www.editada.org

*International Journal of Combinatorial Optimization Problems and Informatics, 15(5), Dec 2024, 75-83. ISSN: 2007-1558. https://doi.org/10.61467/2007.1558.2024.v15i5.559*

# Convolutional Neural Networks to Decode Images in a Wavefront Coding Imaging System

*José Manuel Reyes Alfaro, Carina Toxqui Quitl, María Angélica Espejel Rivera, Alfonso Padilla Vivanco, and Enrique González Amador.*

Universidad Politécnica de Tulancingo, Laboratorio de Visión por Computadora, Calle Ingenierías 100, Huapalcalco, 43625, Hidalgo, México.

manuel.reyes1831027@upt.edu.mx, carina.toxqui@upt.edu.mx

**Abstract.** Wavefront coding technique has been used to extend the depth of focus in an optical imaging system. An optical element called a phase mask allows coded images to be obtained since the point spread function remains almost invariant in an axial range. Subsequently, a computational technique is required to decode the acquired images. An optical-computational technique is proposed to use a phase mask for the coding stage and a convolutional neural network for the final restoration. Comparative results are made between cubic and trefoil profile phase masks and decoded using the traditional Wiener filter and a convolutional neural network. Image quality evaluation is done using the peak signal-to-noise ratio.

**Keywords:** Wavefront coding, Trefoil phase mask, Convolutional neural network, Depth of focus, Image restoration.

Article Info
*Received May 10, 2024.*
*Accepted Nov 20, 2024.*

## 1 Introduction

Hybrid optical–computational systems with extended depth of focus have been used for different applications where high resolution is required, such as microscopy, biometrics, ophthalmology, and object detection [1,2]. Dowski and Cathey [1] incorporate a phase mask (PM) as an optical element at the exit pupil of the imaging system, which codes the wavefront. In this way, the optical system is partially invariant to defocus. However, the quality of the encoded image obtained requires a computational process for its decoding. Traditionally, restoration algorithms are based on different filters such as Inverse [3], Wiener [3], spectrum equalization, and recently machine learning techniques [4, 5].

In the computer vision field, since real imaging conditions are usually imperfect, the digital image represents only a degraded version of the original scene. Most efforts are focused on addressing the problems caused by occlusion, deformation, and small-size imaging. The deep architectures developed for this task are based on CNNs, Generative Adversarial Networks (GANs) and autoencoders [6,7].

Haoyuan Du et al. [4] investigate a framework of neural networks for restoring blurring images in WFC. They take the classical cubic phase mask in the wavefront coding step. Nevertheless, it is well known that the cubic phase mask introduces artifacts in the final image [8]. Reyes-Alfaro, et al. [9] employ a trefoil phase mask in a wavefront coding imaging system and a CNN architecture as a deblurring algorithm. They show that the decoded image quality evaluation remains constant at different defocus values for the CNN architecture as a deblurring algorithm. This work presents an optical-computational technique that uses cubic and trefoil profile

phase masks (PM) for encoding and a convolutional neural network for image deblurring. As we can see in Figure 1, the proposed method does not require information about the peak and valley of the point spread function (PSF), as traditional image restoration algorithms do [1,3].

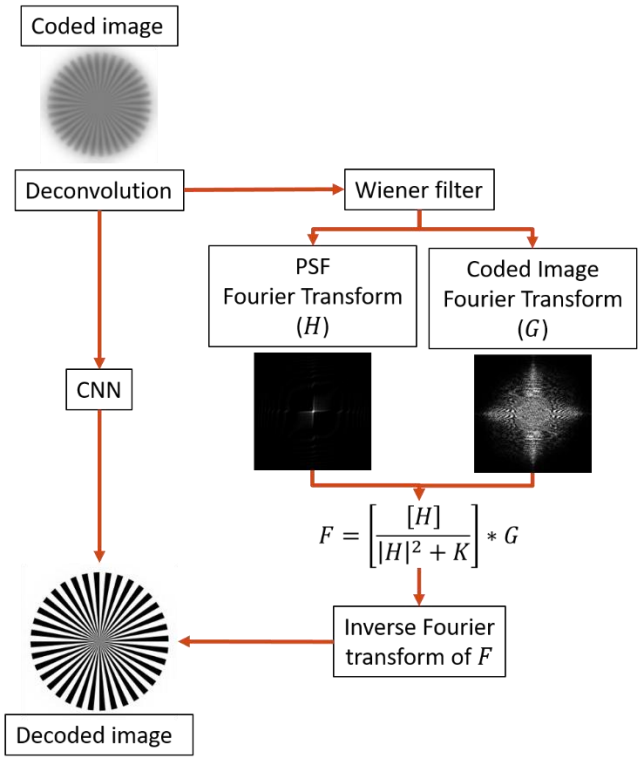


Figure 1 Restored images using the traditional Wiener filter and the proposed trained network.

