

**UNIVERSIDAD DE VALLADOLID**

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE  
TELECOMUNICACIÓN

**TRABAJO FIN DE GRADO**

GRADO EN INGENIERÍA DE TECNOLOGÍAS DE  
TELECOMUNICACIÓN

**Procesado de retinografías  
basado en Deep Learning  
para la ayuda al diagnóstico  
de la Retinopatía Diabética**

Autora:

**D<sup>a</sup>. María Herrero Tudela**

Tutores:

**Dra. D<sup>a</sup>. María García Gadañón**

**D. Roberto Romero Oraá**

Valladolid, 23 de Septiembre de 2020



---

**TÍTULO:                    Procesado de retinografías basado  
en Deep Learning para la ayuda al  
diagnóstico de la Retinopatía  
Diabética**

**AUTOR:                    D<sup>a</sup>. María Herrero Tudela**

**TUTORA:                   Dra. D<sup>a</sup>. María García Gadañón  
D. Roberto Romero Oraá**

**DEPARTAMENTO:        Departamento de Teoría de la Señal  
y Comunicaciones e Ingeniería  
Telemática**

---

**TRIBUNAL**

---

**PRESIDENTE:         Dr. D. Roberto Hornero Sánchez**

**SECRETARIO:         Dr. D. Jesús Poza Crespo**

**VOCAL:                 Dra. D<sup>a</sup>. María García Gadañón**

**SUPLENTE:            Dr. D. Carlos Gómez Peña**

**SUPLENTE:            Dr. D. Miguel López-Coronado  
Sánchez-Fortún**

---

---

**FECHA:                 23 de Septiembre de 2020**

**CALIFICACIÓN:**

---



*A mis padres*



## **AGRADECIMIENTOS**

Son muchas las personas que han contribuido a la realización de este Trabajo Fin de Grado. Es por este motivo que considero necesario utilizar esta página para mostrarles mi más sincero agradecimiento.

Durante todo este tiempo he tenido la suerte de contar con el apoyo de los miembros del Grupo de Ingeniería Biomédica de la Universidad de Valladolid. Muchas gracias por la ayuda y los consejos que me habéis dado. Quiero hacer mención especial a mi tutor Rober. Muchas gracias por resolver gran parte de mis dudas, darme consejos y facilitarme tanto el trabajo. También quiero mostrar toda mi gratitud y reconocimiento a mi tutora María García por guiarme en la realización de este trabajo. A ellos les debo haber podido completar este trabajo.

Durante mi etapa universitaria, he de mencionar a todos mis compañeros y compañeras con los que he compartido parte de esta experiencia. Muchas gracias a todos, pero especialmente a Laura Lesarri por haber hecho más llevaderos tantos momentos de trabajo.

Tampoco puedo olvidar la inestimable ayuda que me han prestado muchas personas en el plano personal. A mis amigos, les agradezco su apoyo y paciencia. A mi familia le agradezco su absoluto apoyo sin el cual nunca hubiera podido llegar hasta aquí. En especial, quiero agradecer a mis padres su amor, su dedicación, su confianza en mí en todo momento y sus ánimos cuando nada salía bien. Finalmente, me gustaría agradecer a Dani su apoyo, cariño y comprensión, tanto en los momentos buenos como en los momentos malos.





## RESUMEN

La Retinopatía Diabética (RD) es una complicación de la diabetes y es la causa más frecuente de ceguera en la población laboral activa de los países desarrollados. Sin embargo, cuando se trata de forma precoz, más del 90% de la pérdida de visión se puede prevenir. Las retinografías capturadas durante exámenes oculares regulares son el método estándar para detectar RD. No obstante, el aumento de los casos de diabetes a nivel mundial y la falta de especialistas dificultan el diagnóstico. Las imágenes de fondo de ojo generalmente se obtienen usando cámaras de fondo de ojo en condiciones de luz y ángulos variados. Por lo tanto, estas imágenes son propensas a una iluminación no uniforme, contraste deficiente, bajo brillo y falta de nitidez, lo que provoca imágenes borrosas. Estas imágenes borrosas o con falta de iluminación podrían afectar el diagnóstico clínico. Por lo tanto, mejorar estas imágenes de calidad insuficiente puede ser muy útil para evitar diagnósticos erróneos en sistemas de cribado automáticos o manuales.

Recientemente, el aprendizaje automático, especialmente las técnicas basadas en *Deep Learning*, han supuesto una revolución en el campo de la reconstrucción de imágenes. Por ello, en este trabajo, se propone un método de mejora de calidad de retinografías basado en redes de generativas antagónicas (*Generative Adversarial Network*, GAN). El modelo está formado por dos redes neuronales convolucionales: una red neuronal que actúa como generador de imágenes sintéticas con el objetivo de engañar a una red discriminadora que está entrenada para distinguir las imágenes generadas de alta calidad de las imágenes reales. Este modelo puede funcionar con imágenes de gran resolución, lo que lo hace ampliamente beneficioso para las imágenes clínicas. En este trabajo, la mejora de calidad de la imagen de fondo de ojo abarca una fase de corrección de la nitidez y una segunda fase de corrección de la iluminación.

Para el desarrollo y validación del método propuesto, se utilizó una base de datos propia de 1000 imágenes. Dichas imágenes se dividieron en un conjunto de entrenamiento con 800 imágenes de entrenamiento y un conjunto de test con 200 imágenes, de las cuales la mitad tenían calidad insuficiente para su análisis. Sobre ellas, se aplicó un método con varias etapas. En primer lugar, se abordó la mejora de imágenes borrosas empleando una red profunda de tipo GAN. En segundo lugar, se abordó la mejora de imágenes con falta de iluminación, también a través de una red GAN. Cualitativamente, los resultados obtenidos son satisfactorios. Asimismo, se abordó la evaluación cuantitativa de los resultados desde dos perspectivas: evaluación con referencia y evaluación sin referencia. Para la evaluación sin referencia, se utilizan las métricas *Blind/Referenceless Image Spatial Quality Evaluator* (BRISQUE), *Natural Image Quality Evaluator* (NIQE) y entropía. En cuanto a la evaluación con una imagen de referencia, se utilizaron la relación señal a ruido (*Peak Signal-to-Noise Ratio*, PSNR) y el índice de similitud estructural (*Structural Similarity Index Measure*, SSIM). La evaluación con referencia sirve como guía para comparar las imágenes de buena calidad que han sido degradadas intencionadamente. Por otra parte, la evaluación sin referencia es necesaria para evaluar la mejora que el método produce sobre imágenes de mala calidad ya que, de partida, no se dispone de una versión de buena calidad de dichas imágenes. En la fase de mejora de nitidez y sobre las imágenes de test buena calidad, los resultados obtenidos muestran una mejora del 6.22%, 3.33% y 3.26% en términos de PSNR, SSIM y entropía,

respectivamente. No obstante, las medidas BRISQUE y NIQE no presentan una mejora. En esta misma etapa, pero sobre las imágenes de test mala calidad los resultados muestran un 31.80%, 4.27% y 3.89% de mejora en términos de BRISQUE, NIQE y entropía respecto a la imagen original real. Asimismo, en la fase de mejora de imágenes con falta de iluminación, los resultados sobre el conjunto de imágenes de buena calidad muestran una mejora del 156.81%, 14.59%, 3.12% y 2.28% en términos de PSNR, SSIM, BRISQUE y NIQE; mientras que la entropía no presenta una mejoría. En esta fase, y sobre el conjunto de imágenes de mala calidad los resultados reflejan una mejora del 50.62% y un 8.33% en términos de BRISQUE y entropía. Sin embargo, en este grupo de imágenes, la medida NIQE no mejora. Finalmente, se ha llevado a cabo un último experimento con ambas redes en serie. En primer lugar, las imágenes atraviesan la red que corrige la iluminación, y posteriormente se corrige su nitidez con la segunda red. Sobre las imágenes de test de buena calidad se ha conseguido un 4.84%, 5.68%, 3.38% y 2.57% de mejora respecto de la imagen original en términos de PSNR, SSIM, NIQE y entropía, aunque no se observa mejora en términos de BRISQUE. En este último experimento, y sobre las imágenes de test de mala calidad se ha obtenido un 88.95%, 21.17% y 2.46% de mejora en términos de BRISQUE, NIQE y entropía. Los resultados obtenidos muestran que el método propuesto podría ser utilizado como primera etapa dentro de sistemas automáticos de análisis de retinografías para la ayuda al diagnóstico de diversas enfermedades oculares.

## **PALABRAS CLAVE**

Análisis de retinografías; Deep Learning; Mejora de calidad; Procesado digital de imágenes; Redes generativas antagónicas.

## ABSTRACT

Diabetic Retinopathy (DR) is a complication of diabetes and the leading cause of blindness worldwide. However, when treated early, more than 90% of vision loss can be prevented. Color fundus photography has been the standard method for detecting DR. However, the growing incidence of diabetes and the lack of specialists make diagnosis difficult. Fundus images are generally obtained using fundus cameras in varied light conditions and angles. Thence, these images are prone to non-uniform illumination, poor contrast, low brightness and lack of sharpness resulting in blurry images. These blurry or poor illuminated images could affect clinical diagnosis. Therefore, improving these poor-quality images can be very helpful in avoiding misdiagnosis in automatic or manual screening systems.

Recently, machine learning, especially deep learning techniques, have brought revolution to image super resolution reconstruction. For this reason, in this work, we propose a retinal fundus image enhancement method based on Generative Adversarial Networks (GAN). The model is composed of two convolutional neural networks: a neural network that acts as a generator of synthetic images with the aim of tricking a discriminating network that is trained to distinguish high-quality generated images from real images. This model can work with high resolution images, which makes it widely beneficial for clinical images. In this work, the fundus image enhancement method includes both the sharpness correction and the lighting correction.

The proposed technique was evaluated in a proprietary database of 200 images, of which half were of insufficient quality. A method with several stages was applied to them. Firstly, blurry image enhancement was addressed by a GAN network. Secondly, the improvement of images with lack of lighting was addressed, also through a GAN network. To evaluate the retinal image enhancement performance, visual and quantitative evaluation were carried out. Two kinds of image quality assessment were adopted: full-reference and no-reference evaluation. For no-reference assessment, Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), Natural Image Quality Evaluator (NIQE) and Entropy were chosen to assess each enhanced image and its original blurry retinal image. As to full-reference assessment, Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) were used. SSIM and PSNR give the comparison between the enhanced image and the original image. Quantitatively, in the blurred image improvement phase using good quality images, the results obtained show that it is possible to achieve an improvement of 6.22%, 3.33% and 3.26% in terms of PSNR, SSIM and entropy. However, the BRISQUE and NIQE measures do not show an improvement. In this same stage, but on the images of poor quality, the results show a 31.80%, 4.27% and 3.89% improvement in terms of BRISQUE, NIQE and entropy with respect to the real original image. Likewise, in the improvement phase of images with lack of lighting, the results on the set of good quality images show an improvement of 156.81%, 14.59%, 3.12% and 2.28% in terms of PSNR, SSIM, BRISQUE and NIQE; while entropy does not improve. In this phase, using the set of poor-quality images, the results reflect an improvement of 50.62% and 8.33% in terms of BRISQUE and entropy. However, in this group of images the NIQE measure does not improve. Finally, a last experiment was carried out with both networks. First, the images passed through the GAN network that

corrected their lighting, and then their sharpness was corrected with the second GAN network. On the good quality test images, the results obtained show an improvement of 4.84%, 5.68%, 3.38% and 2.57% in terms of PSNR, SSIM, NIQE and entropy, although the BRISQUE measure does not improve. In this last experiment, and on the poor-quality test images, the results show an improvement of 88.95%, 21.17% and 2.46% in terms of BRISQUE, NIQE and entropy. The results indicate that the proposed method could be used as a first stage in automatic retinography analysis systems to aid in the diagnosis of various eye diseases.

## **KEYWORDS**

Deep Learning; Digital Image Processing; Generative Adversarial Networks; Retinal Fundus Image Enhancement; Retinal Imaging.

## **GLOSARIO DE TÉRMINOS Y ACRÓNIMOS**

**ANN.** Red neuronal artificial (*Artificial Neural Network*)

**BD.** Base de Datos

**BRISQUE.** *Blind/Referenceless Image Spatial Quality Evaluator*

**CBIR.** Recuperación de imágenes basada en contenido (*Content-Based Image Retrieval*)

**CLAHE.** Ecuación de histograma adaptativo limitado por contraste (*Contrast-Limited Adaptive Histogram Equalization*)

**CNN.** Red neuronal convolucional (*Convolutional Neural Network*)

**DM.** Diabetes Mellitus

**ECV.** Enfermedades cardiovasculares

**EMD.** Edema macular diabético

**EPR.** Epitelio pigmentario de la retina

**ERM.** Minimización del riesgo empírico (*Empirical Risk Minimization*)

**FEDER.** Fondo Europeo de Desarrollo Regional

**FOV.** Campo de visión (*Field of View*)

**GAN.** Red Generativa Antagónica (*Generative Adversarial Network*)

**HSV.** Espacio de color Matiz-Saturación-Valor (*Hue-Saturation-Value*)

**IOBA.** Instituto de Oftalmobiología Aplicada

**IRMA.** Anomalía microvascular intrarretiniana (*Intraretinal microvascular abnormality*)

**JPEG:** *Joint Photographic Experts Group*

**MA.** Microaneurisma

**MLP.** Perceptrón Multilineal (*Multilinear Perceptron*)

**MRI.** Imagen por resonancia magnética (*Magnetic Resonance Imaging*)

**MSCN.** *Mean Subtracted Contrast Normalized*

**MSE.** Error cuadrático medio (*Mean Squared Error*)

**NICE.** Instituto Nacional para la Salud y el Cuidado de la Excelencia (*National Institute for Health and Care Excellence*.)

**NIQE.** *Natural Image Quality Evaluator*

**NN.** Red neuronal (*Neural Network*)

**NSS.** *Natural Scene Statistic*

**OCT.** Tomografía de coherencia óptica (*Optical Coherence Tomography*)

**OCTA.** Angiografía OCT (*Optical Coherence Tomography Angiography*)

**OMS.** Organización Mundial de la Salud

**PSNR.** *Peak Signal-to-Noise Ratio*

**RD.** Retinopatía Diabética

**ReLU.** Unidad Lineal Rectificada (*Rectified Linear Unit*)

**RGB.** Espacio de color Rojo-Verde-Azul (*Red-Green-Blue*)

**SGD.** Descenso de gradiente estocástico (*Stochastic Gradient Descent*)

**SLT.** Tomografía de láser de barrido (*Laser Scanning Tomography*)

**SR.** Superresolución

**SSIM.** Índice de similitud estructural (*Structural Similarity Index Measure*)

**TC.** Tomografía Computarizada

**TFG.** Trabajo Fin de Grado

**VDSR.** *Very Deep Super Resolution*

**ZAF.** Zona avascular de la fovea

# ÍNDICE GENERAL

<b>CAPÍTULO 1. INTRODUCCIÓN .....</b>	<b>1</b>
1.1 INTRODUCCIÓN .....	3
1.2 BIOINGENIERÍA E INGENIERÍA BIOMÉDICA .....	3
1.2 IMAGEN MÉDICA. EVOLUCIÓN .....	3
1.3 DEEP LEARNING EN IMAGEN MÉDICA .....	4
1.4 IMÁGENES DE LA RETINA. RETINOGRAFÍAS .....	7
1.5 DIABETES MELLITUS .....	11
1.6 RETINOPATÍA DIABÉTICA .....	14
1.7 DIAGNÓSTICO DE LA RETINOPATÍA DIABÉTICA .....	16
1.7.1 DIAGNÓSTICO ESTÁNDAR Y TIPOS DE IMÁGENES DE LA RETINA .....	16
1.7.2 PROCESADO AUTOMÁTICO DE RETINOGRAFÍAS .....	17
1.8 HIPÓTESIS DE TRABAJO .....	18
1.9 OBJETIVOS DEL TFG .....	19
1.10 METODOLOGÍA EMPLEADA .....	20
1.11 ESTRUCTURA DEL DOCUMENTO .....	20
<b>CAPÍTULO 2. REVISIÓN DEL ESTADO DE LA TÉCNICA .....</b>	<b>23</b>
2.1 INTRODUCCIÓN .....	25
2.2 MÉTODOS BASADOS EN SUPERRESOLUCIÓN DE IMÁGENES .....	25
2.3 MÉTODOS BASADOS EN CORRECCIÓN DE CONTRASTE E ILUMINACIÓN .....	27
<b>CAPÍTULO 3. MATERIALES Y MÉTODOS .....</b>	<b>29</b>
3.1 INTRODUCCIÓN .....	31
3.2 DESCRIPCIÓN DE LA BASE DE DATOS EMPLEADA .....	31
3.3 MÉTODO DESARROLLADO .....	33
3.3.1 PREPROCESADO .....	33
3.3.2 GENERACIÓN DE IMÁGENES DE CALIDAD INSUFICIENTE .....	33
3.4 DEEP LEARNING EN MEJORA DE LA CALIDAD DE IMAGEN .....	36
3.4.1 REDES NEURONALES .....	36
3.4.2 REDES NEURONALES CONVOLUCIONALES .....	39
3.4.3 GENERATIVE ADVERSARIAL NETWORKS .....	50
<b>CAPÍTULO 4. RESULTADOS .....</b>	<b>57</b>
4.1 INTRODUCCIÓN .....	59
4.2 FORMA DE EVALUACIÓN .....	59
4.3 RESULTADOS PARA LA MEJORA DE CALIDAD DE IMÁGENES BORROSAS .....	63

4.3.1	FASE DE ENTRENAMIENTO. RENDIMIENTO DE LA GAN. ....	63
4.3.2	FASE DE TEST .....	64
4.4	RESULTADOS PARA LA MEJORA DE CALIDAD DE IMÁGENES OSCURAS .....	68
4.4.1	FASE DE ENTRENAMIENTO. RENDIMIENTO DE LA GAN. ....	68
4.4.2	FASE DE TEST. ....	71
4.5	RESULTADOS PARA AMBAS REDES EN SERIE .....	73
4.5.1	RESULTADOS SOBRE LAS IMÁGENES DE TEST DE BUENA CALIDAD .....	75
4.5.2	RESULTADOS SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD.....	76
<b>CAPÍTULO 5. DISCUSIÓN.....</b>		<b>79</b>
5.1	INTRODUCCIÓN .....	81
5.2	CORRECCIÓN DE LA NITIDEZ DE IMÁGENES DE RETINA .....	81
5.3	CORRECCIÓN DE LA ILUMINACIÓN DE IMÁGENES DE RETINA.....	84
5.4	COMPARACIÓN CON ESTUDIOS PREVIOS .....	86
<b>CAPÍTULO 6. CONCLUSIONES Y LÍNEAS FUTURAS.....</b>		<b>89</b>
6.1	INTRODUCCIÓN .....	91
6.2	CONTRIBUCIONES ORIGINALES .....	91
6.3	CONCLUSIONES .....	92
6.4	LIMITACIONES Y LÍNEAS FUTURAS .....	93
<b>BIBLIOGRAFÍA .....</b>		<b>95</b>



# ÍNDICE DE FIGURAS

<b>FIGURA 1. 1</b> ESTRUCTURA DEL OJO HUMANO (ADAPTADO DE STEWART, 2017).....	7
<b>FIGURA 1. 2</b> IMAGEN DE RETINA EN LA QUE SE MARCAN LA PAPILA, LA MÁCULA, LA FÓVEA Y LOS VASOS SANGUÍNEOS. ....	8
<b>FIGURA 1. 3</b> CORTE DEL GLOBO OCULAR DONDE SE REPRESENTA LA SITUACIÓN DE LOS CONOS Y BASTONES (FUENTE: <a href="https://www.aaopt.org/eye-health/anatomy/photoreceptors">HTTPS://WWW.AAOPT.ORG/EYE-HEALTH/ANATOMY/PHOTORECEPTORS</a> ).....	9
<b>FIGURA 1. 4</b> MAPA QUE MUESTRA LA CANTIDAD TOTAL CALCULADA DE ADULTOS (ENTRE 20 Y 79 AÑOS) CON DIABETES EN 2019. (FUENTE: FEDERACIÓN INTERNACIONAL DE DIABETES, 2019) .....	14
<b>FIGURA 1.5</b> RETINOGRAFÍA DE UN SUJETO CON RD, DONDE APARECEN MICROANEURISMAS, MICROHEMORRAGIAS Y EXUDADOS.....	15
<b>FIGURA 1.6</b> METODOLOGÍA DE TRABAJO. ....	20
<b>FIGURA 3. 1</b> RETINÓGRAFO NO MIDRIÁTICO TOPCON TRC-NW400.....	31
<b>FIGURA 3. 2</b> ESQUEMA DE CAPTURA DE LOS DOS CAMPOS. EN AMARILLO SE REPRESENTA EL CAMPO CENTRADO EN PAPILA Y EN VERDE EL CAMPO CENTRADO EN MÁCULA (FUENTE: <a href="https://www.opticianonline.net/cet-archive/188">HTTPS://WWW.OPTICIANONLINE.NET/CET-ARCHIVE/188</a> ).....	32
<b>FIGURA 3. 3</b> (A) IMAGEN ORIGINAL, <i>I<sub>ORIG</sub></i> . (B) IMAGEN FILTRADA, <i>I<sub>FILTER</sub></i> . ....	35
<b>FIGURA 3. 4</b> (A) IMAGEN ORIGINAL, <i>I<sub>ORIG</sub></i> . (B) IMAGEN OSCURECIDA, <i>I<sub>DARKER</sub></i> . ....	36
<b>FIGURA 3. 5</b> MODELO NO LINEAL DE NEURONA ARTIFICIAL (FUENTE: GARCÍA GADAÑÓN, 2008).....	37
<b>FIGURA 3. 6.</b> EJEMPLO DE RED NEURONAL (ADAPTADO DE ZHOU, S.KEVIN; RUECKERT, DANIEL; FICHTINGER, 2019) .....	38
<b>FIGURA 3. 7</b> (A) IMAGEN ORIGINAL. (B) IMAGEN VOLTEADA. ....	47
<b>FIGURA 3. 8</b> (A) IMAGEN ORIGINAL. (B) IMAGEN ROTADA $\pi/2$ . ....	48
<b>FIGURA 3. 9</b> MODELO DE MEJORA DE CALIDAD DE IMAGEN PROPUESTO. ....	51
<b>FIGURA 3. 10</b> ARQUITECTURA DE LA RED DEL GENERADOR CON SU TAMAÑO DE KERNEL (K), EL NÚMERO DE FILTROS (N) Y EL STRIDE (S) PARA CADA CAPA CONVOLUCIONAL. (ADAPTACIÓN DE AHIRWAR, 2019) .....	52
<b>FIGURA 3. 11</b> ARQUITECTURA DE LA RED DEL DISCRIMINADOR CON SU TAMAÑO DE KERNEL (K), EL NÚMERO DE FILTROS (N) Y EL STRIDE (S) PARA CADA CAPA CONVOLUCIONAL (ADAPTACIÓN DE AHIRWAR, 2019)).....	53
<b>FIGURA 4. 1</b> PÉRDIDA DEL GENERADOR DE LA RED QUE CORRIGE LA NITIDEZ. ....	63
<b>FIGURA 4. 2</b> PÉRDIDA DEL DISCRIMINADOR DE LA RED QUE CORRIGE LA NITIDEZ. ....	63
<b>FIGURA 4. 3.</b> RESULTADOS DESPUÉS DE 275 ÉPOCAS DE LA RED QUE CORRIGE LA NITIDEZ (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA.....	64
<b>FIGURA 4. 4.</b> RESULTADOS DESPUÉS DE 5000 ÉPOCAS. DE LA RED QUE CORRIGE LA NITIDEZ (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA.....	65
<b>FIGURA 4. 5.</b> RESULTADOS DESPUÉS DE 30000 ÉPOCAS DE LA RED QUE CORRIGE LA NITIDEZ (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA.....	65
<b>FIGURA 4. 6.</b> RESULTADOS DE LA RED QUE CORRIGE LA NITIDEZ SOBRE ALGUNAS IMÁGENES DE TEST DE BUENA CALIDAD. COLUMNA DE LA IZQUIERDA: IMAGEN DEGRADADA. COLUMNA CENTRAL: IMAGEN ORIGINAL. COLUMNA DE LA DERECHA: IMAGEN GENERADA. ....	66
<b>FIGURA 4. 7</b> RESULTADOS VISUALES SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA NITIDEZ (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA.....	67
<b>FIGURA 4. 8</b> RESULTADOS VISUALES SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA NITIDEZ AÑADIENDO UN FILTRO GAUSSIANO CON VARIANZA BAJA. COLUMNA DE LA IZQUIERDA: IMAGEN ORIGINAL. COLUMNA DE LA DERECHA: IMAGEN GENERADA.....	68
<b>FIGURA 4. 9</b> PÉRDIDA DEL GENERADOR DE LA RED QUE CORRIGE LA ILUMINACIÓN. ....	69
<b>FIGURA 4. 10</b> PÉRDIDA DEL DISCRIMINADOR DE LA RED QUE CORRIGE LA ILUMINACIÓN.....	69
<b>FIGURA 4. 11</b> RESULTADOS DESPUÉS DE 275 ÉPOCAS DE LA RED QUE CORRIGE LA ILUMINACIÓN (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA .....	70
<b>FIGURA 4. 12</b> RESULTADOS DESPUÉS DE 5000 ÉPOCAS DE LA RED QUE CORRIGE LA ILUMINACIÓN (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA. ....	70
<b>FIGURA 4. 13</b> RESULTADOS DESPUÉS DE 30000 ÉPOCAS DE LA RED QUE CORRIGE LA ILUMINACIÓN (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA. ....	70
<b>FIGURA 4. 14</b> RESULTADOS SOBRE IMÁGENES DEL CONJUNTO DE TEST DE BUENA CALIDAD DE LA RED QUE CORRIGE LA ILUMINACIÓN. COLUMNA DE LA IZQUIERDA: IMAGEN DEGRADADA. COLUMNA CENTRAL: IMAGEN ORIGINAL. COLUMNA DE LA DERECHA: IMAGEN GENERADA .....	72

