

Universal and High-Fidelity Resolution Extending for Fluorescence Microscopy Using a Single-Training Physics-Informed Sparse Neural Network ¹

Mauricio Ramos

Facultad de Ingeniería
Universidad de la República

Seminario PINN versión 2024

¹<https://doi.org/10.34133/icomputing.0082>.

Contexto

- ▶ El paper propone un framework al que denomina DPS (deep-physics-informed sparsity) cuyo objetivo es mejorar la resolución física por lo menos en un factor de 1.67, para una variedad de modalidades de imágenes.
- ▶ El framework contiene varias redes de neuronas que actúan de manera coordinada sobre un flujo de procesamiento y fueron entrenadas con una estrategia de aprendizaje sin supervisión.
- ▶ Se presenta la red Res-U-DBPN, la cual contiene en su función objetivo un término asociado a la PSF² del proceso de captura.
- ▶ El entrenamiento únicamente utiliza una base de datos de imágenes SIM³ con diferentes tipos de muestras biológicas, con el objetivo de mejorar la universalidad de la red.

²Point Spread Function

³Structured Illumination Microscopy

Foco de la exposición

- ▶ Entender los diferentes componentes de las redes que componen el framework así cómo las funciones de costo utilizadas en su entrenamiento.

Image Kernels

input image

output image

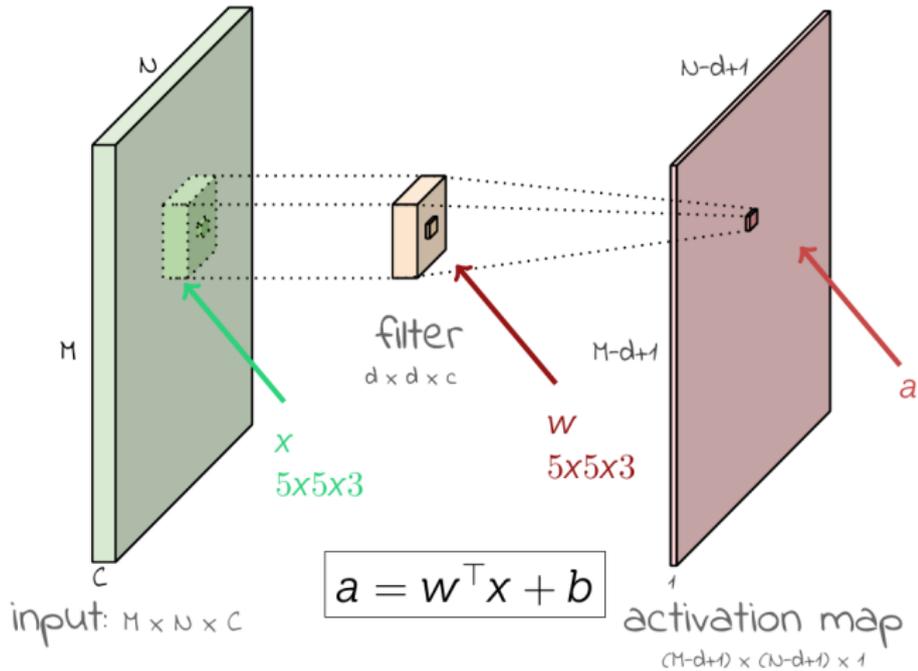
kernel: sharpen

$$\text{Kernel} - h = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

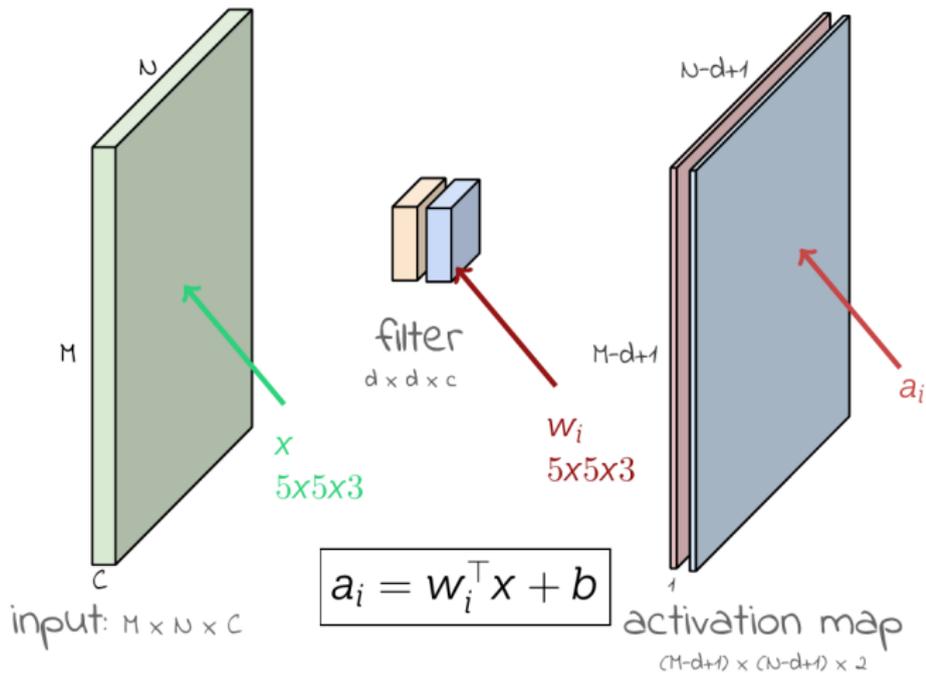
- Explicación visual por Victor Powell ⁴