The content of this document is organized as follows: in Section 2 a Wavefront coding (WFC) system is defined with a cubic and trefoil PM placed in the exit pupil of the system. The architecture and characteristics of the CNN are shown in section 3. The numerical simulation for WFC and the CNN for the deconvolution task is described in section 4. Comparative results of the extension depth of focus (DoF) using the different deconvolution techniques are summarized in Section 4. Finally, in Section 5 the conclusions of this work are given.

## 2 Wavefront Coding

Let $f_o\ (x_1, y_1)$ be a function that describes the intensity of the object whose image will be formed through the optical system, the intensity function of the geometric image is $f_g\ (x, y)$ where $(x_1, y_1)$ and $(x\ , y)$, are the coordinates in the object plane and the image plane, respectively. The intensity pattern $f_i\ (x, y, z)$ can be described as the convolution of the geometric image $f_g\ (x, y)$, and the corresponding point spread function (PSF) $|h(x, y, z)|^2$, i.e.,

$$f_i\ (x, y, z) = f_g\ (x, y) * |h(x, y, z)|^2 \tag{1}$$

The $h(x, y, z)$ in an optical system can be calculated using Fresnel diffraction theory. As can be seen in Figure 2, in a traditional optical system as we move from the focal plane or image, the images present blur. The WFC system uses a PM placed in the exit pupil (ExP) of the optical system, which causes a partially invariant PSF in an axial range, schematized in Figure 3. Therefore, the pupil function $P'$ of the wavefront coding system is given as,

$$P'(x_0, y_0) = P(x_0, y_0)\, exp\left[-i\frac{2\pi}{\lambda}\alpha PM(x_0 + y_0)\right] exp\left[-i\frac{2\pi}{\lambda}\psi(x_0^2 + y_0^2)\right]. \quad (2)$$

Where $P$ represents the pupil function, the first exponential is the phase term generated by the PM and the last one exponential is given by the amount of defocus ($\psi$) of the optical system[9].

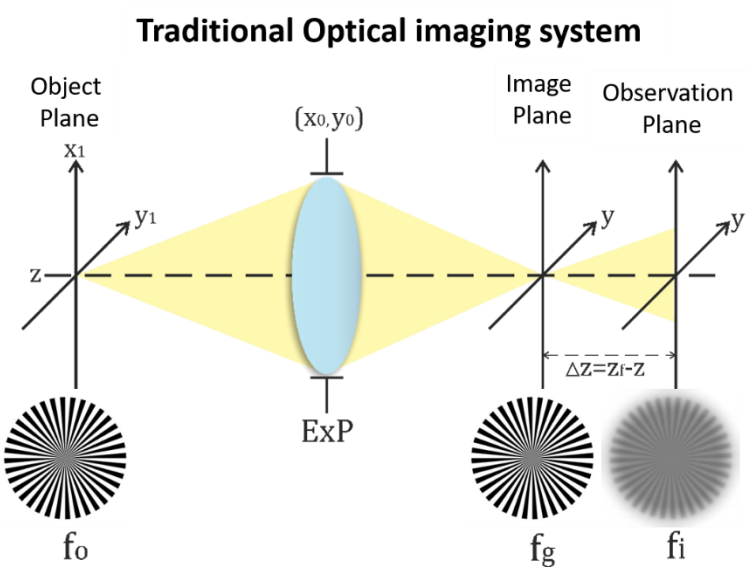


Figure 2. Scheme of an optical imaging system. When the image is observed in a different plane, the image will be out of focus by an amount $\psi$.