<b>FIGURA 4. 15</b> RESULTADOS VISUALES SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA ILUMINACIÓN(A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA.....	73
<b>FIGURA 4. 16</b> RESULTADOS VISUALES SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA ILUMINACIÓN AÑADIENDO UN PORCENTAJE LEVE DE OSCURIDAD. COLUMNA DE LA IZQUIERDA: IMAGEN ORIGINAL. COLUMNA DE LA DERECHA: IMAGEN GENERADA. ....	74
<b>FIGURA 4. 17</b> RESULTADO DE UNA IMAGEN DEL CONJUNTO DE TEST DE BUENA CALIDAD TRAS LA CORRECCIÓN DE SU ILUMINACIÓN. (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA. ....	75
<b>FIGURA 4. 18</b> RESULTADO DE UNA IMAGEN DEL CONJUNTO DE TEST DE BUENA CALIDAD TRAS LA CORRECCIÓN DE SU NITIDEZ. (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA. ....	76
<b>FIGURA 4. 19</b> RESULTADOS VISUALES SOBRE IMÁGENES DEL CONJUNTO DE TEST DE MALA CALIDAD TRAS PASAR POR AMBAS REDES. COLUMNA DE LA IZQUIERDA: IMAGEN ORIGINAL. COLUMNA CENTRAL: IMAGEN TRAS LA CORRECCIÓN DE ILUMINACIÓN. COLUMNA DE LA DERECHA: IMAGEN TRAS LA CORRECCIÓN DE NITIDEZ. ....	77
<b>FIGURA 5. 1</b> COMPARACIÓN DE LA CALIDAD DE LAS IMÁGENES SOBRE UNA IMAGEN DEL CONJUNTO DE TEST DE BUENA CALIDAD. (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA. ....	81
<b>FIGURA 5. 2</b> COMPARACIÓN CON ZOOM DE LA NITIDEZ DE UNA RETINOGRAFÍA PERTENECIENTE AL CONJUNTO DE TEST DE BUENA CALIDAD. (A) IMAGEN DEGRADADA CON ZOOM EN LA PAPILA. (B) IMAGEN ORIGINAL CON ZOOM EN LA PAPILA. (C) IMAGEN GENERADA CON ZOOM EN LA PAPILA. (D) IMAGEN DEGRADADA CON ZOOM EN LOS EXUDADOS. (E) IMAGEN ORIGINAL CON ZOOM EN LOS EXUDADOS. (F) IMAGEN GENERADA CON ZOOM EN LOS EXUDADOS. ....	82
<b>FIGURA 5. 3</b> EFECTO DE LA CORRECCIÓN DE LA NITIDEZ DE UNA IMAGEN DE MALA CALIDAD DEL CONJUNTO DE TEST. (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA. ....	83
<b>FIGURA 5. 4</b> EFECTO DE LA CORRECCIÓN DE LA NITIDEZ EN UNA IMAGEN DE MALA CALIDAD DEL CONJUNTO DE TEST A LA QUE SE HA APLICADO UN FILTRO GAUSSIANO DE VARIANZA BAJA. (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA.....	83
<b>FIGURA 5. 5</b> DETALLE DEL EFECTO DE LA CORRECCIÓN DE LA NITIDEZ EN UNA IMAGEN DE MALA CALIDAD DEL CONJUNTO DE TEST. (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA. ....	84
<b>FIGURA 5. 7</b> EFECTO DE LA CORRECCIÓN DE LA ILUMINACIÓN SOBRE UNA IMAGEN DEL CONJUNTO DE TEST DE MALA CALIDAD. (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA. ....	85
<b>FIGURA 5. 6</b> EFECTO DE LA CORRECCIÓN DE LA ILUMINACIÓN DE LAS IMÁGENES. (A) IMAGEN DEGRADADA. (B) IMAGEN ORIGINAL. (C) IMAGEN GENERADA. ....	85
<b>FIGURA 5. 8</b> EFECTO DE LA CORRECCIÓN DE LA ILUMINACIÓN SOBRE UNA IMAGEN DEL CONJUNTO DE TEST DE MALA CALIDAD CON ILUMINACIÓN ADECUADA. (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA.....	86
<b>FIGURA 5. 9</b> EFECTO DE LA CORRECCIÓN DE LA ILUMINACIÓN SOBRE UNA IMAGEN CON ILUMINACIÓN ADECUADA A LA QUE SE HA AÑADIDO UN PORCENTAJE LEVE DE OSCURIDAD. (A) IMAGEN ORIGINAL. (B) IMAGEN GENERADA.....	86

# ÍNDICE DE TABLAS

<b>TABLA 3. 1</b> SEPARACIÓN DE LAS IMÁGENES PARA LA TAREA DE MEJORA DE LA CALIDAD EN UN CONJUNTO DE ENTRENAMIENTO Y UN CONJUNTO DE TEST .....	33
<b>TABLA 3. 2</b> ARQUITECTURA VGG19. ....	45
<b>TABLA 3. 3</b> ESPECIFICACIÓN DEL BLOQUE PRE-RESIDUAL. ....	52
<b>TABLA 3. 4</b> ESPECIFICACIÓN DEL BLOQUE RESIDUAL. ....	52
<b>TABLA 3. 5</b> ESPECIFICACIÓN DEL BLOQUE POST-RESIDUAL.....	53
<b>TABLA 3. 6</b> ESPECIFICACIÓN DE LA CAPA DE CONVOLUCIÓN.....	53
<b>TABLA 3. 7</b> ESPECIFICACIÓN DE LOS BLOQUES DE LA ARQUITECTURA DE LA RED DEL DISCRIMINADOR.....	54
<b>TABLA 4. 1</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DEL CONJUNTO DE TEST DE BUENA CALIDAD DE LA RED QUE CORRIGE LA NITIDEZ. ....	65
<b>TABLA 4. 2</b> RESULTADOS DE LA EVALUACIÓN CON REFERENCIA SOBRE LAS IMÁGENES DEL CONJUNTO DE TEST DE BUENA CALIDAD DE LA RED QUE CORRIGE LA NITIDEZ. ....	65
<b>TABLA 4. 3</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DEL CONJUNTO DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA NITIDEZ. ....	66
<b>TABLA 4. 4</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DEL CONJUNTO DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA NITIDEZ AÑADIENDO UN FILTRO GAUSSIANO CON $\sigma = 0.4$ .....	68
<b>TABLA 4. 5</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE EL CONJUNTO DE IMÁGENES DE TEST DE BUENA CALIDAD DE LA RED QUE CORRIGE LA ILUMINACIÓN. ....	71
<b>TABLA 4. 6</b> RESULTADOS DE LA EVALUACIÓN CON REFERENCIA SOBRE EL CONJUNTO DE IMÁGENES DE TEST DE BUENA CALIDAD DE LA RED QUE CORRIGE LA ILUMINACIÓN. ....	71
<b>TABLA 4. 7</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DEL CONJUNTO DE TEST DE MALA CALIDAD DE LA RED QUE CORRIGE LA ILUMINACIÓN. ....	72
<b>TABLA 4. 8</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DEL CONJUNTO DE TEST DE MALA CALIDAD AÑADIENDO UN PORCENTAJE LEVE DE OSCURIDAD.....	73
<b>TABLA 4. 9</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DE TEST DE BUENA CALIDAD PARA AMBAS REDES EN SERIE.....	75
<b>TABLA 4. 10</b> RESULTADOS DE LA EVALUACIÓN CON REFERENCIA SOBRE LAS IMÁGENES DE TEST DE BUENA CALIDAD PARA AMBAS REDES EN SERIE.....	75
<b>TABLA 4. 11</b> RESULTADOS DE LA EVALUACIÓN SIN REFERENCIA SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD PARA AMBAS REDES EN SERIE.....	76
<b>TABLA 5. 1</b> COMPARACIÓN CUANTITATIVA ENTRE EL MÉTODO PROPUESTO Y OTROS MÉTODOS.....	87



# CAPÍTULO 1

---

## INTRODUCCIÓN

1.1	INTRODUCCIÓN .....	3
1.2	BIOINGENIERÍA E INGENIERÍA BIOMÉDICA .....	3
1.2	IMAGEN MÉDICA. EVOLUCIÓN. ....	3
1.3	DEEP LEARNING EN IMAGEN MÉDICA .....	4
1.4	IMÁGENES DE LA RETINA. RETINOGRAFÍAS. ....	7
1.5	DIABETES MELLITUS.....	11
1.6	RETINOPATÍA DIABÉTICA.....	14
1.7	DIAGNÓSTICO DE LA RETINOPATÍA DIABÉTICA .....	16
1.7.1	DIAGNÓSTICO ESTÁNDAR Y TIPOS DE IMÁGENES DE LA RETINA .....	16
1.7.2	PROCESADO AUTOMÁTICO DE RETINOGRAFÍAS .....	17
1.8	HIPÓTESIS DE TRABAJO .....	18
1.9	OBJETIVOS DEL TFG .....	19
1.10	METODOLOGÍA EMPLEADA.....	20
1.11	ESTRUCTURA DEL DOCUMENTO .....	20



## 1.1 INTRODUCCIÓN

Este primer capítulo sirve como presentación de las principales ideas que se desarrollan a lo largo del presente Trabajo Fin de Grado (TFG). Se definen los conceptos en los que se encuentra enmarcado, tales como Ingeniería Biomédica, Retinografías, Retinopatía Diabética (RD), la hipótesis de trabajo y los objetivos del TFG. También se describe la metodología empleada y la estructura del documento.

## 1.2 BIOINGENIERÍA E INGENIERÍA BIOMÉDICA

La Real Academia Española define el prefijo 'bio' como 'vida' u 'organismo vivo'. Cuando la física y la química se han aplicado a los seres vivos, esta unión ha sido denominada con los términos biofísica y bioquímica. Por lo tanto, el matrimonio de la disciplina de la biología y la ingeniería se llama 'bioingeniería'. Igualmente, la unión de la medicina y la ingeniería se denomina ingeniería biomédica (Sawhney, 2007).

La bioingeniería tiene como objetivo estudiar y aplicar los principios y métodos, tanto de las ciencias exactas como de la ingeniería, a la solución de los problemas que presentan las ciencias biológicas y médicas. En 1977, el comité de educación del *Group of Engineering in Medicine and Biology* del *Institute of Electrical and Electronics Engineers* (IEEE) estableció la siguiente clasificación de la bioingeniería:

- a) Bioingeniería o Ingeniería Biológica.
- b) Ingeniería Biomédica o Ingeniería Médica.
- c) Ingeniería Clínica.

De esta clasificación extraemos que la ingeniería biomédica o ingeniería médica se encuentra contenida en la ingeniería biológica. Se trata de una disciplina que busca controlar las enfermedades (Mompín Poblet, 1988). Los ingenieros biomédicos aplican principios de ingeniería eléctrica, química, óptica, mecánica y de otro tipo para comprender, modificar o controlar los sistemas biológicos (es decir, humanos y animales). Por lo tanto, los ingenieros biomédicos son miembros del equipo de atención médica que buscan nuevas soluciones a los problemas de salud que enfrenta la sociedad moderna (Enderle, John D. & Bronzino, 2012).

## 1.2 IMAGEN MÉDICA. EVOLUCIÓN.

El procesamiento de imágenes tiene múltiples aplicaciones en la ciencia médica. Las técnicas de imagen biomédica tienen un papel destacado tanto en el campo de diagnóstico como en el terapéutico, permitiendo mejorar la atención médica de los pacientes. La terapia guiada mediante imágenes ha reducido riesgo de errores humanos al proporcionar una mayor precisión en la detección de enfermedades y procedimientos quirúrgicos (Ilangovan et al., 2016).

La historia de la imagen médica comenzó en 1895 con el descubrimiento de los rayos X por Roentgen. La amplia gama de aplicaciones de este descubrimiento se reconoció muy rápidamente y, en 1896, ya se establecieron departamentos de rayos X

dentro de muchos hospitales (Elliott, 2005). Posteriormente, en 1946, Felix Bloch y Edward Purcell descubrieron la resonancia magnética nuclear (RMN), que fue el primer paso para el descubrimiento de la imagen por resonancia magnética (*Magnetic Resonance Imaging*, MRI). Con la llegada de la RMN, Raymond Vahan Damadian propuso el primer escáner de RMN en 1969, con el que descubrió que el tumor se puede distinguir de las células normales (Ilangovan et al., 2016).

La segunda gran revolución en imagen comenzó en 1972 con el anuncio de Godfrey Hounsfield de un escáner tomográfico de rayos X asistido por ordenador, lo que se conoce como tomografía computerizada (CT). La imagen de CT es una reconstrucción en dos dimensiones (2D) de una sección transversal de la anatomía del paciente que proporciona buen contraste entre diferentes tipos de tejidos blandos en el cuerpo, así como contraste entre el hueso y todo el tejido blando (Guy, Chris & Ffytche, 2005). El éxito médico de la CT estimuló el desarrollo y aceleró la introducción de la RMN para uso hospitalario a principios de la década de 1980 (Guy, Chris & Ffytche, 2005).

El objetivo de todos estos procedimientos es mejorar la representación de las imágenes del cuerpo humano para su posterior análisis, contribuyendo al avance de la medicina al optimizar la detección y diagnóstico de ciertas enfermedades (Guy, Chris & Ffytche, 2005).

### 1.3 DEEP LEARNING EN IMAGEN MÉDICA

Una imagen se puede definir como una función bidimensional  $f(x, y)$ , donde  $x$  e  $y$  son coordenadas espaciales y la amplitud de  $f$  en cualquier par de coordenadas  $(x, y)$  es la intensidad o el nivel de gris de la imagen en ese punto. Si  $x$ ,  $y$ , y los valores de intensidad de  $f$  son todos finitos y discretos, la imagen es digital. El campo del procesamiento de imágenes digitales se refiere al procesamiento de imágenes digitales usando un ordenador. Una imagen digital consta de un número finito de elementos llamados píxeles, cada uno con una ubicación y un valor (Gonzalez, Rafael C. & Woods, 2018).

En el ámbito de la medicina, las imágenes se utilizan para obtener información sobre los órganos internos del cuerpo o el esqueleto para determinar el estado físico del paciente. Una imagen puede mostrar daños en los órganos que no se pueden visualizar externamente (Jennings et al., 1995).

Tan pronto como fue posible escanear y cargar imágenes médicas en un ordenador, se crearon sistemas para el análisis automatizado. Inicialmente, desde la década de 1970 hasta la década de 1990, el análisis de imágenes médicas se centraba en el procesado de píxeles y modelado matemático (líneas de ajuste, círculos y elipses) para construir sistemas que resolvieran tareas particulares (Litjens et al., 2017). El siguiente objetivo fue conseguir que los sistemas aprendiesen las características que representan los datos para el problema en cuestión. Este concepto se encuentra en la base de muchos algoritmos de aprendizaje profundo: modelos (redes) compuestos de muchas capas que transforman los datos de entrada, en este caso imágenes, en variables de salida como, por ejemplo, la ausencia o presencia de enfermedad (Litjens et al., 2017). Uno de los modelos más populares para el análisis de imágenes son las redes neuronales convolucionales (*Convolutional Neural Network*, CNN). Sin embargo, el entrenamiento de CNNs requiere de extensas bases de datos que normalmente son difíciles de conseguir. En este contexto, las redes generativas antagónicas (*Generative Adversarial Networks*, GANs) consiguen



solucionar este problema gracias al aumento de datos. Las GANs están compuestas por dos redes neuronales que se entrenan en un proceso en el que una red genera imágenes sintéticas y la otra red discrimina entre las imágenes reales y las sintéticas repetidamente (Frid-Adar et al., 2018). Este tipo de redes es el que se ha utilizado en el desarrollo de este TFG.

En los últimos años, técnicas basadas en *Deep Learning* han sido utilizadas para el análisis de distintos tipos de imágenes médicas, mostrando buenos resultados en diversas aplicaciones que se detallan a continuación (Litjens et al., 2017):

- **Clasificación**

- **Clasificación de imágenes:** La clasificación de imágenes fue una de las primeras áreas en las que el aprendizaje profundo hizo una contribución importante al análisis de imágenes médicas. En la clasificación se tiene una o varias imágenes como entrada con una única variable de diagnóstico como salida, por ejemplo, enfermedad presente o ausente.
- **Clasificación de objetos o lesiones:** La clasificación de objetos generalmente se enfoca en la clasificación de una pequeña parte, previamente identificada, de la imagen médica en dos o más clases, por ejemplo, la clasificación de nódulos en una tomografía de tórax (M. Kim et al., 2019) o la clasificación de lesiones en la piel (Y. Li & Shen, 2018). Para muchas de estas tareas, se requiere tanto la información local sobre la apariencia de la lesión como información global sobre la ubicación de la lesión para conseguir una clasificación precisa.

- **Detección**

- **Localización de órganos, regiones y puntos de referencia:** La localización anatómica de objetos, como órganos o puntos de referencia, sirve como método de preprocesado en las tareas de segmentación o en el trabajo clínico para la planificación de la terapia o intervención.
- **Detección de objetos o lesiones:** La detección de objetos de interés o lesiones en imágenes es una parte clave del diagnóstico. Típicamente, las tareas consisten en la localización e identificación de pequeñas lesiones en una imagen completa. Hay una larga tradición de investigación en sistemas de detección asistidos por ordenador diseñados para detectar automáticamente lesiones, mejorando la precisión de la detección o disminuyendo el tiempo de clasificación de los profesionales.

- **Segmentación**

Consiste en la separación de las regiones de interés de la imagen del fondo de la misma y de otras regiones. Es el proceso de particionar una imagen en múltiples regiones que comparten atributos similares, permitiendo la localización y cuantificación (Lundervold & Lundervold, 2019). La segmentación de órganos y otras subestructuras en imágenes médicas permite el análisis cuantitativo de parámetros clínicos relacionados con el volumen y la forma, como, por ejemplo, en el análisis cardíaco o cerebral (Akkus et al., 2017; Avendi et al., 2016). En el campo de la RD se ha utilizado para la

segmentación de los vasos sanguíneos de la retina con el objetivo de facilitar el diagnóstico de la enfermedad (Jebaseeli et al., 2019).

- **Registro**

El registro de imágenes se puede definir como la búsqueda de una transformación espacial que permita establecer correspondencias entre puntos homólogos de las imágenes que se desean registrar. De esta forma se puede seguir la evolución de una determinada enfermedad, si se toman imágenes de la misma zona en diferentes instantes de tiempo. (Maurer, Calvin R.; Fitzpatrick, 1993)

- **Recuperación de imágenes basada en contenido**

La recuperación de imágenes basada en contenido (*Content-based image retrieval*, CBIR) es una técnica en la que se analiza el contenido de la imagen en bases de datos masivas, ofreciendo la posibilidad de identificar historias de casos similares y comprender trastornos raros, teniendo como finalidad mejorar la atención al paciente. El desafío al que se enfrentan los investigadores en el desarrollo de estos métodos es cómo extraer representaciones precisas de características de la información a nivel de píxel y asociar dichas características con conceptos significativos. La gran capacidad que poseen los modelos CNN para aprender características en múltiples niveles de abstracción ha propiciado el uso de los mismos en la comunidad CBIR (Litjens et al., 2017).

- **Generación y mejora de imágenes**

Se ha propuesto una variedad de métodos de generación y mejora de imágenes que utilizan arquitecturas profundas, que van desde normalizar o mejorar la calidad de la imagen hasta el descubrimiento de patrones, que pueden indicar la presencia de alguna enfermedad.

En la generación de imágenes, se utilizan CNN de dos dimensiones (2D) o de tres dimensiones (3D) para convertir una imagen de entrada en otra. Estos sistemas se entrenan con un conjunto de datos en el que están presentes tanto la entrada como la salida deseada, definiendo las diferencias entre la salida generada y la deseada como una función de pérdida (Litjens et al., 2017).

Dos de las clases más comunes de algoritmos de procesamiento de imágenes son la restauración de imágenes y la mejora de imágenes. La restauración de imágenes es un proceso que busca recuperar una imagen que ha sido degradada de alguna manera. La mejora de imagen es un proceso imagen a imagen que crea nuevas características en imágenes que no estaban presentes en la imagen original. La mejora de la imagen generalmente se usa para generar imágenes visualmente más agradables. En el contexto de la imagen médica, ambos procesos se utilizan para mejorar las imágenes buscando facilitar un análisis posterior para aplicaciones clínicas o científicas (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019).

La síntesis y la superresolución de imágenes son casos especiales de restauración y mejora de imágenes. La síntesis de imágenes, donde se crean nuevas imágenes cuyo contraste no se había adquirido, puede considerarse como perteneciente a cualquiera de las categorías, restauración o mejora, según su uso. La superresolución, donde se crean imágenes cuya resolución es mejor que la observada, se encuentra en la categoría de

mejora de imagen y puede implementarse utilizando técnicas de síntesis de imagen (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

#### 1.4 IMÁGENES DE LA RETINA. RETINOGRAFÍAS.

El conocimiento anatómico de los ojos, las órbitas, el campo visual, los nervios craneales superiores y las vías centrales que regulan el movimiento ocular es una condición necesaria para la correcta interpretación de las enfermedades oftálmicas (Riordan-Eva, Paul; Cunningham, 2012).

El ojo humano mide aproximadamente de 22 a 27 mm de diámetro anteroposterior. El globo ocular consta de tres capas primarias, siendo estas: (1) la capa de soporte más externa del ojo, que incluye la córnea y la esclerótica; (2) la capa media superior del ojo, que constituye la capa vascular central del globo, que abarca el iris, el cuerpo ciliar y la coroides; y (3) la capa interior del ojo, formada por la retina, el humor acuoso y el humor vítreo (Kels et al., 2015).

Centrándonos en la retina humana, esta constituye el tejido ocular más complejo y su estructura se encuentra altamente organizada. Recibe la imagen visual a través del sistema óptico del ojo y convierte la energía luminosa en señales eléctricas que se transmiten a la corteza visual a través del nervio óptico. Es en la corteza visual donde se perciben los atributos estructurales (forma, color y contraste) y espaciales (posición, profundidad y movimiento) (Riordan-Eva, Paul; Cunningham, 2012).

Anatómicamente, la retina es una membrana delgada, transparente y delicada que, junto con otros elementos constituye la parte inicial del nervio óptico. Está situada entre la membrana hialoidea del cuerpo vítreo, por delante, y la coroides, por detrás (Figura 1.1). La retina cubre el interior del globo ocular, hasta la ora serrata. La ora serrata es la unión entre la retina y el cuerpo ciliar. Esta unión marca la transición entre el área no fotosensible del cuerpo ciliar y el área fotosensible de la retina. (Alemañy Martorell, Jaime; Villar Valdés, 2003).

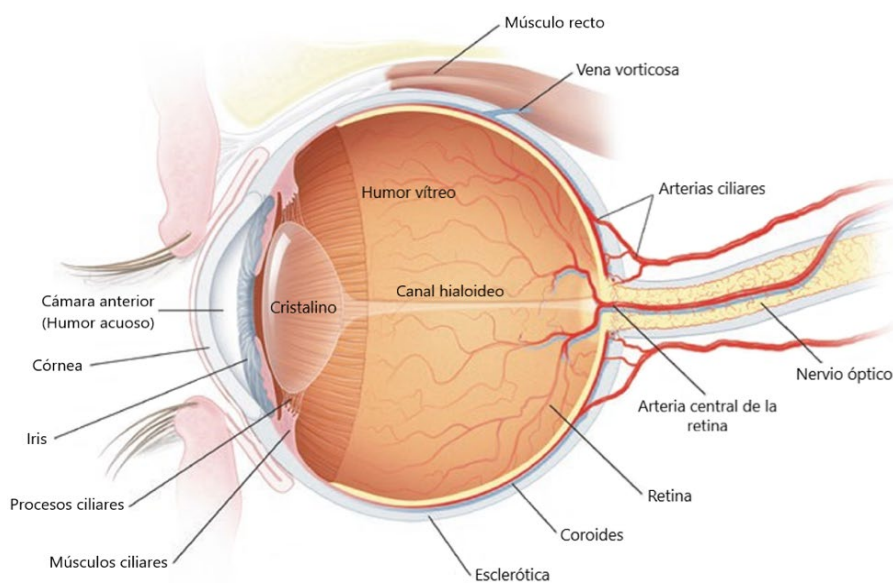


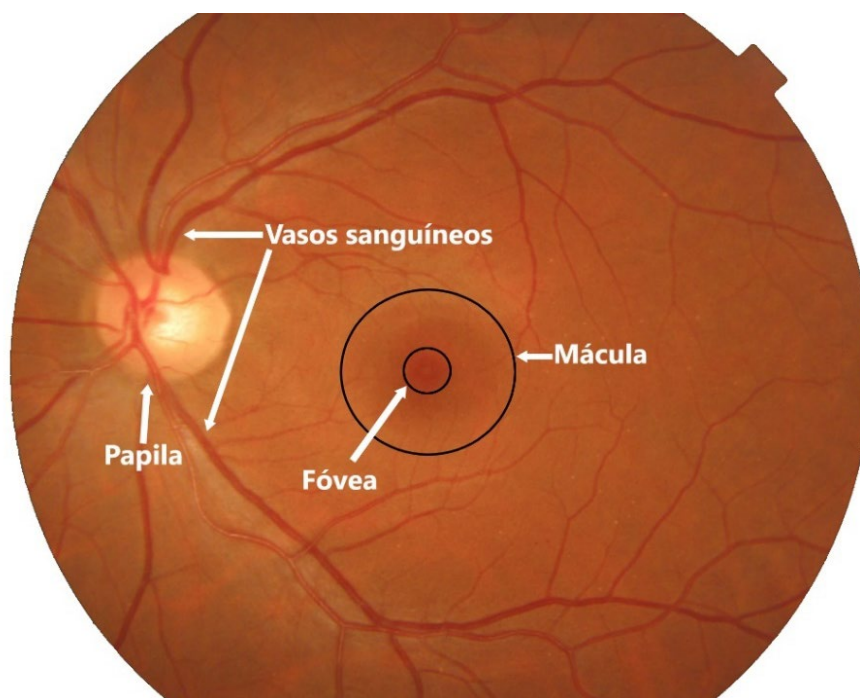
Figura 1. 1 Estructura del ojo humano (adaptado de Stewart, 2017).

Está compuesta por un conjunto de capas de células y de sinapsis que se pueden dividir en una capa epitelial externa, la cual se conoce como epitelio pigmentario de la retina (EPR), y una capa interna sensorial, conocida como retina neurosensorial o neurorretina (Boyd, 2010).

Dentro de la retina neurosensorial, la zona definida como mácula lútea (Figura 1.2), es la porción de la retina posterior que contiene pigmento xantófilo, de color amarillo. La porción central de 1,5 mm de diámetro de la mácula está ocupada por la fovea, que, por su anatomía y composición en fotorreceptores, está especializada para la elevada agudeza espacial y la visión del color. Dentro de la fovea existe una región que carece de vasos retinianos conocida como zona avascular de la fovea (ZAF). Con frecuencia, se considera que el centro geométrico de la ZAF es el centro de la mácula (American Academy of Ophtalmology, 2012).

Por otra parte, nos encontramos con el disco óptico o papila, área circular en la parte posterior del interior del ojo donde los axones de las neuronas ganglionares de la retina salen del ojo para formar la cabeza del nervio óptico (Figura 1.2). El nervio óptico transporta los impulsos nerviosos sensoriales desde más de un millón de células ganglionares de la retina hacia el cerebro (Grisan, 2005). Bajo examen oftalmológico, se aprecia la papila como un disco circular o ligeramente ovalado. De la papila parten los principales vasos sanguíneos que se encargan de nutrir, junto con los vasos sanguíneos coroideos, la retina neurosensorial (Abràmoff et al., 2010).

