate.
- **Augmentation:** Synchronized *Random Crop* ( $256 \times 256$ ) and *Random Horizontal Flip* (50% probability).
- **Normalization:** Conversion to `float32` tensors in  $(C, H, W)$  format, normalized to  $[0, 1]$ .
- **Validation:** Deterministic *Center Crop* ( $256 \times 256$ ) without augmentation.

#### 3.3 Traditional Methods Pipeline

Deterministic image-by-image processing implemented for baseline methods:

- **Scope:** Conducted on the 1,111 test images.
- **Processing:** Fixed *Center Crop* ( $256 \times 256$ ).
- **Conversion:** Tensors permuted to NumPy ( $H, W, C$ ) format and cast to `float64` to ensure numerical stability during deconvolution.

### 4 Implementation

Development was centered on **PyTorch** for deep learning and **scikit-image** for metrics. Experiments were conducted on **NVIDIA Tesla T4** GPUs via Kaggle.

## 4.1 Deep Learning Architectures

Four architectures were evaluated over 50 epochs:

- **DeblurringSimple:** Single-layer baseline ( $5 \times 5$  kernel).
- **DeblurCNN:** 5-layer residual network with global skip connections ( 150K parameters).
- **UNet:** Encoder-decoder with 4 downsampling stages and feature-map skip connections ( 7.7M parameters).
- **DeblurGANv2:** Generator featuring 8 residual blocks and Sigmoid activation ( 600K parameters).

## 4.2 Traditional Methods

Deterministic baselines were implemented using **SciPy** and **NumPy**:

- **Wiener Deconvolution:** Frequency-domain filtering with a  $15 \times 15$  kernel.
- **Richardson-Lucy:** Bayesian iterative deconvolution (50 iterations,  $\sigma = 1.0$ ).
- **Stochastic Deconvolution:** MCMC-based restoration utilizing a diagonal motion PSF.

## 5 Evaluation and Results

### 5.1 Evaluation Protocol

Restoration quality was assessed using **PSNR** (pixel fidelity) and **SSIM** (structural similarity). Quantitative metrics displayed in the comparative charts (Figs. 2 and 5) represent **averages computed over 20 randomly sampled images** from the GoPro test set.

### 5.2 Deep Learning Performance

Quantitative results for the neural models are detailed in Table 2. The **DeblurCNN** achieved the highest global performance with a Test PSNR of **24.40 dB**.

Model	Best Val PSNR (dB)	Test PSNR (dB)
DeblurringSimple	25.21	23.80
UNet	25.22	23.84
<b>DeblurCNN</b>	<b>25.67</b>	<b>24.40</b>
DeblurGANv2	25.34	23.97

Table 2: Final results summary for deep learning architectures evaluated on the full test set.

To analyze specific restoration behavior, a secondary evaluation was conducted on a randomized 20-image sample (Table 3). Despite a high blurred baseline of 27.08 dB, only the **DeblurCNN** achieved a **positive improvement** (+0.11 dB). Figure 2 visualizes these averaged metrics.

Model	Avg. PSNR (dB)	Improvement (dB)
Blurred Baseline	27.08	—
DeblurringSimple	26.75	-0.34
UNet	25.96	-1.12
<b>DeblurCNN</b>	<b>27.20</b>	<b>0.11</b>
DeblurGANv2	26.50	-0.59

Table 3: Average PSNR metrics and restoration improvement over a 20-image random sample.

Qualitative assessment is provided through visual comparisons of a **single representative test image**. Figure 3 demonstrates that the residual connections in DeblurCNN effectively recover spatial details that purely encoder-decoder models fail to preserve in this specific case.

### 5.3 Model Convergence

Stable convergence was observed over 50 epochs. Early stopping prevented overfitting, with validation PSNR plateauing notably for the DeblurCNN model, as shown in Figure 4.