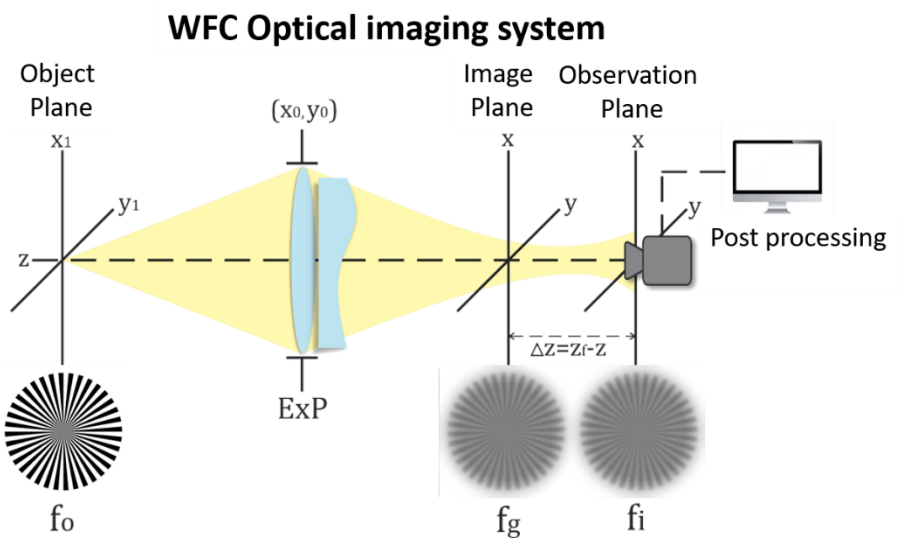


Figure 3. Scheme of an optical imaging system with WFC. The PM placed at the exit pupil of the system achieves a partially invariant PSF in an axial range of the optical axis. A second decoding stage is required in order to deblur the image.

## 3 Convolutional Neural Network

In deep learning algorithms, feature extraction is key to their performance. Traditional models require a stage external to the network for this process, while deep learning transfers this task to the neural network itself. Convolutional neural networks (CNN) are useful in image, audio and video processing. Its operation is based on identifying characteristics of the input data and passing them on to a neural network for processing. The kernel is like a type of filter that is applied to images, since it allows us to identify potentially important features or patterns, among some of these you can find edge detection, focus, component isolation, etc.

The CNN attempts to learn a mapping function between the degraded image $g(x, y)$ and the ideal image $f(x, y)$. Its architecture consists of convolution layers, each of them representing a level of feature extraction [10]. The input layer is described as:

$$f_1(x, y) = \sigma(W_1 * g_1(x, y) + b_1) \tag{3}$$

where $W_1$, and $b_1$ represent the filters and bias learned in this layer, respectively. $f_1$ denotes the extracted feature map and $\sigma$ is a nonlinear function, which can be of the sigmoid, hyperbolic tangent or ReLU (Rectified Linear Unit) type. In the architecture of this network, the ReLU activation function is used to filter values by applying a simple mathematical operation. Takes the input value $x$ and returns the maximum of $0$ and $x$. In other words, if the input value is positive or zero, ReLU returns the input value itself; otherwise, it returns $0$. Mathematically, it can be defined as $relu(x) = max(0, x)$.

In image restoration, CNN networks have shown promising results for noise removal and other types of degradation [11]. In [11] a CNN was designed to learn the deconvolution operation without needing to know the cause of visual artifacts. The input images to the network were affected with noise, saturation and compression artifacts. In [6], to improve image quality and suppress noise, they proposed two models: a CNN and a CGAN (Conditional Generative Adversarial Networks). As input data to the network they used images encoded with a WFC optical system that uses a Cubic PM. Additionally, they added additive white Gaussian noise.

The CNN model proposed in this work was trained using the MSE (Mean Square Error) cost function and the ADAM optimizer as an update rule, with an initial learning rate of $0.0001$ and a batch size of 32. A each convolutional layer is followed by a batch normalization layer and a ReLU activation function. It contains four convolutional layers, in the first and second layer 64 9x9 filters were used, in the third 32 1x1 filters and in the last layer a 5x5 filter, as can be seen in Figure 4.

