

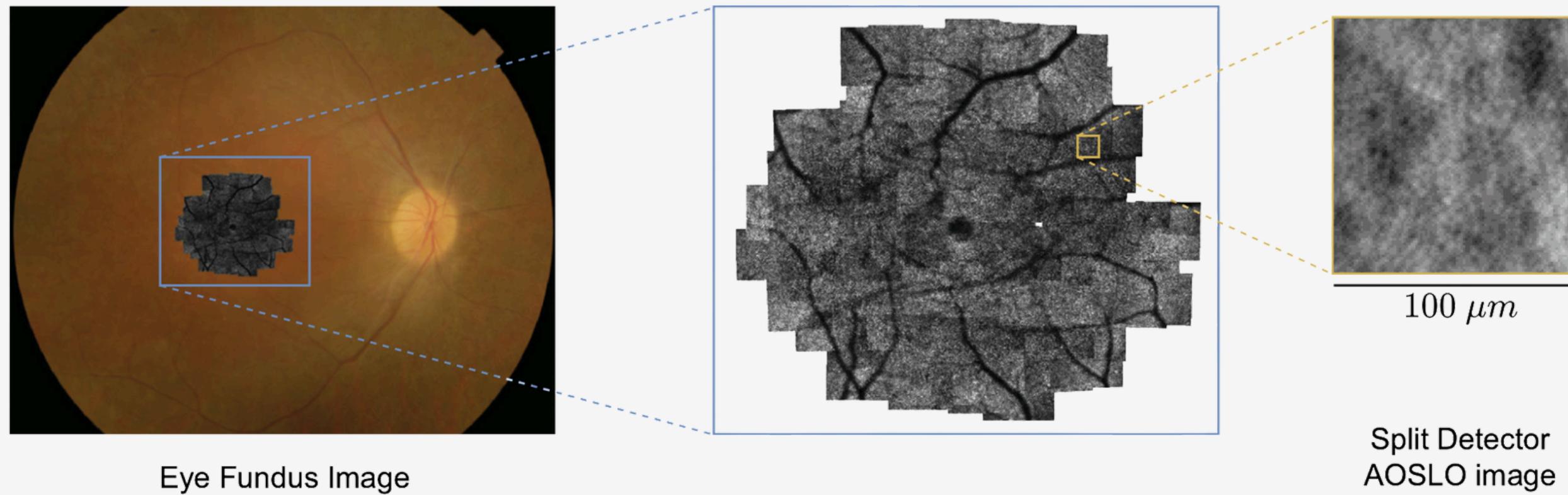
# **ESTIMACIÓN DE DENSIDAD DE CONOS EN IMÁGENES AOSLO MEDIANTE PROCESAMIENTO DE IMÁGENES Y APRENDIZAJE PROFUNDO**

Nicolay Cortez

Tutor de tesis: Santiago Toledo Cortés

# Introducción

AOSLO: Adaptive Optics Scanning Laser Ophthalmoscopy



# Introducción

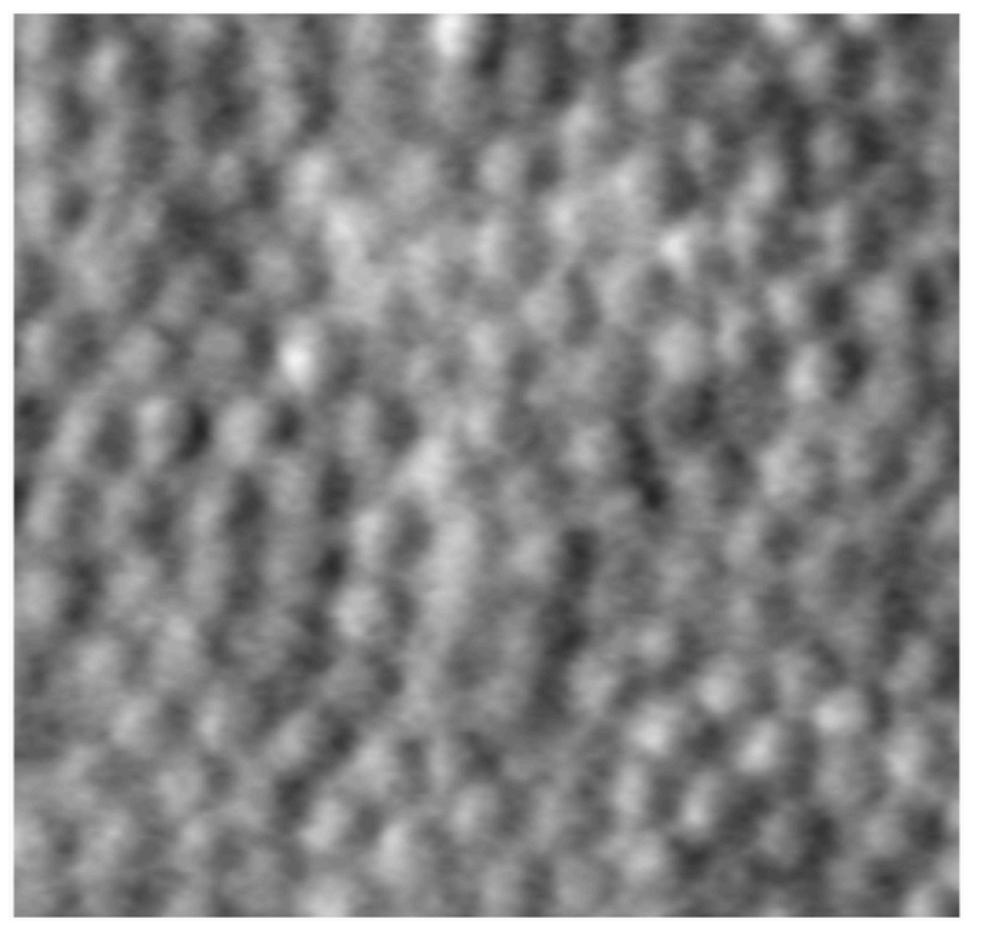
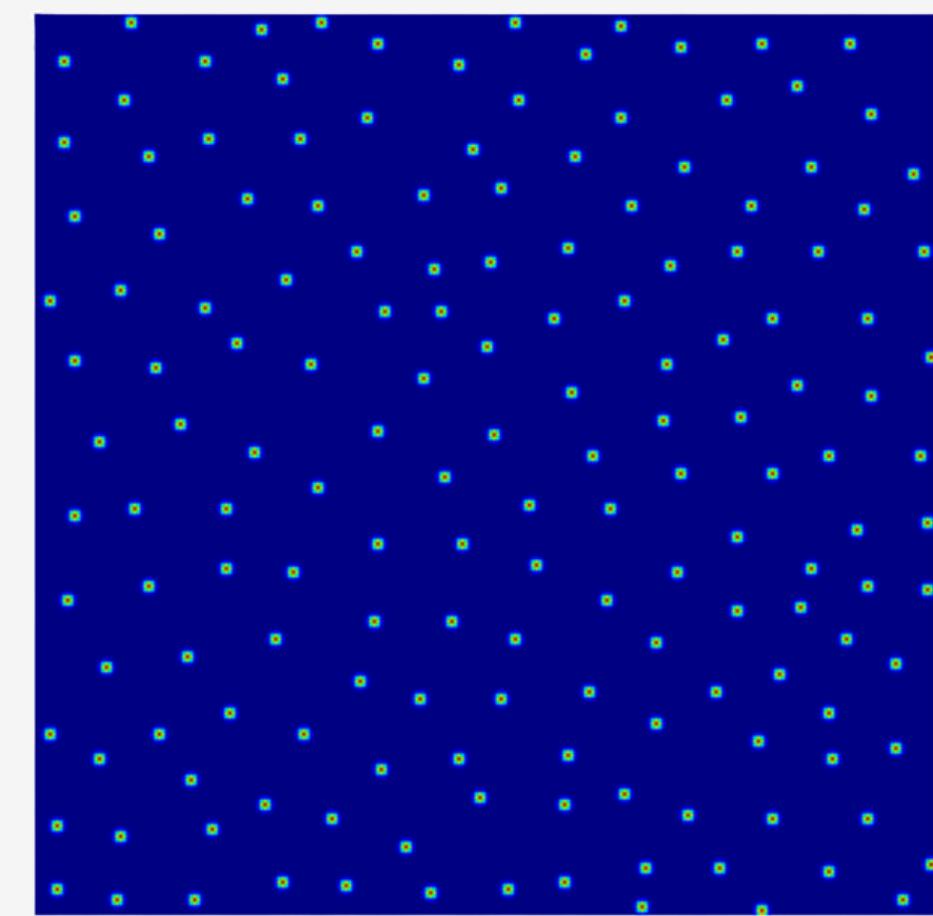
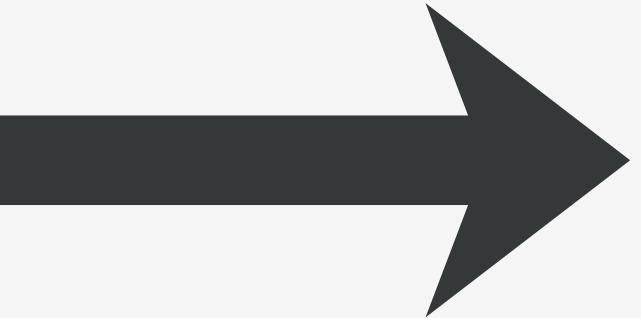