Cuando la luz entra en el ojo, antes de llegar a la retina debe atravesar la córnea, el humor acuoso, la pupila, el cristalino y el humor vítreo. Para producir una imagen enfocada, la luz que entra en el ojo debe refractarse y dirigirse hacia la retina. Esta refracción ocurre cuando la luz pasa a través de la córnea y el cristalino. Este proceso se llama acomodación (Watson & Lowe, 2019).



**Figura 1. 2** Imagen de retina en la que se marcan la papila, la mácula, la fóvea y los vasos sanguíneos.

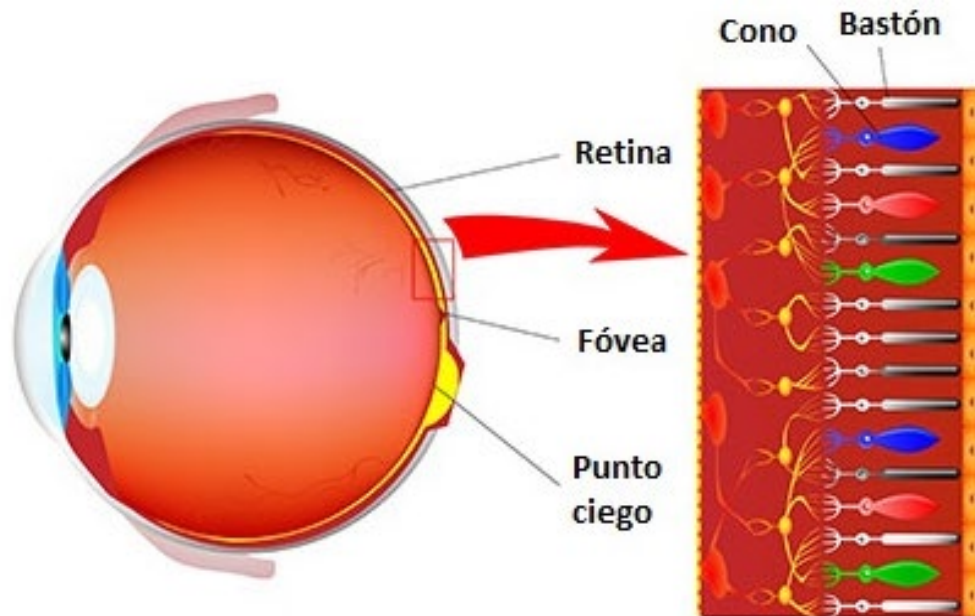
En la primera capa de la retina, donde se convierte la luz en señales neuronales, podemos distinguir dos tipos de fotorreceptores: bastones y conos. Existe también un tercer tipo de fotorreceptor: las células ganglionares sensibles a la luz, que permite las respuestas reflejas cuando los niveles de luz son muy altos (Boyd, 2010).

Tanto en los seres humanos como en otros animales, los bastones son los fotorreceptores más sensibles a la luz. Así, en condiciones de poca luz solo los bastones están activos. Sin embargo, a mayores niveles de luz como ocurre durante el día, la visión está dominada principalmente o completamente por los conos (Hadjikhani & Tootell, 2000).

Como los humanos tendemos a ser más activos durante el día, la mayor parte de nuestra experiencia visual es mediada por los conos. Este dominio de los conos ocurre a pesar de un sesgo anatómico prominente: la retina humana es un 95% de bastones y solo un 5% de conos (Hadjikhani & Tootell, 2000).

Los conos y bastones están ubicados encima del epitelio pigmentario, la capa más profunda de la retina (Figura 1.3). Por tanto, la luz que incide sobre la retina debe atravesar todo el grosor de la misma para llegar a los fotorreceptores.

Tanto la densidad como la distribución de los conos y bastones es diferente. Existe una alta concentración de conos en la fóvea, predominantemente sensibles al rojo y al verde, con una densidad mayor de 140.000 conos/mm<sup>2</sup>. Los bastones tienen su máxima densidad en una zona situada a 20° de la fóvea, donde alcanzan una densidad de aproximadamente 160.000 bastones/mm<sup>2</sup>. El número de conos y bastones disminuye rápidamente al alejarse del centro, es decir, la periferia casi no contiene conos ni bastones (American Academy of Ophthalmology, 2012).



**Figura 1. 3** Corte del globo ocular donde se representa la situación de los conos y bastones (Fuente: <https://www.aao.org/eye-health/anatomy/photoreceptors>)

La anatomía externa del ojo es visible y puede examinarse con instrumentos simples. Con dispositivos más complejos, se puede observar el interior del ojo, ya que este es visible a través de la córnea. De esta manera obtendremos una visión directa de los vasos sanguíneos y los tejidos del sistema nervioso central (retina y nervio óptico), haciendo posible, mediante un examen ocular, la identificación de efectos derivados de enfermedades infecciosas, autoinmunes y vasculares (Riordan-Eva, Paul; Cunningham, 2012).

Existen varias técnicas para obtener imágenes de la retina, y entre las más utilizadas están las siguientes (Holz, Frank G.; Spaide, 2010; Riordan-Eva, Paul; Cunningham, 2012; Stewart, 2017):

- **Oftalmoscopia directa.** Con el oftalmoscopio manual directo, se pueden obtener imágenes monoculares. Es un instrumento portátil que permite obtener imágenes detalladas del disco óptico y de la vasculatura retiniana. El oftalmólogo puede ajustar tanto el color y el tamaño del área iluminada, como el enfoque del oftalmoscopio
- **Oftalmoscopia indirecta.** Esta técnica requiere de la dilatación de la pupila. El oftalmoscopio indirecto es un instrumento que el examinador coloca en la cabeza y permite la visión binocular a través de un conjunto de lentes fijas. La parte frontal del dispositivo tiene una lámpara de luz con la que se puede ajustar el brillo y que se dirige hacia el ojo del paciente. Con su mano, el examinador sostiene una lente convexa a unos centímetros del ojo del paciente para enfocar la luz en la retina, de tal manera que se forma una imagen de la misma en el espacio entre el paciente y el examinador.
- **Angiografía con fluoresceína.** La fluoresceína es un colorante cuyas moléculas emiten luz verde al ser estimuladas por luz azul. En las fotografías, este colorante resalta los detalles vasculares y anatómicos del fondo del ojo. La angiografía con fluoresceína identifica defectos de perfusión en la vasculatura retiniana.
- **Autofluorescencia.** Este procedimiento se basa en la propiedad autofluorescente que tiene la retina. Esta cualidad consiste en la emisión de una luz en el espectro de longitudes de onda de los 500 a los 700 nm. La autofluorescencia de la retina se debe, principalmente a un componente llamado lipofucsina, un pigmento de color amarillento que, bajo condiciones normales, forma parte del EPR. Por lo tanto, los cambios en este epitelio pigmentario serán perfectamente visibles con esta prueba.
- **Tomografía de coherencia óptica (*Optical Coherence Tomography*, OCT).** Es una modalidad utilizada para examinar y medir en tres dimensiones las estructuras intraoculares. La OCT del segmento posterior permite el análisis detallado de la papila, la capa de fibras nerviosas de la retina y la mácula. Para el segmento anterior se utiliza un aparato de OCT distinto, el cual proyecta un haz de luz infrarroja de mayor longitud de onda (1300 nm). Este instrumento posibilita mediciones e imágenes de alta resolución de córnea, iris y otras estructuras intraoculares.

- **Angiografía OCT** (*Optical Coherence Tomography Angiography*, OCTA). Es una tecnología recientemente introducida que permite visualizar los vasos sanguíneos coroidales y retinianos sin el uso de un tinte inyectado. La OCTA detecta defectos de perfusión en los plexos capilares retinianos superficiales y profundos y, en última instancia, puede reemplazar la angiografía con fluoresceína para evaluar la enfermedad vascular retiniana.
- **Técnicas de imagen con láser (para disco óptico y retina)**. Tecnologías como la tomografía de láser de barrido (*Scanning Laser Tomography*, SLT) permiten obtener imágenes y cuantificarlas con detalles microscópicos del disco óptico y la capa de fibras nerviosas que lo rodea. Con la SLT se registran reflejos de un haz láser de barrido a distintas profundidades de tejido, de modo que se obtiene una serie de 64 secciones transversales coronales de tomografía, perpendiculares al eje del disco óptico. Mediante un ordenador y programas específicos se despliegan esos datos en imágenes tridimensionales y es posible realizar un análisis similar al efectuado con imágenes de OCT.
- **Retinografía**. Se denomina retinografía a la obtención de una fotografía del fondo de ojo. Las retinografías pueden ser de varios tipos, en función de la necesidad de usar colirios midriáticos o no, y del ángulo que se obtiene con la foto en una única captura. La midriasis es el aumento del diámetro o dilatación de la pupila, que es provocado mediante la aplicación de un colirio llamado tropicamida (Martínez Rubio et al., 2012). Por tanto, los retinógrafos también serán de dos tipos: midriáticos y no midriáticos, aunque actualmente la mayoría de los retinógrafos usados son no midriáticos. Estos permiten obtener imágenes del fondo de ojo sin necesidad de dilatar al paciente evitando las molestias de la midriasis provocada: elevada sensibilidad a la luz y visión borrosa. Además, recientemente se han incorporado al diagnóstico los retinógrafos de campo amplio, con los que es posible conseguir imágenes de más de 200° del fondo de ojo en una única captura (Ortiz Pérez, 2015).

## 1.5 DIABETES MELLITUS

La Diabetes Mellitus (DM) es un grupo heterogéneo de trastornos caracterizados por niveles elevados de glucosa en sangre, condición que se conoce como hiperglucemia, y otras anomalías metabólicas (Yanoff, Myron; Sassani, 2019). Se trata de una afección de carácter grave y crónico que ocurre porque el cuerpo no puede producir insulina, la produce en una cantidad insuficiente, o no puede utilizar de manera eficaz la insulina que produce. La insulina es una hormona que se produce en el páncreas y que permite que la glucosa del torrente circulatorio entre en las células del cuerpo, donde se convierte en energía. Además, es fundamental para el metabolismo de las proteínas y las grasas. (Federación Internacional de Diabetes, 2019)

La DM puede desarrollarse rápidamente como ocurre en la diabetes tipo I, o seguir síntomas que son vagos y no apuntan a ninguna enfermedad específica, como en el caso de la diabetes tipo II (Kharroubi, 2015).

Hay tres signos y síntomas comunes en la DM, todos ellos estrechamente relacionados con la hiperglucemia (Moini, 2019):

- **Polidipsia:** sed excesiva. Se produce por la deshidratación intracelular, provocada por el aumento de la glucosa en sangre.
- **Polifagia:** hambre excesiva. Esto suele ser un síntoma de la diabetes tipo I, y no tanto de la diabetes tipo II. Generalmente ocurre debido a la necesidad de nutrir las células, así como al agotamiento de las proteínas celulares, grasas y glúcidos.
- **Poliuria:** micción excesiva. Dado que la glucosa es una molécula pequeña y osmóticamente activa, la elevada glucosa en sangre da como resultado que la cantidad de glucosa que puede filtrarse por los glomérulos de los riñones supere la cantidad que puede ser reabsorbida por los túbulos renales. El resultado es la glucosuria y pérdidas significativas de agua a través de la orina.

Si el déficit de insulina no es controlado a largo plazo, muchos de los órganos del cuerpo pueden resultar dañados, provocando complicaciones de salud como enfermedades cardiovasculares (ECV), lesiones de los nervios (neuropatía), enfermedades renales (nefropatía) y afecciones oculares (retinopatía). No obstante, con un tratamiento apropiado de la diabetes, estas complicaciones se pueden atrasar o prevenir totalmente. (Federación Internacional de Diabetes, 2019)

Los casos de DM se pueden clasificar principalmente en dos tipos:

- **Diabetes tipo I**

La diabetes tipo I está causada por una reacción autoinmunitaria en la que el sistema inmunitario del organismo ataca a las células beta ( $\beta$ ) del páncreas que producen insulina. Como consecuencia, estas células ya no pueden secretar insulina o la cantidad que producen es insuficiente provocando que los niveles de glucosa en sangre se eleven provocando hiperglucemia. La diabetes tipo I es una enfermedad autoinmune, en la cual los linfocitos CD8-T citotóxicos atacan y destruyen los islotes pancreáticos (Moini, 2019).

La DM tipo I se denominaba anteriormente DM dependiente de insulina o de inicio juvenil debido a que, aunque puede aparecer a cualquier edad, se desarrolla con más frecuencia en niños y adolescentes, representando entre el 5% y el 10% de todos los casos de diabetes. Las personas que padecen diabetes tipo I requieren inyecciones de insulina para mantener el nivel de glucosa dentro de los valores apropiados (Kharroubi, 2015).

El grupo conocido como *American Diabetes Association Expert Committee* presentó la siguiente clasificación de la DM tipo I (American Diabetes Association, 1997).

- **Tipo IA:** También llamada inmuno-mediada, debido a la destrucción autoinmune de las células  $\beta$  pancreáticas. Supone más del 95% de los casos. (Abiru et al., 2002)
- **Tipo IB:** Se trata de una forma rara de diabetes tipo I de origen desconocido, aunque menos severa que la diabetes tipo I autoinmune. Se produce pérdida idiopática de secreción de insulina, debido a la destrucción no autoinmune de las células  $\beta$ . Supone menos del 5% de los casos.



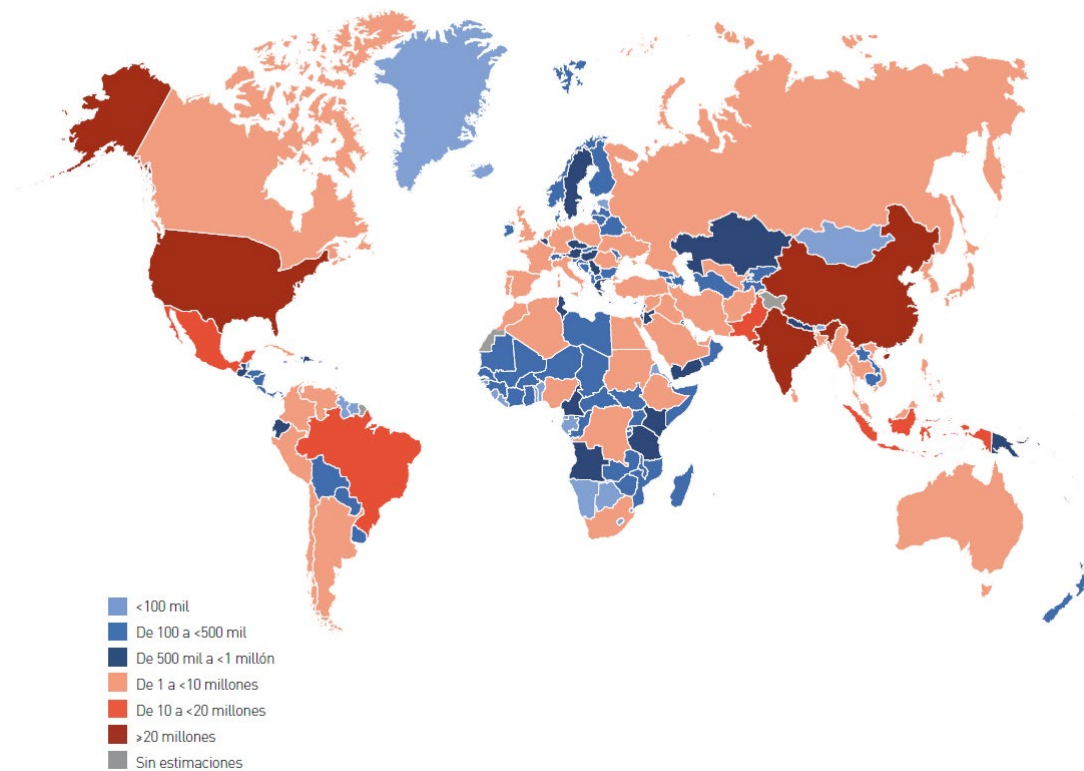
- **Diabetes tipo II**

La diabetes tipo II también es conocida como diabetes no insulino dependiente o diabetes de inicio en la edad adulta. Resulta de una deficiencia de insulina relativa, a diferencia de la deficiencia absoluta de la misma observada en la diabetes tipo I (Moini, 2019).

La hiperglucemia en la diabetes tipo II se produce como resultado de la incapacidad de las células del cuerpo de responder totalmente a la insulina, lo que se conoce como “resistencia a la insulina” (Federación Internacional de Diabetes, 2019). La diabetes tipo II ocurre cuando la glucosa no entra a las células y se acumula en la sangre. Debido a esto, el páncreas produce más insulina buscando que la glucosa ingrese en las células, ocasionando que el nivel de azúcar en sangre se eleve demasiado (Halban et al., 2014). Además, la secreción de insulina disminuye con el aumento de la demanda de insulina durante un largo intervalo de tiempo debido a la destrucción gradual de las células  $\beta$ , lo que podría provocar que algunos de los pacientes con diabetes tipo II se vuelvan dependientes de la insulina (Druet et al., 2006). Los estudios han demostrado que los riesgos de desarrollar diabetes tipo 2 son mayores en personas con antecedentes familiares, son inactivos y tienen sobrepeso u obesidad (Dabelea et al., 2014). A medida que aumenta el índice de masa corporal (IMC), también lo hacen los riesgos de desarrollar diabetes (Moini, 2019).

La diabetes tipo II representa aproximadamente entre el 90% y el 95% de todos los casos totales de diabetes (Yanoff, Myron; Sassani, 2019). Este tipo de diabetes puede aparecer con síntomas similares a los de la diabetes tipo I, pero, en general, la aparición del tipo II es mucho menos drástica y es probable que ocurra sin síntomas, imposibilitando determinar el momento exacto de la aparición de la diabetes tipo II. Como consecuencia, el período prediagnóstico es, a menudo, prolongado, haciendo que entre un tercio y la mitad de las personas con diabetes tipo II no reciban el diagnóstico correspondiente. Cuando no se identifica la enfermedad en el momento de la aparición de la misma, en el instante del diagnóstico pueden estar ya presentes ciertas complicaciones como la RD (Federación Internacional de Diabetes, 2019), enfermedad que se ha convertido en la principal causa de la pérdida de visión en países desarrollados (Ong et al., 2004).

La diabetes representa una epidemia mundial. En la actualidad, 351.7 millones de personas en edad activa (20 - 64 años) tienen diabetes diagnosticada o sin diagnosticar. Se prevé que este número aumente a 417.3 millones para 2030 y a 486.1 millones para el año 2045 (Federación Internacional de Diabetes, 2019). En la Figura 1.4 se muestra un mapa mundial con la cantidad total calculada de adultos (de entre 20 y 79 años) con diabetes en 2019. En Europa, se estima que el número de adultos de entre 20 y 79 años con diabetes es de 59.3 millones, lo que representa el 8.9% de la población regional en este grupo de edad. Esto incluye a 24.2 millones de adultos con diabetes no diagnosticada (Federación Internacional de Diabetes, 2019). Este aumento del número de casos de DM es atribuible al envejecimiento de la población, al desarrollo económico y al aumento de la urbanización, lo que provoca un estilo de vida sedentario y un mayor consumo de alimentos relacionados con la aparición de la obesidad (Federación Internacional de Diabetes, 2019).



**Figura 1. 4** Mapa que muestra la cantidad total calculada de adultos (entre 20 y 79 años) con diabetes en 2019. (Fuente: Federación Internacional de Diabetes, 2019)

## 1.6 RETINOPATÍA DIABÉTICA

La RD es la causa más frecuente en nuevos casos de ceguera diagnosticada entre 20 y 74 años de edad en los países desarrollados (Moini, 2019). La hiperglucemia crónica, hipertensión, hipercolesterolemia y tabaquismo son los factores de riesgo para el desarrollo y progreso de la RD. Las personas jóvenes con diabetes tipo I no desarrollan RD al menos de tres a cinco años después del inicio de la enfermedad. Los diabéticos de tipo II pueden tener RD en el momento del diagnóstico lo que puede ser la manifestación presentadora. (Riordan-Eva, Paul; Cunningham, 2012)

No existe un método de prevención actual para la RD. Sin embargo, cuando se trata adecuadamente, más del 90% de la pérdida de visión se puede prevenir. El enfoque principal es la identificación temprana de la condición, clasificándola correctamente, y tratamiento oportuno (Moini, 2019). Antes de que la visión sea afectada, ocurren cambios fácilmente detectables. La detección sistemática de la RD debe efectuarse en los tres años a partir del diagnóstico de diabetes tipos I y II, y cada año en lo sucesivo en ambos tipos. La fotografía digital del fondo del ojo ha demostrado ser un método eficaz para la detección de esta enfermedad. (Riordan-Eva, Paul; Cunningham, 2012)

Existen distintos grados de severidad de la RD, por lo que es necesaria una clasificación de la misma. Debido a esto, el grupo conocido como *Global Diabetic Retinopathy Project Group* (GDRPG), elaboró una clasificación publicada en 2003 por la Academia Americana de Oftalmología (Wilkinson CP, Ferris FL 3RD, Klein RE, Lee

PP, Agardh CD, 2003) que se ha convertido en el estándar de trabajo para una práctica clínica de calidad.

La severidad de la RD se divide en 5 niveles definidos por la observación oftalmoscópica del fondo de ojo (Aliseda & Berastegui, 2008):

- **Sin retinopatía aparente**

No se observan lesiones características al examen oftalmoscópico.

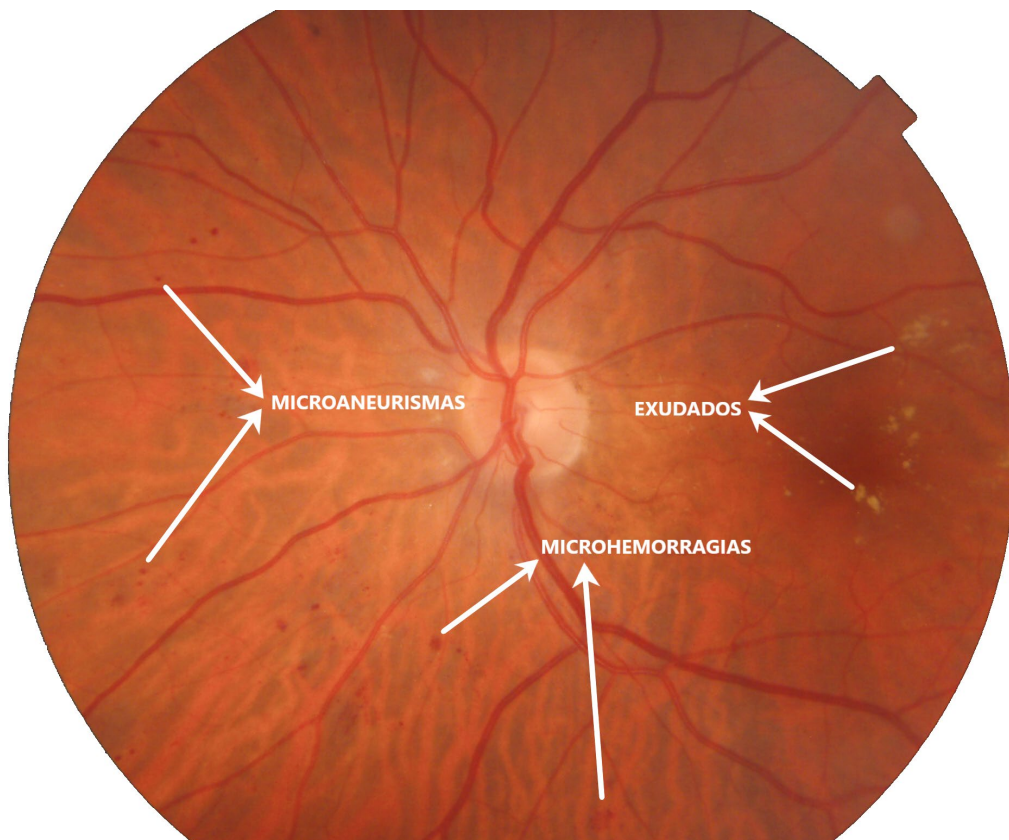
- **RD no proliferante leve**

Se encuentran pequeñas protuberancias en los pequeños vasos sanguíneos en la retina, como primera alteración apreciable oftalmoscópicamente. Estas protuberancias llamadas microaneurismas (MAs) son dilataciones de la pared de los capilares y aparecen como puntos rojos pequeños de bordes muy nítidos. Puede hacer que los vasos goteen pequeñas cantidades de sangre en la retina.

- **RD no proliferante moderada**

En esta etapa, los vasos sanguíneos en las retinas se hinchan. Se observa la aparición de algunas hemorragias retinianas, pudiendo existir exudados duros o lipídicos y blandos o algodonosos, todo ello acompañado de dilataciones venosas (Figura 1.5). Es posible que las venas no lleven la sangre tan bien como solían hacerlo.

Los exudados duros son depósitos lipídicos intrarretinianos profundos de color amarillento y bien definidos. Los exudados blandos o algodonosos son engrosamientos



**Figura 1.5** Retinografía de un sujeto con RD, donde aparecen microaneurismas, microhemorragias y exudados.

isquémicos de la capa de fibras nerviosas, con una forma más redondeada, bordes más difusos y color blanco, que tapan los vasos retinianos.

Estos cambios pueden conducir al edema macular diabético (EMD). Esto sucede cuando se acumula sangre y otros líquidos en la mácula.

- **RD no proliferante severa**

Puede presentarse cualquiera de las siguientes alteraciones: hemorragias intrarretinianas severas, dilataciones venosas o anomalías microvasculares intrarretinianas (*Intraretinal Microvascular Abnormality*, IRMA). Los IRMA son vasos finos tortuosos anormales anexos a la arquitectura vascular retiniana, que en ocasiones son difíciles de identificar en la exploración oftalmoscópica directa. En esta etapa, los vasos sanguíneos se bloquean aún más. Esto significa que llega menos sangre a ciertas áreas de la retina. La falta de sangre provoca que se envíe una señal a estas áreas retinianas para crear nuevos vasos sanguíneos formando tejido cicatricial. Si estos vasos sanguíneos se cierran por completo, se produce isquemia macular, que puede conducir a una visión borrosa con manchas oscuras, elevándose el riesgo de pérdida visual severa.

- **RD proliferante**

En esta etapa crecen nuevos vasos sanguíneos en la retina y en el líquido gelatinoso que llena los ojos. Este crecimiento se llama neovascularización. Estos vasos son delgados y débiles, que a menudo sangran provocando la creación de tejido cicatricial. A medida que el tejido cicatricial se hace más pequeño, puede alejar la retina de la parte posterior del ojo, lo que se conoce como desprendimiento de retina. Esto puede conducir a la pérdida permanente de la visión directa y lateral.

