The Alexander Kofkin
Faculty of Engineering
Bar-Ilan University

Noisier2Noise

Learning to Denoise from Unpaired Noisy Data

Samy Nehmad
Yaniv Hajaj

Introduction to Image Denoising

Introduction

Image denoising is an important process used in various fields such as **healthcare, space exploration, and photography**. Its main goal is to clarify images by removing 'noise' (random variations in brightness or color that distort the true content).

Challenges with Traditional Methods

Traditional denoising techniques require large datasets of paired noisy and clean images. However, **obtaining clean images is often impractical**, especially in real-world scenarios like space photography and medical imaging.

Innovation with Noise2Noise

Noise2Noise improved image denoising by using two noisy versions of the same image instead of clean data. This method effectively trains denoising networks without clean images, but it **requires two noisy captures**, which isn't always possible.

Innovation with Noisier2Noise

Noisier2Noise addresses these challenges by **using only a single noisy image and a statistical noise model**, eliminating the need for clean images. This method allows for effective denoising even when clean data is unavailable.

The Noisier2Noise Method

01

Problem Statement

The Noisier2Noise method was developed to overcome the limitations of previous denoising techniques by **requiring only a single noisy image and a statistical model of the noise**. It is particularly useful in situations where **multiple noisy images or clean images are impractical to obtain**.

02

Methodology

Noisier2Noise adds synthetic noise to an already noisy image and trains a neural network to predict the original noisy image. **The network learns to distinguish between the original and synthetic noise**, allowing it to reconstruct a cleaner image.

the system describe as follows:

$$X \rightarrow^{+N} Y \rightarrow^{+M} Z$$

The Noisier2Noise algorithm employs a neural network that **learns to map a noisier input Z to a less noisy target Y by minimizing the expected mean squared error**, represented mathematically as:

$$E_z[\| f(Z;\theta) - Y \|_2].$$

Given that the network never sees N or M in isolation, **the ideal approach of simply subtracting M from Z isn't feasible.**

the system describe as follows:

$$X \rightarrow^{+N} Y \rightarrow^{+M} Z$$

The Noisier2Noise Method

Given that the network never sees N or M in isolation, **the ideal approach of simply subtracting M from Z isn't feasible.**

Instead, the best strategy is to predict $E[Y \mid Z]$ By leveraging the relationship $\mathrm{E}[\mathrm{Y} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} + \mathrm{N} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{N} \mid \mathrm{Z}]$ and noting the independence and identical distribution of M and N, we derive:

$$E[Y \mid Z] = E[X + N \mid Z] = E[X \mid Z] + E[N \mid Z] \quad \rightarrow^{\mathrm{M,N} \sim \mathrm{A\ iid}}$$

$$2\mathrm{E}[\mathrm{Y} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{N} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{N} \mid \mathrm{Z}] \rightarrow^{\mathrm{E}[\mathrm{N} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{M} \mid \mathrm{Z}]}$$

$$2\mathrm{E}[\mathrm{Y} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + (\mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{N} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{M} \mid \mathrm{Z}]) \quad \rightarrow$$

$$2\mathrm{E}[\mathrm{Y} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{X} + \mathrm{N} + \mathrm{M} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{E}[\mathrm{Z} \mid \mathrm{Z}] = \mathrm{E}[\mathrm{X} \mid \mathrm{Z}] + \mathrm{Z}$$

Thus, we can extract an estimate of the clean image $E[X \mid Z]$ by calculating:

$$\boldsymbol{2E[Y \mid Z] - Z}$$

We can therefore recover an estimate of the clean image by **doubling our network's output and subtracting the noisier version.**

Improvements on the Method

Overlap Image Technique

We introduced an overlapping technique during testing to improve the denoising process. This technique involves generating multiple predictions of the denoised image and then aggregating these predictions using mean or median operations to produce a final result.

our system describe as follows:

$$X \rightarrow^{+N} Y \rightarrow^{+M} \begin{matrix} Z_1 \\ Z_2 \\ \cdot \\ \cdot \\ \cdot \\ Z_{k-1} \\ Z_k \end{matrix}$$

Our experiments and Results

We experimented with different overlap techniques to refine the denoising process:

01

Mean

The mean aggregation method was tested as a baseline for combining multiple predictions during overlap.