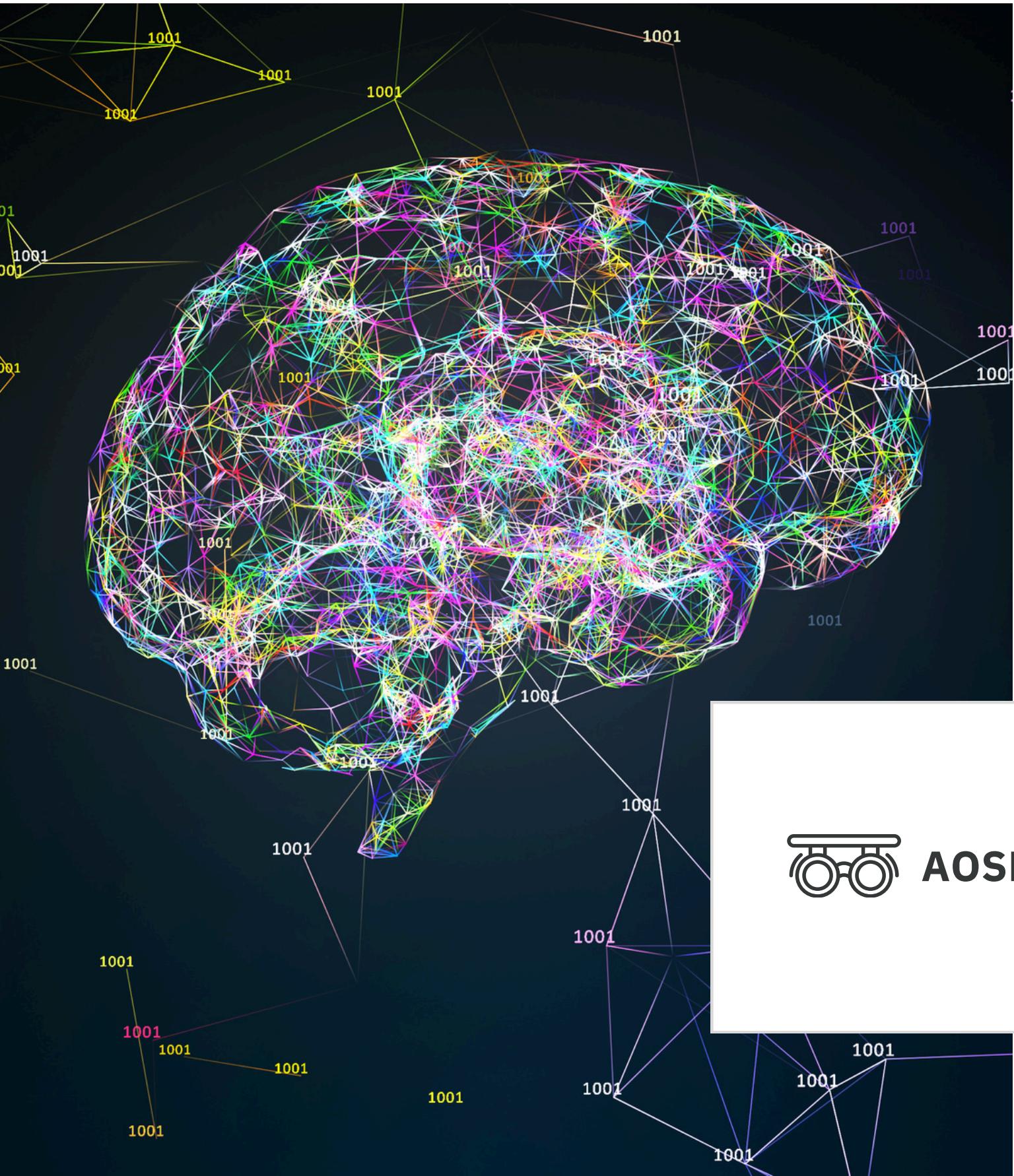
Imagen AOSLO



Anotaciones manuales

# Motivación

- **El Problema:** La detección temprana de enfermedades degenerativas (Stargardt, Retinitis Pigmentosa) depende de la densidad de conos. Relación también con la ubicación de los conos.
- **El Desafío:** El conteo manual es lento y propenso a errores.
- **Meta:** Democratizar el análisis AOSLO mediante modelos eficientes que funcionen en estaciones de trabajo clínicas estándar.