Deep Learning Models - Performance Metrics Comparison

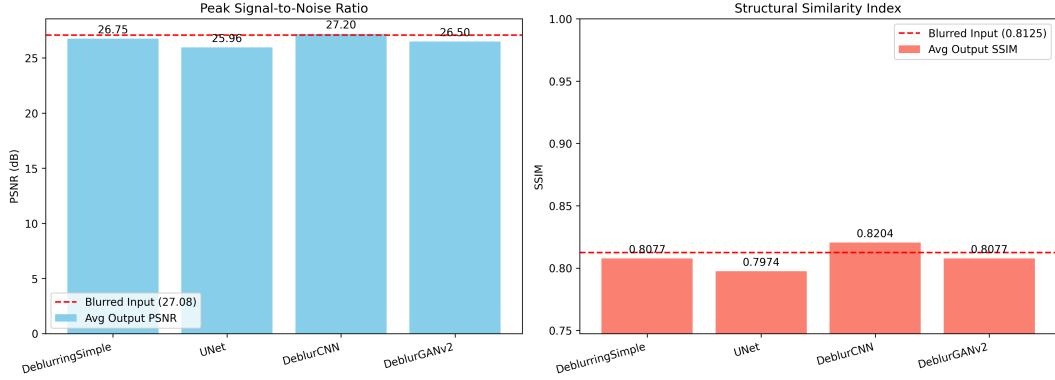


Figure 2: Averaged PSNR and SSIM performance across the 20-image sample subset.

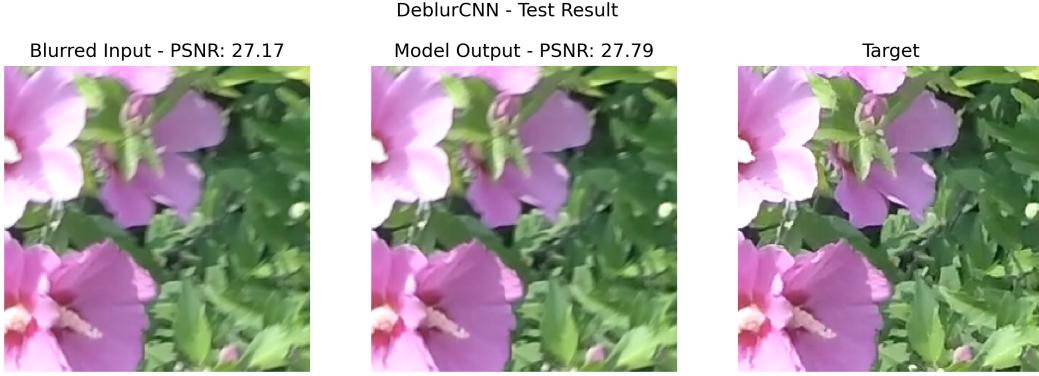


Figure 3: Visual restoration results for a single representative image using DeblurCNN.

#### 5.4 Traditional Methods Performance

Traditional restoration filters were evaluated on a randomized 20-image sample to establish a performance baseline. As shown in Table 4 and in Figure 5, all deterministic methods resulted in a **negative PSNR improvement**, effectively degrading the image quality relative to the blurred input.

Method	Avg. PSNR (dB)	Avg. SSIM	Improvement (dB)
Blurred Baseline	23.89	0.7173	—
Wiener	21.94	0.5753	-1.95
<b>Richardson-Lucy</b>	<b>22.99</b>	<b>0.6502</b>	<b>-0.90</b>
Stochastic	20.25	0.5221	-3.64

Table 4: Quantitative evaluation of traditional methods (Averaged over 20 images).

The failure of these methods is attributed to the complexity of non-uniform motion blur in the GoPro dataset. Visual evidence (Fig. 6) confirms that assuming a static blur kernel leads to severe **ringing artifacts** and high-frequency noise amplification, particularly evident in the Stochastic and Wiener approaches.

#### 5.5 Comparative Analysis Summary

The performance analysis across the 20-image randomized sample highlights the following technical observations:

- **DeblurCNN Robustness:** This was the only architecture to achieve a positive restoration delta (+0.11 dB).
- **High-Capacity Model Limitations:** More complex models like **UNet** resulted in negative deltas (-1.12 dB).