## 1.7 DIAGNÓSTICO DE LA RETINOPATÍA DIABÉTICA

### 1.7.1 DIAGNÓSTICO ESTÁNDAR Y TIPOS DE IMÁGENES DE LA RETINA

Para prevenir la ceguera causada por la RD, es necesario que todos los pacientes diabéticos se sometan a exámenes oculares periódicos (Javitt et al., 1990; Wilkinson et al., 2003). La Organización Mundial de la Salud (OMS) recomienda que la periodicidad de estos exámenes sea anual (World Health Organization, 2005). Además, se debe establecer un estándar común para clasificar y tratar esta enfermedad (López Gálvez, 2004).

Como parte del programa de detección, se deben examinar las dos retinas con los siguientes objetivos (García Gadañón, 2008):

- Detectar los signos visibles de RDNP: EXs, MAs, HEs, etc.
- Determinar el grado de DR en función del número y localización de las lesiones detectadas.
- Evaluar la posibilidad de que el paciente sufra complicaciones derivadas de los signos clínicos detectados, como edema de mácula o hemorragia vítrea.
- Monitorizar la evolución de la enfermedad.
- Determinar el tratamiento más adecuado para cada caso.

Las fotografías de fondo de color siguen siendo el método estándar para detectar RD. Además, los programas de cribado fotográfico se han vuelto particularmente importantes en los países en desarrollo donde un número insuficiente de especialistas y el aumento del número de casos de diabetes impiden la realización de exámenes regulares del fondo de ojo (Stewart, 2017).

Si bien existen varios protocolos para la detección y caracterización de la RD, no se ha determinado cuál de ellos es el más adecuado para establecer un programa de cribado (García Gadañón, 2008). El protocolo ETDRS (*Early Treatment Diabetic Retinopathy Study*), con 7 fotos estándar de pacientes con midriasis y evaluadas por personal cualificado está aceptado como estándar para la detección y graduación de la RD (Early Treatment Diabetic Retinopathy Study Research Group, 1991). Sin embargo, su aplicación en la práctica clínica está limitada por la complejidad del proceso de adquisición, que requiere de personal altamente cualificado, la dificultad de evaluación de la imagen por parte de expertos y la incomodidad del proceso de adquisición de las imágenes para el paciente (Grisan, 2005). Por estos motivos, este protocolo generalmente no se usa para la detección. La oftalmoscopia indirecta proporciona una alternativa adecuada a las fotografías tomadas bajo el protocolo ETDRS (García Gadañón, 2008). Por esta razón, y debido a la mejora de la calidad de los retinógrafos, se han propuesto y verificado otros protocolos de adquisición más sencillos, como el protocolo EURODIAB, el protocolo JVN (*Joslin Vision Network*), el protocolo de cinco campos o el protocolo de único campo en color (García Gadañón, 2008).

En la actualidad, el cribado de la RD requiere que los expertos examinen manualmente las imágenes de retina buscando las lesiones oculares relacionadas con dicha enfermedad. A medida que aumenta la incidencia de diabetes, el número de imágenes que deben revisar los oftalmólogos crece considerablemente. Además, la escasez de oftalmólogos profesionales, especialmente en áreas rurales, dificulta el diagnóstico temprano de la RD. El desarrollo de herramientas automatizadas para la detección de la RD permitiría examinar un mayor número de imágenes en un tiempo más reducido lo que reduciría la carga de trabajo de los expertos.

### **1.7.2 PROCESADO AUTOMÁTICO DE RETINOGRAFÍAS**

Según la OMS, una vez que a un paciente se le diagnostica diabetes, es necesario que se someta a exámenes oculares cada año para detectar los signos clínicos de la RD de manera temprana (World Health Organization, 2005). Sin embargo, se estima que el porcentaje de pacientes diabéticos que no se someten a exámenes oculares regulares supera el 50% (Lee et al., 2003). Entre las razones que provocan que los pacientes no se sometan a estos exámenes se encuentra que el paciente no sea capaz de percibir los síntomas de la enfermedad en su etapa inicial. Cuando la enfermedad está más avanzada y la visión disminuye, es cuando el paciente percibe la enfermedad. Otras condiciones que dificultan la detección precoz de la RD son los factores socio-demográficos, las limitaciones geográficas y la falta de profesionales que realicen exámenes oculares periódicos para diagnosticar y tratar correctamente la RD (García Gadañón, 2008).

Una forma de garantizar la suficiente sensibilidad y especificidad en la detección de signos clínicos relacionados con la RD es utilizar imágenes digitales de la retina como parte del protocolo de inspección (Lin et al., 2002). Sin embargo, con el aumento de la

incidencia de diabetes, la cantidad de imágenes que los médicos deben verificar es muy grande. Además, la escasez de oftalmólogos profesionales, especialmente en las zonas rurales, hace que el alto coste de los exámenes oculares dificulte el diagnóstico temprano de la RD. En este contexto, el uso de sistemas informáticos para ayudar al diagnóstico de la RD puede proporcionar muchos beneficios. Principalmente, permitiría examinar una gran cantidad de imágenes en menor tiempo y de una manera más objetiva que con técnicas dependientes del observador. Esto puede ser una herramienta que ayude al diagnóstico de la enfermedad permitiendo reducir la carga de trabajo de los expertos (Patton et al., 2006).

Entre sus principales aplicaciones, el procesado automático de retinografías permite evaluar y mejorar la calidad de la imagen, identificar las distintas estructuras normales de la retina, detectar las lesiones características de la RD y, con ellas, estimar el grado de severidad de la enfermedad. Todas estas aplicaciones formarían parte de un sistema automático de diagnóstico de la RD. Sin embargo, el desarrollo de estos métodos no es una tarea sencilla. Las características físicas del paciente, como el color de su piel o iris, y el proceso de adquisición de imágenes, hacen que las imágenes de la práctica clínica sean muy variables. Por este motivo, los métodos automatizados suelen incluir una etapa de preprocesado, cuyo objetivo es mejorar la calidad de la imagen, permitiendo una mejor identificación de las lesiones características de la RD o la detección de las estructuras anatómicas (Zhou et al., 2018).

## 1.8 HIPÓTESIS DE TRABAJO

La vista es uno de los sentidos humanos más importantes porque percibimos visualmente la mayor parte de la información recibida. Para procesar esta información, el cerebro necesita una señal de luz para penetrar el ojo a través del cristalino y dar forma a la retina. Este complejo proceso puede ser fácilmente perturbado por diversas enfermedades que impiden la correcta interpretación de la señal visual, entre las que se incluye la RD (Kajan, S.; Goga, J.; Lacko, K. and Pavlovičová, 2020).

La mitad de los 463 millones de adultos que viven con diabetes hoy desconocen que tienen la afección y, por lo tanto, corren un alto riesgo de desarrollar RD (Federación Internacional de Diabetes, 2019). Tras la detección temprana de la RD, el curso del daño ocular puede ralentizarse o incluso detenerse quirúrgicamente, pero, sin ningún tratamiento, causará un daño irreversible o ceguera en una etapa avanzada. Los síntomas de la enfermedad comienzan a aparecer solo en etapas posteriores, cuando el paciente se da cuenta de la visión deteriorada. Es por eso que el diagnóstico temprano es un aspecto clave para prevenir la pérdida permanente de la visión (Kajan, S.; Goga, J.; Lacko, K. and Pavlovičová, 2020).

Los detalles de pequeños hitos anatómicos y patologías son críticos para el análisis preciso de la enfermedad. Sin embargo, muchas imágenes médicas reales sufren de nitidez limitada, debido al equipo de imagen y parámetros de imagen (Bing et al., 2019). Tal baja resolución de imágenes médicas impide la precisión en la detección o segmentación de pequeños puntos de referencia anatómicos y patologías e impide el diagnóstico preciso de algunas enfermedades graves en su etapa inicial.

El preprocesado es una cuestión clave en el análisis automatizado de la RD. Diversos estudios han demostrado que entre el 11.9% y el 20.8% de los pacientes tiene

imágenes de al menos un ojo que no se pueden analizar clínicamente debido a la insuficiente calidad de la imagen. Las principales causas de esta insuficiente calidad son la iluminación no uniforme, el reducido contraste y los movimientos del ojo. La aplicación de las técnicas de preprocesamiento puede mitigar o incluso eliminar estos problemas (Besenczi et al., 2016).

Un sistema de mejora de calidad de retinografías puede ser una herramienta útil para la ayuda al diagnóstico de la RD. Para obtener un diagnóstico correcto tanto por parte de los sistemas automáticos como de los oftalmólogos, es necesario que las retinografías tengan una calidad suficiente para su análisis (Xiong et al., 2017a; Zhou et al., 2018). En este contexto, las últimas técnicas de *Deep Learning*, como las GAN, pueden ser muy efectivas para desarrollar un método automático de mejora de la calidad en imágenes de fondo de ojo.

En este TFG se parte de la hipótesis de que el desarrollo de un sistema automático basado en GANs puede ayudar a mejorar la calidad de las imágenes y reducir el número de imágenes de mala calidad que pasan analizarse en un sistema automático más general de ayuda al diagnóstico. Con el método desarrollado se pretende mejorar tanto la precisión de los sistemas automáticos de análisis de retinografías como la evaluación por parte de los especialistas.

## 1.9 OBJETIVOS DEL TFG

El objetivo de este trabajo ha sido el estudio y desarrollo de un método para la mejora de calidad de imágenes de fondo de ojo basado en *Deep Learning* que sirva como etapa de preprocesado para la ayuda al diagnóstico de la RD. Para conseguir dicho objetivo general, se han planteado los siguientes objetivos específicos:

1. Revisión de la bibliografía disponible relacionada con el análisis de retinografías mediante técnicas de *Deep Learning*, comprensión de los conceptos necesarios para la realización de este trabajo e identificación de las técnicas empleadas en dichos estudios previos. De esta revisión surgen las bases sobre las que se fundamenta el algoritmo desarrollado.
2. Elaboración de una base de datos (BD) de retinografías que incluye imágenes de diferentes niveles de calidad. De esta manera se pueden evaluar convenientemente los resultados del modelo desarrollado.
3. Desarrollo de métodos de procesado digital de imágenes aplicados a la mejora de calidad de imágenes de retina a través del lenguaje de programación Python, y más concretamente, mediante las librerías *Keras* y *TensorFlow*.
4. Comprobación del funcionamiento del método desarrollado en la BD de retinografías propia y obtención de los resultados del estudio.
5. Análisis los resultados obtenidos por el algoritmo de mejora de calidad de retinografías desarrollado.
6. Extracción de conclusiones.

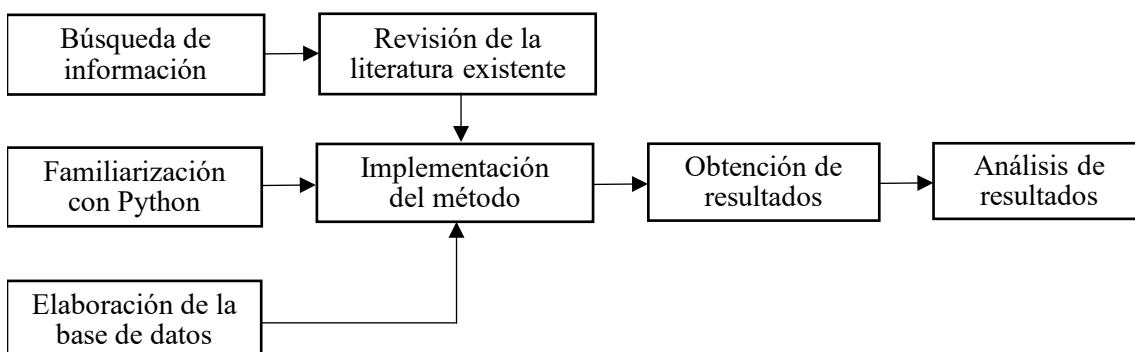
## 1.10 METODOLOGÍA EMPLEADA

Para la consecución del objetivo perseguido en este TFG, se ha seguido la metodología mostrada en la Figura 1.6. En este diagrama de bloques se representan las diferentes etapas de este trabajo, que se detallan a continuación.

1. Búsqueda de información para comprender el problema a resolver. Esto incluye conocer la información más relevante relativa a la DM y a la RD, así como de los métodos de mejora de calidad de imágenes desarrollados en trabajos previos.
2. Revisión de la literatura existente orientada al preprocesado de retinografías basado en redes GAN.
3. Elaboración de una BD de retinografías para la evaluación del método implementado.
4. Familiarización con el lenguaje de programación Python y en especial con las librerías *Keras* y *TensorFlow* utilizadas en el contexto de *Deep Learning*.
5. Desarrollo del método seleccionado incluyendo su implementación en Python.
6. Procesado de las imágenes y obtención de resultados. Para la obtención de resultados, ha sido necesario, en primer lugar, buscar los parámetros óptimos del método empleando un subconjunto de las imágenes de la BD lo que llamamos conjunto de entrenamiento. Una vez determinada la configuración más adecuada de acuerdo con los datos de entrenamiento, se obtienen los resultados sobre un conjunto diferente de imágenes, no empleadas antes, lo que se conoce como conjunto de test.
7. Análisis de los resultados y extracción de conclusiones. Documentación del trabajo desarrollado.

## 1.11 ESTRUCTURA DEL DOCUMENTO

El documento se divide en seis capítulos. En este primer capítulo de introducción se han expuesto los conceptos básicos en los que se encuentra enmarcado el TFG junto con el problema a resolver, así como la hipótesis de trabajo, los objetivos y la metodología a seguir. Se complementa con los capítulos que se describen a continuación:



**Figura 1.6** Metodología de trabajo.



En el capítulo 2 “Revisión del estado de la técnica” se presentan algunos de los trabajos previos en relación a la mejora de calidad de imágenes de retina, con el objetivo de situar este trabajo en el contexto adecuado.

En el capítulo 3 “Materiales y métodos” se presenta la BD utilizada en este TFG, así como los métodos aplicados en la mejora de la calidad de retinografías.

En el capítulo 4 “Resultados” se presentan los resultados más significativos del método desarrollado sobre la BD empleada. Por una parte, se presentarán resultados en términos de rendimiento de la red. Por otra parte, se presentan los resultados finales en términos de métricas de calidad de imagen, de forma que se tenga una visión completa de la utilidad del método.

En el capítulo 5 “Discusión” se interpretan los resultados expuestos en el capítulo anterior. A continuación, se muestra una comparación con los resultados obtenidos en otros estudios previos.

En el capítulo 6 “Conclusiones y líneas futuras” se recogen las conclusiones más relevantes, así como las aportaciones del trabajo realizado. Asimismo, se describen las limitaciones y las posibles líneas de trabajo futuro que se podrían desarrollar.



## CAPÍTULO 2

---

### REVISIÓN DEL ESTADO DE LA TÉCNICA

2.1	INTRODUCCIÓN .....	25
2.2	MÉTODOS BASADOS EN SUPERRESOLUCIÓN DE IMÁGENES .....	25
2.3	MÉTODOS BASADOS EN CORRECCIÓN DE CONTRASTE E ILUMINACIÓN .....	27



## 2.1 INTRODUCCIÓN

En este TFG, con la revisión del estado de la técnica se analizan diferentes técnicas aplicadas a la mejora de la calidad de imágenes de retina. La literatura sobre la tarea de mejorar la calidad de retinografías es escasa, aunque pueden encontrarse algunos métodos aplicados a esta modalidad de imagen. Estos incluyen métodos basados en interpolación bicúbica y filtrado de Lanczos (Duchon, 1979), así como métodos más complejos basados en operadores estadísticos para recrear imágenes (Sun et al., 2008).

Recientemente, han surgido algunas implementaciones basadas en *Deep Learning*, que basan su funcionamiento en arquitecturas CNN. Esta técnica difiere de las técnicas tradicionales de *machine learning* en que no aprende con características extraídas por el desarrollador, sino que aprende características de manera autónoma (Saha et al., 2018). Para desarrollar estos sistemas, se entrenan CNNs con extensas bases de datos de retinografías que disponen de anotaciones de calidad. La red CNN aprende de manera automática las características de las imágenes y las clasifica (Zhao et al., 2018). Dentro de estas redes, se encuentran las conocidas como *Super-Resolution Convolutional Neural Network* (Dong et al., 2016), *Denoising Convolutional Neural Network*, *Low Light Net* o *Staged Low Light Net* (Kaur & Kaur, 2018).

De forma general, los métodos de mejora de calidad de imágenes de retina se pueden dividir en dos tipos:

- i) Métodos de superresolución (SR) de imágenes.
- ii) Métodos de corrección de contraste e iluminación.

Tanto la iluminación insuficiente o desigual como la falta de nitidez provocan que los detalles de diagnóstico sean indetectables (Qureshi et al., 2019). Por este motivo, esta revisión se ha centrado en métodos que permitan corregir ambos problemas, encontrando en los modelos SR de imagen basados en GAN los mejores resultados.

En este capítulo, se presenta una visión general sobre las técnicas existentes en visión artificial e imágenes médicas para dar una idea de cómo el aprendizaje automático puede aplicarse a la reconstrucción de imágenes.

## 2.2 MÉTODOS BASADOS EN SUPERRESOLUCIÓN DE IMÁGENES

El término resolución se refiere a la cantidad de pequeños detalles que se pueden observar en una imagen. La resolución de un dispositivo de imágenes depende de la calidad de sus componentes ópticos, de grabación y de visualización (Thapa et al., 2014). En este contexto, la superresolución de imágenes puede considerarse como una mejora de la resolución, cuando la adquisición de imágenes limita la resolución, o como una restauración de la resolución, cuando la transmisión de imágenes ha provocado una pérdida de datos y, en consecuencia, resolución degradada (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

El objetivo de los algoritmos de SR es estimar una imagen de alta resolución a partir de una imagen de entrada de baja resolución. En la mayoría de estos modelos, la imagen de baja resolución se obtiene aplicando un filtro gaussiano seguido de una operación de disminución de resolución (Ledig et al., 2017). Esta fase preliminar es necesaria durante el entrenamiento del modelo ya que es necesario simular el tipo de imágenes que se van a utilizar durante la fase de test. El objetivo es mejorar la calidad de una imagen que carece de suficiente nitidez o iluminación, por lo que es preciso que el entrenamiento se realice con imágenes que cumplen esas características.

Los métodos de SR basados en redes neuronales constan de varias etapas. Primero, en la fase de entrenamiento, se extraen las características de las imágenes de alta resolución y su versión de baja resolución. En esta etapa, se pueden usar distintos tipos de modelos de extracción de características, como valores de luminancia, coeficientes DCT, coeficientes *wavelet*, coeficientes de contorno, derivados de gradiente, histogramas, etc. En la segunda etapa, conocida como fase de test, se extraen las características de las imágenes de baja resolución de entrada. Finalmente, se buscan las características extraídas de las imágenes de test que coinciden con las características del conjunto de entrenamiento y se selecciona la mejor solución del conjunto de entrenamiento (Thapa et al., 2014).

En la literatura se pueden encontrar varios métodos basados en SR. Por ejemplo, Kim y Kwon (J. Kim et al., 2016) propusieron un método de SR basado en ejemplos, en el que cada característica de baja resolución de la imagen de test se comparó con las características del conjunto de entrenamiento almacenado y, utilizando un método de búsqueda de vecino más cercano, se seleccionó el píxel más cercano y su correspondiente píxel de alta resolución. Sin embargo, este enfoque a menudo da como resultado una imagen borrosa que puede corregirse mediante técnicas de postprocesamiento (Planck et al., 2008).

Otros autores, como Mathieu et al. (2016) y Denton et al. (2015), abordaron este problema mediante el empleo de GANs. Este tipo de redes también se ha utilizado en el aprendizaje mediante técnicas no supervisadas en Radford et al. (2016). La idea de usar redes de tipo GAN para aprender un mapeo de características se puede utilizar para la transferencia de textura de una imagen (C. Li & Wand, 2016) o para generar el contenido de las imágenes que se ha perdido en el proceso de adquisición (Yeh et al., 2017). Con objetivo de minimizar el error, el estudio de Johnson et al. (2016) y Bruna et al. (2016) propone el uso de características extraídas de una red VGG preentrenada. Más específicamente, estos autores formulan una función de pérdida basada en la distancia euclídea entre mapas de características extraídas de la red VGG19.

En el ámbito médico, en los últimos 30 años se han publicado una gran cantidad de trabajos para mejorar la resolución de imágenes médicas reales (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019). Los primeros métodos de mejora de la resolución, como las interpolaciones cúbicas básicas y sus variantes, generalmente sufren pérdidas de los detalles de bordes (Akhtar & Azhar, 2010). Basado en la regresión de vectores de soporte multinúcleo, Jebadurai y Peter propusieron un algoritmo SR para imágenes de retina (Jebadurai & Peter, 2018). Sin embargo, este método no era capaz de generar imágenes de buena calidad en el caso de un factor de escalado elevado.

Entre los métodos que utilizan la SR en imágenes médicas destacan los métodos basados en ejemplos que aprovechan la información de imágenes de alta resolución junto con una imagen de baja resolución y generan una imagen que se aproxima a la imagen de alta resolución original (Rousseau, 2014). Sin embargo, estos métodos dependen demasiado de datos externos que puede que no siempre estén disponibles, lo que presenta una gran desventaja (Mahapatra & Bozorgtabar, 2019).

Debido a los grandes logros del *Deep Learning* en visión artificial, se han desarrollado varias técnicas de SR para imágenes médicas. En este campo, y basado en la red VGG, Kim et al. (2016) presentaron un método SR altamente preciso con llamado *Very Deep Super Resolution* (VDSR). Dong et al. (2016) introdujeron por primera vez las CNN a la SR de imagen única, y propusieron el modelo SRCNN (*Super-Resolution Convolutional Neural Network*). Basado en la estructura básica de las CNN, se propone un método SR para imágenes médicas en escala de grises en (Liu et al., 2018a).

Recientemente, debido al buen desempeño de las GAN en la producción de imágenes realistas, los modelos de SR de imagen basados en este tipo de red están creciendo. Por ejemplo, SRGAN (Ledig et al., 2017), Neural Enhance (Sajjadi et al., 2017) y ESRGAN (Wang et al., 2019) son todos modelos de SR basados en GAN. Las GANs se pueden utilizar para mejorar la resolución de imágenes médicas (Zhao et al., 2018). Estas redes también se pueden utilizar para mejorar las imágenes en términos de color y nitidez (Chuang et al., 2018). Específicamente, Mahapatra et al. propusieron un algoritmo de SR para imágenes de retina utilizando redes adversas progresivas generativas (*Progressive GAN*, P-GAN) (Mahapatra et al., 2019). Otros métodos de SR de imágenes de fondo de ojo también basados en el marco GAN incluyen a Bing et al. (2019), quienes diseñaron una función de pérdida que permitía entrenar el modelo reduciendo las pérdidas de la red, lo que resultó en imágenes de retina muy similares a las originales. En la literatura también podemos encontrar el trabajo de Mahapatra & Bozorgtabar (2019), que propusieron un método de SR de imágenes de retina basado también en GANs que permitía aumentar las imágenes de retina en un factor de 32. Posteriormente, estas imágenes generadas se utilizaban para la segmentación de la red vascular. Finalmente, Zhao et al. (2019) propusieron un modelo de mejora de la imagen de la retina basado en GANs. Su objetivo fue mejorar una imagen borrosa y producir una imagen mejorada con la misma calidad que la imagen retiniana original.

### **2.3 MÉTODOS BASADOS EN CORRECCIÓN DE CONTRASTE E ILUMINACIÓN**

La correcta adquisición del color en la imagen retiniana es esencial para que los oftalmólogos realicen un buen diagnóstico. Debido a razones importantes como el proceso de adquisición y las enfermedades de la retina, la mayoría de las imágenes de retina pueden mostrar poca iluminación, desenfoque y bajo contraste, lo que impide la identificación de la afección retiniana subyacente. El realce de las retinografías es necesario para que las estructuras oculares y las posibles lesiones oculares sean más visibles con respecto al fondo de la retina (Pizer et al., 1987; Winder et al., 2009).

La corrección de la iluminación en imágenes ha sido estudiada en los últimos años, y se han propuesto varios métodos (Gupta & Tiwari, 2019; He et al., 2011; Polesel et al., 2000; Starck et al., 2003; Xiong et al., 2017). Para las imágenes médicas, los métodos se centran en tareas específicas. El método basado en la ecualización de histograma

adaptativo limitado por contraste (*Contrast-Limited Adaptive Histogram Equalization*, CLAHE), es ampliamente utilizado para mejorar el contraste de las imágenes de retina de baja calidad. Esto se debe a que una de las principales características de la imagen retiniana en color es el contraste de los vasos sanguíneos. El contraste de una imagen es una combinación del rango de los valores de intensidad y la diferenciación de estos valores de píxel máximos y mínimos. El objetivo de la mejora de la imagen mediante la manipulación del histograma, es obtener la distribución uniforme de la intensidad. Una imagen con contraste bajo tiene un rango de intensidad estrecho. La ecualización de histograma extiende la distribución de los valores de intensidad, ajustando la intensidad de la imagen original (Setiawan et al., 2013).

Otro método de ajuste de luminosidad y contraste fue propuesto por Zhou et al. (2018). En este estudio reforzaron la luminosidad con una matriz de ganancia de luminancia, que se obtiene mediante la corrección gamma del canal de valor en el espacio de color *Hue-Saturation-Value* (HSV), con el objetivo de mejorar los canales R, G y B respectivamente. Posteriormente, realizaron el contraste de la imagen en el canal de luminosidad del espacio de color CIELAB ( $L^*a^*b^*$ ) mediante CLAHE. Otros autores eliminaron la opacidad utilizando la transformada de Fourier y, posteriormente, aplicaron CLAHE para mejorar el contraste (Mitra, Roy, Roy, & Kumar, 2018).

En la literatura también se pueden encontrar métodos basados en *Deep Learning*. Entre ellos, el modelo de Shen et al. (2017) se basa en la teoría de Retinex y una CNN. En este modelo destaca la realización de una transformación logarítmica multiescala y una función de restauración de color. La transformación logarítmica multiescala mejora la imagen oscura original mediante sumas ponderadas de múltiples transformaciones logarítmicas, acelerando así la convergencia de la red. La función de restauración de color elimina principalmente las distorsiones de color y las zonas grises. Además de mejorar la imagen oscura original, este modelo también corrige el color de la imagen. Sin embargo, provoca un efecto halo en ciertas imágenes. Otros estudios previos incluyen métodos que utilizan una CNN para mejorar las imágenes con poca luz y, posteriormente, se aplica un modelo para mejorar el contraste (Tao et al., 2017). Este modelo también propone una estimación efectiva de la luz ambiental y logra un equilibrio entre el contraste y la preservación de la textura. En el estudio de Chuang et al. (2018) se propone un método de mejora del color de la imagen basado en GANs. El método solo requiere un conjunto de fotografías con el color adecuado en la fase de entrenamiento para producir imágenes con iluminación adecuada como salida en la fase de test.