Aprendizaje  
profundo

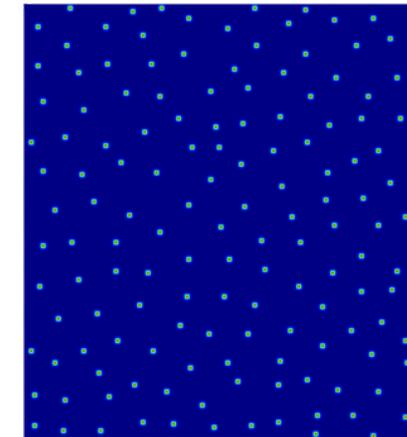
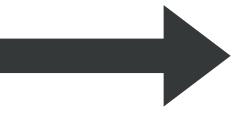
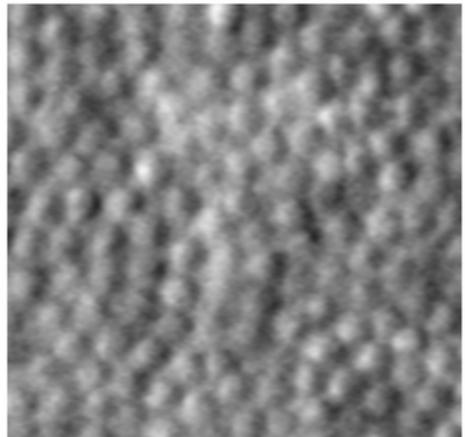


Diagnóstico  
enfermedades  
oculares

# Problema

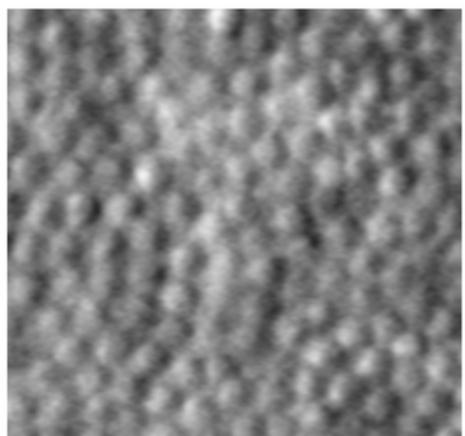
## Problema 1

### Regresión de mapa de densidad



## Problema 2

### Regresión para conteo de conos



**155 conos**

**Problem 1 (Density Map Regression):** Learn a mapping  $f_{\theta} : \mathcal{X} \rightarrow \mathcal{Y}_{\text{map}} \subset \mathbb{R}^{256 \times 256}$ , where  $\mathcal{Y}_{\text{map}} \subset \mathbb{R}^{256 \times 256}$  is the space of continuous density maps. For each training example  $(\mathbf{x}_i, \mathbf{y}_i)$ , the label  $\mathbf{y}_i = \mathbf{D}_i \in \mathbb{R}^{256 \times 256}$  is a ground truth density map encoding the spatial distribution of cone photoreceptors

- **Output del modelo:** Mapa de densidad de  $256 \times 256 \times 1$  con activación ReLU
- **Ground truth:** Mapa de densidad anotado ( $256 \times 256 \times 1$ )
- **Función de Pérdida:** MSE (Mean Squared Error) píxel por píxel

$$\mathcal{L}_{\text{density}}(\theta) = \frac{1}{n} \sum_{k=1}^n \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \left( D_k(i, j) - \hat{D}_{\theta}(\mathbf{x}_k)(i, j) \right)^2$$

**Problem 2 (Direct Count Regression):** Learn a mapping  $g_{\phi} : \mathcal{X} \rightarrow \mathcal{Y}_{\text{count}} \subset \mathbb{R}^+$ , where  $\mathcal{Y}_{\text{count}}$  is the space of non-negative scalar cone counts. For each training example  $(\mathbf{x}_i, \mathbf{y}_i)$ , the label  $\mathbf{y}_i = C_i \in \mathbb{R}^+$  is a scalar value representing the total cone count

- **Output del modelo:** Escalar (conteo total de conos)
- **Ground truth:** Conteo total de conos calculado
- **Función de Pérdida:** MSE sobre el conteo total

$$\mathcal{L}_{\text{count}}(\phi) = \frac{1}{n} \sum_{k=1}^n \left( C_k - \hat{C}_{\phi}(\mathbf{x}_k) \right)^2$$

# Preguntas de investigación

RQ1: ¿Pueden las arquitecturas ligeras de aprendizaje profundo (<600K parámetros) alcanzar una precisión comparable a la de los modelos de última generación (2,06M+ parámetros) para la estimación de la densidad de conos en imágenes AOSLO?

RQ2: ¿Cómo se comparan los enfoques de regresión de mapas de densidad (Problema 1) y regresión directa del conteo (Problema 2) en términos de precisión, eficiencia computacional y utilidad clínica?

RQ3: ¿Cuáles son las decisiones clave de diseño arquitectónico (profundidad del codificador-decodificador, conexiones de salto, representaciones de salida) que equilibran la complejidad del modelo con la precisión de la estimación?

# Objetivos

## Objetivo general:

Desarrollar, optimizar y evaluar técnicas de Deep Learning ligero para la estimación de densidad de conos en imágenes AOSLO.

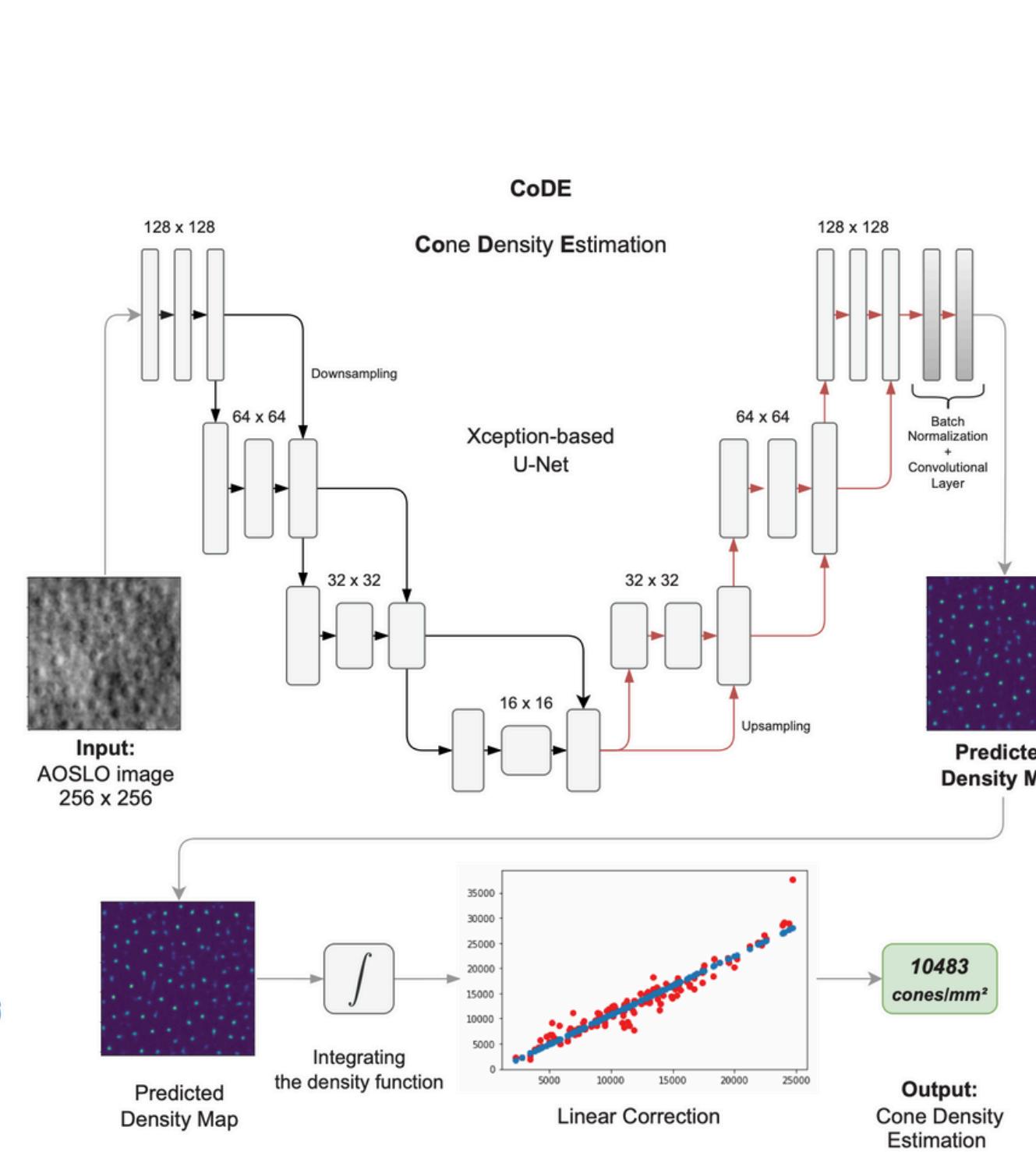
## Objetivos específicos:

Explorar y analizar conjuntos de datos de imágenes AOSLO y revisar los modelos existentes de aprendizaje profundo para la estimación de la densidad de conos.

Entrenar modelos de aprendizaje profundo utilizando conjuntos de datos grandes y diversificados para estimar con precisión la densidad de conos en imágenes AOSLO.

Evaluar de forma sistemática los modelos propuestos para mejorar la precisión del conteo de conos en imágenes AOSLO.

# Revisión Literatura



**Figure 2.** CoDE model architecture for cone density estimation on split detector AOSLO images. The original image is input into a modified Xception-based U-Net<sup>21,25</sup> to generate a density map of the cones in the original image. The integral over this density map is linearly corrected to provide an accurate estimation of the number of cones in the image.

## Deep Density Estimation for Cone Counting and Diagnosis of Genetic Eye Diseases From Adaptive Optics Scanning Light Ophthalmoscope Images

Santiago Toledo-Cortés, Adam M. Dubis, Fabio A. González y Henning Müller.

**tvst**  
Artificial Intelligence  
**Deep Density Estimation for Cone Counting and Diagnosis of Genetic Eye Diseases From Adaptive Optics Scanning Light Ophthalmoscope Images**  
Santiago Toledo-Cortés<sup>1,2</sup>, Adam M. Dubis<sup>3,4</sup>, Fabio A. González<sup>2</sup>, and Henning Müller<sup>5,7</sup>

<sup>1</sup> Department of TI and Process Optimization, Faculty of Engineering, Universidad de La Sabana Campus Puente del Común km 7, Chía, Colombia  
<sup>2</sup> MindLab Research Group, Universidad Nacional de Colombia, Bogotá, Colombia  
<sup>3</sup> Moorfields Eye Hospital NHS Foundation Trust, London, Institute of Ophthalmology, University College London, London, UK  
<sup>4</sup> Global Business School for Health, University College London, London, UK  
<sup>5</sup> Institute of Information Systems, HES-SO (University of Applied Sciences and Arts Western Switzerland), Sierre, Switzerland  
<sup>6</sup> Medical Faculty, University of Geneva, Switzerland  
<sup>7</sup> The Sense research and innovation center, Sion and Lausanne, Switzerland

**Purpose:** Adaptive optics scanning light ophthalmoscope (AOSLO) imaging offers a microscopic view of the living retina, holding promise for diagnosing and researching eye diseases like retinitis pigmentosa and Stargardt's disease. The technology's clinical impact of AOSLO hinges on early detection through automated analysis tools.

**Methods:** We introduce Cone Density Estimation (CoDE) and CoDE for Diagnosis (CoDED). CoDE is a deep density estimation model for cone counting that estimates a density function whose integral is equal to the number of cones. CoDED is an integration of CoDE with deep image classifiers for diagnosis. We use two AOSLO image datasets to train and evaluate the performance of cone density estimation and classification models for retinitis pigmentosa and Stargardt's disease.

**Results:** Bland-Altman plots show that CoDE outperforms state-of-the-art models for cone density estimation. CoDED reported an F1 score of  $0.770 \pm 0.04$  for disease classification, outperforming traditional convolutional networks.

**Conclusion:** CoDE shows promise in classifying the retinitis pigmentosa and Stargardt's disease cases from a single AOSLO image. Our preliminary results suggest the potential role of analyzing patterns in the retinal cellular mosaic to aid in the diagnosis of genetic eye diseases.