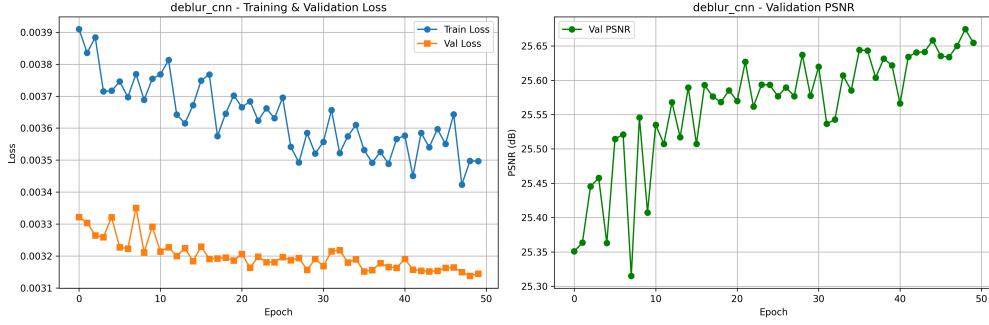


Figure 4: Training and validation PSNR curves for the best performing model.

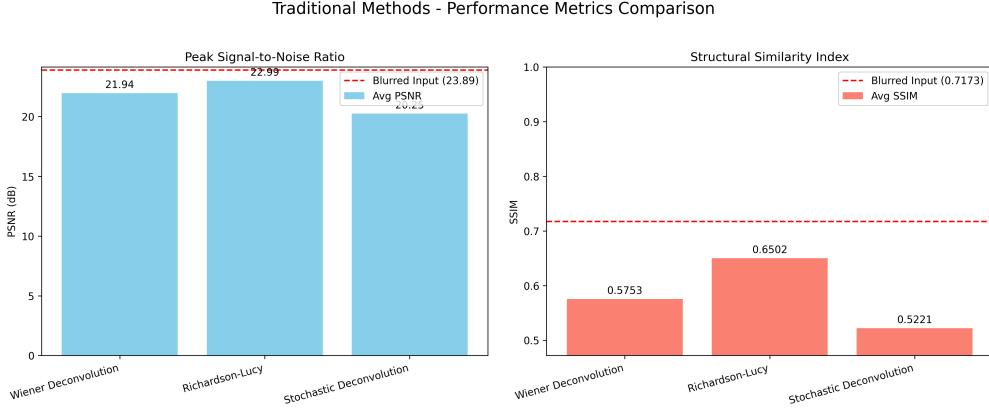


Figure 5: Average PSNR and SSIM metrics for traditional methods across the 20-image sample.

- **Traditional Methodology Failure:** Deterministic filters were consistently unable to model the non-uniform blur, resulting in significant quality degradation (up to  $-3.64$  dB) and visible ringing artifacts.

## 6 Conclusions

This work benchmarked several deep learning and traditional signal processing approaches for image deblurring. While deep learning demonstrated a clear advantage in handling dynamic scenes, the results represent a baseline for further optimization.

### Main Findings:

- **Architectural Performance:** Neural architectures, led by the **DeblurCNN** at **24.40 dB Test PSNR**, outperformed traditional mathematical inversion methods by learning spatially-variant mappings.
- **Determinism vs. Data-Driven:** Traditional filters are fundamentally limited by static Point Spread Function (PSF) assumptions, which lead to high-frequency noise amplification in real-world motion scenarios.
- **Restoration Challenges:** The frequent occurrence of negative PSNR deltas in the evaluation samples underscores that deblurring remains a highly sensitive task where model choice is critical.

**Future Work:** To improve upon these results, future efforts should focus on:

1. Exploring **Transformer-based architectures** to better capture long-range spatial dependencies.
2. Integrating **Perceptual and Adversarial losses** to improve visual sharpness beyond pixel-level MSE metrics.
3. Investigating **Real-time optimization** to allow for practical deployment in consumer photography or autonomous systems.

### Visual Comparison of Deblurring Methods



Figure 6: Visual comparison showing artifacts introduced by traditional restoration filters.

## References

- [1] Scikit image development team. Image deconvolution: Richardson-lucy and wiener filters, 2024. Accessed on: December 23, 2025.
- [2] Adwyth Darsan R. Gopro image deblurring dataset, 2023.

## 7 Annexes

### 7.1 Training and Validation Diagnostics

This section presents the comprehensive diagnostic curves for all deep learning architectures evaluated in this study. The curves illustrate the stability of the **PSNR** metric across 50 training epochs on the **Kaggle Tesla T4** infrastructure.