Por otra parte, hay métodos basados en *Deep Learning* dedicados a mejorar el contraste de las imágenes. Por ejemplo, algunos autores han propuesto un método para identificar y eliminar las regiones de bajo contraste en una imagen mediante el uso de CNN. Los resultados muestran la recuperación de los detalles de las regiones de bajo contraste (Han et al., 2018). Finalmente, en la literatura más reciente podemos encontrar el artículo de Qureshi et al., (2019) en el que propusieron un método de corrección de color de imágenes de retina basado en métodos clásicos como CLAHE.



# CAPÍTULO 3

---

## MATERIALES Y MÉTODOS

3.1	INTRODUCCIÓN .....	31
3.2	DESCRIPCIÓN DE LA BASE DE DATOS EMPLEADA .....	31
3.3	MÉTODO DESARROLLADO .....	33
3.3.1	PREPROCESADO .....	33
3.3.2	GENERACIÓN DE IMÁGENES DE CALIDAD INSUFICIENTE.....	33
3.4	DEEP LEARNING EN MEJORA DE LA CALIDAD DE IMAGEN .....	36
3.4.1	REDES NEURONALES.....	36
3.4.2	REDES NEURONALES CONVOLUCIONALES.....	39
3.4.3	GENERATIVE ADVERSARIAL NETWORKS.....	50



### 3.1 INTRODUCCIÓN

En este capítulo se presentan las imágenes empleadas en este TFG, junto con los métodos aplicados en la tarea de mejora de la calidad de imágenes retinianas. El objetivo de este TFG es el desarrollo de un método de mejora de calidad de retinografías fiable y que produzca resultados que puedan ser analizados por especialistas. Este método debe ser capaz tanto de corregir tanto la borrosidad como la iluminación de las retinografías.

### 3.2 DESCRIPCIÓN DE LA BASE DE DATOS EMPLEADA

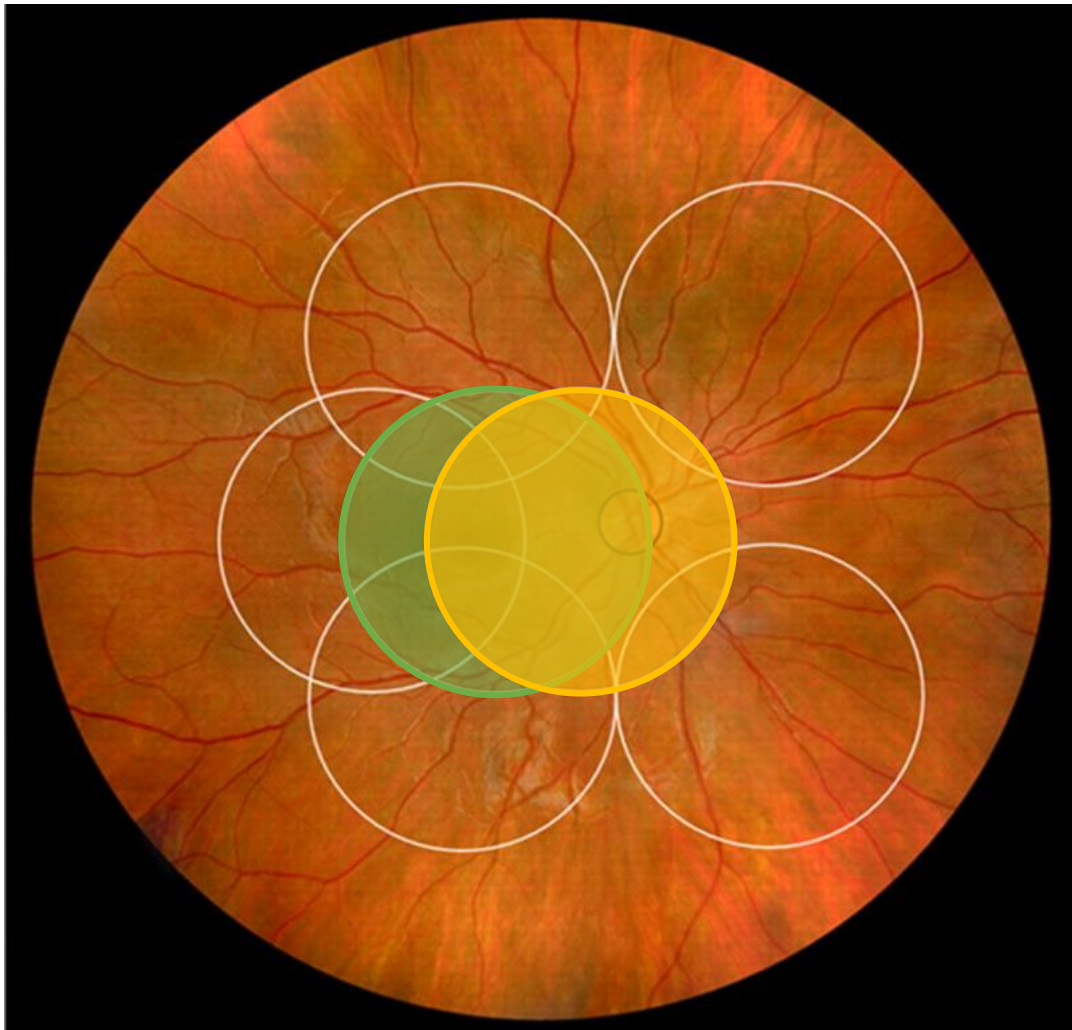
Para el desarrollo de este TFG se elaboró una BD privada de 1000 imágenes proporcionadas por el Instituto de Oftalmobiología Aplicada (IOBA) de la Universidad de Valladolid y el Hospital Clínico Universitario de Valladolid. Estas imágenes pertenecen al proyecto “Análisis Automático de Imágenes de Fondo de Ojo (A2IFO)”. Este proyecto fue financiado por el Ministerio de Economía y Competitividad y cofinanciado por la Unión Europea con cargo al Fondo Europeo de Desarrollo Regional (FEDER) dentro del objetivo temático "Promover el desarrollo tecnológico, la innovación y una investigación de calidad". Para la captura de estas imágenes, se utilizó el retinógrafo no midriático automático TRC-NW400 (Figura 3.1). Se trata de un retinógrafo de diseño compacto que tiene un campo de visión (*Field of View*, FOV) de 45°.



**Figura 3. 1** Retinógrafo no midriático TOPCON TRC-NW400.

La resolución de las imágenes que componen la BD es de 1956x1934 píxeles en formato JPEG con una profundidad de 24 bits. Esto implica que cada píxel podrá tomar 256 valores de intensidad diferentes en el rango 0 – 255, para cada uno de los tres canales de color RGB.

El protocolo de captura de imágenes utilizado ha sido el de dos campos del programa de cribado de la RD del Reino Unido. El *National Service Framework for Diabetes* del Reino Unido adoptó este protocolo en 2003, al considerar que, bajo midriasis, este protocolo resulta útil y aceptable tanto para los profesionales sanitarios como para los pacientes (Department of Health, 2002). Este protocolo se basa en la captura de fotografías del fondo de ojo de dos campos de la retina: uno centrado en mácula y con la papila situada en el lado nasal y otro centrado en papila y con la mácula en el lado temporal de la fotografía (Boucher et al., 2003). Ambos se muestran en la Figura 3.2. Se trata de un protocolo que cumple con los requisitos exigidos por el Instituto Nacional para la Salud y el Cuidado de la Excelencia (*National Institute for Health and Care Excellence, NICE*) para su utilización en un programa de cribado (Boucher et al., 2003).



**Figura 3. 2** Esquema de captura de los dos campos. en amarillo se representa el campo centrado en papila y en verde el campo centrado en mácula (fuente: <https://www.opticianonline.net/cet-archive/188>)

En total, la BD completa contiene un total de 1000 imágenes, de las cuales 100 son de calidad insuficiente y son de utilidad únicamente para comprobar la efectividad del método en la fase de test. Las restantes 900 imágenes de la BD presentan una calidad adecuada. Durante el entrenamiento, es preciso tener tanto las imágenes originales de buena calidad, como su versión en mala calidad generada sintéticamente, ya que el método se basa en un mapeo de las características de estas imágenes. Entrenar con imágenes de mala calidad sin tener su correspondiente imagen de buena calidad no logrará que la red aprenda cómo tratar estas imágenes.

Para validar la efectividad del sistema de mejora de la calidad, es recomendable que las poblaciones de cada clase estén equilibradas (Bishop, 1995). Por este motivo, en la fase de test, se trabajó con 100 imágenes de buena calidad y con 100 imágenes de insuficiente calidad. De este modo, el conjunto de entrenamiento se componía de 800 imágenes y las 200 restantes formaron el conjunto de test (Tabla 3.1).

### 3.3 MÉTODO DESARROLLADO

La mejora de calidad de imágenes de fondo de ojo es de utilidad en la ayuda al diagnóstico de la RD. Como se ha mostrado en el capítulo 2, se han desarrollado multitud de técnicas para la corrección de la nitidez e iluminación de imágenes de fondo de ojo. Entre ellas, son especialmente relevantes para este estudio aquellas que emplean redes de tipo GAN. En este trabajo, las GANs se emplean como base de un algoritmo de mejora de calidad de retinografías. Este método consta de cuatro etapas: preprocesado, generación de imágenes de calidad insuficiente, corrección de nitidez y corrección de iluminación. En las siguientes subsecciones se explica cada una de las etapas del método desarrollado.

#### 3.3.1 PREPROCESADO

Las imágenes fueron previamente redimensionadas a 512 x 512 píxeles con el objetivo de acelerar el proceso de entrenamiento y sin afectar significativamente a los resultados. Además, la gran mayoría de imágenes utilizadas en otros estudios de la literatura mencionados emplean imágenes de tamaño similar (Liu et al., 2018b; Zhao et al., 2018).

#### 3.3.2 GENERACIÓN DE IMÁGENES DE CALIDAD INSUFICIENTE

En el análisis de imágenes mediante redes de tipo GAN, como paso previo a la tarea de mejora de su calidad, es habitual procesarlas con el objetivo de obtener sus correspondientes imágenes de mala calidad, ya que se trata de un proceso imagen-imagen.

		Número de imágenes
<b>Conjunto de entrenamiento</b>	Calidad suficiente	800
	Calidad insuficiente	200
<b>Conjunto de test</b>	Calidad suficiente	100
	Calidad insuficiente	100
<b>Total</b>		1000

**Tabla 3. 1** Separación de las imágenes para la tarea de mejora de la calidad en un conjunto de entrenamiento y un conjunto de test

Esto quiere decir que en la fase de entrenamiento necesitamos tener tanto las imágenes originales de buena calidad, como las imágenes generadas de mala calidad (Takano & Alaghband, 2019). En este caso, la red del generador recibe como entrada una imagen de mala calidad y tiene la tarea de mejorar su calidad minimizando también un error de reconstrucción de píxeles (Armanious et al., 2018). El funcionamiento más detallado de las GANs se especifica en la sección 3.4.3.

En esta fase, ha sido necesario aplicar dos operaciones preliminares con objetivo de degradar la calidad las imágenes. La primera, aplicar un filtro gaussiano de media aleatoria para degradar la calidad de la imagen. El "desenfoco gaussiano" implica la convolución con un núcleo (*kernel*), descrita por una función gaussiana, con los píxeles de la imagen. La convolución en el dominio discreto viene dada por (Oppenheim et al., 1997):

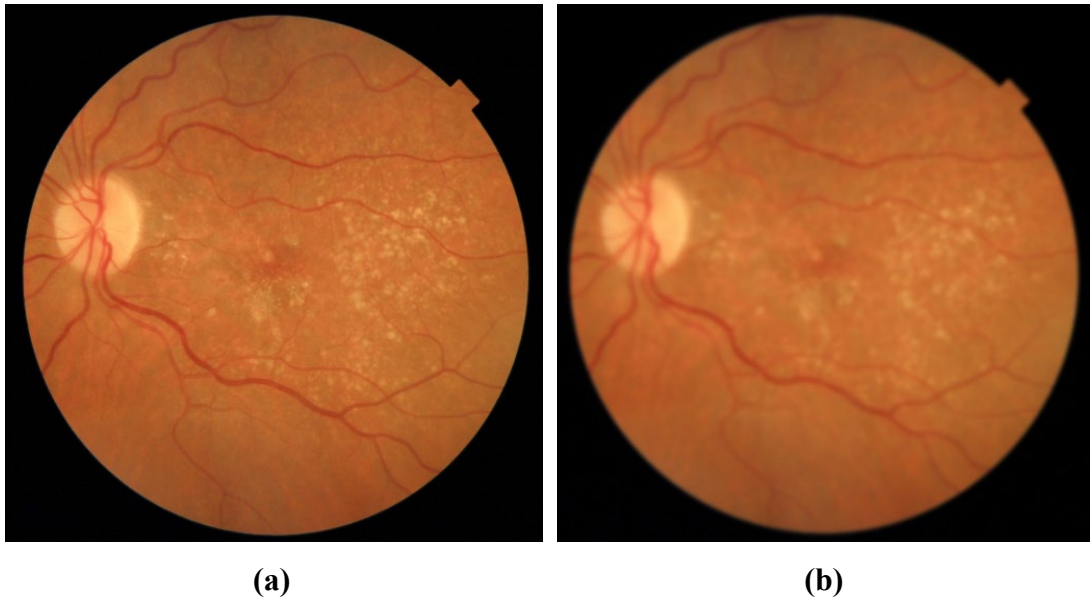
$$f * g[n] = \sum_{m=-\infty}^{\infty} f[m] \cdot g[n - m], \quad (3.1)$$

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (3.2)$$

donde  $f$  es la imagen y  $g$  es la función gaussiana. En la generación de las imágenes de este TFG, los valores de la distribución gaussiana se usan para construir una matriz de convolución que se aplica a la imagen original. El nuevo valor de cada píxel es un promedio ponderado de los vecinos de ese píxel. Por lo tanto, el valor del píxel original recibe el mayor peso (teniendo el valor gaussiano más alto) y los píxeles vecinos reciben pesos más pequeños a medida que aumenta su distancia al píxel original (en una vecindad de tamaño 4x4).

La convolución con el *kernel* gaussiano es un filtro paso bajo, que suaviza los bordes de la imagen (Gonzalez & Woods, 2000). El parámetro que se ha utilizado para definir el filtro es la varianza de la distribución gaussiana ( $\sigma$ ), ya que afecta a los resultados del filtrado (Gonzalez & Woods, 2000). Al aumentar la varianza del filtro gaussiano, los bordes se suavizan, de forma que se pierde contraste y la calidad se degrada (Gedraite & Hadad, 2011). En la Figura 3.3 se muestra la imagen original,  $I_{orig}$ , así como la imagen resultante de la aplicación del filtro gaussiano,  $I_{filter}$ . En este TFG el valor para el parámetro  $\sigma$  ha sido escogido aleatoriamente para cada iteración en el rango [1.5, 2]. Este rango se ha elegido experimentalmente con el objetivo de abarcar distintos grados de borrosidad. Si se incrementa o decrementa más ese valor, el resultado es una imagen demasiado borrosa, o, por el contrario, demasiado nítida.

En una segunda etapa, se creó una máscara negra circular del mismo radio de la retina de la imagen original para, posteriormente, superponer ambas imágenes. La máscara se ha creado buscando la mínima distancia entre el centro y el extremo de la imagen. Esta distancia conformará el radio de la máscara circular. A continuación, se calcula la distancia euclídea desde el centro a cada píxel y, si esa distancia es menor que



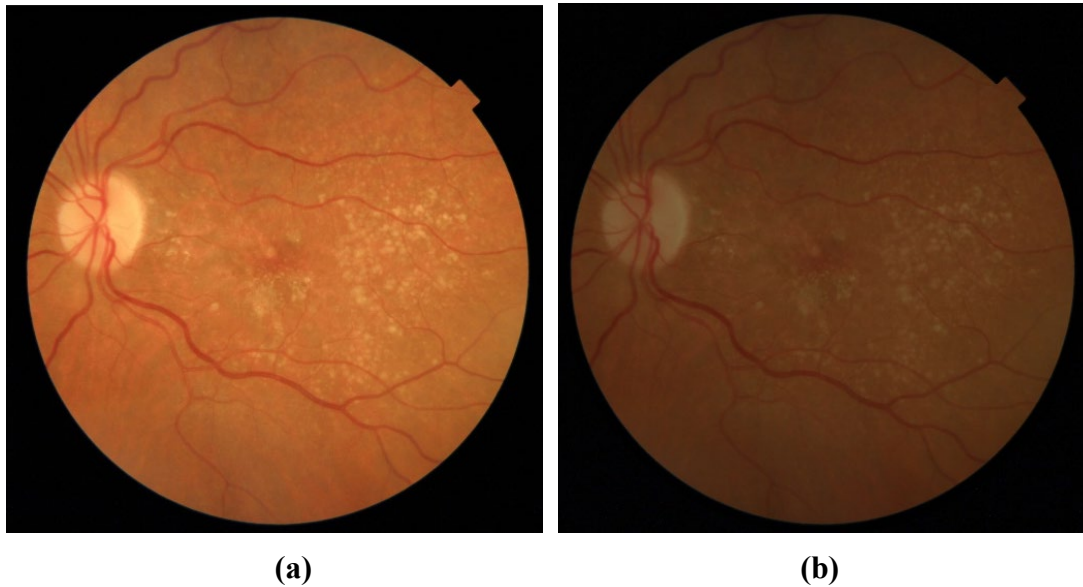
**Figura 3. 3** (a) Imagen original,  $I_{orig}$ . (b) Imagen filtrada,  $I_{filter}$ .

el radio de la circunferencia, se marca como incluido en la máscara, y si es mayor, se excluye de la máscara. La superposición de dos imágenes del mismo tamaño, da lugar a una tercera imagen con el mismo tamaño que las primeras. En ella, el valor de cada pixel es la suma de los valores correspondientes a cada pixel de las imágenes que se van a superponer. Los valores de los píxeles de salida vienen dados por:

$$R(m,n) = peso \cdot P(m,n) + Q(m,n) \quad (3.3)$$

donde  $P$ ,  $Q$  y  $R$  representan las matrices de la imagen original, la imagen negra y la imagen resultante, respectivamente. Todas estas imágenes son del mismo tamaño. Los parámetros  $m$  y  $n$  indican la posición del pixel considerado (Qidwai & Chen, 2009; Vernon, 1991). El término *peso* hace referencia a la opacidad que se da a la imagen original, por lo que cuanto menor sea este valor, más oscura será la imagen resultante. En este trabajo, el valor del peso se eligió de forma aleatoria para cada iteración del método dentro del rango  $[0.4, 0.6]$  ya que, si se decrementa o incrementa más este valor, el resultado es una imagen demasiado oscura o, por el contrario, demasiado clara. De esta manera se obtiene una imagen más oscura que la original. Es necesario mencionar que el rango empleado se ha elegido experimentalmente para abarcar distintos grados de oscuridad. La Figura 3.4 muestra la imagen original,  $I_{orig}$ , junto con la imagen oscurecida,  $I_{darker}$ , resultado de la superposición de la imagen original con la imagen negra.

El objetivo de este TFG es mejorar la calidad de imágenes por lo que, durante el entrenamiento, es necesario utilizar imágenes de calidad reducida para que la red aprenda cómo mejorar su calidad. Por este motivo, no se han incluido imágenes de buena calidad en la fase de entrenamiento.



**Figura 3. 4** (a) Imagen original, *Iorig*. (b) Imagen oscurecida, *Idarker*.

### 3.4 CORRECCIÓN DE NITIDEZ E ILUMINACIÓN MEDIANTE DEEP LEARNING

Dentro del *Deep Learning*, existen varios tipos de arquitecturas neuronales que se usan comúnmente en diversas aplicaciones de aprendizaje automático. Para el desarrollo de este TFG, las CNNs son las más adecuadas, ya que son redes que se utilizan en la visión artificial para el análisis de imagen (Aggarwal, 2018).

Esta sección presenta cómo el aprendizaje automático puede mejorar la calidad de retinografías generando imágenes de alta calidad a partir de datos incompletos, como son imágenes de mala calidad.

#### 3.4.1 REDES NEURONALES

Para comprender las CNN, es necesario comprender primero las redes neuronales básicas (*Neural Networks*, NN). El interés por las NN surge con el objetivo de emular el comportamiento del cerebro humano (Haykin, 1999). Los modelos de NN tratan de resolver ciertas tareas de forma eficiente mediante la interconexión de elementos computacionales simples (Lippmann, 1987) que se construyen a partir de (Hush & Horne, 1992):

- Modelos matemáticos de las células nerviosas individuales.
- Modelos arquitecturales que describen la interconexión entre las neuronas.

Se piensa que las neuronas realizan funciones simples, aunque se suelen modelar como sistemas no lineales y variantes en el tiempo. Los modelos de arquitectura empleados en las NNs comúnmente agrupan las neuronas en capas conectadas en una estructura masivamente paralela (Hush & Horne, 1992). Por lo tanto, la unidad básica de la arquitectura será la neurona artificial. Una o varias capas compondrán una NN. Por último, una NN (o un conjunto de ellas), junto con las interfaces de entrada y salida, constituirán el sistema global de proceso (García Gadañón, 2008).



Las NNs se pueden diseñar para resolver diferentes problemas a través de un proceso de aprendizaje o entrenamiento. Durante el entrenamiento, se presentan una serie de ejemplos a la entrada de la NN (conjunto de entrenamiento) a partir de los cuales la red aprende la tarea que debe realizar. El procedimiento computacional que se emplea en el proceso de aprendizaje se denomina algoritmo de entrenamiento, y cuya función es ajustar los pesos sinápticos de la NN de forma ordenada para conseguir el objetivo de diseño deseado (García Gadañón, 2008; Haykin, 1999).

El elemento esencial de procesamiento de la NN es la neurona artificial, la cual es un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas, proporciona una única respuesta o salida. En la Figura 3.5 se muestra el modelo genérico de una neurona artificial de etiqueta  $k$ , en el que podemos identificar los siguientes elementos (García Gadañón, 2008; Haykin, 1999; Martín & Sanz, 2006):

- **Conjunto de entradas**,  $x_m(t)$ .
- **Pesos sinápticos** de la neurona  $k$ ,  $w_{km}$ , que representan la intensidad de la interacción entre cada neurona presináptica  $m$  y la neurona postsináptica  $k$ .
- **Regla de propagación**,  $\sigma[w_{km}, x_m(t)]$ , que proporciona el valor del potencial postsináptico de la neurona  $k$ ,  $v_k(t)$ , en función de sus pesos y entradas.
- **Función de activación o función de transferencia**,  $\varphi_k[a_k(t-1), v_k(t)]$ , de la neurona  $k$  en función de su estado anterior  $a_k(t-1)$  y de su potencial postsináptico actual. Su función es limitar la amplitud de la salida de la neurona.
- **Función de salida**,  $F_k[a_k(t)]$ , que proporciona la salida actual de la neurona  $k$ ,  $y_k(t) = F_k[a_k(t)]$  en función de su estado de activación. Es decir, matemáticamente, una neurona es una no linealidad aplicada a una función afín (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019). Las características de entrada  $x = (x_1, x_2, \dots, x_n)$  se pasan a través de una función afín compuesta con una no linealidad  $\varphi$ :

$$y(x) = \varphi(\sum_i W_i x_i + b) = \varphi(W \cdot x + b) \quad (3.4)$$

con pesos dados  $\mathbf{W}$  y sesgo  $b$ .

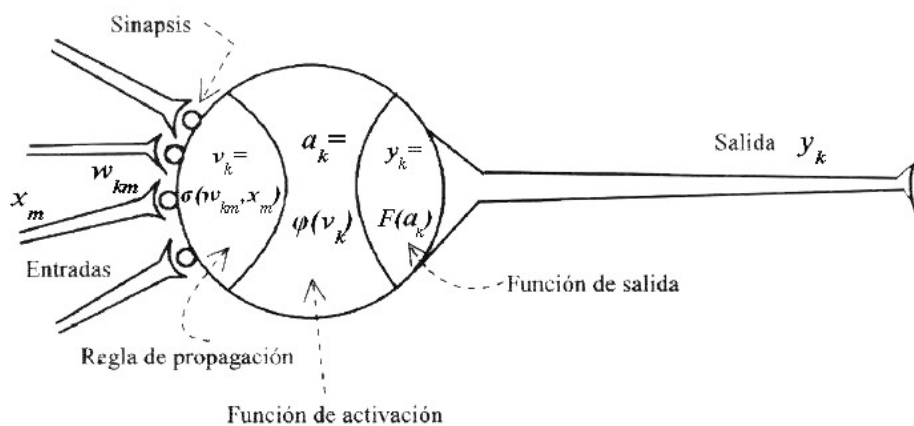


Figura 3. 5 Modelo no lineal de neurona artificial (Fuente: García Gadañón, 2008).

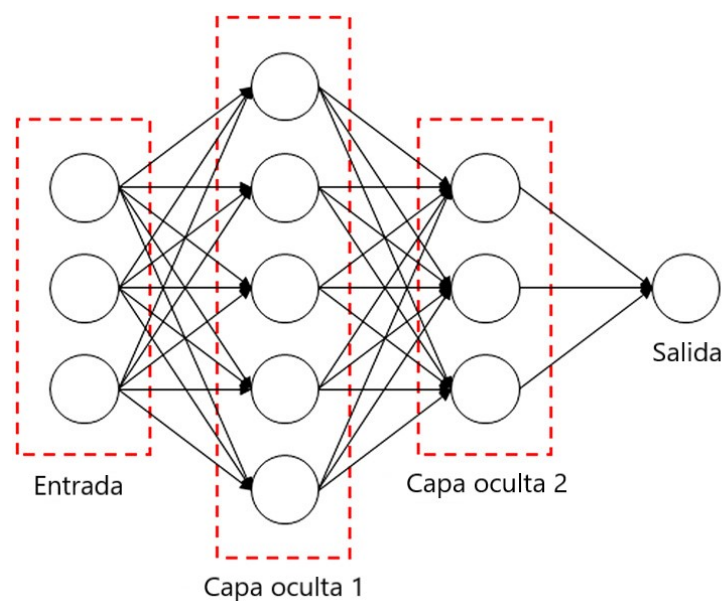
El *sesgo* o *bias* se trata de un peso externo cuya función es incrementar o disminuir la entrada a la función de activación, dependiendo de si esta es positiva o negativa, respectivamente (Haykin, 1999).