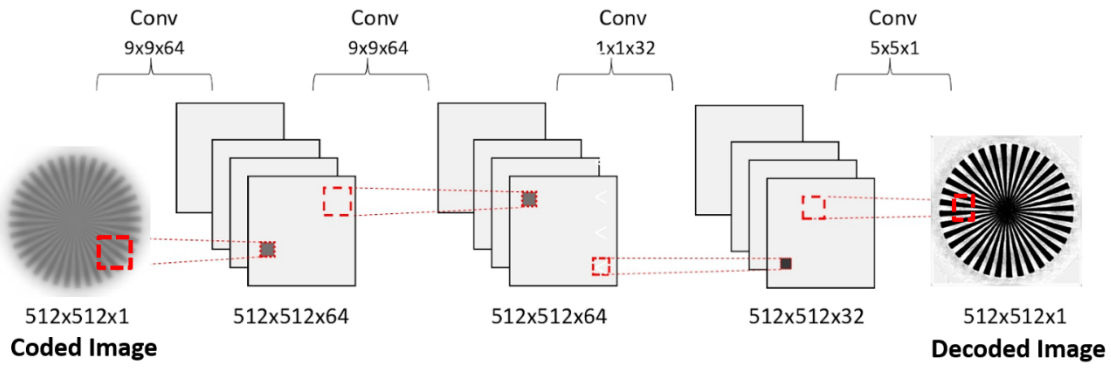


Figure 4. CNN architecture for deconvolution of images coded with a WFC optical system.

## 4 Proposed Method

The optical system schematized in Figure 3 was numerically simulated on the Matlab. As an input image the Siemens star, or spoke target with a dimension of 512×512 pixels is used as test object. A digital image in a traditional optical system outside the image plane presents degradation due to defocus aberration as shown in Figure 2, the numerical simulation of the images obtained is given through equation (2).

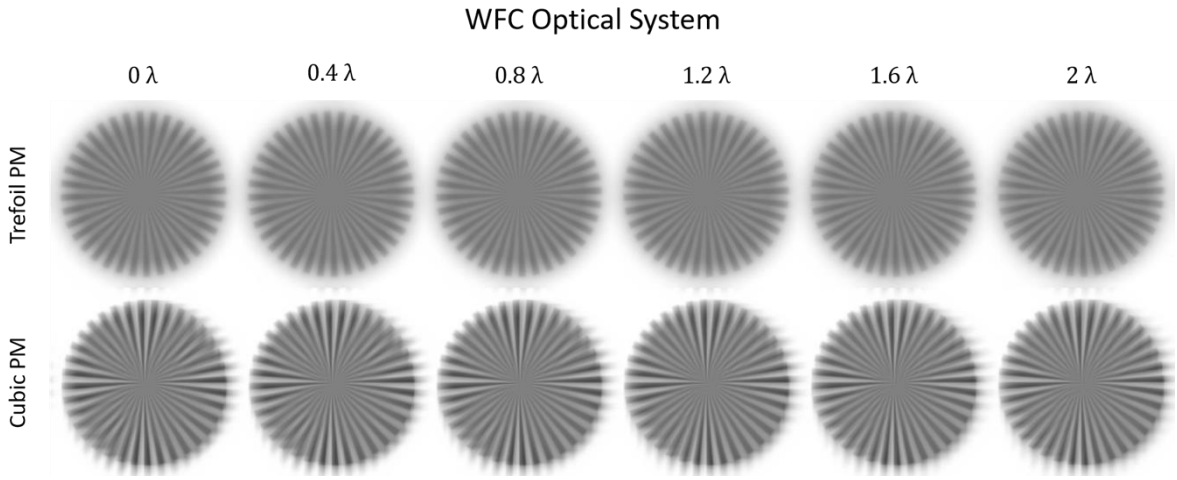


Figure 5. Images generated by a WFC system, affected by defocus $\psi = 0\lambda$ up to $\psi = 2\lambda$. The diffraction-limited image is in the plane of focus with $\psi = 0\lambda$.

The coded images provided by an optical system with WFC trefoil and cubic PM are shown in Figure 5. As can be seen, since the PSF of this system remains almost invariant in an axial range, the generated images present the same visibility.

Figure 6 shows the Modulation Transfer Function (MTF) curves of a) a diffraction-limited optical system. The optical system with WFC and b) Cubic PM affected by defocus aberration, and c) Trefoil PM. The MTF remains almost unchanged at defocus values from $\psi = 0\lambda$ to $\psi = 2\lambda$. Furthermore, because the MTF values are different from zero at high frequencies, it is possible to recover the information in the decoding stage.