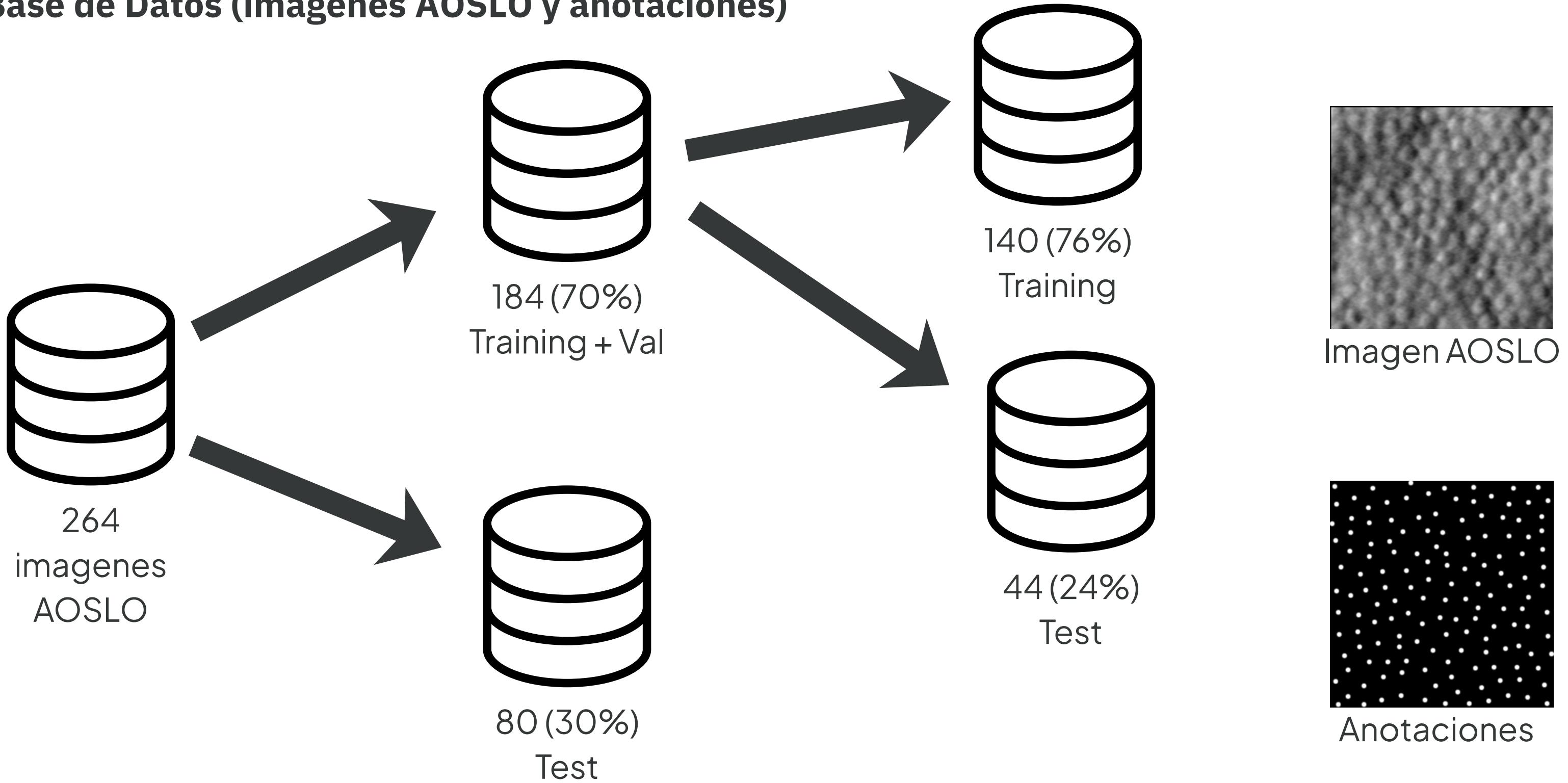
**Translational Relevance:** Our study explores the potential of deep density estimation models to aid in the analysis of AOSLO images. Although the initial results are encouraging, more research is needed to fully realize the potential of such methods in the treatment and study of genetic retinal pathologies.

**Introduction**  
**Motivation**  
The world of ophthalmology has been transformed by our ability to image the back of the eye. Inspection of the ocular fundus allows specialists to detect signs of degenerative diseases that may even extend beyond the visual system, such as retinopathy caused by diabetes mellitus. In general, the density and regularity of the pattern of photoreceptor cells in the retina is affected in diverse ways by different diseases.<sup>1</sup> However, even with the best clinical cameras, the loss of hundreds of thousands of retinal cells cannot be quantified by the time the macroscopic changes of disease are detected.<sup>2</sup>

Copyright 2023 The Authors  
tvst.arvojournals.org | ISSN: 2164-2591  
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

# Dataset

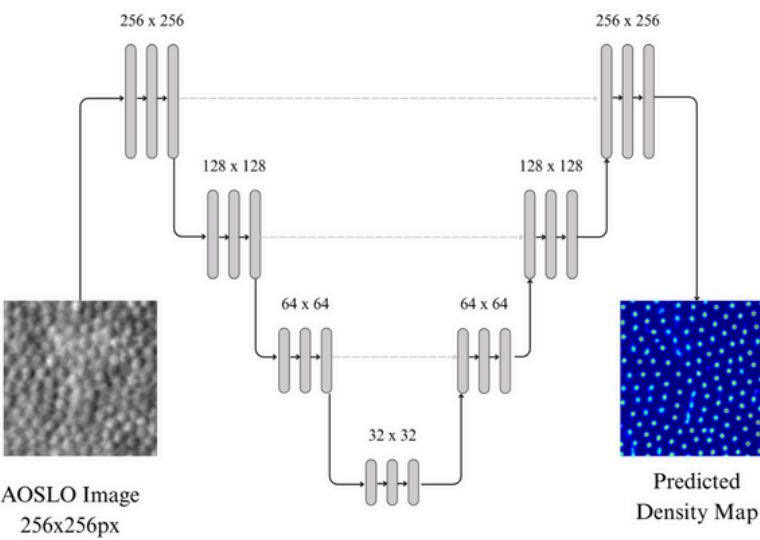
## Base de Datos (imágenes AOSLO y anotaciones)



# Arquitecturas

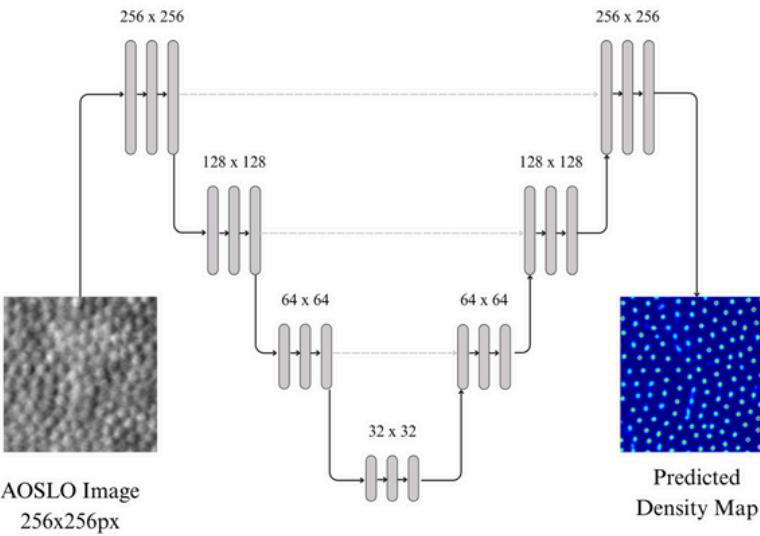
## Problema 1 Regresión de mapa de densidad

### Modelo A



- Encoder:** Three downsampling blocks with 16, 32, and 64 filters respectively
- Bottleneck:** 128 filters with 0.5 dropout for regularization
- Decoder:** Three upsampling blocks with skip connections
- Output:** Single channel density map with ReLU activation
- Total parameters:** 538,609 (significant reduction from baseline)

### Modelo B



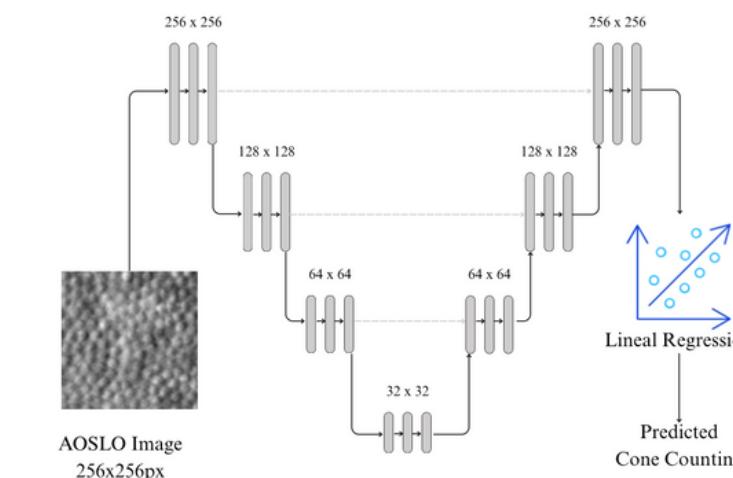
- Encoder:** Three downsampling blocks with 16, 32, and 64 filters respectively  
**Bottleneck:** 128 filters with 0.5 dropout for regularization  
**Decoder:** Three upsampling blocks with skip connections  
**Output:** Single channel density map with ReLU activation  
**Total parameters:** 538,609 (significant reduction from baseline)