02

Median

The median technique, which takes the middle value of the predictions, was also evaluated. This method provided a good balance between preserving details and reducing noise.

03

Trimmed Mean

The trimmed mean technique aims to enhance denoising quality by discarding the highest and lowest predictions before averaging, minimizing the impact of extreme values.

$$\hat{X} = \frac{1}{k}\sum_{i=1}^{k} f(Z_i)$$

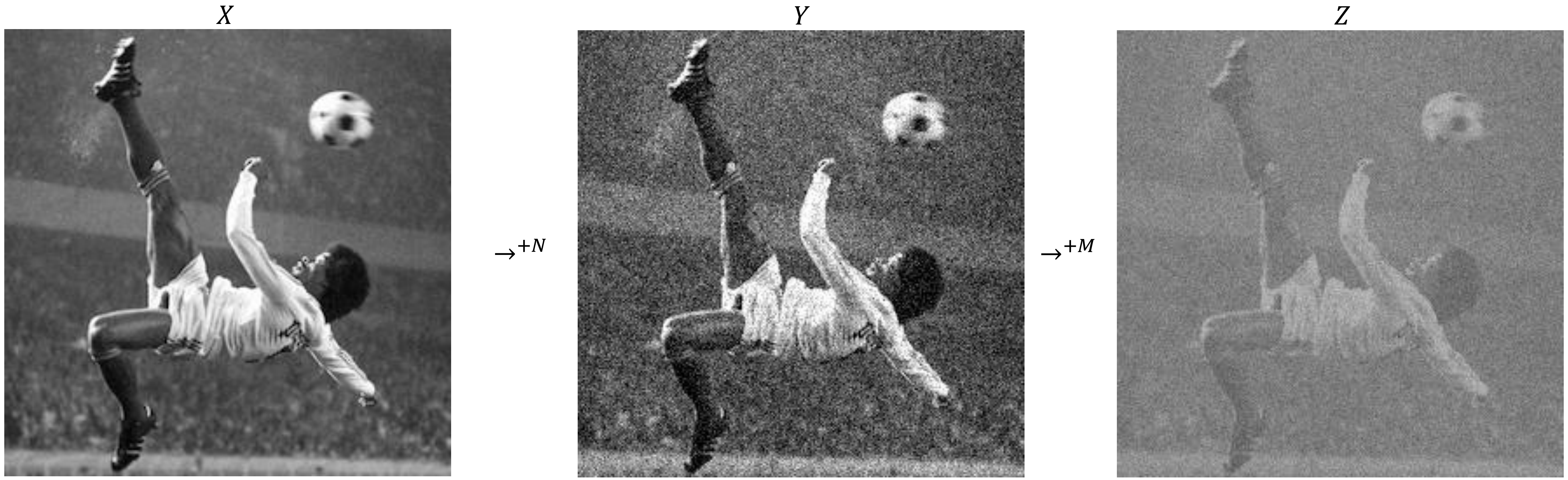
$$\hat{X} = \mathrm{median}\{f(Z_1), f(Z_2), \ldots, f(Z_k)\}$$

$$\hat{X} = \frac{1}{(1-2T)k}\sum_{i\in\text{Trimmed Set}} f(Z_i)$$

* "Trimmed Set" refers to the middle $(1-2T)K$ % of predictions.

Our experiments and Results

We experimented with different overlap techniques to refine the denoising process (Gaussian):