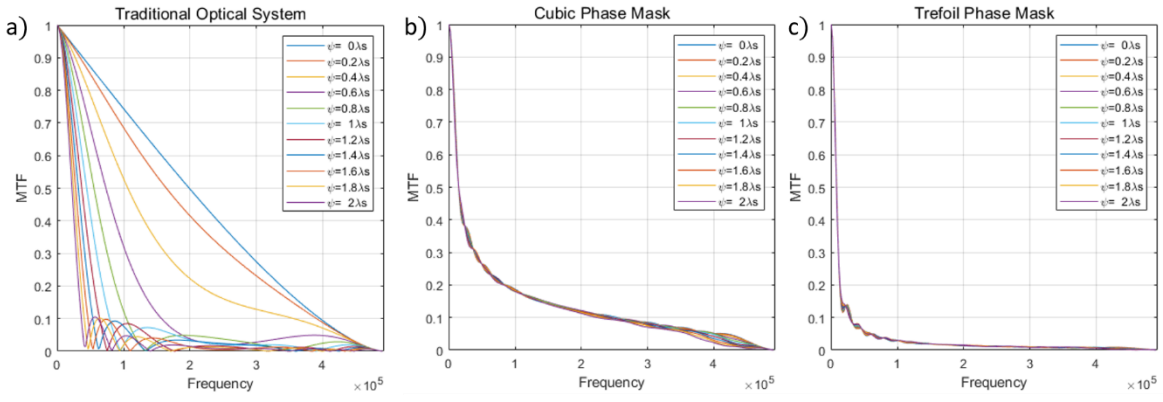


Figure 6. MTF curves with defocus $\psi = 0, 0.2\lambda, \ldots, 2\lambda$, and wavelength $\lambda = 650nm$.

A deconvolution stage of the blur images will be carried out through the classic Wiener filter and the CNN described in section 3. The Wiener filter algorithm is considered effective in this task, by making use of the Optical Transfer Function (OTF) of the optical system, as follows [1,6,12],

$$F'(u,v) = \left[\frac{H(u,v)}{H(u,v)^2+K}\right] * G(u,v). \quad (4)$$

The constant $K$ represents the ratio between the noise power spectrum and the image power spectrum, which in this work is estimated at $0.0001$. $F'(u,v)$ and $G(u,v)$ are the Fourier transform of the restored image $f(x,y)$ and the blurred $g(x,y)$, respectively.

## 4 Experimental Results

An algorithm was implemented to numerically simulate a WFC system using a Cubic [1,4,13] and Trefoil PM [8] with strength $\alpha = 3\lambda$. To evaluate the restoration of the coded images, a defocus aberration ψ=0λ, 0.2λ, ..., 2λ is added. In the deconvolution processes a) the Wiener filter and a b) CNN are used.

For image deblurring, a CNN is used. Two data sets are created that includes 240 images, 189 images for training and 51 images for validation. Each network is trained on the data set generated by the point spread function with different defocus values. The CNN model was trained using the MSE loss function and the ADAM optimizer as the update rule, with an initial learning rate of $10^{-4}$. Each convolutional layer is followed by a batch normalization layer and a ReLU activation function.

Figure 7 shows the image acquired at a defocus value of $0\lambda$, and $1.2\lambda$, in a) image without the WFC technique, it is observed that it is no longer visible in detail, b) the coded image obtained by the system with WFC with cubic PM and c) with trefoil PM.