⁴setosa.io/ev/image-kernels

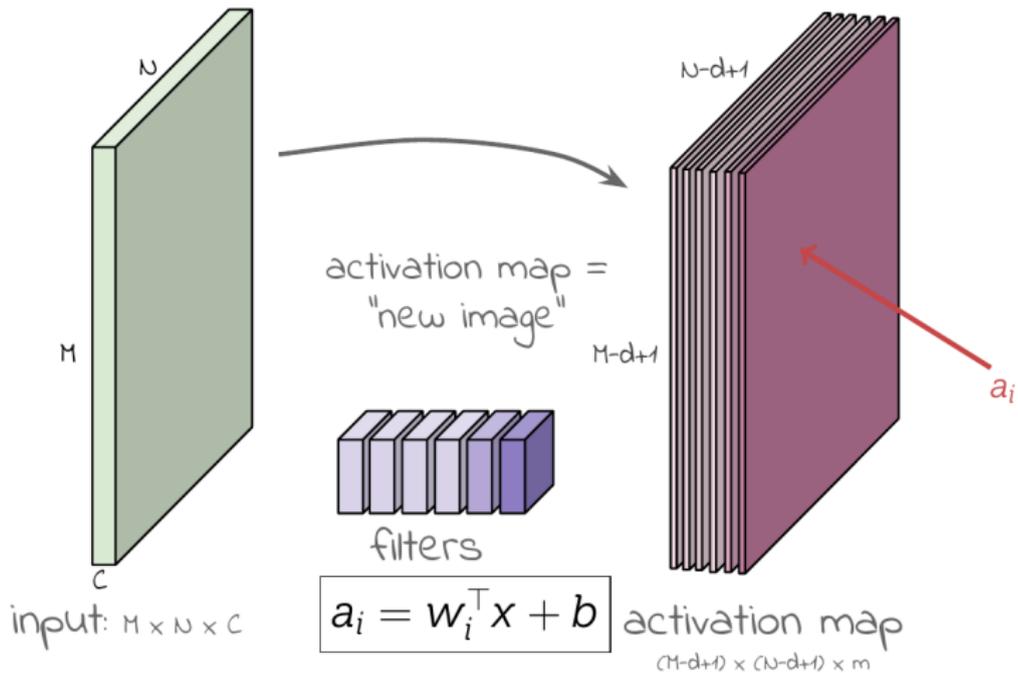
Convolución 2D



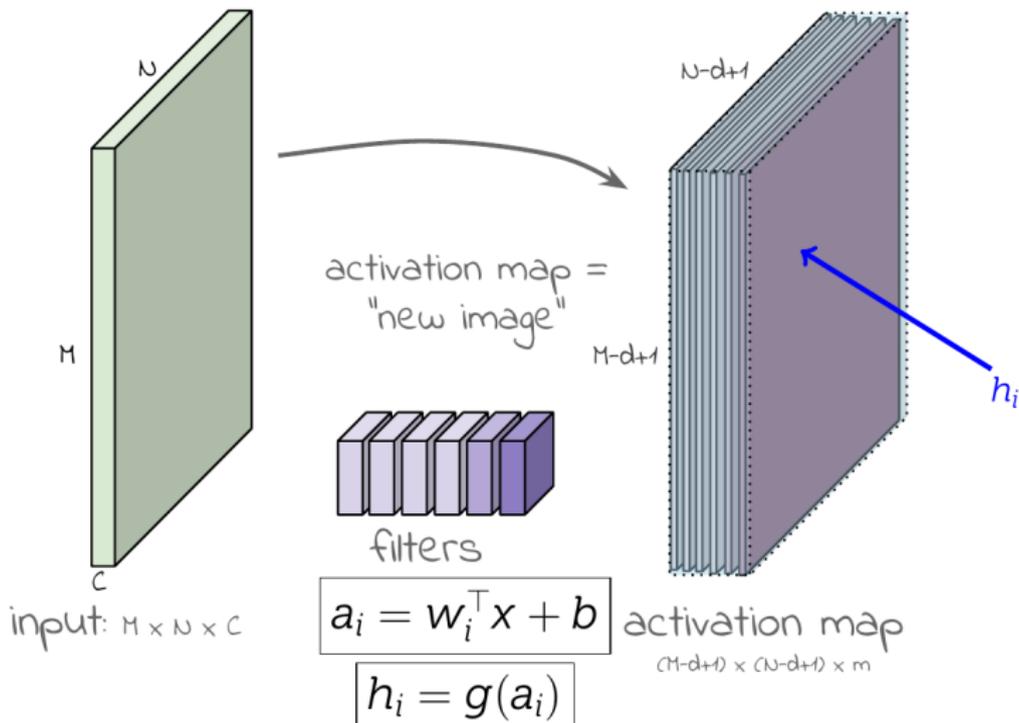
Convolución 2D



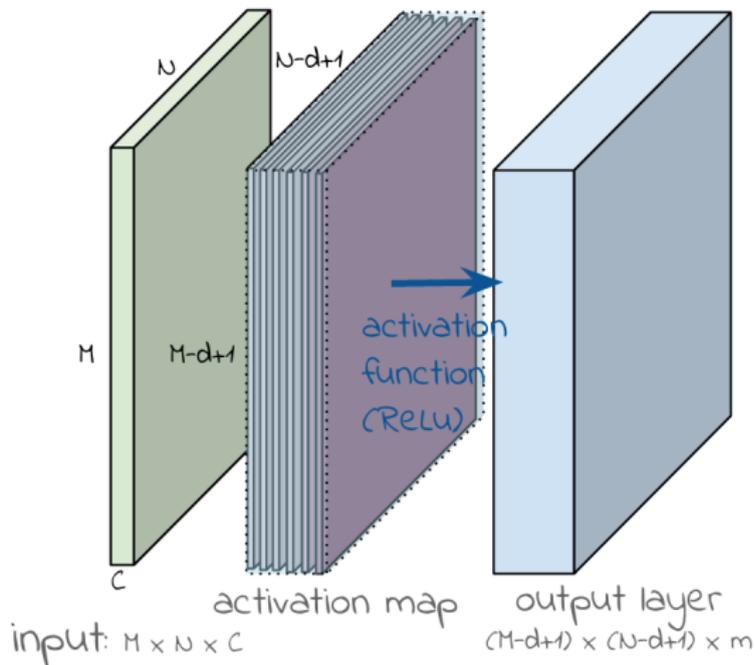
Convolución 2D



Convolución 2D

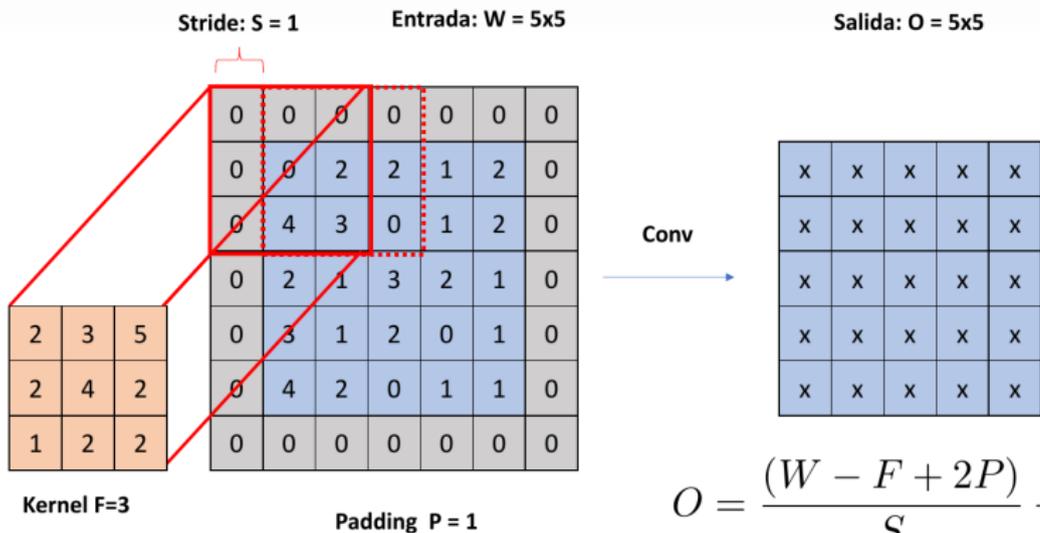


Convolución 2D

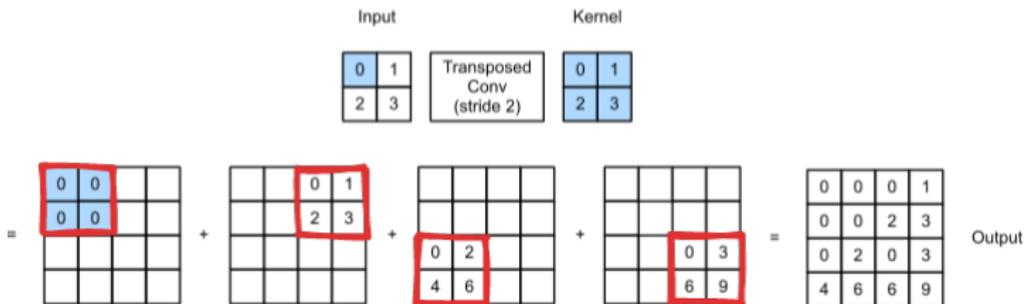
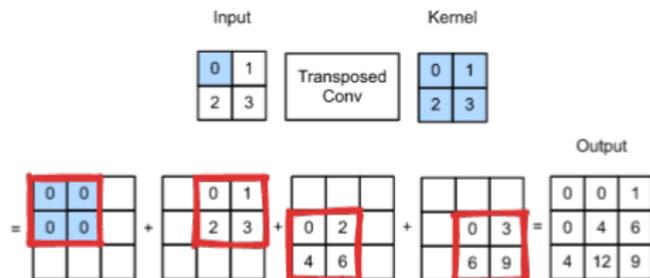


- ▶ ¿Qué elementos controlan número de features/embedding/layers?