En definitiva, una NN puede modelarse como una colección de neuronas que están conectadas en un gráfico acíclico. Es decir, la salida de algunas de las neuronas se convierte en entradas a otras neuronas, y los ciclos en los que la salida de una neurona se asigna a una entrada intermedia anterior están prohibidos. Las NN constan de una capa de entrada, una o más capas ocultas y una capa de salida. A diferencia de las capas ocultas, la capa de salida generalmente no tiene una función de activación (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

Este tipo de redes generalmente se denominan Redes Neuronales Artificiales (*Artificial Neural Networks*, ANNs). Si queremos ser más explícitos sobre el número de capas, podríamos referirnos a una red como una red de  $N$  capas, donde  $N$  es el número de capas, excluyendo la capa de entrada. Un ejemplo de este tipo de redes se muestra en la Figura 3.6. Para usar una NN, necesitamos encontrar los valores adecuados para los parámetros ( $W$ ,  $b$ ) y definir una función para asignar la salida de la NN a una predicción. Estos parámetros son los llamados parámetros entrenables, y el número de estos parámetros sirve como una métrica para el tamaño (o capacidad) de la NN (Aggarwal, 2018).

#### 3.4.1.1 FUNCIÓN DE PÉRDIDA

Dependiendo de la tarea a llevar a cabo, las NN se pueden entrenar de forma supervisada o sin supervisión. Para el tipo de tareas que se encuentran con mayor frecuencia en las imágenes médicas, generalmente se trabaja con modelos discriminativos, lo que significa que tenemos dos espacios de objetos:  $X$  ("las entradas") e  $Y$  ("las etiquetas") y nos gustaría diseñar una red neuronal  $f_{\theta}$  con los parámetros  $\theta$ , que



**Figura 3. 6.** Ejemplo de red neuronal (adaptado de Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019)

genere un objeto en  $Y$  dado un objeto en  $X$ , o  $f_\theta: X \rightarrow Y$ . Para hacer esto, tenemos un conjunto de entrenamiento de tamaño  $m$  a nuestra disposición  $(X_i, Y_i)_{i=1}^m$  donde  $X_i$  es la entrada; en este caso, una imagen, e  $Y_i$  es la etiqueta correspondiente. Es decir,  $Y_i$  es la respuesta que esperamos del modelo  $g_\theta$  con  $X_i$  como entrada (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

En este contexto, el entrenamiento implica la modificación de los pesos sinápticos de la red empleando muestras de entrenamiento etiquetadas. Cada muestra de entrenamiento consiste en una única señal de entrada y su correspondiente etiqueta. Durante el entrenamiento se presenta a la entrada de la red un conjunto de muestras tomadas de forma aleatoria del espacio de entrada y los pesos de la misma se modifican de forma que se minimice la diferencia, de acuerdo con algún criterio estadístico bien definido, entre la etiqueta (respuesta que se desea obtener de la red) y la respuesta que realmente se obtiene para cada señal de entrada. La red aprende construyendo un mapeo entre el espacio de entrada y salida (García Gadañón, 2008).

Para formular el problema de entrenamiento, suponemos que hay una función de pérdida de valor real no negativo  $L$  que puede medir cuanto se aleja la predicción de la hipótesis  $y_{pred}$  del valor real  $y$ . El riesgo asociado con la hipótesis  $g_\theta$  se define como la expectativa de la función de pérdida (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019)

$$R(g_\theta) = E_{(X,y) \sim P(x,y)} [L(g_\theta(X), y)] \quad (3.5)$$

El objetivo del aprendizaje es encontrar parámetros  $\theta$  tales que  $R$  sea mínimo, es decir, queremos encontrar  $\theta^*$  tales que  $\theta^* = \arg \min_{\theta \in R^m} R(g_\theta)$  donde  $m$  es el número de parámetros de la red neuronal (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

En general, el riesgo  $R(g_\theta)$  no se puede calcular porque la distribución  $P(x, y)$  es desconocida. En este caso, reemplazamos la expectativa con el promedio sobre el conjunto de entrenamiento para obtener el *riesgo empírico* como una aproximación a la expectativa (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019):

$$R_{emp}(g_\theta) = \frac{1}{m} \sum_{i=1}^m L(g_\theta(X_i), y_i) \quad (3.6)$$

El proceso de minimizar el riesgo empírico se llama minimización del riesgo empírico (*Empirical Risk Minimization*, ERM).

### 3.4.2 REDES NEURONALES CONVOLUCIONALES

Para comprender las GAN es necesario comprender primero las CNNs, ya que es la base en la que se fundamentan las GAN. Las CNN son un caso especial de redes neuronales de alimentación directa, es decir, los datos atraviesan la red desde la capa de entrada hacia la capa de salida (Aggarwal, 2018). Estas redes son muy similares a las NN presentadas anteriormente en el sentido de que están formadas por neuronas con pesos y sesgos que se ajustan. La diferencia esencial es que la arquitectura CNN asume implícitamente que las entradas son como imágenes, lo que nos permite codificar ciertas

propiedades en la arquitectura (Vieira & Ribeiro, 2018). Esto hace que la función de transferencia sea más eficiente, reduciendo enormemente el número de parámetros y, por lo tanto, hace que la red sea más fácil de optimizar y menos dependiente del tamaño de los datos.

A diferencia de las redes neuronales simples, las capas de CNN tienen neuronas dispuestas en pocas dimensiones: número de canales, anchura, altura y número de filtros en el caso 2D. Una CNN consiste también en una secuencia de capas, donde cada capa transforma las activaciones o salidas de la capa anterior a través de una función diferenciable. Los bloques más comunes en las arquitecturas CNN son la capa de convolución, la capa de agrupación y las capas completamente conectadas. En esencia, estas capas actúan como extractores de características, reductores de la dimensión y clasificadores, respectivamente (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

### 3.4.2.1 CONVOLUCIONES

Matemáticamente, la convolución  $(x * w)(a)$  de las funciones  $x$  y  $w$  se define en todas las dimensiones como (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019)

$$(x * w)(a) = \int x(t)w(a - t)da, \quad (3.7)$$

donde  $a$  está en  $R^n$  para cualquier  $n \geq 1$ .

En la terminología de las CNN,  $x$  se llama entrada,  $w$  se llama núcleo o *kernel*, y la salida a menudo se denomina activación o mapa de características. En la ecuación (3.13) se modelan la entrada y el *kernel* como funciones continuas. Sin embargo, debido a la naturaleza discreta de los sensores de imagen, este no será el caso en la práctica y es más realista suponer que el parámetro  $t$  es discreto. Si suponemos que éste es el caso, entonces podemos definir la convolución discreta como (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019):

$$(x * w)(a) = \sum_a x(t)w(t - a), \quad (3.8)$$

donde  $a$  recorre todos los valores en el espacio y puede estar en cualquier dimensión. En *Deep Learning*, normalmente  $\mathbf{x}$  es una matriz de datos multidimensional y el *kernel*  $w$  involucra parámetros que se pueden aprender y usualmente tiene un soporte finito, es decir, únicamente hay determinados valores para los cuales  $w(a)$  no es cero. Esto significa que podemos implementar (3.14) como una suma finita. La definición en (3.14) es independiente de la dimensión, pero en las imágenes médicas, como las retinografías, trabajaremos principalmente con convoluciones bidimensionales, por lo que la convolución de las funciones  $I$  y  $K$  se define como (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019):

$$(I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n), \quad (3.9)$$

Además, las convoluciones (3.13) y (3.14) son conmutativas, lo que significa que  $I * K = K * I$ , de modo que también podemos escribir (3.15) como

$$(I * K)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n) \quad (3.10)$$

Hay que tener en cuenta que, en realidad, no se aplica un solo *kernel*, sino que habrá varios. El conjunto de varios *kernels* se conoce como filtros (Aggarwal, 2018).

Una observación es que la operación de convolución reduce el tamaño de la capa  $l+1$  en comparación con el tamaño de la capa  $l$ . Este tipo de reducción de tamaño no es deseable en general, ya que tiende a perder información a lo largo de los bordes de la imagen (Aggarwal, 2018). Este problema se puede resolver mediante el uso del concepto *padding* (relleno). El *padding* permite añadir  $(F_l - 1)/2$  píxeles alrededor de los bordes del mapa de características para mantener la dimensión espacial. Existen dos opciones de *padding* (Aggarwal, 2018):

- Rellenar la imagen con ceros.
- Eliminar la parte de la imagen donde el filtro no encaja. Esto se llama relleno válido que mantiene solo una parte válida de la imagen.

Además, hay otras formas en que la convolución puede reducir la dimensión espacial de la imagen. El enfoque anterior realiza la convolución en cada posición en la ubicación espacial del mapa de características. Sin embargo, no es necesario realizar la convolución en cada posición espacial en la capa. Se puede reducir la dimensión espacial de la imagen a través de la convolución usando la noción de *strides* (Aggarwal, 2018). *Strides* hace referencia al número de desplazamientos de píxeles sobre la matriz de entrada. Cuando *stride* es uno, movemos los filtros un píxel por vez. Cuando *stride* es dos, movemos los filtros dos píxeles a la vez y así sucesivamente. Como resultado, el uso de *strides* provocará una reducción de cada dimensión espacial de la capa (Aggarwal, 2018). Es más común usar un *stride* de 1, aunque ocasionalmente también se usa un *stride* de 2, como ocurre en el caso de este TFG.

### 3.4.2.2 CAPAS

Las capas que se han utilizado para la realización del método mostrado este TFG son las siguientes:

- **Capa de agrupación**

El objetivo de una capa de agrupación es producir una estadística resumida de su entrada y reducir las dimensiones espaciales del mapa de características, pero sin perder información esencial (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019). Para esto, la capa de agrupación máxima (*max pooling*) proporciona los valores máximos en cada conjunto de vecinos de cada entidad de entrada, mientras que la capa de agrupación promedio (*average pooling*) calcula los valores promedio. La forma más común de *max pooling* utiliza *stride* 2 junto con el tamaño de *kernel* 2, que corresponde a la división espacial del mapa de características en una cuadrícula regular de bloques cuadrados con el lado 2 y toma el valor máximo o el promedio sobre dichos bloques para cada característica de entrada (Aggarwal, 2018).

- **Capas completamente conectadas**

La capa completamente conectada con  $n$  dimensiones de entrada y  $m$  dimensiones de salida se define en base a la matriz de pesos  $\mathbf{W}$  y el vector sesgo  $\mathbf{b}$ . La salida de la capa está determinada por la matriz de peso  $W \in M_{m,n}(R)$  que tiene  $m$  filas y  $n$  columnas, y el vector de sesgo  $b \in R^m$ . Dado el vector de entrada  $\mathbf{x} \in R^n$ , la salida de una capa completamente conectada con la función de activación  $f$  se define como (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019)

$$FC(x) := f(W \cdot \mathbf{x} + b) \in R^m \quad (3.11)$$

En la fórmula anterior,  $W \cdot \mathbf{x}$  es el producto matricial y la función  $f$  se aplica componente a componente.

Las capas totalmente conectadas se usan como capas finales en problemas de clasificación, donde unas pocas de ellas, normalmente una o dos, se unen en la parte superior de una CNN. Para esto, la salida CNN se convierte en un solo vector (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

- **Capas de *Batch Normalization***

La técnica conocida como *batch normalization* permite abordar los problemas que hacen que los gradientes de activación en capas sucesivas se reduzcan o aumenten en magnitud. Además, otro problema importante en el desarrollo de redes profundas es el hecho de que los parámetros cambien durante el entrenamiento y, por lo tanto, las activaciones de variables ocultas también cambien (Aggarwal, 2018). Cambiar las entradas de las primeras capas a las capas posteriores provoca una convergencia más lenta durante el entrenamiento porque los datos de entrenamiento para las capas posteriores no son estables. El método *batch normalization* puede reducir este efecto (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

La idea de esta técnica es agregar "capas de normalización" adicionales entre capas ocultas que resistan este tipo de comportamiento mediante la creación de características con una variación algo similar. Además, cada unidad en las capas de normalización contiene dos parámetros adicionales  $\beta_i$  y  $\gamma_i$  que regulan el nivel de precisión de normalización en la unidad  $i$ -ésima (Aggarwal, 2018).

- **Capa convolucional**

Esencialmente, una capa de convolución usa un núcleo convolucional como filtro de entrada. Por lo general, hay muchos de esos filtros. Durante un paso hacia adelante, se calcula el mapa de activación del filtro en ese punto para obtener la activación en dicho punto (Aggarwal, 2018). Esto significa que, al entrenar una CNN de este tipo, la red aprenderá filtros que capturan algún tipo de información visual, como un borde o una orientación y, finalmente, en una capa superior de la red, patrones completos. En cada capa de convolución, tenemos un conjunto completo de dichos filtros, cada uno de los cuales producirá un mapa de activación por separado. Estos mapas de activación se apilan para obtener el mapa de salida (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

### 3.4.2.3 FUNCIONES DE ACTIVACIÓN

Las funciones de activación son esenciales para el diseño de una CNN. Se trata de funciones no lineales, lo que puede tener un gran impacto en la velocidad de entrenamiento de la red. En este TFG se han utilizado las siguientes (Aggarwal, 2018; Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019):

- **Sigmoidea.** Esta no linealidad se define como

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad x \in \mathbb{R} \quad (3.12)$$

Se puede observar que  $\sigma(x) \in (0, 1)$  para todo  $x \in \mathbb{R}$ . Además,  $\sigma$  es una función monótona creciente,  $\lim_{x \rightarrow \infty} \sigma(x) = 1$  y  $\lim_{x \rightarrow -\infty} \sigma(x) = 0$ .

Esto hace que la función de activación sigmoidea sea adecuada cuando el objetivo es producir salidas contenidas en el rango  $[0, 1]$ , como probabilidades o imágenes normalizadas. También se puede demostrar que  $\lim_{x \rightarrow \infty} \sigma'(x) = \lim_{x \rightarrow -\infty} \sigma'(x) = 0$ .

Este hecho implica que la no linealidad sigmoidea puede conducir a gradientes de fuga: cuando la entrada  $x$  a la función sigmoidea está lejos de cero, la neurona se saturará y el gradiente de  $\sigma(x)$  con respecto a  $x$  estará cerca a cero, lo que dificultará la optimización sucesiva. Esta es la razón por la cual las no linealidades sigmoideas rara vez se usan en las capas intermedias de las CNN.

- **Tanh.** La función tangente hiperbólica de sigmoide (*tanh*) se define como

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad x \in \mathbb{R} \quad (3.13)$$

Se puede demostrar que  $\tanh(x) \in (-1, 1)$  para todo  $x \in \mathbb{R}$ . Además,  $\tanh$  es una función monótona creciente,  $\lim_{x \rightarrow \infty} \tanh(x) = 1$  y  $\lim_{x \rightarrow -\infty} \tanh(x) = -1$ . Similar a la no linealidad sigmoidea, esta no linealidad puede conducir a gradientes de fuga y rara vez se usa en capas intermedias de CNN.

- **ReLU.** La función conocida como unidad lineal rectificadora (*Rectified Linear Unit*, ReLU) se define como

$$\text{ReLU}(x) = \max(0, x), \quad x \in \mathbb{R} \quad (3.14)$$

Se puede ver que  $\text{ReLU}'(x) = 1$  para  $x > 0$  y que  $\text{ReLU}'(x) = 0$  para  $x < 0$ . La no linealidad de ReLU generalmente conduce a una convergencia más rápida en comparación con las no linealidades sigmoidea o *tanh*, y generalmente funciona bien en CNN con una estrategia de inicialización de peso y una tasa de aprendizaje elegidas adecuadamente.

- **Leaky ReLU.** La función ReLU con fugas se define utilizando un parámetro adicional  $\alpha \in (0, 1)$ :

$$\phi(v) = \begin{cases} \alpha \cdot v & v \leq 0 \\ v & \text{en otro caso} \end{cases} \quad (3.15)$$

Aunque  $\alpha$  es un hiperparámetro elegido por el usuario, también es posible aprenderlo.

#### 3.4.2.4 CNN PREENTRENADA EMPLEADA: VGG19

La arquitectura VGG fue propuesta en 2014 por Simonyan y Zisserman (Simonyan & Zisserman, 2014). Se trata de una arquitectura CNN formada por cinco bloques que contienen una cierta cantidad de capas convolucionales, una función de activación no lineal y una capa *max pooling*, seguidos por 3 capas completamente conectadas y una capa *Softmax*. Los dos primeros bloques contienen dos capas convolucionales de 64 y 128 filtros respectivamente; el bloque intermedio contiene tres capas convolucionales de 256 filtros, y los dos últimos contienen tres capas convolucionales de 512 filtros cada uno (Simonyan & Zisserman, 2014).

Dentro de las distintas opciones de arquitecturas VGG, en este TFG se ha utilizado la arquitectura VGG19 porque ha demostrado muy bien rendimiento en imágenes genéricas y, en particular, en imágenes médicas (Litjens et al., 2017; Lundervold & Lundervold, 2019). VGG19 presenta 19 capas entrenables: 16 capas convolucionales y 3 capas completamente conectadas. Las capas convolucionales se basan en convoluciones  $3 \times 3$  con *stride* 1 y activaciones ReLU. Las operaciones de *max pooling* entre los bloques convolucionales se realizan con un *kernel* de  $2 \times 2$  y un *stride* de 2 (Simonyan & Zisserman, 2014). La arquitectura VGG19 se resume en la Tabla 3.2.

Es necesario mencionar que una opción para comenzar a entrenar una red es usar una CNN preentrenada, utilizar capas completamente conectadas con pesos inicializados aleatoriamente para una tarea de clasificación particular, y posteriormente ajustar la red resultante en un conjunto de datos particular (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019). Esta metodología para entrenar una red forma parte del concepto que se conoce como *transfer learning* y se explica más detalladamente en la sección 3.4.2.5. En este TFG se va a utilizar una red VGG19 preentrenada ya que, en la mayoría de estudios previos, ésta es la arquitectura empleada (Ahirwar, 2019; Ledig et al., 2017; Mitra, Roy, Roy, & Setua, 2018; Zhao et al., 2019).

#### 3.4.2.5 Transfer Learning. Inicialización de pesos y redes preentrenadas

Si dos neuronas en el mismo lugar de un gráfico computacional tienen exactamente el mismo sesgo y peso, siempre obtendrán exactamente los mismos gradientes y, por lo tanto, nunca podrán aprender características distintas, lo que provocaría la pérdida de algún poder expresivo de la red (Lu et al., 2017).

Al iniciar el entrenamiento de la red neuronal, es una práctica común inicializar los pesos aleatoriamente con media y varianza cero dependiendo del tamaño de entrada de la neurona, mientras que los sesgos aún se inicializan con ceros. La estrategia de



Tipo	N filtros/parámetros
Conv2D	64
Conv2D	64
Max Pool	-
Conv2D	128
Conv2D	128
Max Pool	-
Conv2D	256
Conv2D	256
Conv2D	256
Conv2D	256
Max Pool	-
Conv2D	512
Conv2D	512
Conv2D	512
Conv2D	512
Max Pool	-
Conv2D	512
Conv2D	512
Conv2D	512
Conv2D	512
Max Pool	-
Fully-Conn	4096
Fully-Conn	4096
Fully-Conn	1000
Softmax	-

**Tabla 3. 2** Arquitectura VGG19.

inicialización más reciente se desarrolló particularmente para las activaciones de ReLU, y sugiere inicializar los pesos mediante una distribución gaussiana de media cero cuya desviación estándar sea igual a  $\sqrt{\frac{2}{n}}$ , donde  $n$  se determina de la siguiente manera (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019):

- Al inicializar una capa completamente conectada,  $n$  es igual al número de características de entrada de la capa.
- Al inicializar, una capa convolucional bidimensional de dimensión  $k \times k$  con  $m$  mapas de características de entrada,  $n$  es igual al producto  $k^2 \cdot m$ .

Una segunda opción para comenzar a entrenar es usar una CNN preentrenada, utilizar capas completamente conectadas con pesos inicializados aleatoriamente, y luego ajustar la red resultante a un conjunto de datos particular. Esto forma parte de lo que se conoce como *transfer learning*.

Las arquitecturas de red actuales están formadas por una gran cantidad de parámetros y requieren un gran tiempo de entrenamiento y grandes conjuntos de datos. Un método comúnmente utilizado en imágenes médicas, especialmente en aquellos casos en los que es difícil adquirir grandes conjuntos de datos, es utilizar el concepto de *transfer learning* (Do & Ng, 2006). La idea de *transfer learning* (Do & Ng, 2006) es que las redes se entrenan previamente en grandes conjuntos de datos disponibles y posteriormente, la red se ajusta a una cantidad limitada de datos dependientes de la aplicación, de modo que las redes se adaptan a la aplicación actual (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

Esta técnica permite utilizar un modelo de código abierto que ha sido entrenado previamente en un gran conjunto de datos y usar sus parámetros como punto de partida para posteriormente continuar entrenando el modelo con el conjunto de datos que tengamos para una tarea determinada (Do & Ng, 2006). Si el conjunto de datos original con el que se entrenó la red neuronal preentrenada es suficientemente grande, entonces la jerarquía espacial de las características aprendidas por la red preentrenada permite a este modelo preentrenado actuar como un modelo genérico cuyas características pueden resultar útiles para resolver muchos problemas diferentes de visión artificial (Do & Ng, 2006). Este es el caso de este TFG, en el que se ha empleado una red VGG19 preentrenada. La necesidad de usar *transfer learning* en este TFG se debe al ahorro de tiempo de entrenamiento, al mejor rendimiento de las redes neuronales al utilizar redes que han sido entrenadas previamente y al hecho de no disponer de un gran número de imágenes de entrenamiento.

#### 3.4.2.6 Estandarización y aumento de datos

Antes de proporcionar los datos a la NN para el entrenamiento, generalmente se realiza un preprocesado de los mismos. Las dos formas más importantes de preprocesado son la estandarización de datos y el aumento del conjunto de datos (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

Los modelos de NN basados en *Deep Learning* aprenden una asignación de variables de entrada a una variable de salida. Como tal, la escala y la distribución de los datos extraídos del dominio de entrada pueden ser diferentes para cada variable. Las variables de entrada pueden tener diferentes unidades que, a su vez, puede significar que las variables tengan diferentes escalas. Las diferencias en las escalas entre las variables de entrada pueden aumentar la dificultad del problema que se está modelando. Por este motivo, es necesario que tanto los datos de entrada como los de salida se escalen a las mismas unidades (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019). En este TFG, antes de alimentar la red, ha sido necesario normalizar los valores. Los colores de los píxeles tienen valores comprendido entre 0 y 255, por lo que es necesario hacer una transformación de cada pixel:  $\frac{valor}{255}$  obteniendo, de esta forma, un valor entre 0 y 1.

Por otra parte, se sabe que las NN profundas se benefician de grandes cantidades de datos de entrenamiento, y es una práctica común ampliar artificialmente el conjunto de datos mediante un proceso llamado *aumento* (Aggarwal, 2018). Este proceso tiene como objetivo proporcionar un conjunto de datos de entrenamiento más grande para el algoritmo. En un escenario de aprendizaje supervisado, se nos da un conjunto de datos  $D$

que consiste en pares  $(x_j, y_j)$  de una muestra de entrenamiento  $x_j \in R^d$  y la etiqueta correspondiente  $y_j$ . Dado el conjunto de datos  $D$ , hay que diseñar transformaciones  $T_1, T_2, \dots, T_n: R^d \rightarrow R^d$  que preservan la etiqueta en el sentido de que para cada muestra  $(x_j, y_j) \in D$  y cada transformación  $T_i$  el vector resultante  $T_i x_j$  se vea como una muestra de  $D$  con la etiqueta  $y_j$ . Estas nuevas muestras resultantes con etiquetas asignadas se agregan al conjunto de datos de entrenamiento y se realiza la optimización de la forma habitual (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019).

En general, la elección de la transformación utilizada para el aumento de datos depende del conjunto de los mismos, pero hay algunas estrategias comunes para el aumento de datos en el campo de visión artificial. En este TFG se emplean las siguientes (Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, 2019):

- **Voltear.** La imagen  $x$  se refleja en una o dos dimensiones, produciendo una o dos muestras adicionales. Se puede ver un ejemplo de esta transformación en la Figura 3.7.
- **Rotación aleatoria.** Una imagen  $x$  se gira mediante un ángulo aleatorio  $\phi$  (a menudo limitado al conjunto  $\phi \in [\frac{\pi}{2}, \pi, \frac{3\pi}{2}]$ ). Un ejemplo de esta transformación se puede observar en la Figura 3.8.

Para aumentar el conjunto de datos de entrenamiento, en este TFG se eligió aleatoriamente una de las siguientes operaciones: rotar  $90^\circ$ ,  $180^\circ$  o  $270^\circ$ , o voltear la imagen. Con estas operaciones se consigue aumentar el conjunto de imágenes de entrenamiento, aumentando la precisión del método y evitando el sobreentrenamiento. Es decir, se evita que el modelo se ajuste a unas características muy específicas de los datos de entrenamiento y no responda adecuadamente sobre los conjuntos de datos de test (Aggarwal, 2018).

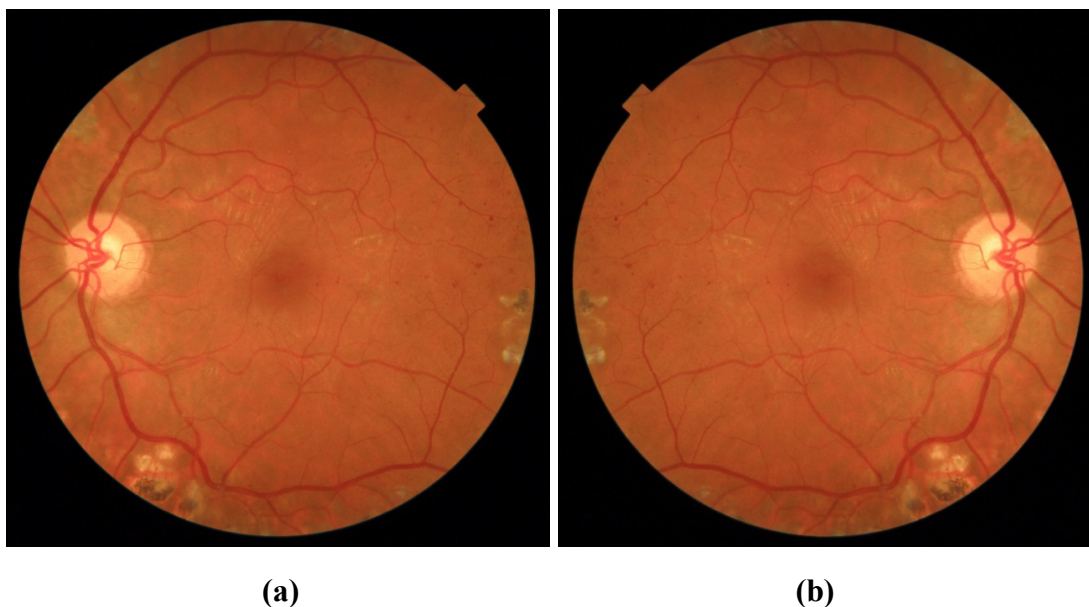
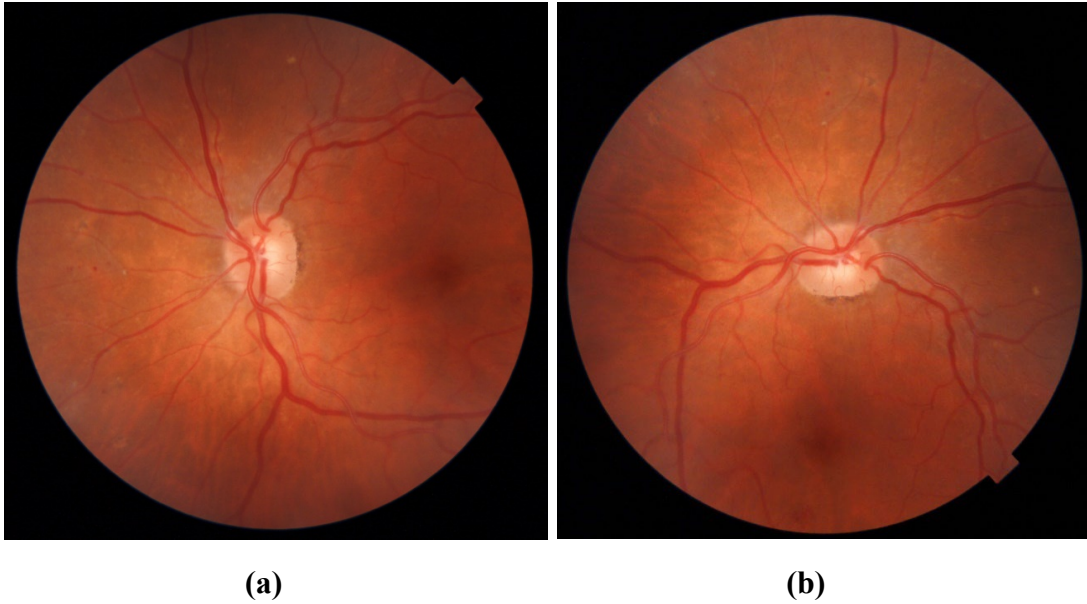


Figura 3. 7 (a) Imagen original. (b) Imagen volteada.