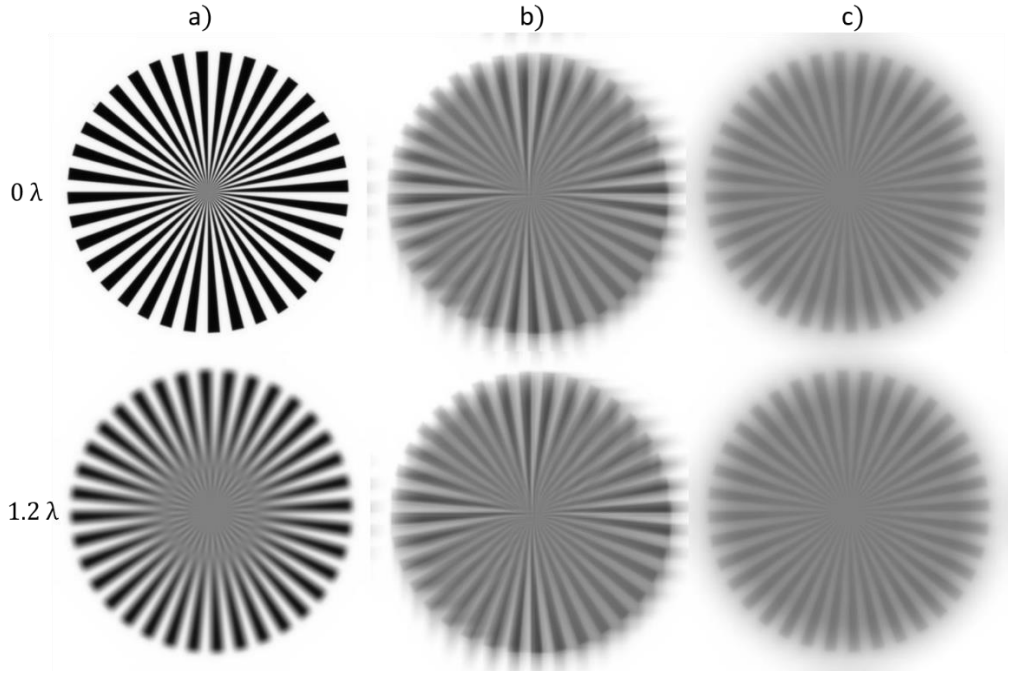


Figura 7. Images affected with defocus $\psi = 0\ \lambda$ and $1.2\lambda$, in a) image without PM, in b) image coded by cubic PM, and c) coded by trefoil PM.

Figure 8 shows the results with a) Wiener filter and in b) with the proposed CNN. Allowing the depth of focus of the optical system to be extended.

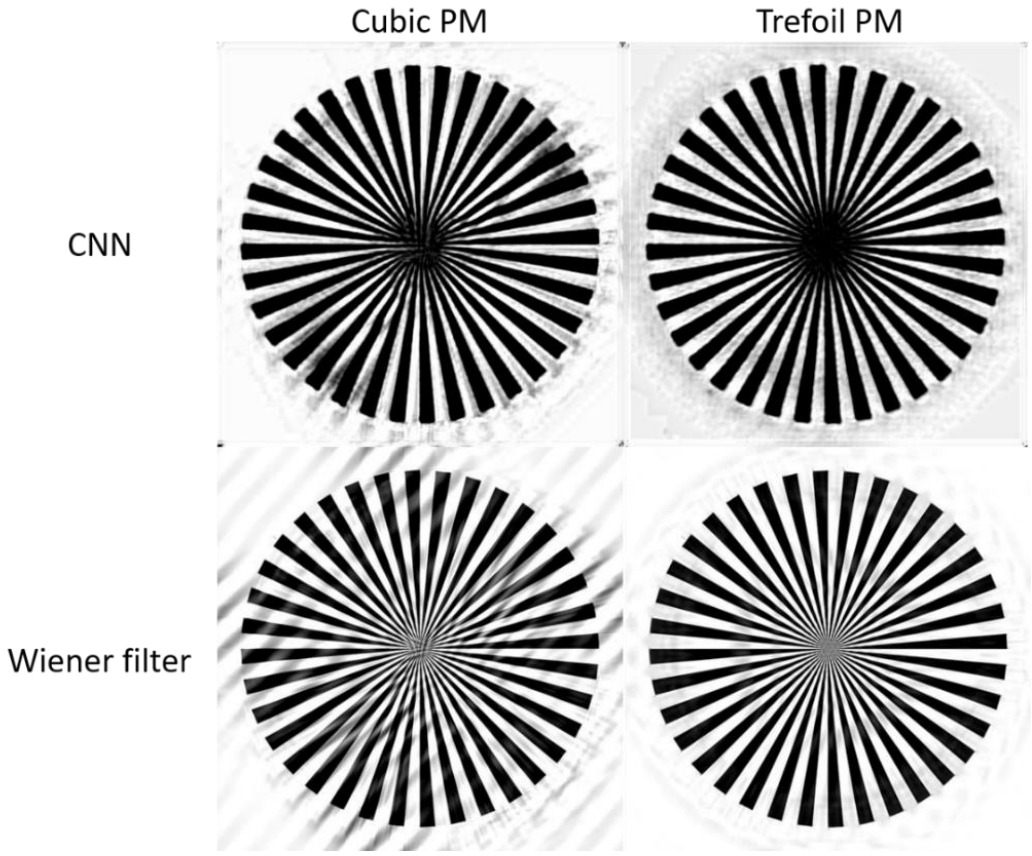


Figure 8. Restored images applying Wiener filter and proposed CNN, for an image obtained at a defocus value of $1.2\lambda$, with the phase masks mentioned in the coding stage.

Figure 9 shows the results obtained with both methods, the restoration is carried out in 6 images acquired every $0.4\lambda$ from the aforementioned defocus values. . In the traditional optical system, as can be seen, the image is blurred as the distance from the observation plane increases. The quality of the restored images was evaluated with PSNR and the results are presented in Figure 10.