- ▶ ¿Y cómo modificamos tamaño de las features a la salida?

Hiperparámetros de capa de Convolución 2D



Convolución 2D Transpuesta

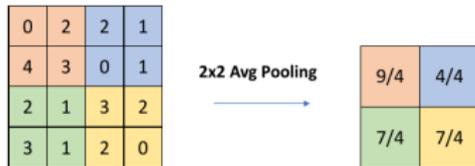
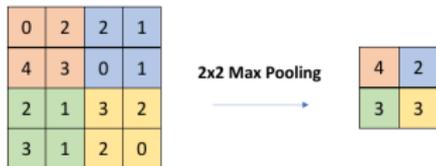


► ¿Se les ocurre un uso?

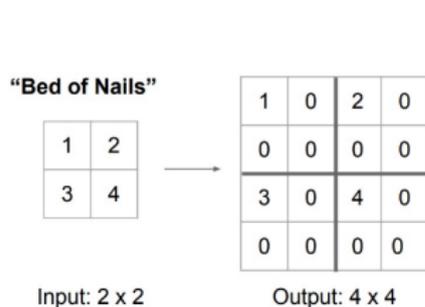
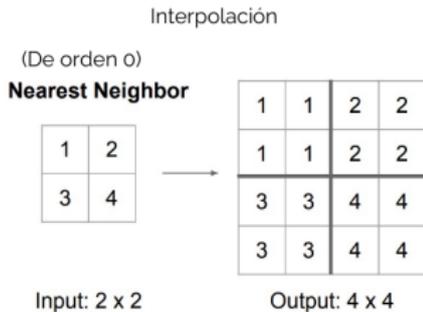
► $(W - 1)S - 2P + F = O$

Capas de downSampling y upSampling

Pooling

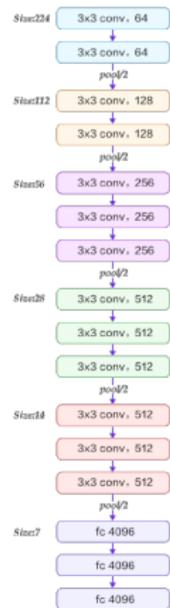
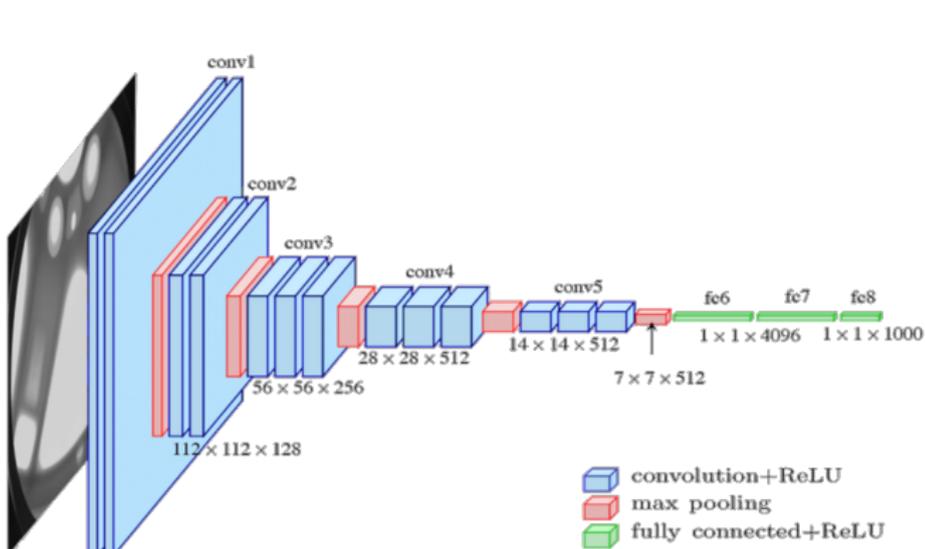