prediction k=1

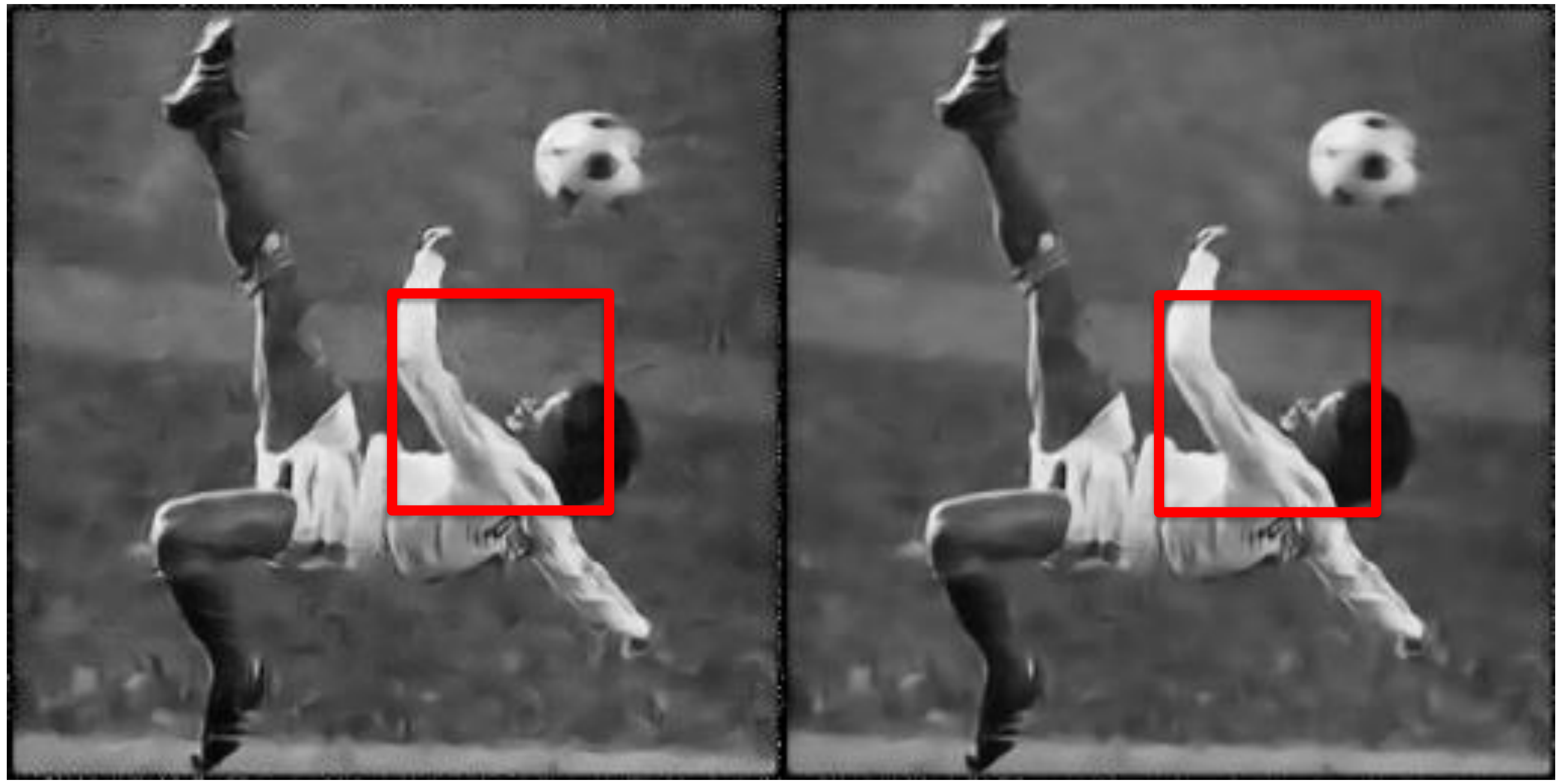
Mean k=30

Median k=30

Trimmed Mean k=30

Visual Results

Comparative visual results demonstrated that Noisier2Noise, especially with the **overlap technique, provided superior denoising, with clearer and more detailed images**.