**Figura 3. 8** (a) Imagen original. (b) Imagen rotada  $\frac{\pi}{2}$ .

### 3.4.2.7 Optimizadores y tasa de aprendizaje

La CNN busca encontrar la solución que maximiza o minimiza una determinada función objetivo sujeta a una serie de restricciones (García Gadañón, 2008). Como se discutió anteriormente, el objetivo de optimización al entrenar NNs es minimizar el riesgo empírico  $R_{emp}$ . Esto se hace, en primer lugar, calculando el gradiente  $\nabla_{\theta} R_{emp}$  del riesgo en una muestra pequeña de datos de entrenamiento con respecto a los parámetros de la NN  $\theta$  utilizando *backpropagation*, y, en segundo lugar, actualizando los pesos de la NN  $\theta$ . Esta actualización, en su forma más básica de descenso de gradiente estocástico, está dada por la fórmula (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019):

$$\theta := \theta - \eta \nabla_{\theta} R_{emp}, \quad (3.16)$$

donde  $\eta > 0$  es el hiperparámetro llamado *tasa de aprendizaje*. Una extensión común de este algoritmo es la adición de un término adicional denominado *momentum*, que, en teoría, debería acelerar la convergencia del algoritmo en partes planas de la superficie de error. Es decir, esta técnica permite que la red responda no sólo al gradiente local, sino también a las tendencias de la superficie de error. De este modo, si la curvatura de la superficie de error es pequeña, la tasa de aprendizaje efectiva se incrementará. Por el contrario, en zonas de elevada curvatura, los pesos se ajustarán en una pequeña cantidad y el movimiento hacia el mínimo se hará sin oscilaciones divergentes (García Gadañón, 2008).

En este caso, el algoritmo recuerda la dirección de actualización anterior y lo combina con el gradiente recién calculado para determinar la nueva dirección de actualización (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019):

$$\delta\theta := \alpha \cdot \delta\theta - \eta \nabla_{\theta} R_{emp}, \quad (3.17)$$

$$\theta := \theta + \delta\theta$$

Las variaciones más recientes del descenso de gradiente estocástico son métodos adaptativos como RMSProp y Adam (Kingma & Ba, 2015). Adam, en particular, es una buena opción una vez que se elige una adecuada tasa de aprendizaje inicial, y es el optimizador que se ha escogido en este TFG siguiendo la arquitectura empleada en estudios anteriores (Ahirwar, 2019; Bing et al., 2019; Ledig et al., 2017; Mahapatra et al., 2019).

La elección de la tasa de aprendizaje depende en gran medida de la arquitectura de la NN. En general, tiene sentido elegir el tamaño del lote (*batch*) lo más grande posible dada la arquitectura de red y el tamaño de la imagen, y luego elegir la mayor tasa de aprendizaje posible que permita un aprendizaje estable. Si el error sigue oscilando (en lugar de disminuir constantemente), se recomienda reducir la tasa de aprendizaje inicial (Lu et al., 2017). Además, es común cambiar la tasa de aprendizaje durante el entrenamiento dependiendo del número actual de épocas (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019). Una época es un hiperparámetro que se define antes de entrenar el modelo. Al dividir la BD en *batches*, cuando se termina de procesar cada partición en la fase de entrenamiento se dice que ha concluido una época. A mayor tamaño de *batch*, más rápido se entrenará el modelo. Pero si este tamaño es demasiado grande, necesitará más épocas de entrenamiento. En este TFG se ha escogido una tasa de aprendizaje de 0.0001, un tamaño de *batch* de 1 y un número de épocas de 30000. Para la elección de estos parámetros se han seguido las indicaciones de otros autores en estudios previos (Ahirwar, 2019; Bing et al., 2019; Ledig et al., 2017; Mahapatra et al., 2019).

#### 3.4.2.8 Regularización

La regularización, en general, abarca un amplio número de técnicas destinadas a reducir el sobreajuste de los modelos basados en redes neuronales. En este TFG se utiliza la técnica conocida como *dropout*. Introducido en (Srivastava et al., 2014), *dropout* es un método de regularización muy potente y simple para NNs. Durante el entrenamiento, *dropout* se implementa manteniendo una neurona activa con cierta probabilidad  $p \in (0, 1)$  mientras se divide la activación de salida por  $p$ , y de lo contrario se establece en cero (Achille & Soatto, 2018). Por ejemplo, si asignamos un *dropout* entre la capa 3 y 4 de nuestro modelo, y decimos que el parámetro  $p$  es de 0.4 estamos diciendo que todos los valores que van a salir de la capa 3 en tiempo de entrenamiento tienen un 40% de probabilidad de llegar a la capa 4, mientras que el resto serán forzados a ser 0. A menudo, las probabilidades  $p$  se eligen de manera que las primeras capas se mantengan intactas con probabilidades cercanas o iguales a 1, mientras que la probabilidad de mantener la neurona activa disminuye para las capas más profundas, no siendo recomendable utilizarlo entre la última capa oculta y la capa de salida. Al eliminar de manera aleatoria valores entre las capas, se pueden crear modelos grandes sin que esto impacte la capacidad del modelo de generalizar a valores no vistos en el entrenamiento (Achille & Soatto, 2018; Srivastava et al., 2014).

### 3.4.3 GENERATIVE ADVERSARIAL NETWORKS

Las GANs consisten en dos CNN (Goodfellow et al., 2014). La primera red, el generador, intenta generar muestras sintéticas, pero perceptivamente convincentes,  $x \in p_{sintéticas}$ , que parezcan haber sido extraídas de una de distribución de datos real,  $p_{data}$ . Por lo tanto, la primera red transforma los vectores de entrada  $z$  extraídos de una distribución  $p_z$  en nuevas muestras, es decir,  $x = G(z)$ . La segunda red, el discriminador, tiene acceso tanto a las muestras reales de  $p_{data}$ , como a las muestras generadas por el generador e intenta discriminar entre ambos tipos de muestras. Las GANs se entrenan resolviendo el siguiente problema de optimización que el discriminador intenta maximizar y el generador intenta minimizar (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019):

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (3.18)$$

donde  $G$  es el generador,  $D$  es el discriminador,  $V(D, G)$  es la función objetivo,  $p_{data}$  es la distribución de muestras reales y  $p_z$  es una distribución de la que se extraen los vectores de entrada. La capa final de la red discriminadora contiene una función de activación sigmoidea, de modo que  $D(x), D(G(z)) \in [0, 1]$ . Al maximizar la función anterior, el discriminador minimiza el error de sus predicciones con respecto a los valores objetivo: 1 y 0 para muestras reales y generadas, respectivamente. Por el contrario, el generador intenta minimizar la posibilidad de que el discriminador prediga un 0 para muestras generadas. Por lo tanto, la pérdida del generador depende directamente del rendimiento del discriminador (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019).

La expresión anterior es una formulación general. En este TFG, la GAN desarrollada se entrena resolviendo el siguiente problema de optimización.

$$\min_{\theta_G} \max_{\theta_D} E_{I^{alta\ calidad} \sim p_{train}(I^{alta\ calidad})} [\log D_{\theta_D}(I^{alta\ calidad})] + E_{I^{baja\ calidad} \sim p_G(I^{baja\ calidad})} [\log(1 - D_{\theta_D}(G_{\theta_G}(I^{baja\ calidad}))) \quad (3.19)$$

La idea general detrás de esta formulación es que permite entrenar un modelo generativo  $G$  con el objetivo de engañar a un discriminador diferenciable  $D$  que está entrenado para distinguir imágenes generadas de alta calidad de imágenes reales (Zhou, S.Kevin; Rueckert, Daniel; Fichtinger, 2019). Con este enfoque, nuestro generador puede aprender a crear soluciones que son muy similares a las imágenes reales y, por lo tanto, difíciles de clasificar por el discriminador. El objetivo final es conseguir que el error del generador sea nulo, y que la probabilidad de que el discriminador sepa de dónde procede la muestra sea igual a 0.5, es decir, que no sea capaz de diferenciar ambas distribuciones. El esquema general del modelo propuesto en este TFG se puede observar en la Figura 3.9.

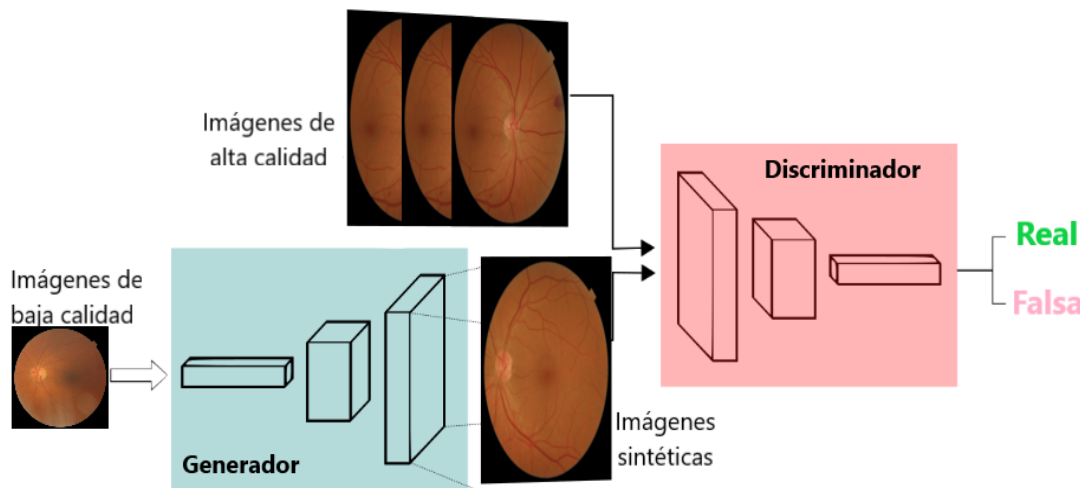


Figura 3.9 Modelo de mejora de calidad de imagen propuesto.

### 3.4.3.1 Arquitectura del generador

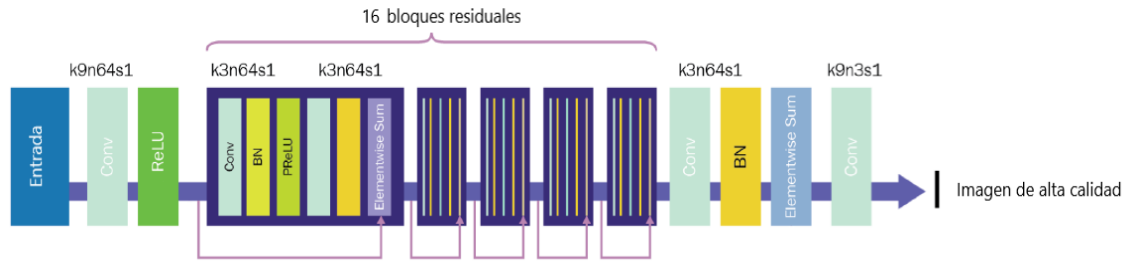
Para generar imágenes sintéticas de alta calidad que sean similares a imágenes reales, es necesario entrenar una red generadora. La arquitectura de esta red está basada en la arquitectura descrita en Ledig et al., (2017), aunque se han incluido algunas modificaciones. En concreto, desaparecen todos los bloques de interpolación.

La arquitectura utilizada para la red del generador en este TFG se muestra en la Figura 3.10. En ella se pueden distinguir los siguientes bloques:

- El bloque pre-residual.
- 16 bloques residuales.
- El bloque post-residual.
- La capa final de convolución.

A continuación, se detallan estos bloques uno por uno:

- El **bloque pre-residual**: este bloque contiene una capa de convolución 2D y ReLU como función de activación. La configuración del bloque se muestra en la Tabla 3.3.
- El **bloque residual**: el bloque residual contiene dos capas de convolución 2D. Ambas capas tienen a continuación una capa de normalización con un valor de *momentum* igual a 0.8. La configuración de cada bloque residual se detalla en la Tabla 3.4. La última capa de este bloque, llamada capa de adición, calcula la suma del tensor de entrada al bloque y la salida de la última capa de normalización.
- El **bloque post-residual**: este bloque también contiene una única capa de convolución 2D y ReLU como la función de activación. La capa de convolución tiene a continuación una capa de normalización con un valor de *momentum* de 0.8. La configuración del bloque post-residual se resume en la Tabla 3.5.
- La **capa final de convolución**: la última capa es una capa de convolución 2D que usa *tanh* como la función de activación. Genera una imagen de un tamaño de (512,



**Figura 3. 10** Arquitectura de la red del generador con su tamaño de kernel ( $k$ ), el número de filtros ( $n$ ) y el  $stride$  ( $s$ ) para cada capa convolucional. (adaptación de Ahirwar, 2019)

512, 3) siendo 3 el número de canales utilizados al tratarse de una imagen en color. La configuración de la última capa es la mostrada en la Tabla 3.6. Las columnas “forma de entrada” y “forma de salida” indican la altura, anchura y profundidad de la imagen al atravesar cada una de las capas.

Nombre de la capa	Hiperparámetros	Forma de entrada	Forma de salida
<i>2D convolution layer</i>	<i>Filters=64, kernel_size=3, strides=1, padding='same', activation='relu'</i>	(512, 512, 3)	(512, 512, 64)

**Tabla 3. 3** Especificación del bloque pre-residual.

Nombre de la capa	Hiperparámetros	Forma de entrada	Forma de salida
<i>2D convolution layer</i>	<i>Filters=64, kernel_size=3, strides=1, padding='same', activation='relu'</i>	(512, 512, 64)	(512, 512, 64)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(512, 512, 64)	(512, 512, 64)
<i>2D convolution layer</i>	<i>Filters=64, kernel_size=3, strides=1, padding='same'</i>	(512, 512, 64)	(512, 512, 64)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(512, 512, 64)	(512, 512, 64)
<i>Addition layer</i>	<i>None</i>	(512, 512, 64)	(512, 512, 64)

**Tabla 3. 4** Especificación del bloque residual.



Nombre de la capa	Hiperparámetros	Forma de entrada	Forma de salida
<i>2D convolution layer</i>	<i>Filters=64, kernel_size=3, strides=1, padding='same'</i>	(512, 512, 64)	(512, 512, 64)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(512, 512, 64)	(512, 512, 64)

Tabla 3. 5 Especificación del bloque post-residual.

Nombre de la capa	Hiperparámetros	Forma de entrada	Forma de salida
<i>2D convolution layer</i>	<i>Filters=3, kernel_size=9, strides=1, padding='same', activation='tanh'</i>	(512, 512, 64)	(512, 512, 3)

Tabla 3. 6 Especificación de la capa de convolución.

### 3.4.3.2 Arquitectura del discriminador

Para discriminar imágenes reales de muestras de alta calidad generadas, es necesario entrenar una red discriminadora. El diseño del discriminador está basado en la arquitectura descrita en Ledig et al., (2017). La arquitectura del discriminador utilizado en este TFG se muestra en la Figura 3.11.

La red discriminadora también es una CNN profunda. Contiene ocho bloques de convolución seguidos de dos capas densas completamente conectadas. Cada bloque de convolución tiene a continuación una capa de normalización. Las dos capas densas situadas al final de la red funcionan como un bloque de clasificación. La última capa predice la probabilidad de que una imagen pertenezca al conjunto de imágenes reales o al conjunto de imágenes falsas. La configuración detallada de la red discriminadora se muestra en la Tabla 3.7.

### 3.4.3.3 La función objetivo del entrenamiento. Funciones de pérdida

Para entrenar las GANs, hay una función objetivo o función de pérdida que debemos minimizar para entrenar el modelo. La función objetivo de las GAN para SR de imágenes se denomina función de pérdida perceptiva, que es una suma ponderada de las siguientes dos funciones de pérdida (Ahrwar, 2019):

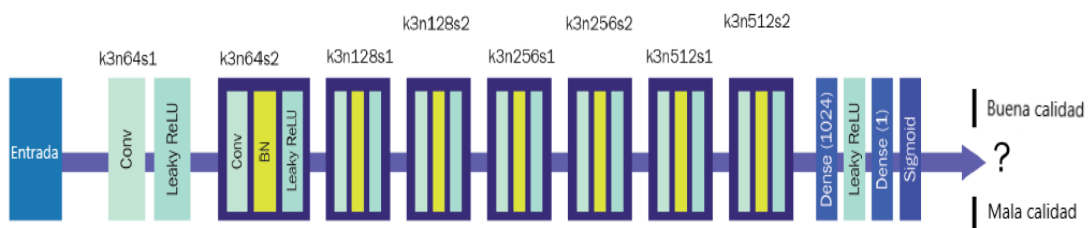


Figura 3. 11 Arquitectura de la red del discriminador con su tamaño de *kernel* (k), el número de filtros (n) y el *stride* (s) para cada capa convolucional (adaptación de Ahrwar, 2019)).

<b>Nombre de la capa</b>	<b>Hiperparámetros</b>	<b>Forma de entrada</b>	<b>Forma de salida</b>
<i>Input layer</i>	<i>None</i>	(512, 512, 3)	(512, 512, 3)
<i>2D convolution layer</i>	<i>filters=64, kernel_size=3, strides=1, padding='same', activation='leakyrelu'</i>	(512, 512, 3)	(512, 512, 64)
<i>2D convolution layer</i>	<i>filters=64, kernel_size=3, strides=2, padding='same', activation='leakyrelu'</i>	(512, 512, 64)	(256, 256, 64)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(256, 256, 64)	(256, 256, 64)
<i>2D convolution layer</i>	<i>filters=128, kernel_size=3, strides=1, padding='same', activation='leakyrelu'</i>	(256, 256, 64)	(256, 256, 128)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(256, 256, 128)	(256, 256, 128)
<i>2D convolution layer</i>	<i>filters=128, kernel_size=3, strides=2, padding='same', activation='leakyrelu'</i>	(256, 256, 128)	(64, 64, 256)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(64, 64, 256)	(64, 64, 256)
<i>2D convolution layer</i>	<i>filters=256, kernel_size=3, strides=1, padding='same', activation='leakyrelu'</i>	(64, 64, 256)	(64, 64, 512)
<i>Batch normalization layer</i>	<i>momentum=0.8</i>	(64, 64, 512)	(64, 64, 512)
<i>2D convolution layer</i>	<i>filters=256, kernel_size=3, strides=2, padding='same', activation='leakyrelu'</i>	(64, 64, 512)	(32, 32, 512)

**Tabla 3. 7** Especificación de los bloques de la arquitectura de la red del discriminador.

- Pérdida de contenido:
  - Pérdida basada en el error cuadrático medio (*Mean Squared Error*, MSE) de los píxeles.
  - Pérdida de VGG.
- Pérdida adversa.

### Pérdida de MSE en píxeles

La pérdida de contenido es el error cuadrático medio calculado entre cada valor de píxel de la imagen real y cada valor de píxel de la imagen generada. Esta pérdida calcula cómo

de diferentes son las imágenes generadas de las imágenes reales. La fórmula para calcular la pérdida de MSE en píxeles es la siguiente (Ahirwar, 2019):

$$l_{MSE} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{altacalidad} - G_{\theta_G}(I^{bajacalidad})_{x,y})^2. \quad (3.20)$$

Aquí,  $G_{\theta_G}(I^{bajacalidad})$  representa una imagen de alta resolución generada por la red generadora.  $I_{x,y}^{altacalidad}$  representa una imagen de alta resolución muestreada del conjunto de datos real.

### Pérdida VGG

La pérdida de VGG representa otra pérdida de contenido, que se aplica sobre imágenes generadas e imágenes reales. Las capas intermedias de una red VGG19 previamente capacitada funcionan como extractores de características y se pueden usar para extraer mapas de características de las imágenes generadas y las imágenes reales. La pérdida de VGG se basa en estos mapas de características extraídos. Se calcula como la distancia euclidiana entre los mapas de características de la imagen generada y la imagen real. La fórmula para la pérdida de VGG es la siguiente (Ahirwar, 2019):

$$l_{VGG/i,j} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\Phi_{i,j}(I^{altacalidad})_{x,y} - \Phi_{i,j}(G_{\theta_G}(I^{bajacalidad}))_{x,y})^2. \quad (3.21)$$

Aquí,  $\Phi_{i,j}$  representa el mapa de características generado por la red VGG19.  $\Phi_{i,j}(I^{altacalidad})$  representa los mapas de características extraídos de la imagen real y  $\Phi_{i,j}(G_{\theta_G}(I^{bajacalidad}))$  representa los mapas de características extraídos de la imagen generada de alta resolución.

Cualquiera de las pérdidas de contenido anteriores se puede utilizar para entrenar las GAN. En este TFG se utiliza la pérdida VGG.

### Pérdida adversa

La pérdida adversa se calcula sobre las probabilidades devueltas por la red discriminadora. En las redes tipo GAN, la red discriminadora se alimenta con imágenes

generadas por la red generadora. La pérdida de confrontación puede representarse mediante la ecuación (Ahirwar, 2019):

$$l_{Gen} = \sum_{n=1}^N -\log D_{\theta_D}(G_{\theta_G}(I^{bajacalidad})), \quad (3.22)$$

donde  $G_{\theta_G}(I^{bajacalidad})$  es la imagen generada y  $D_{\theta_D}(G_{\theta_G}(I^{bajacalidad}))$  representa la probabilidad de que la imagen generada sea una imagen real.

### **Pérdida total**

La función de pérdida total de esta red es una suma ponderada de la pérdida de contenido y la pérdida de confrontación, que se calcula como (Ahirwar, 2019):

$$l_{GAN} = 1.0 * l_{VGG/i,j} + 0.001 * l_{Gen} \quad (3.23)$$

Donde la pérdida total está representada por  $l_{GAN}$ .  $l_{VGG/i,j}$  es la pérdida de contenido provocada por la red VGG19. Finalmente,  $l_{Gen}$  representa la pérdida adversa.

Al minimizar el valor de la pérdida, la red del generador intenta engañar al discriminador. A medida que disminuye el valor de la pérdida de percepción, la red del generador comienza a generar imágenes más realistas.

# CAPÍTULO 4

---

## RESULTADOS

4.1	INTRODUCCIÓN .....	59
4.2	FORMA DE EVALUACIÓN.....	59
4.3	RESULTADOS PARA LA MEJORA DE CALIDAD DE IMÁGENES BORROSAS .....	63
4.3.1	FASE DE ENTRENAMIENTO. RENDIMIENTO DE LA GAN. ....	63
4.3.2	FASE DE TEST .....	64
4.4	RESULTADOS PARA LA MEJORA DE CALIDAD DE IMÁGENES OSCURAS .....	68
4.4.1	FASE DE ENTRENAMIENTO. RENDIMIENTO DE LA GAN. ....	68
4.4.2	FASE DE TEST. ....	71
4.5	RESULTADOS PARA AMBAS REDES EN SERIE .....	73
4.5.1	RESULTADOS SOBRE LAS IMÁGENES DE TEST DE BUENA CALIDAD .....	75
4.5.2	RESULTADOS SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD.....	76



## 4.1 INTRODUCCIÓN

Una vez que se ha descrito la metodología empleada en este trabajo, se presentan en este capítulo los principales resultados que ofrece el sistema de mejora de calidad de retinografías desarrollado en este TFG. Se han obtenido resultados sobre imágenes borrosas e imágenes oscuras. Para evaluar los resultados, se utilizaron dos tipos de métricas: una evaluación de referencia completa, en la que se compara la imagen original con la imagen generada, y una evaluación sin referencia, en la que se calculan ciertas medidas tanto sobre la imagen original como sobre la generada, que se detallan a lo largo de este capítulo.

## 4.2 FORMA DE EVALUACIÓN

El modelo basado en GAN propuesto se ha implementado en *Keras* en CentOS Linux 7 con CUDA 10.1, CUDNN 7 y TITAN X (Pascal)/PCIe/SSE2. Como se ha comentado en la sección 3.2, el conjunto de imágenes de la BD utilizada se ha dividido en dos subconjuntos: entrenamiento y test. Hay dos subconjuntos dentro de la BD de test. El primero contiene únicamente imágenes de buena calidad, mientras que el segundo solo contiene imágenes borrosas o poco iluminadas. En el primer caso, se ha degradado la calidad de estas imágenes de buena calidad para estimar su versión de alta calidad del mismo modo que se hizo en la fase de entrenamiento.

El conjunto de entrenamiento se utiliza para que la red aprenda las características óptimas para la generación de imágenes realistas. Una vez finalizada la fase de entrenamiento, el algoritmo se validó sobre las imágenes del conjunto de test para obtener los resultados del estudio.