For the cubic PM case, the reduction of artifacts and the higher contrast in the deblurring image is evident when the CNN is used. In the trefoil PM the artifacts are less, but the contrast is higher with the CNN.

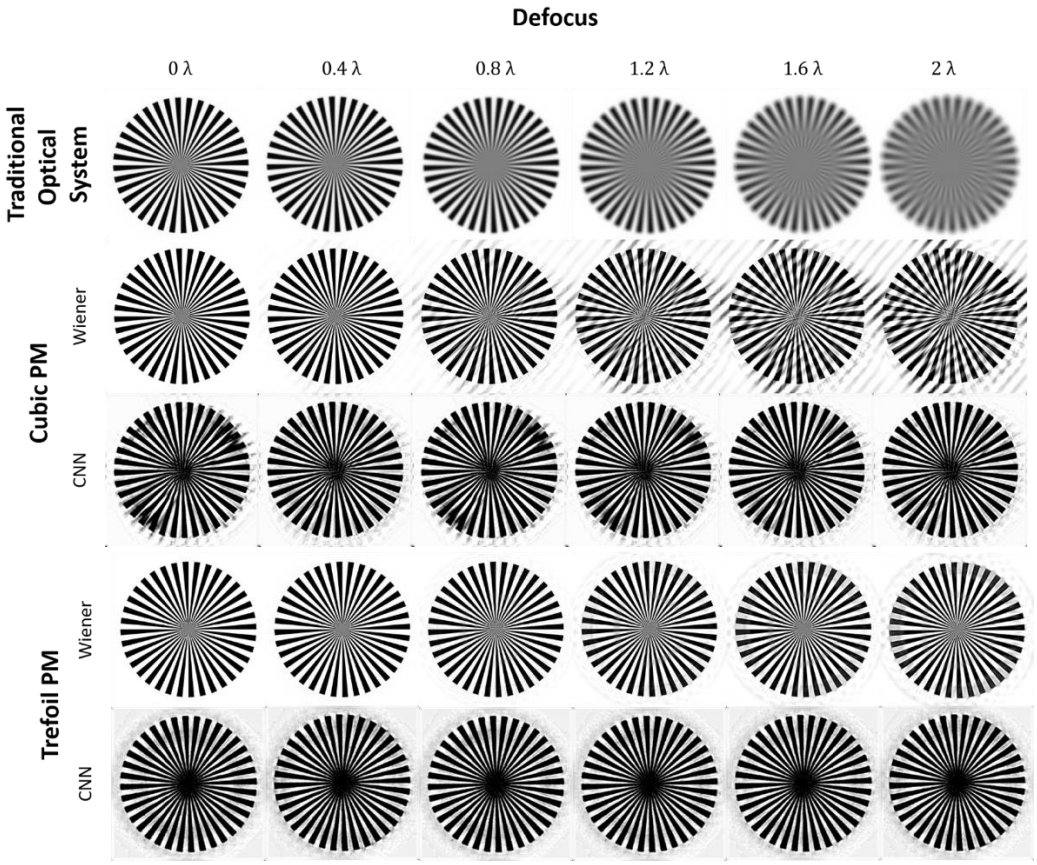


Figure 9. Experimental results with different DoF, coded by a WFC optical system through Cubic and trefoil PM, and restored through Wiener filter and proposed CNN.

| | PSNR | | | | | |
|---|---|---|---|---|---|---|
| | Cubic PM | | | Trefoil PM | | |
| | Coded | CNN | Wiener Filter | Coded | CNN | Wiener Filter |
| 0 λ | 8.942994118 | 16.23788452 | 38.49230957 | 9.088680267 | 17.99086952 | 38.49516678 |
| 0.2 λ | 8.944410324 | 16.33903885 | 36.67863464 | 9.08801651 | 18.02503586 | 37.70360947 |
| 0.4 λ | 9.145870209 | 15.2272234 | 11.54454327 | 9.037004471 | 16.26637077 | 16.05162811 |
| 0.6 λ | 8.94906044 | 16.62407303 | 30.0450592 | 9.086105347 | 18.08459473 | 33.87414551 |
| 0.8 λ | 8.95746994 | 16.96260071 | 25.23714828 | 9.084156036 | 18.14818573 | 30.2693615 |
| 1 λ | 8.9692173 | 17.23745346 | 21.93663025 | 9.08080101 | 18.15312004 | 27.52833939 |
| 1.2 λ | 8.986392975 | 17.42596436 | 19.29980278 | 9.077386856 | 18.11362648 | 24.9847126 |
| 1.4 λ | 9.007278442 | 17.44797897 | 17.1163578 | 9.071504593 | 17.93992424 | 22.73608017 |
| 1.6 λ | 9.031961441 | 17.23760033 | 15.30467987 | 9.063552856 | 17.68677902 | 20.79616547 |
| 1.8 λ | 9.063294411 | 16.77382088 | 13.80402184 | 9.056847572 | 17.31952667 | 19.06200409 |
| 2 λ | 9.10147953 | 16.08874893 | 12.56099796 | 9.04780674 | 16.85459518 | 17.47319984 |