prediction k=1 Mean k=30

prediction k=1 75X75 pixels Mean k=30 75X75 pixels

Performance Metrics

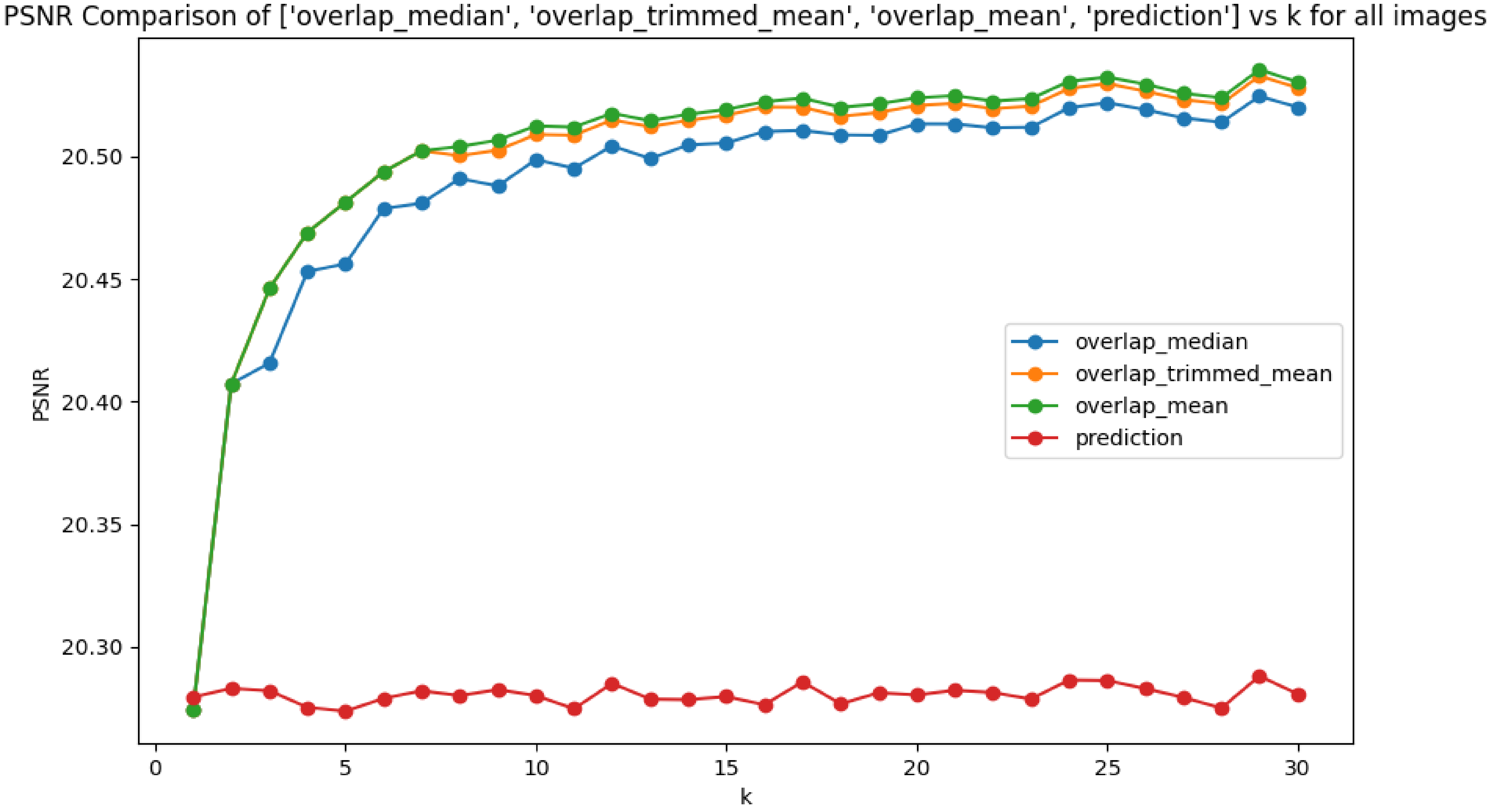
Graphs: PSNR

$$\mathrm{PSNR} = 10 \times \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$$

PSNR **(Peak signal-to-noise ratio)** A higher PSNR value suggests a lower level of error and noise, indicating better image quality. **High PSNR values are generally associated with images that preserve fine details closely** resembling the original.

* MAX is the maximum possible pixel value of the image (255 for 8-bit grayscale image).