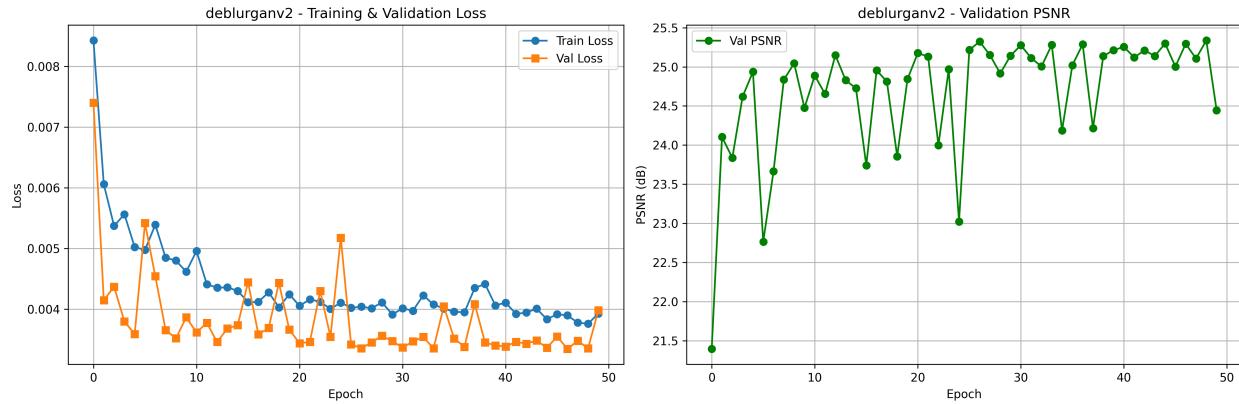


Figure 7: Training and validation PSNR curves for DeblurGANv2 model

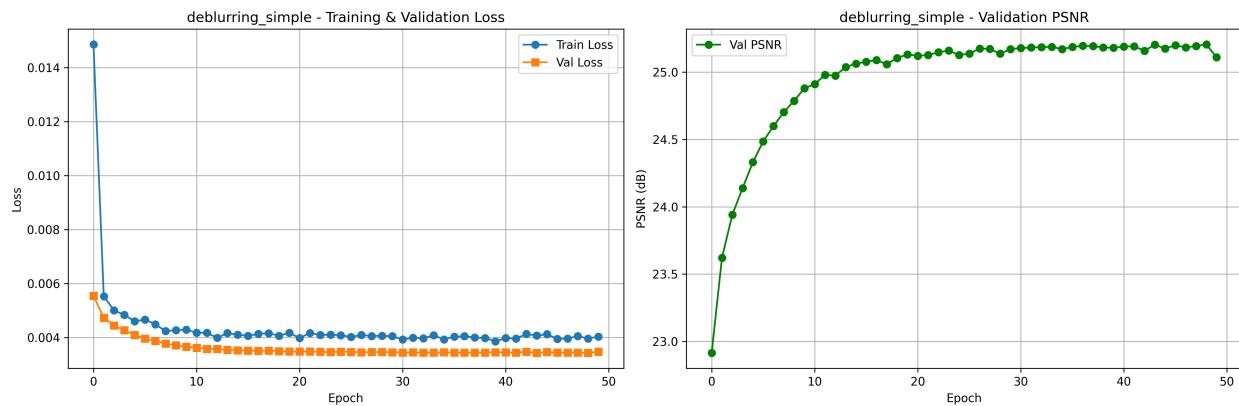


Figure 8: Training and validation PSNR curves for DeblurringSimple model

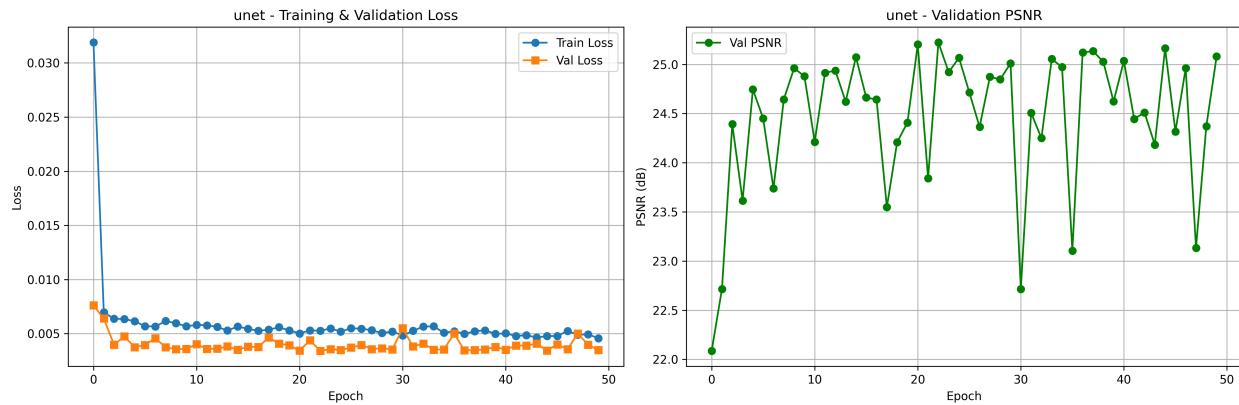


Figure 9: Training and validation PSNR curves for UNet model

## 7.2 Supplementary Qualitative Comparisons

The figures below provide additional visual outputs for a **single test image