Same base architecture as Model A  
Additional linear layer:  $\hat{C}_{corrected} = \alpha \cdot \hat{C}_{raw} + \beta$

+ Regresión Lineal Posterior

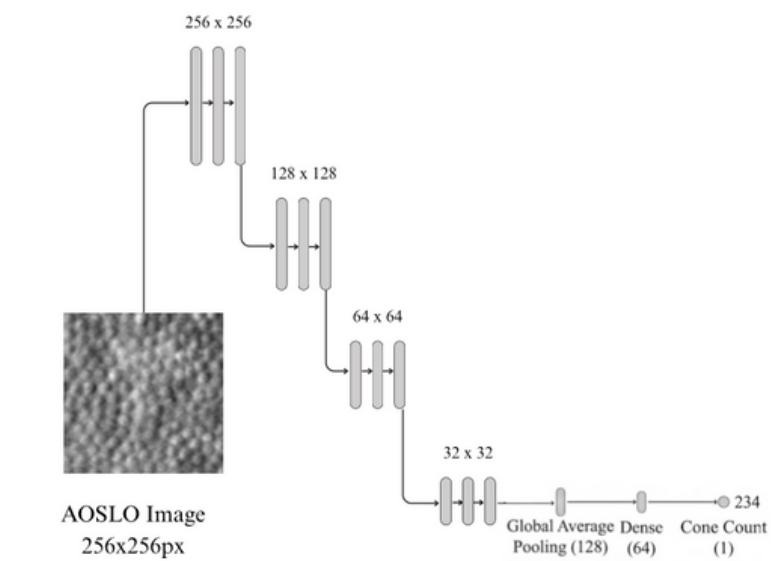
## Problema 2 Regresión para conteo de conos

### Modelo C



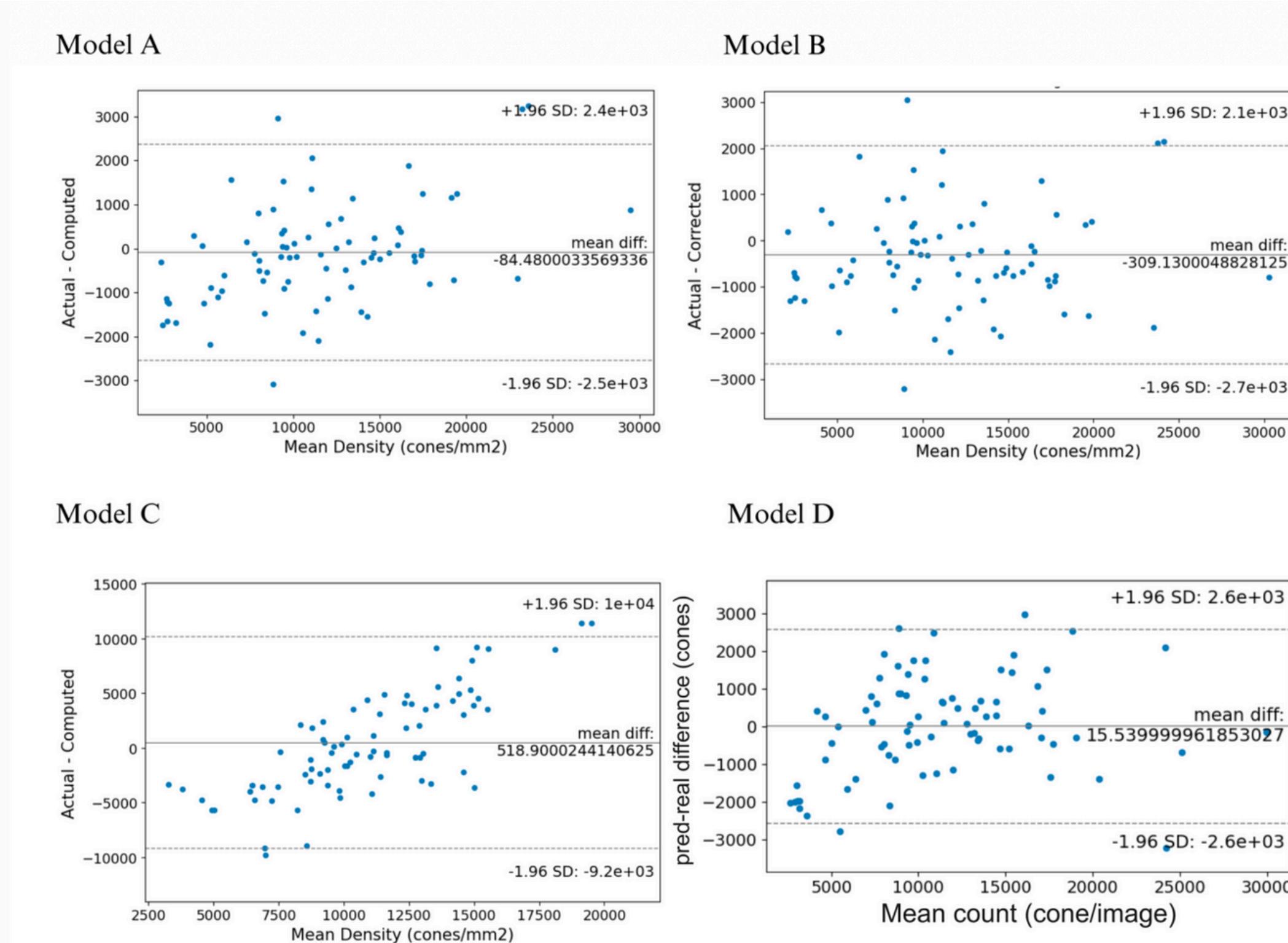
- Base U-Net architecture producing intermediate density map
- Global sum layer (differentiable):  $\hat{C} = \sum_{i,j} \hat{D}_{i,j}$
- Dense layer with single neuron for final count
- Trained with count-level MSE loss on scalar outputs

### Modelo D



- Encoder-only architecture
- Global average pooling after final encoder stage
- Dense layers for direct count regression
- Minimal parameters but no spatial interpretability

# Resultados



- Modelo A (arriba-izq): Muestra un sesgo sistemático moderado (mean bias) de -84.5 conos. Sin embargo, sus límites de acuerdo son amplios ( $\pm 2.5 \times 10^3$ ), lo que indica una varianza considerable en las predicciones individuales a pesar de un error medio aceptable.
- Modelo B (arriba-der): Exhibe una sobreestimación sistemática severa, evidenciada por un sesgo medio negativo de -309.1 conos, manteniendo una dispersión similar a la del Modelo A.
- Modelo C (abajo-izq): Presenta el desempeño más deficiente. Se observa un claro sesgo proporcional, donde el error aumenta linealmente con la densidad de conos. Sus límites de acuerdo son inaceptablemente amplios (de  $-9.2 \times 10^3$  a  $+1.0 \times 10^4$ ), lo que indica que este enfoque no es fiable.
- Modelo D (abajo-der): Demuestra la mejor calibración de todos los experimentos. Su sesgo medio es mínimo (+15.5 conos), lo que indica la mayor exactitud sistemática. Aunque sus límites de acuerdo ( $\pm 2.6 \times 10^3$ ) son comparables a los del Modelo A, la ausencia de un sesgo significativo lo convierte en el modelo más robusto."