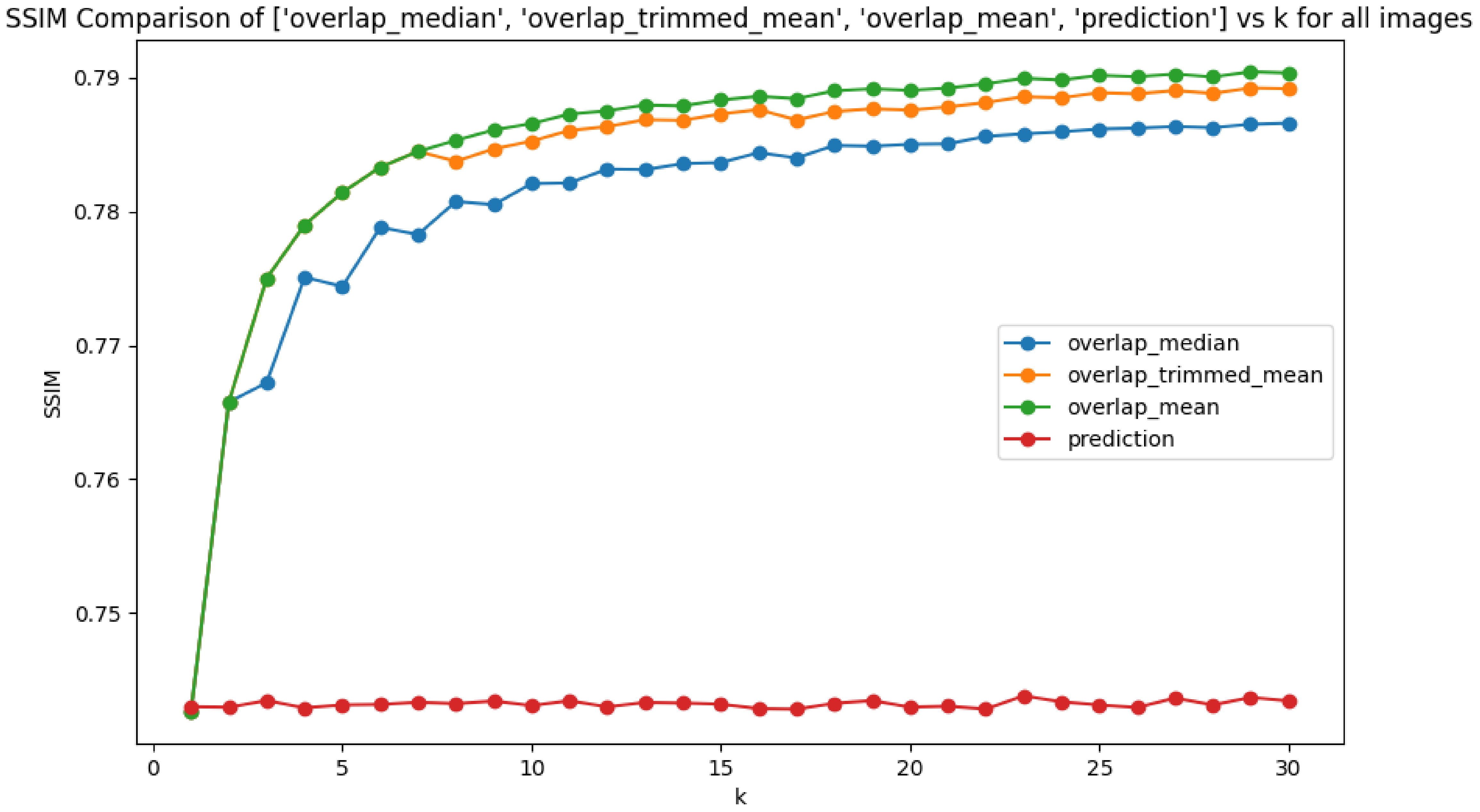
* MSE is the mean squared error between the original and the processed image.



Graphs: SSIM

Performance Metrics

SSIM **(structural similarity index measure)** measures the visual impact of an image's characteristics: brightness, contrast, and structure. It **compares local patterns of pixel intensities** that have been adjusted for brightness and contrast. **SSIM values range between -1 and 1, where 1 indicates perfect similarity.** A high SSIM value suggests that the structure, brightness, and contrast of the image are well-preserved compared to the original. **This metric strongly correlates with visual and perceptual quality as experienced by human observers.**



Conclusions

01

Summary of Findings

Noisier2Noise successfully denoises images **using only a single noisy realization and a statistical noise model. It eliminates the need for clean training data**, making it suitable to a wide range of real-world scenarios.

02

Limitations

The method's effectiveness **depends heavily on the accuracy of the noise model** used during training. Additionally, the computational **complexity of the overlapping** technique may limit its ability to handle larger-scale datasets.

03

Future Work

To improve Noisier2Noise, future research could **explore adaptive noise modeling**, more **efficient processing techniques,** and improved handling of **different noise types**.

04

Final Thoughts

Noisier2Noise represents a significant advancement in image denoising, particularly in **situations where clean data is unavailable.** While there are challenges to address, such as noise model dependency and computational efficiency, the **method shows great promise for practical applications.**