unPooling



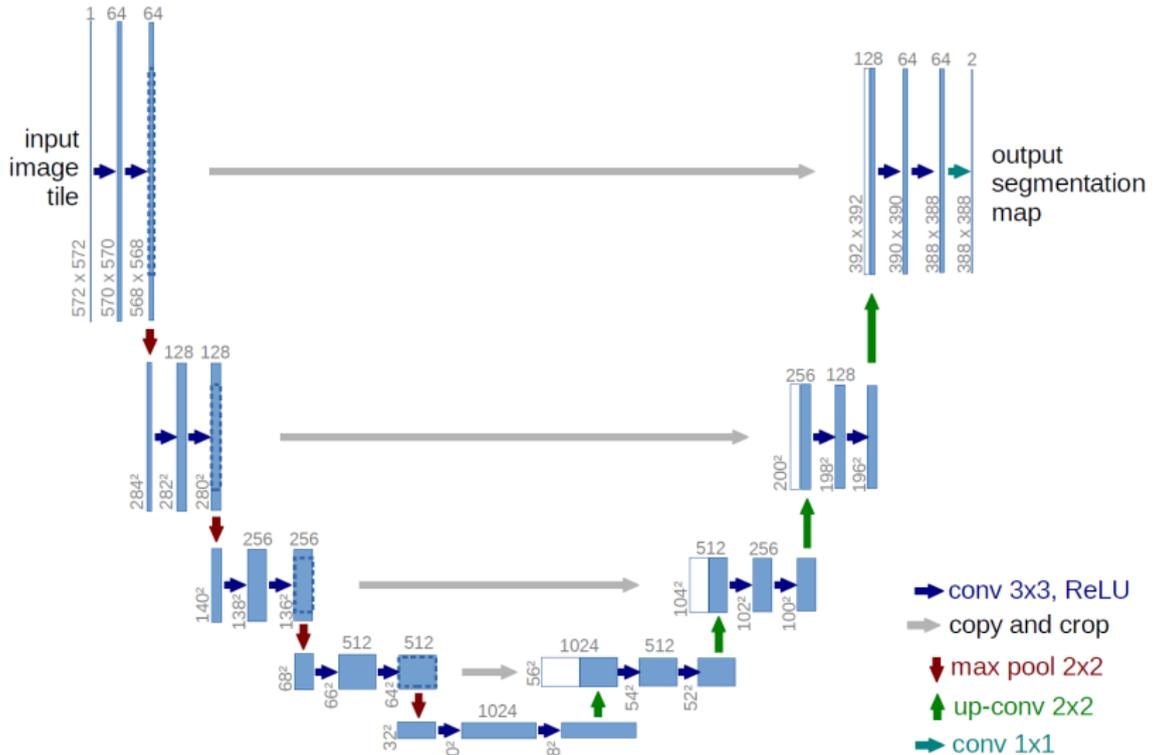
Otra opción: **Interpolación Bilineal**

VGG16-D

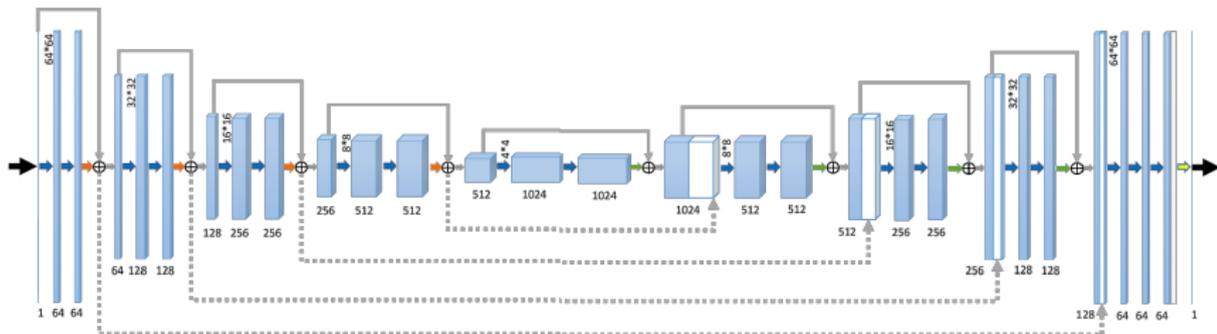
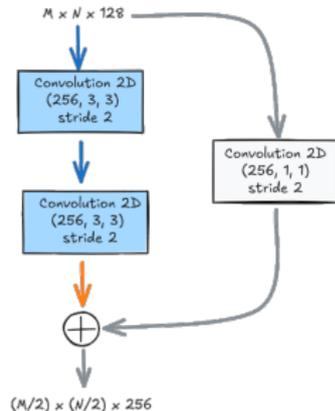
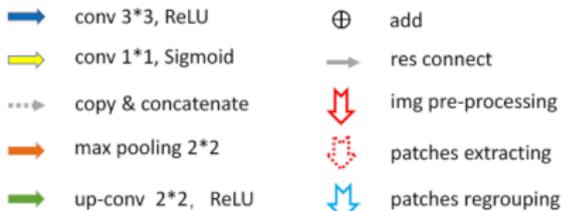


► ILSVRC'14 (ImageNet Large Scale Visual Recognition Challenge)