# Resultados

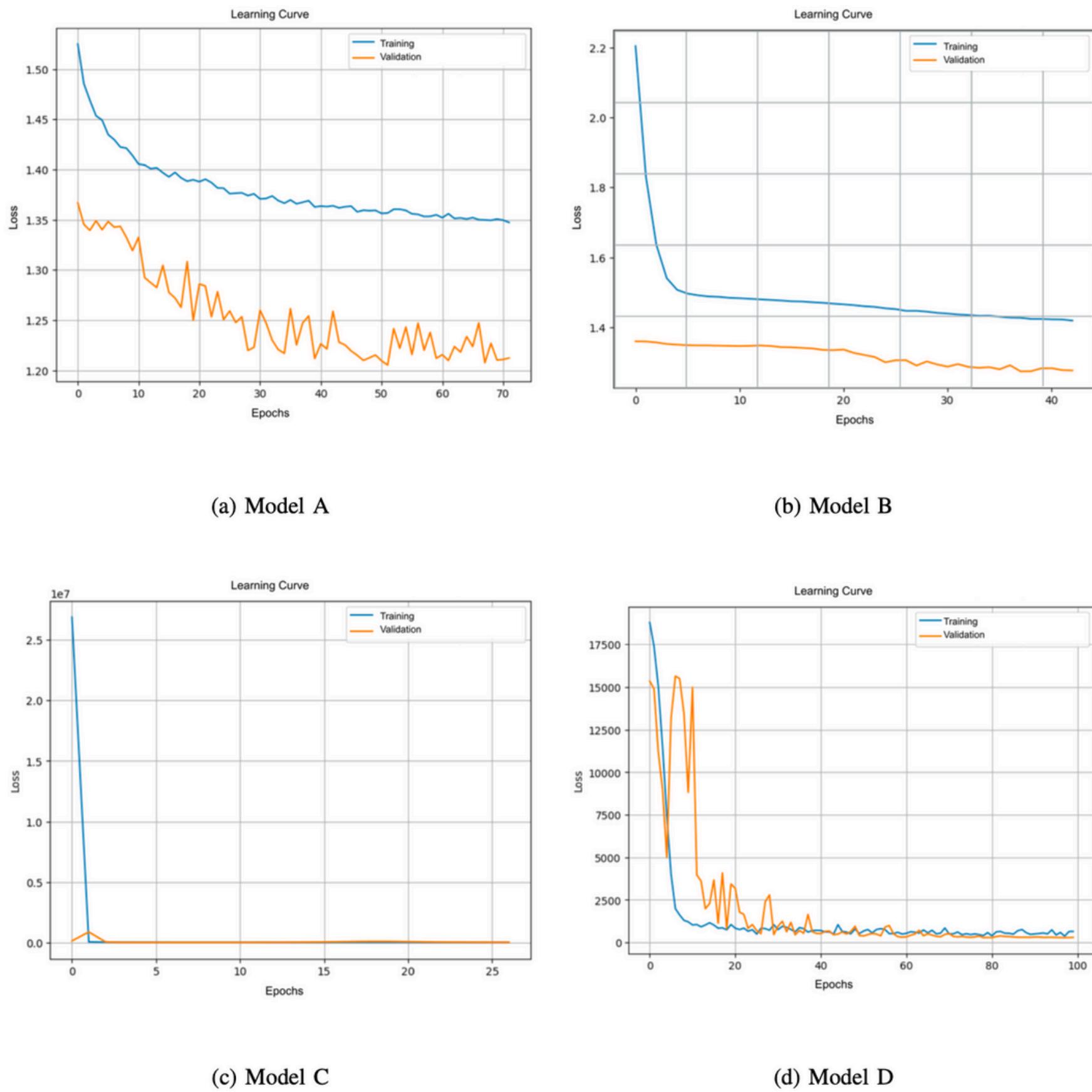


TABLE II  
PERFORMANCE METRICS AND MODEL COMPLEXITY COMPARISON (LOWER IS BETTER FOR ERROR METRICS).

Model	MAE (cones)	RMSE (cones)	MAPE (%)	Params	Size (MB)
<b>Baseline</b>	<b>1534.57</b>	—	—	<b>2.06M</b>	<b>16.0</b>
Model A	<b>921.62</b>	1255.19	13.64	538K	6.4
Model B	<b>977.05</b>	1253.57	12.80	538K	6.4
Model C	3884.54	4970.76	55.29	~540K	6.4
Model D	<b>1043.95</b>	1317.45	17.08	380K	3.6

*Best improvement:* Model A achieves 40% lower MAE with 74% fewer parameters

- (a) El Modelo A demuestra una convergencia estocástica estable sin una brecha de divergencia, lo que confirma que la arquitectura ligera previene el sobreajuste (overfitting) a pesar del pequeño tamaño de lote (batch size).
- (b) El Modelo B exhibe una pérdida de validación consistentemente menor que la de entrenamiento, una firma característica de la fuerte regularización por Dropout (0.5) que previene eficazmente la memorización.
- (c) El Modelo C ilustra un colapso de la optimización (nótese la escala de 1e7), lo que indica que la supervisión mediante conteo escalar por sí sola es insuficiente para guiar los pesos espaciales de la U-Net.
- (d) El Modelo D muestra una alta volatilidad inicial, típica de la regresión directa sin prioris espaciales, pero finalmente logra una convergencia suave y estable.