Figura 10. Quality measurements of restored images given by a WFC system.

## 5 Conclusions and Directions for Further Research

A CNN architecture is presented for image deconvolution in WFC applications with a cubic phase mask, which has demonstrated increased contrast in the decoding image and greater extension of the DoF. On the other hand, the trefoil PM reduces the presence of artifacts in the decoded image.

Coded images were used without any preprocessing to train the CNN. Its performance is compared with those obtained through the classic Wiener filter since it is the most used for decoding. Wiener filter depends on the optical transfer function of the optical system. Because of this, a CNN that performs deconvolution is proposed.

PSNR is used to evaluate the quality of the restored image. The results obtained from deconvolution with CNN are promising. This involves further research into other deep learning models that need to be taken into account to detect blurred regions or predict the degree of blur.

## 6 Acknowledgements

M. Reyes-Alfaro y M.A. Espejel-Rivera thanks to Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCyT); with CVU no. 321005, and CVU no. 325373.

## References

1. Dowski, E., & Cathey, W. (1995). Extended depth of field through wave-front coding. *Applied Optics*, *34*, 1859–1866.
2. Sánchez, A. S., González-Amador, E., Padilla-Vivanco, A., Toxqui-Quitl, C., Arines, J., & Acosta, E. (2023). Jacobi-Fourier phase masks solution for presbyopia. In *Current Developments in Lens Design and Optical Engineering XXIV* (Vol. 12666, pp. 81–87). SPIE.
3. Gonzalez, C., & Woods, R. E. (2002). *Digital Image Processing* (3rd ed.). Prentice Hall.
4. Du, H., et al. (2018). Image restoration based on deep convolutional network in wavefront coding imaging system. *IEEE*, 18.
5. Zhou, Y., Wu, Y., & Gui, X. G. (2023). Extended depth-of-field imaging using multi-scale convolutional neural network wavefront coding. *Electronics*, *12*, 4028.
6. Villar-Corrales, A., Schirrmacher, F., & Riess, C. (2021). Deep learning architectural designs for super-resolution of noisy images. In *ICASSP 2021 - IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1635–1639). Toronto, ON, Canada.
7. Zhang, K., Ren, W., Luo, W., et al. (2022). Deep image deblurring: A survey. *International Journal of Computer Vision*, *130*, 2103–2130.
8. González-Amador, E., Padilla-Vivanco, A., Toxqui-Quitl, C., Arines, J., & Acosta, E. (2020). Jacobi–Fourier phase mask for wavefront coding. *Optics and Lasers in Engineering*, *126*.
9. Reyes-Alfaro, J. M., Toxqui-Quitl, C., Espejel-Rivera, A., González-Amador, E., & Padilla-Vivanco, A. (2023). Convolutional neural network to restore images encoded by a wavefront imaging system. In *Current Developments in Lens Design and Optical Engineering XXIV* (Vol. 12666, Article 126660J). SPIE.
10. Xu, L., Ren, J. S. J., Liu, C., & Jia, J. (2014). Deep convolutional neural network for image deconvolution. In *NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems* (Vol. 1, pp. 1790–1798).
11. Jin, Z., Iqbal, M. Z., Bobkov, D., Zou, W., Li, X., & Steinbach, E. (2020). A flexible deep CNN framework for image restoration. *IEEE Transactions on Multimedia*, *22*(4), 1055–1068.
12. Dong, J., Roth, S., & Schiele, B. (2021). Deep wiener deconvolution: Wiener meets deep learning for image deblurring. *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*.
13. Wach, H. B., Dowski, E., & Cathey, W. (1998). Control of chromatic focal shift through wave-front coding. *Applied Optics*, *37*, 5359–5367.