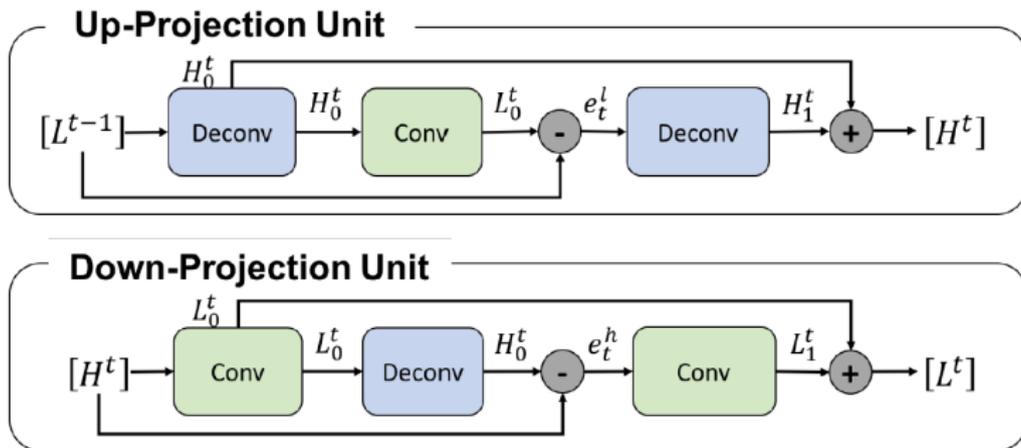
UNet



Res-UNet

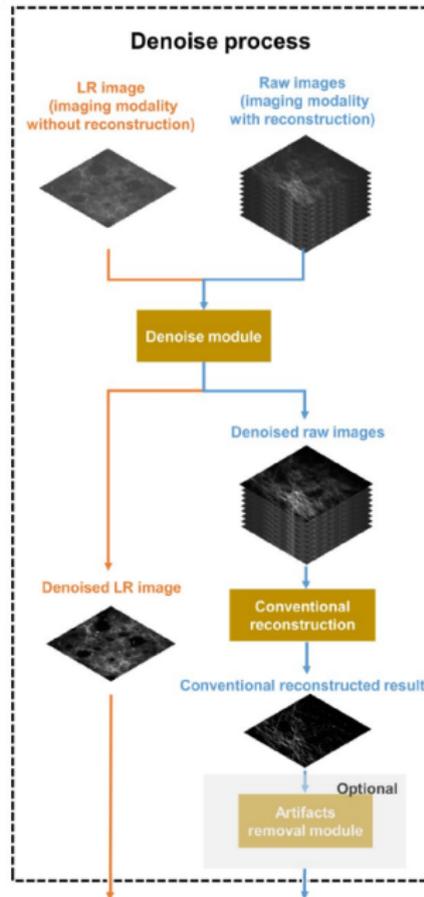
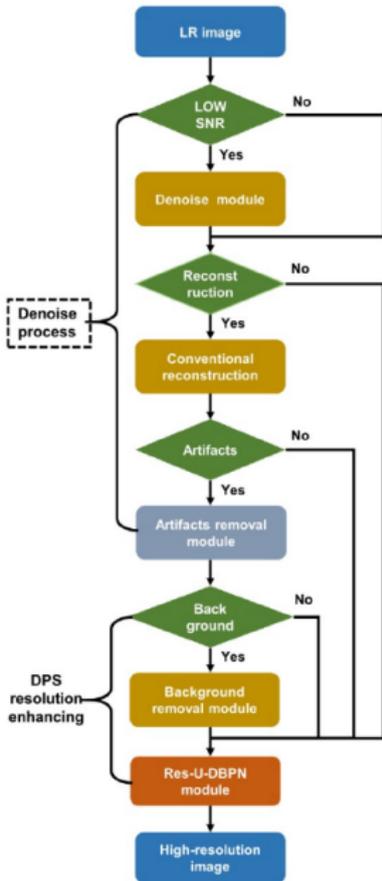


Deep Back-Projection Networks (DBPN)

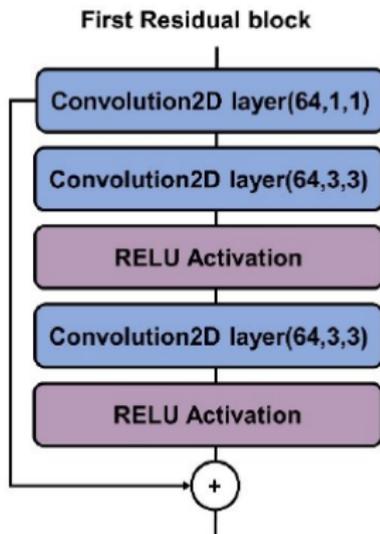
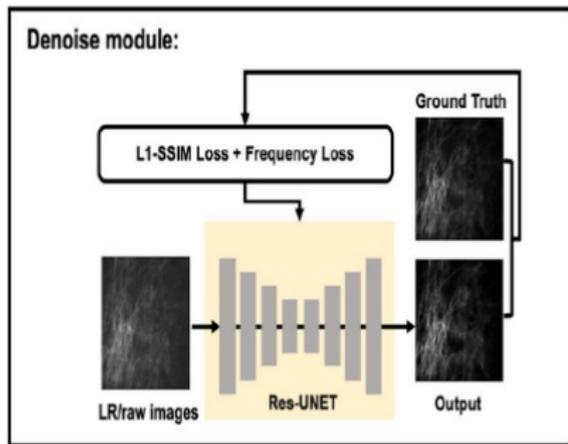


- ▶ La letra L indica features maps de baja resolución y H features maps de alta resolución.
- ▶ Aquí lo que se indica como Deconv son convoluciones 2D transpuestas, las cuales permiten realizar upSampling.

Denoise-DPS Framework



Denoise-DPS Framework - Denoise Module



$$\arg \min \{ \alpha \text{SSIM}(x, y) + \beta \text{L1}(x, y) + \epsilon \text{L1} [\text{FFT}(x) - \text{FFT}(y)] \}.$$

- ▶ x es la salida de la red e y el ground truth.
- ▶ FFT es la transformada rápida de Fourier, y $(\alpha, \beta, \epsilon)$ constantes experimentales.

Structural Similarity Index Measure (SSIM)

$$\text{SSIM}(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y),$$

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1},$$

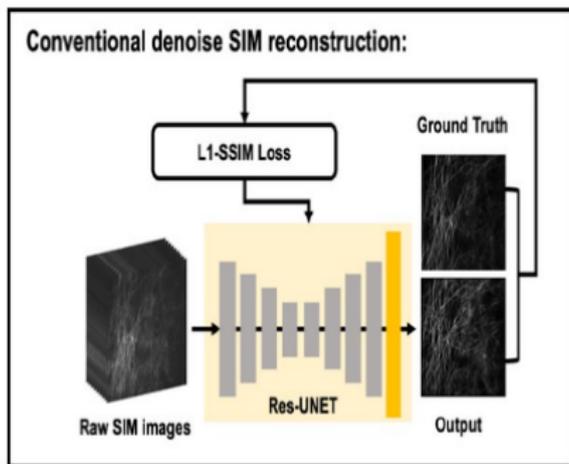
$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2},$$

$$s(x, y) = \frac{\sigma_{xy} + c_2/2}{\sigma_x \cdot \sigma_y + c_2/2}.$$

Donde:

- ▶ l indica luminancia, c contraste y s estructura.
- ▶ μ_x y μ_y son la media de intensidades en píxeles de x e y respectivamente.
- ▶ σ_x^2 y σ_y^2 son la varianza de x e y respectivamente.
- ▶ σ_{xy} es la covarianza de x e y .
- ▶ $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$, dos variables para estabilizar división con denominadores pequeños. Con $k_1 = 0,001$ y $k_2 = 0,03$ por defecto.
- ▶ L es el rango dinámico de los valores de pixel ($2^{\#bits \text{ per pixel}} - 1$).