# Resultados

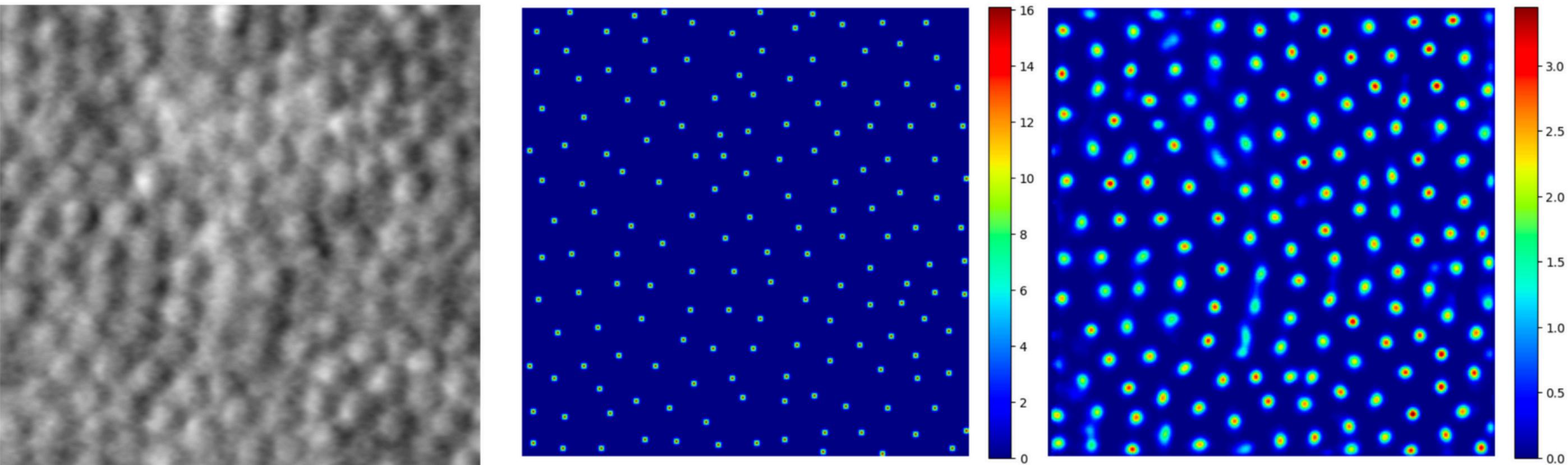


Fig. 5. Qualitative comparison of Model A predictions on representative AOSLO images. Left column: Original adaptive optics scanning laser ophthalmoscopy images showing individual cone photoreceptors as circular structures. Middle column: Ground-truth density maps generated from expert manual annotations, where bright pixels indicate cone centroid locations. The maps preserve the spatial distribution of cones while providing continuous probability surfaces suitable for integration-based counting. Right column: Model A predictions showing remarkable spatial correspondence with ground truth. The predicted density maps successfully capture both the overall cone count and the characteristic spatial patterns.

# Conclusiones



Eficiencia sobre Capacidad Bruta: Se demostró que arquitecturas ligeras (<600K parámetros) pueden superar a modelos complejos del estado del arte, logrando un 40% más de precisión (Modelo A) con una fracción del costo computacional y memoria.



La arquitectura óptima depende del objetivo clínico: se debe elegir un diseño "encoder-decoder" si se requiere validación visual (mapas), o cambiar a "encoder-only" si la prioridad es solo el conteo y la eficiencia.



Trade-off Metodológico: los mapas de densidad ofrecen interpretabilidad para diagnóstico detallado, mientras que el conteo directo ofrece mejor calibración y velocidad.



Viabilidad Clínica y Tiempo Real: La eficiencia computacional alcanzada habilita el procesamiento de imágenes en tiempo real durante el escaneo, eliminando la barrera del conteo manual laborioso y facilitando la adopción clínica.



Generalización en Datos Escasos: En regímenes de pocos datos ( $N=140$ ), las restricciones severas de capacidad y la aumentación de datos demostraron ser mecanismos de regularización más efectivos que aumentar la profundidad del modelo.

# Bibliografia y Referencias

- [1] S. Nakatake, Y. Murakami, J. Funatsu, Y. Koyanagi, M. Akiyama, Y. Momozawa, T. Ishibashi, K. H. Sonoda, and Y. Ikeda, "Early detection of cone photoreceptor cell loss in retinitis pigmentosa using adaptive optics scanning laser ophthalmoscopy," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 257, pp. 1169–1181, 6 2019.
- [2] A. Roorda, F. Romero-Borja, W. J. Donnelly, H. Queener, T. J. Hebert, M. C. W. Campbell, G. R. H. W. Webb, O. Hughes, W. J. Donnelly, F. Romero-Borja, and A. Roorda, "Adaptive optics scanning laser ophthalmoscopy," *Report*, vol. 7, pp. 3–14, 1988. [Online]. Available: <http://www.opt.uh.edu/research/aroorda/aoslo.htm>,
- [3] D. Cunefare, L. Fang, R. F. Cooper, A. Dubra, J. Carroll, and S. Farsiu, "Open source software for automatic detection of cone photoreceptors in adaptive optics ophthalmoscopy using convolutional neural networks," *Scientific Reports*, vol. 7, 12 2017.
- [4] S. Toledo-Cortés, A. M. Dubis, F. A. González, and H. Müller, "Deep density estimation for cone counting and diagnosis of genetic eye diseases from adaptive optics scanning light ophthalmoscope images," *Translational Vision Science I& Technology*, vol. 12, no. 11, p. 25, November 2023. [Online]. Available: <https://doi.org/10.1167/tvst.12.11.25>
- [5] D. Cunefare, A. L. Huckenpahler, E. J. Patterson, A. Dubra, J. Carroll, and S. Farsiu, "Rac-cnn: multimodal deep learning based automatic detection and classification of rod and cone photoreceptors in adaptive optics scanning light ophthalmoscope images," *Biomedical Optics Express*, vol. 10, p. 3815, 8 2019.
- [6] Q. Zhang, K. Sampani, M. Xu, S. Cai, Y. Deng, H. Li, J. K. Sun, and G. E. Karniadakis, "Aoslo-net: A deep learning-based method for automatic segmentation of retinal microaneurysms from adaptive optics scanning laser ophthalmoscope images," 6 2021. [Online]. Available: <http://arxiv.org/abs/2106.02800>
- [7] B. Davidson, A. Kalitzeos, J. Carroll, A. Dubra, S. Ourselin, M. Michaelides, and C. Bergeles, "Automatic cone photoreceptor localisation in healthy and stargardt afflicted retinas using deep learning," *Scientific Reports*, vol. 8, 12 2018.
- [8] W. Xie, J. Noble, and A. Zisserman, "Microscopy cell counting and detection with fully convolutional regression networks," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging Visualization*, pp. 1–10, 05 2016.
- [9] M. Kulyabin, A. Sindel, H. R. Pedersen, S. Gilson, R. Baraas, and A. Maier, "Generalist segmentation algorithm for photoreceptors analysis in adaptive optics imaging," *Electrical Engineering and Systems Science*, p. 168–182, Dec. 2024. [Online]. Available: [http://dx.doi.org/10.1007/978-3-031-78104-9\\_12](http://dx.doi.org/10.1007/978-3-031-78104-9_12)
- J. Amann, A. Blasimme, E. Vayena, D. Frey, V. I. Madai, and the Precise4Q Consortium, "Explainability for artificial intelligence in healthcare: A multidisciplinary perspective," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, p. 310, 11 2020. [Online]. Available: <https://doi.org/10.1186/s12911-020-01332-6>
- A. M. Antoniadi, Y. Du, Y. Guendouz, L. Wei, C. Mazo, B. A. Becker, and C. Mooney, "Current challenges and future opportunities for xai in machine learning-based clinical decision support systems: A systematic review," *Applied Sciences*, vol. 11, no. 11, p. 5088, 5 2021. [Online]. Available: <https://doi.org/10.3390/app11115088>



# Gracias