index**. These side-by-side comparisons show the **Blurred Observation** (left), the **Model Prediction** (center), and the **Sharp Ground Truth** (right).



Figure 10: Visual comparison for a single image sample: DeblurringSimple model.

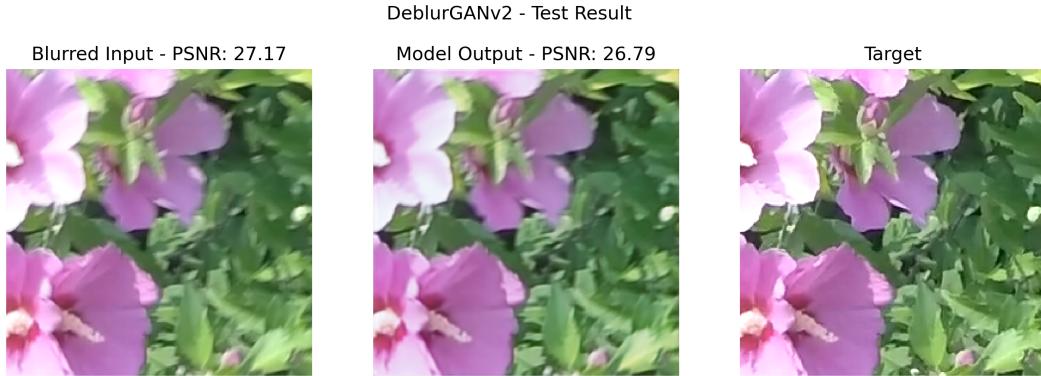


Figure 11: Visual comparison for a single image sample: DeblurGANv2 model.

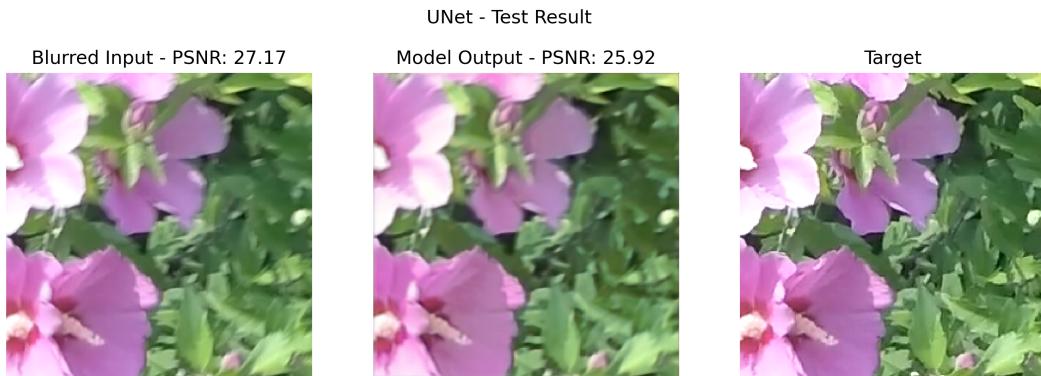


Figure 12: Visual comparison for a single image sample: UNet model.