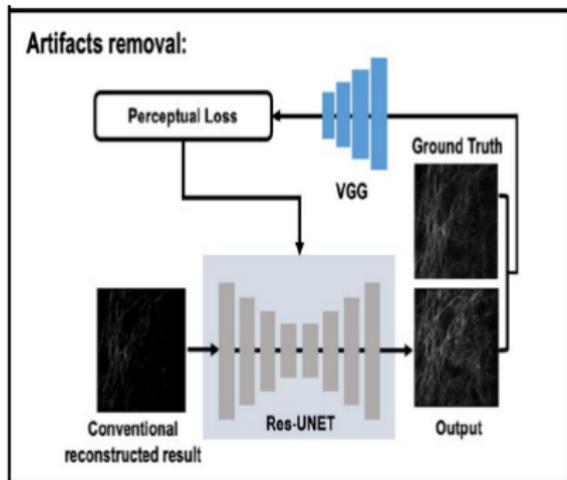
Denoise-DPS Framework - Conventional Denoise SIM Reconstruction



$$\arg \min \{ \alpha \text{SSIM}(x, y) + \beta \text{L1}(x, y) \}.$$

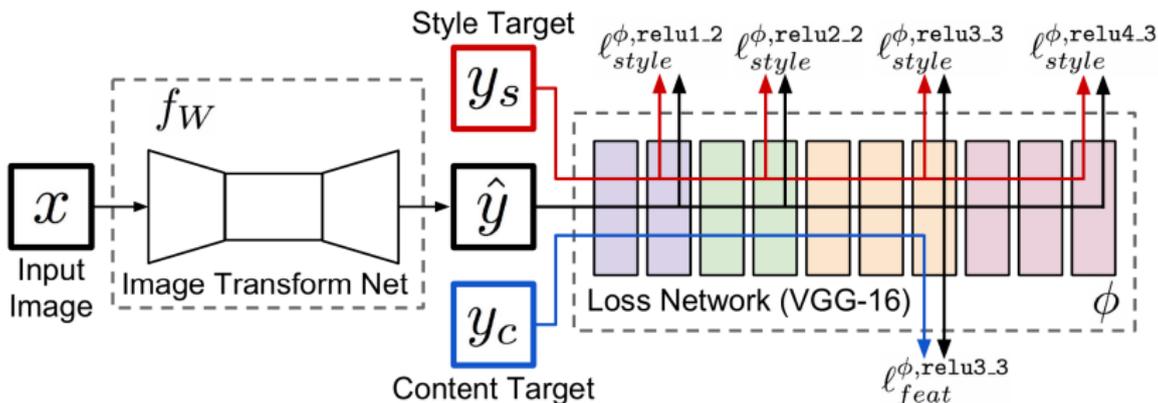
- ▶ x es la salida de la red e y el ground truth.
- ▶ El bloque naranja es una capa extra de upSampling.

Denoise-DPS Framework - Artifacts Removal



- ▶ Intenta mejorar performance sobre todo en imágenes que necesitan post-procesamiento..
- ▶ **¡Utiliza una función de pérdida perceptual!**
- ▶ Solo utiliza la *feature reconstruction loss* para nueve capas de VGG16-D.

Perceptual Loss

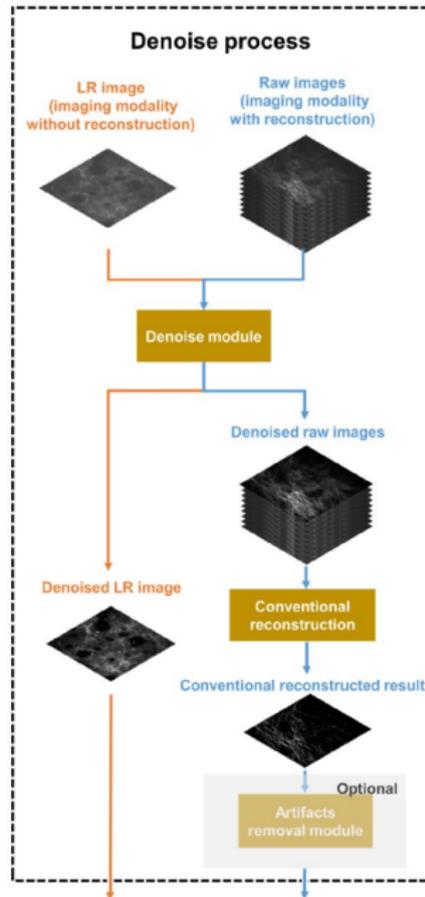
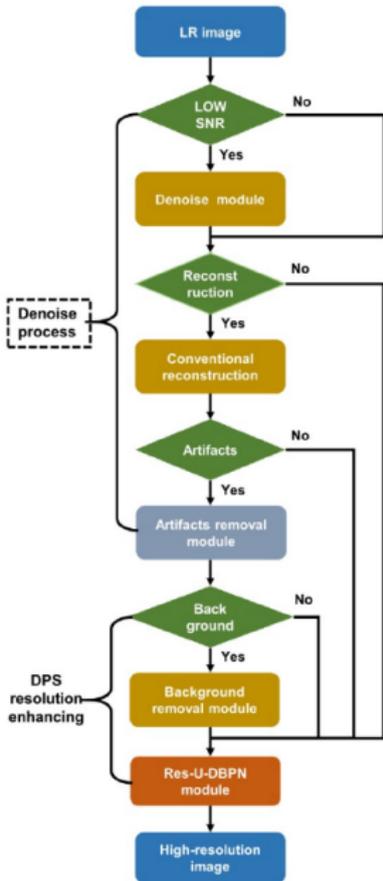


$$\arg \min_W \mathbf{E}_{x, \{y_i\}} \left[\sum_{i=1} \lambda_i l_i(f_W(x), y_i) \right],$$

$$l_{feat}^{\Phi, j}(\hat{y}) = \frac{1}{C_j H_j W_j} \|\Phi_j(\hat{y}) - \Phi_j(y)\|_2^2.$$

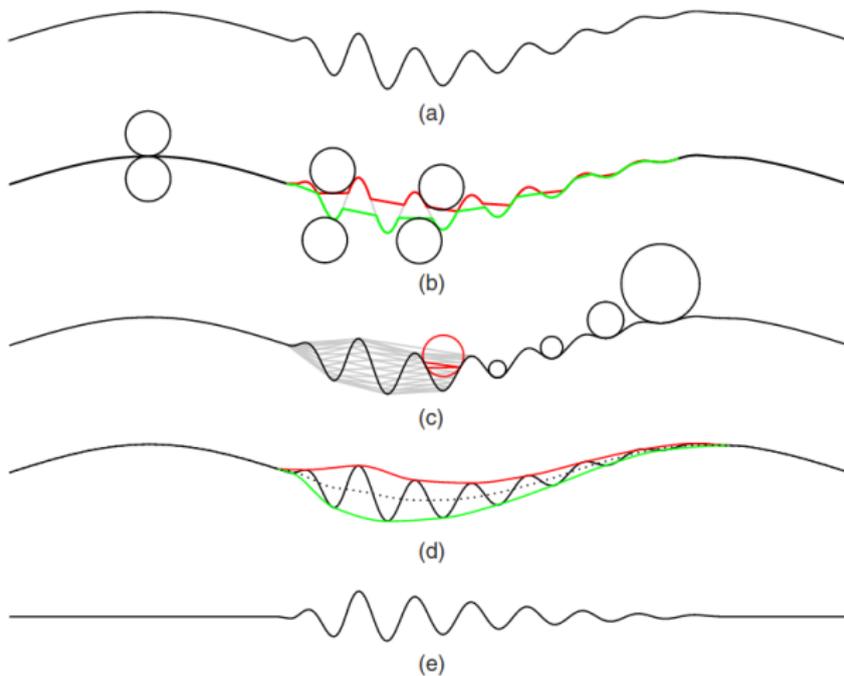
- ▶ $\Phi_j(x)$ es el feature map de la capa j de VGG16, con tamaño $C_j \times H_j \times W_j$.
- ▶ Se definen dos funciones de pérdida perceptual, *feature reconstruction loss* l_{feat}^{Φ} y *style reconstruction loss* l_{style}^{Φ} .

Denoise-DPS Framework



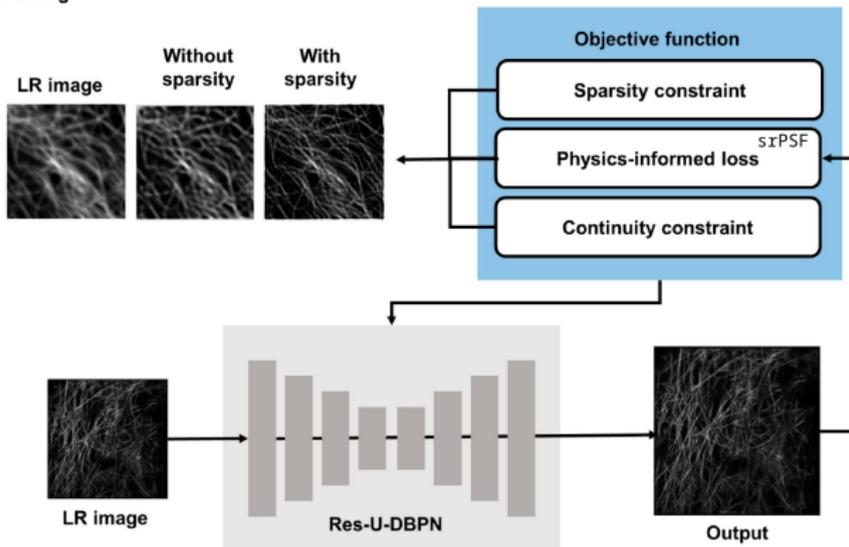
Denoise-DPS - Background removal module

Rolling ball sifting algorithm



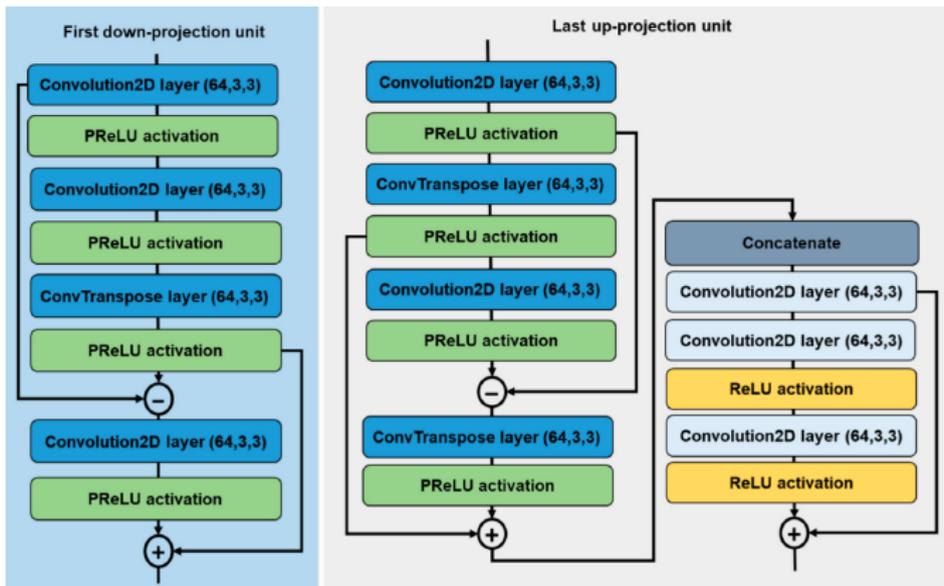
Denoise-DPS - Res-U-DBPN

Training



$$\arg \min \{SSIM(x \otimes srPSF, y) + \beta R_{\text{Hessian}}(x) + \beta \|x\|_1\}$$

Denoise-DPS - Res-U-DBPN



Denoise-DPS - srPSF

$$\arg \min \{ \text{SSIM}(x \otimes \text{srPSF}, y) + \beta R_{\text{Hessian}}(x) + \beta \|x\|_1 \}$$

- ▶ El proceso de captura de una imagen puede ser modelado como:

$$y = x \otimes \text{srPSF} + \eta,$$

donde x es la imagen HR, y la LR y η ruido en adquisición de imagen.

- ▶ srPSF surge de física del sistema de adquisición, por ejemplo para SIM:

$$\text{srPSF} = \left[\frac{2J_1 \left(\frac{2\pi m_f \text{NA} r}{\lambda} \right)}{\frac{2\pi m_f \text{NA} r}{\lambda}} \right]^2,$$

donde J_1 es la función de Bessel de primer orden, λ la longitud de onda de fluorescencia, NA apertura numérica del objetivo del lente, m_f el factor de modulación utilizado para modular NA luego de SIM y r la coordenada radial de la superficie del objeto.

Denoise-DPS - Hessian

$$\arg \min \{ \text{SSIM}(x \otimes \text{srPSF}, y) + \beta R_{\text{Hessian}}(x) + \beta \|x\|_1 \}$$

- ▶ R_{Hessian} penaliza la pérdida de continuidad de las estructuras a lo largo de las distintas dimensiones de la imagen.
- ▶ Asumiendo imagen f para el caso de SIM (información en tres ejes xyt) se puede definir como:

$$R_{\text{Hessian}}(f) = \sum_{\mathbf{r}} (|f_{xx}(\mathbf{r})| + |f_{yy}(\mathbf{r})| + \sigma |f_{tt}(\mathbf{r})| \\ + 2|f_{xy}(\mathbf{r})| + 2\sqrt{\sigma}|f_{xt}(\mathbf{r})| + 2\sqrt{\sigma}|f_{yt}(\mathbf{r})|)$$

- ▶ donde σ es un parámetro introducido para forzar continuidad estructural en el eje t .

Denoise-DPS - L1

$$\arg \min \{ \text{SSIM}(x \otimes \text{srPSF}, y) + \beta R_{\text{Hessian}}(x) + \beta \|x\|_1 \}$$

- ▶ La regularización con norma L1 se utiliza para robustecer entrenamiento y favorecer soluciones esparsas.

Denoise-DPS - PSNR Comparisons

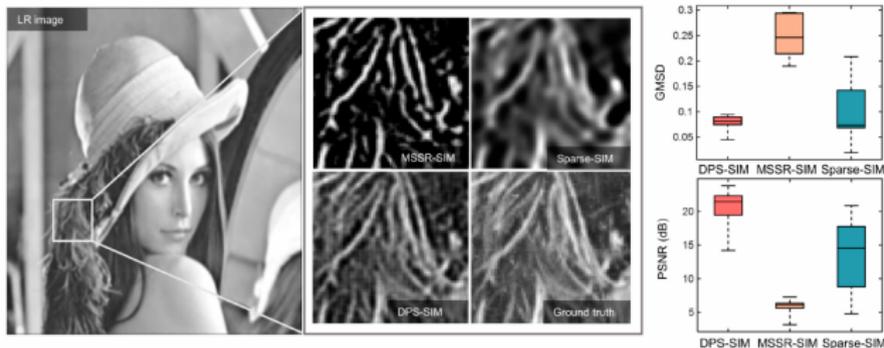
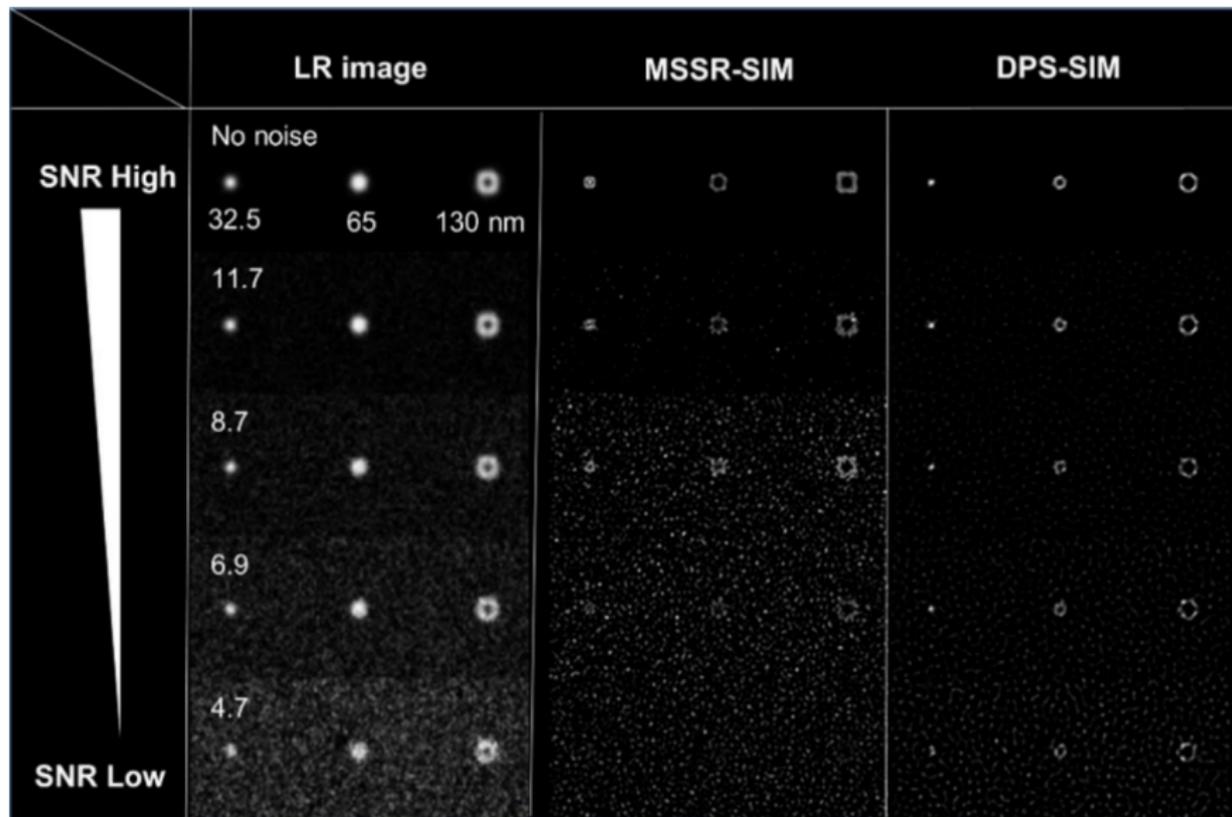


Table 2. PSNR comparisons between the proposed method, Richard–Lucy (RL) deconvolution, and algorithms in [7,8,33,34]. The values in boldface correspond to the optimal results for each test image.

	Barbara	Room	C.man	Lena	Peppers	Pirate
RL	12.83	14.51	11.17	13.90	12.58	15.11
MSSR	4.91	6.25	5.77	5.91	6.25	6.32
Sparse	14.59	16.91	14.50	13.51	7.91	8.77
W-DIP	18.68	17.15	18.11	19.79	26.78	19.02
SelfDeblur	18.76	21.37	19.69	21.71	25.85	22.30
Proposed	19.45	21.78	22.06	21.15	22.57	22.33

Denoise-DPS - SNR



¿Preguntas?

¡Gracias!