Para evaluar el método de mejora de la calidad de imágenes de retina, se lleva a cabo una evaluación visual y cuantitativa. Dentro de la evaluación cuantitativa, se distinguen dos tipos: evaluación con referencia y evaluación sin referencia. Para la evaluación sin referencia, se utilizan las métricas *Blind/Referenceless Image Spatial Quality Evaluator* (BRISQUE), *Natural Image Quality Evaluator* (NIQE) y entropía. Todas estas métricas de calidad sin referencia nos dan una calidad de imagen absoluta pero no la similitud con una imagen de referencia. En cuanto a la evaluación con una imagen de referencia, se utilizaron la relación señal a ruido (*Peak Signal-to-Noise Ratio*, PSNR) y el índice de similitud estructural (*Structural Similarity*, SSIM). Estas dos últimas métricas ofrecen una comparación entre la imagen generada y la imagen original. A continuación, se describen las métricas mencionadas anteriormente:

- ***Peak Signal-to-Noise-Ratio***

La relación señal a ruido (*Peak Signal-to-Noise-Ratio*, PSNR) tiene como objetivo mostrar la calidad de la imagen resultante después de la aplicación del método desarrollado. Un valor de PSNR mayor muestra una mejor calidad de la imagen reconstruida respecto de la degradada. Para obtener el valor de PSNR, es necesario

calcular previamente el error cuadrático medio (*Mean Squared Error*, MSE), como se indica a continuación (Qureshi et al., 2019):

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [X(i,j) - Y(i,j)]^2}{M \cdot N} \quad (4.1)$$

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right) \quad (4.2)$$

En las ecuaciones anteriores,  $X(i, j)$  representa una imagen de entrada con un índice  $i$  y  $j$ , con  $M \times N$  píxeles, mientras que  $Y(i, j)$  denota la imagen mejorada resultante. MAX indica la diferencia máxima entre un valor de imagen original y el valor de MSE.

- ***Structure Similarity Index Measure***

El parámetro conocido como *Structure Similarity Index Measure* (SSIM) es un índice empleado para medir la similitud de la estructura entre una imagen original y una mejorada. Un valor mayor de SSIM indica que la imagen es muy similar a la original (Horé & Ziou, 2010). Así, un valor SSIM igual a la unidad significa la mejor similitud posible entre la imagen original y la generada. La ecuación (4.4) muestra el cálculo del índice SSIM (Qureshi et al., 2019):

$$SSIM(X, Y) = [l(X, Y)]^\alpha \cdot [c(X, Y)]^\beta \cdot [s(X, Y)]^\gamma \quad (4.3)$$

En la ecuación anterior,  $X$  e  $Y$  denotan dos ventanas con las mismas dimensiones que las imágenes original y reconstruida;  $\alpha > 0$ ,  $\beta > 0$ ,  $\gamma > 0$  denotan exponentes constantes utilizados como parámetros de peso; y  $l$ ,  $s$ ,  $c$  representan componentes de luminancia, estructura y contraste, calculados de la siguiente manera (Qureshi et al., 2019):

$$l(X, Y) = \frac{2\mu_x\mu_y + a_1}{\mu_x^2 + \mu_y^2 + a_1} \quad (4.4)$$

$$s(X, Y) = \frac{\sigma_{xy} + a_3}{\sigma_x\sigma_y + a_3} \quad (4.5)$$

$$c(X, Y) = \frac{2\sigma_x\sigma_y + a_2}{\sigma_x^2 + \sigma_y^2 + a_2} \quad (4.6)$$

Los términos  $\mu$  y  $\sigma$  indican el promedio y la varianza de  $X$  e  $Y$ , respectivamente, mientras que  $a_1$ ,  $a_2$  y  $a_3$  se usan como parámetros constantes para evitar la inestabilidad entre el promedio y la varianza de  $X$  e  $Y$ . Estos parámetros se calcularon como



específicamente como  $a = (KL)^2$  (Qureshi et al., 2019). En la expresión anterior,  $L$  denota el rango dinámico de la intensidad de píxeles de la imagen (255 en este caso) y  $K \approx 0.01$  es una constante pequeña (Qureshi et al., 2019). En este TFG se tomaron  $\alpha = \beta = \gamma = 1$  y  $a_3 = \frac{a_2}{2}$  siguiendo el artículo de Qureshi et al., (2019).

- **Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)**

Los coeficientes de luminancia normalizados de las imágenes naturales siguen la distribución de probabilidad de Gauss (Ruderman, 1994). Por lo tanto, la distorsión de la imagen cambiará las características estadísticas del coeficiente normalizado. Al medir el cambio de las características estadísticas, se pueden predecir los tipos de distorsión y evaluar la calidad visual de la imagen. Basado en esta teoría, Mittal et al. propusieron el algoritmo BRISQUE, basado en características estadísticas espaciales de las imágenes (A Mittal et al., 2012). Este método espacial que han desarrollado se puede resumir de la siguiente manera:

### 1) Normalización de píxeles de la imagen

Dada una imagen que posiblemente presente algún grado de distorsión, se calculan luminancias localmente normalizadas a través de la resta media local y la normalización divisiva. La fórmula se puede aplicar a una imagen de intensidad dada  $I(i, j)$  de la siguiente manera (Fu & Wang, 2016):

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + c}, \quad (4.7)$$

donde:

$$\mu(i, j) = \sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} I(i + k, j + l) \quad (4.8)$$

$$\sigma(i, j) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} [I(i + k, j + l) - \mu(i, j)]^2}. \quad (4.9)$$

$I(i, j)$  representa el valor gris de la imagen original,  $I \in \{1, 2, \dots, M\}$ ,  $j \in \{1, 2, \dots, N\}$ .  $M$  y  $N$  son la altura y la anchura de la imagen respectivamente;  $c$  es una constante pequeña que permite evitar la inestabilidad de los resultados calculados cuando el denominador se aproxima a 0. Por otra parte,  $\mu(i, j)$  y  $\sigma(i, j)$  son las medias ponderadas y la varianza, respectivamente. La función  $w = \{w_{k,l} | k = -K, \dots, K, l = -L, \dots, L\}$  es una función de ponderación gaussiana bidimensional circularmente simétrica. Se denomina a los valores de brillo normalizados de  $\hat{I}(i, j)$  como coeficientes *Mean Subtracted Contrast Normalized* (MSCN).

### 2) Extracción de características espaciales

La existencia de distorsión romperá la regularidad entre los coeficientes MSCN adyacentes. Las características de los coeficientes normalizados incluyen las características de distribución de Gauss generalizadas y la correlación de los coeficientes adyacentes (A Mittal et al., 2012).

### 2.1) Características de distribución de Gauss generalizadas:

Estas características se extraen de la siguiente manera (A Mittal et al., 2012):

$$f(x; a, \sigma^2) = \frac{a}{2\beta\tau(\frac{1}{a})} \exp\left(-\left(\frac{|x|}{\beta}\right)^a\right) \quad (4.10)$$

donde  $\beta = \sigma \sqrt{\frac{\tau(\frac{1}{a})}{\tau(\frac{3}{a})}}$ ,  $\tau(\cdot)$  es la función gamma. El parámetro de forma  $a$  controla la forma del modelo Gauss generalizado, y  $\sigma^2$  es la varianza (A Mittal et al., 2012).

### 2.2) Correlación de coeficientes adyacentes:

Adicionalmente, se obtienen las imágenes de correlación de las cuatro direcciones: horizontal, vertical, diagonal y diagonal. Para ello se utiliza la Distribución Gaussiana Asimétrica Generalizada (AGGD) para ajustar las imágenes. Posteriormente, se utiliza el método *fastmatching* (Lasmar et al., 2009) para estimar los parámetros correspondientes a cada dirección. Se utilizan 16 parámetros de estimación de cuatro direcciones como la característica de correlación de los coeficientes adyacentes. Sin embargo, debido a las características estadísticas multiescala de las imágenes naturales, Ruderman descubrió que es más razonable extraer dos características de distribución generalizadas y 16 coeficientes de características de correlación en dos escalas (Ruderman, 1994). Por lo tanto, en total se extrajeron  $(2 + 16) \times 2 = 36$  (A Mittal et al., 2012).

En resumen, el evaluador BRISQUE es un modelo de evaluación de calidad de imagen basado en estadísticas de escena natural (*Natural Scene Statistic*, NSS) que opera en el dominio espacial. No calcula características específicas de distorsión como timbre, desenfoque o bloqueo, sino que utiliza estadísticas de escena de coeficientes de luminancia normalizados localmente para cuantificar posibles pérdidas de 'naturalidad' en la imagen debido a la presencia de distorsiones (A Mittal et al., 2012).

- ***Natural Image Quality Evaluator (NIQE)***

El parámetro NIQE se basa en la construcción de una colección de características estadísticas de calidad basadas en un modelo estadístico de escena natural (*Natural Scene Statistic*, NSS) en el dominio espacial. Estas características se derivan de un conjunto de imágenes naturales sin distorsiones.

La calidad de una imagen de prueba dada se expresa entonces como la distancia entre un ajuste gaussiano multivariado (*Multivariate Gaussian*, MVG) de las características NSS extraídas de la imagen de prueba, y un modelo MVG de las características de calidad extraídas del conjunto de imágenes naturales (Anish Mittal et al., 2013):

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{\left((v_1 - v_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1} (v_1 - v_2)\right)} \quad (4.11)$$

donde  $\mathbf{v1}$ ,  $\mathbf{v2}$  y  $\Sigma_1$ ,  $\Sigma_2$  son los vectores promedio y las matrices de covarianza del modelo MVG natural y el modelo MVG de la imagen distorsionada respectivamente.

- **Entropía**

El objetivo de la entropía es presentar la información de contenido promedio de una imagen. Un valor mayor de entropía muestra una mejor calidad de imagen (Qureshi et al., 2019)

$$ENT = - \sum_{l=0}^{L-1} I(l) \log I(l) \quad (4.12)$$

Donde  $ENT$  es la entropía,  $L$  denota el número de niveles de gris de la imagen,  $I$  es la intensidad del nivel y  $I(l)$  es una función de densidad de probabilidad.

### 4.3 RESULTADOS PARA LA MEJORA DE CALIDAD DE IMÁGENES BORROSAS

#### 4.3.1 FASE DE ENTRENAMIENTO. RENDIMIENTO DE LA GAN

En este TFG, el rendimiento de la GAN propuesta se analizó a través de las pérdidas del generador y discriminador en la fase de entrenamiento. Para ello, fue necesario visualizar ambas pérdidas (Figura 4.1 y Figura 4.2, respectivamente). Analizar el rendimiento de la red en este sentido sirve para conocer en qué momento se puede detener el entrenamiento.

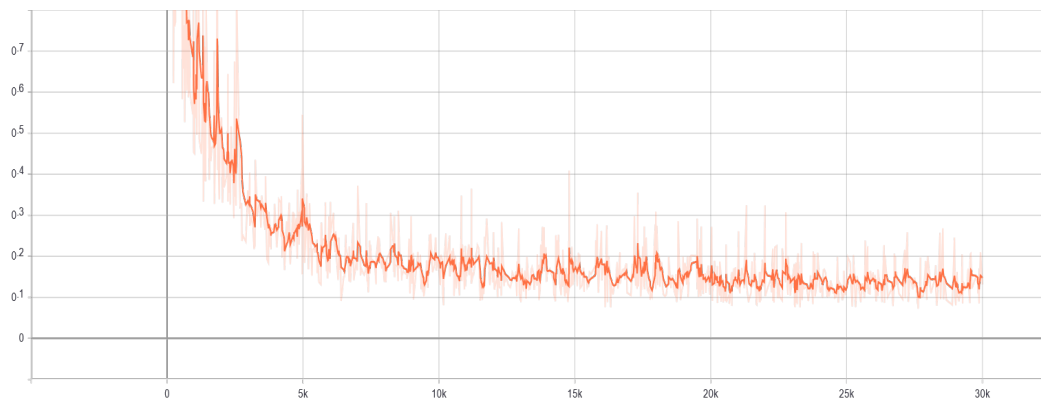


Figura 4.1 Pérdida del generador de la red que corrige la nitidez.

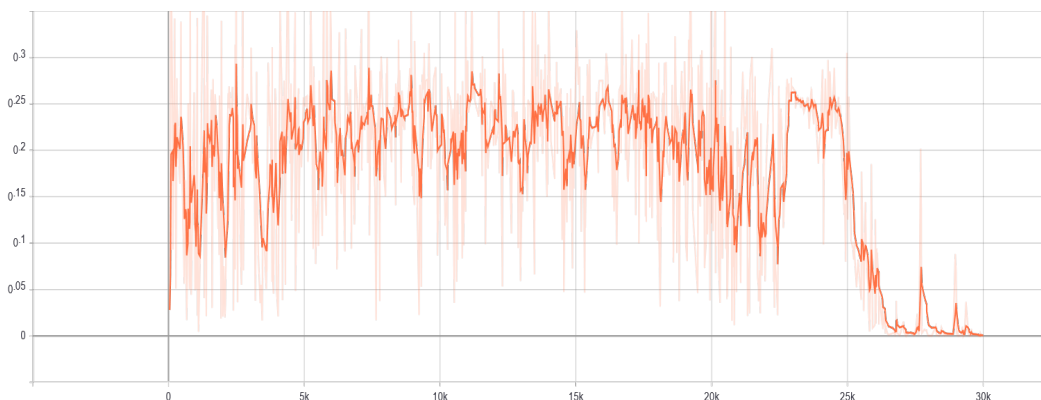


Figura 4.2 Pérdida del discriminador de la red que corrige la nitidez.

Se puede observar cómo la pérdida del generador disminuye rápidamente en las primeras 5000 épocas y se estabiliza en torno a un valor cercano a 0.15. Si nos centramos en la pérdida del discriminador vemos cómo inicialmente tiene un valor mayor, lo que significa que tiende a detectar imágenes generadas como reales. No es hasta épocas superiores a 25000 cuando esta pérdida disminuye. Teniendo esto en cuenta, utilizar un número de épocas cercano a 30000, como se ha escogido en este TFG, es suficiente para obtener unos buenos resultados.

Visualmente, se puede observar cómo, a medida que las épocas avanzan, las imágenes generadas son más nítidas. El valor de las épocas escogido en este TFG ha sido de 30000. La elección de esta cifra se debe a que múltiples autores en estudios previos han escogido valores cercanos a este número (Ahirwar, 2019). En la Figura 4.3 se puede observar la imagen generada junto con las imágenes degradada y original procedente del conjunto de entrenamiento después de 275 épocas de entrenamiento. En las Figuras 4.4 y 4.5 se muestran los mismos resultados obtenidos tras 5000 y 30000 épocas de entrenamiento respectivamente.

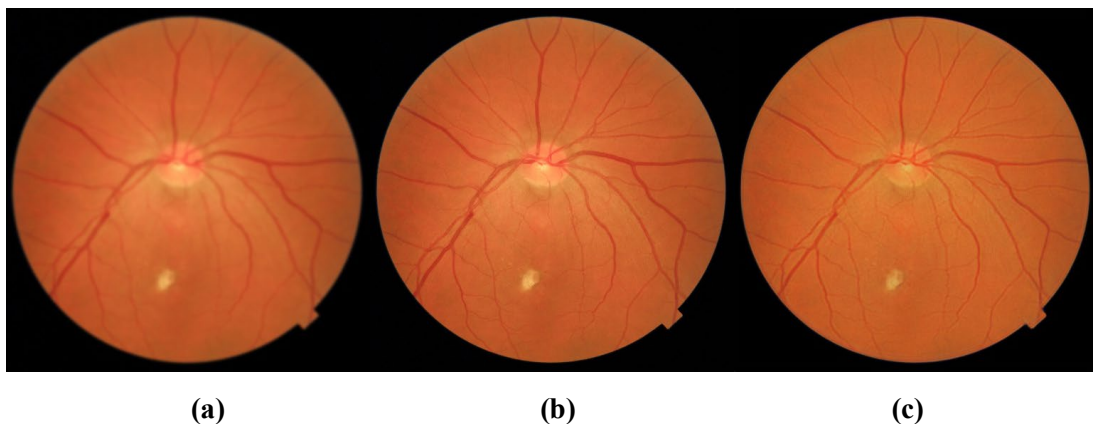
Si nos centramos en la Figura 4.3, podemos apreciar como el color de la retinografía no se ha generado correctamente y los vasos sanguíneos no están definidos completamente. Sin embargo, en la Figura 4.4 el color de la imagen generada es más similar al de la imagen original y los vasos sanguíneos aparecen bien definidos. Esta mayor definición se puede apreciar en los vasos sanguíneos cercanos a la papila.

Finalmente, en la Figura 4.5 vemos como en la última época de entrenamiento el color de la retinografía generada es igual al de la original. Además, en esta imagen se puede observar que la nitidez de los vasos sanguíneos es idéntica en la imagen original y en la generada.

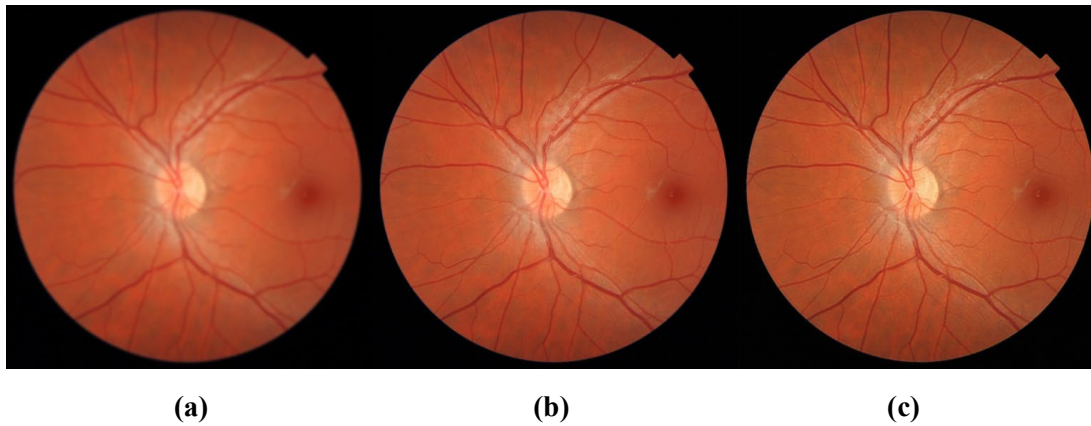
### 4.3.2 FASE DE TEST

#### 4.3.2.1 Resultados sobre las imágenes de test de buena calidad

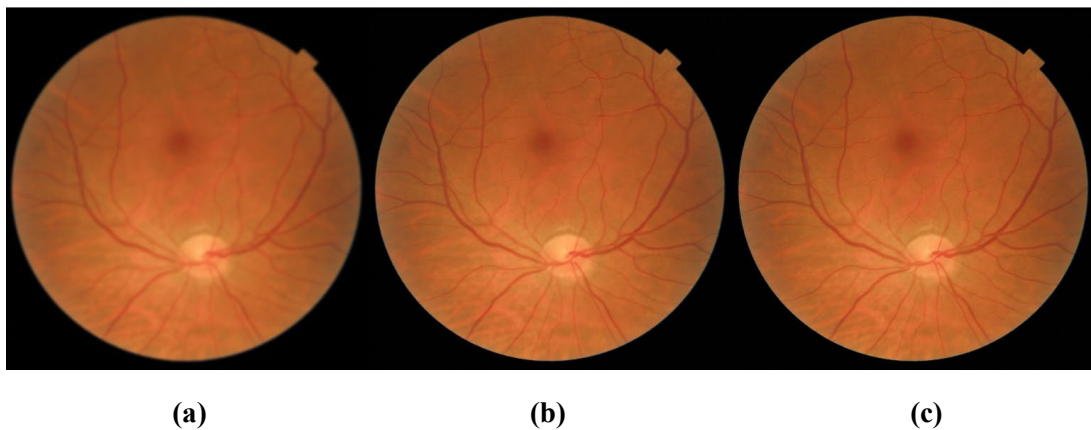
Los resultados de los parámetros explicados anteriormente en el conjunto de test de la BD de buena calidad se muestran en la Tablas 4.1 y 4.2. Estos resultados muestran un 3.33%, 3.26% y 6.22% de mejora en términos de entropía, SSIM y PSNR. Asimismo, se muestran varios ejemplos sobre este conjunto de imágenes en la Figura 4.6.



**Figura 4.3** Resultados después de 275 épocas de la red que corrige la nitidez (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.



**Figura 4.4** Resultados después de 5000 épocas de la red que corrige la nitidez (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.



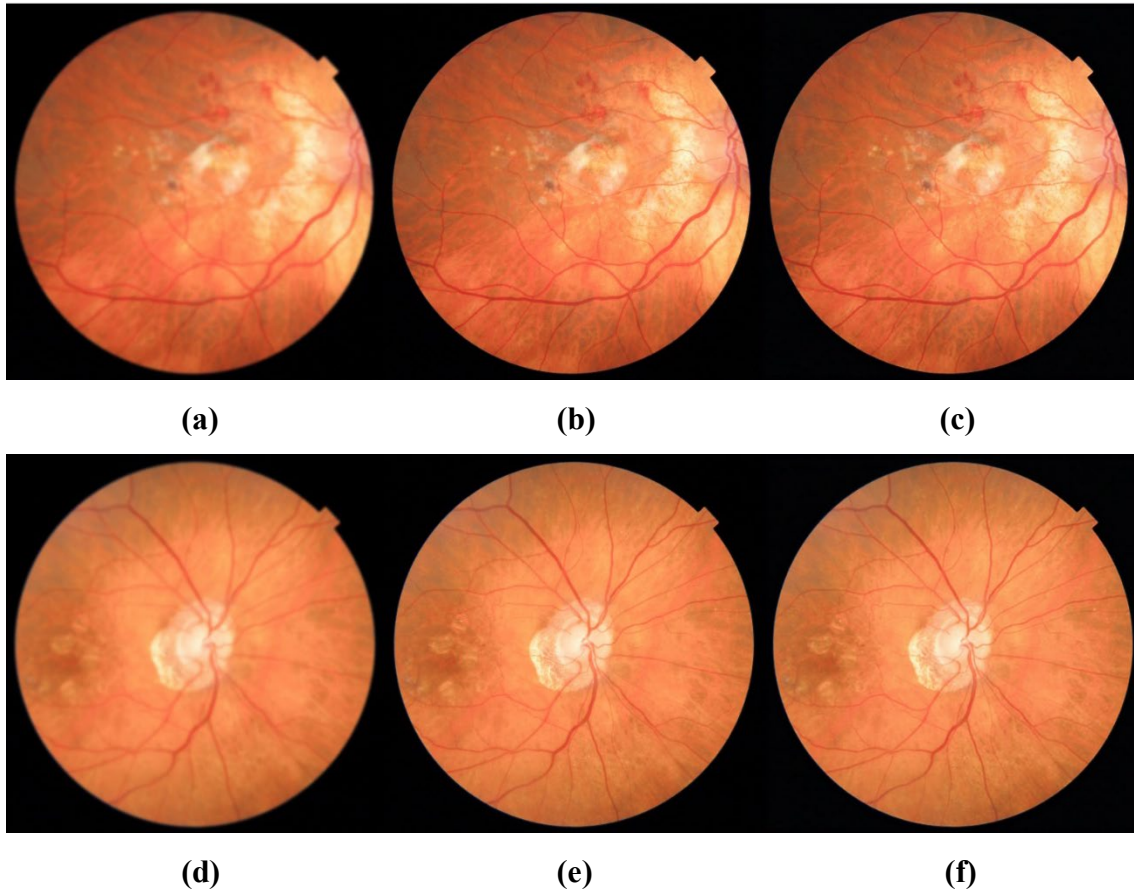
**Figura 4.5** Resultados después de 30000 épocas. de la red que corrige la nitidez (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen degradada	68.55	14.70	3.77
Imagen original	<b>13.45</b>	<b>11.83</b>	3.90
Imagen generada	25.28	12.74	<b>4.03</b>

**Tabla 4. 1** Resultados de la evaluación sin referencia sobre las imágenes del conjunto de test de buena calidad de la red que corrige la nitidez.

	EVALUACIÓN CON REFERENCIA	
	SSIM	PSNR
Imagen original - degradada	0.92	35.81
Imagen original - generada	<b>0.95</b>	<b>38.04</b>

**Tabla 4. 2** Resultados de la evaluación con referencia sobre las imágenes del conjunto de test de buena calidad de la red que corrige la nitidez.



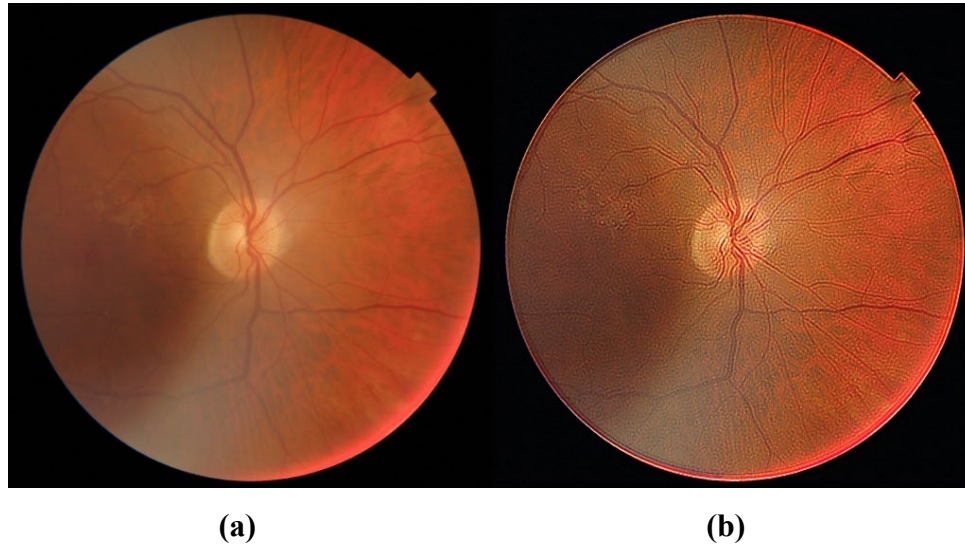
**Figura 4.6** Resultados de la red que corrige la nitidez sobre algunas imágenes de test de buena calidad. Columna de la izquierda: imagen degradada. Columna central: imagen original. Columna de la derecha: imagen generada.

#### 4.3.2.2 Resultados sobre las imágenes de test de mala calidad

Los resultados sobre el conjunto de test de la BD de mala calidad se muestran en la Tabla 4.3. Como se puede observar, aunque los resultados cuantitativos obtenidos en esta tabla muestran una mejora de la imagen generada respecto de la original, en la Figura 4.7 se puede observar que la red no corrige adecuadamente la borrosidad de la imagen original. En la Figura 4.7 (a) se aprecia un cierto grado de borrosidad en la zona central de la imagen, pero los vasos se aprecian correctamente en la mayor parte de la imagen. Sin embargo, tras la utilización del método para la mejora de su calidad, en la Figura 4.7 (b) se aprecia que la imagen generada contiene ruido granular.

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen original	31.03	<b>12.98</b>	3.64
Imagen generada	<b>25.51</b>	15.07	<b>4.15</b>

**Tabla 4.3** Resultados de la evaluación sin referencia sobre las imágenes del conjunto de test de mala calidad de la red que corrige la nitidez.

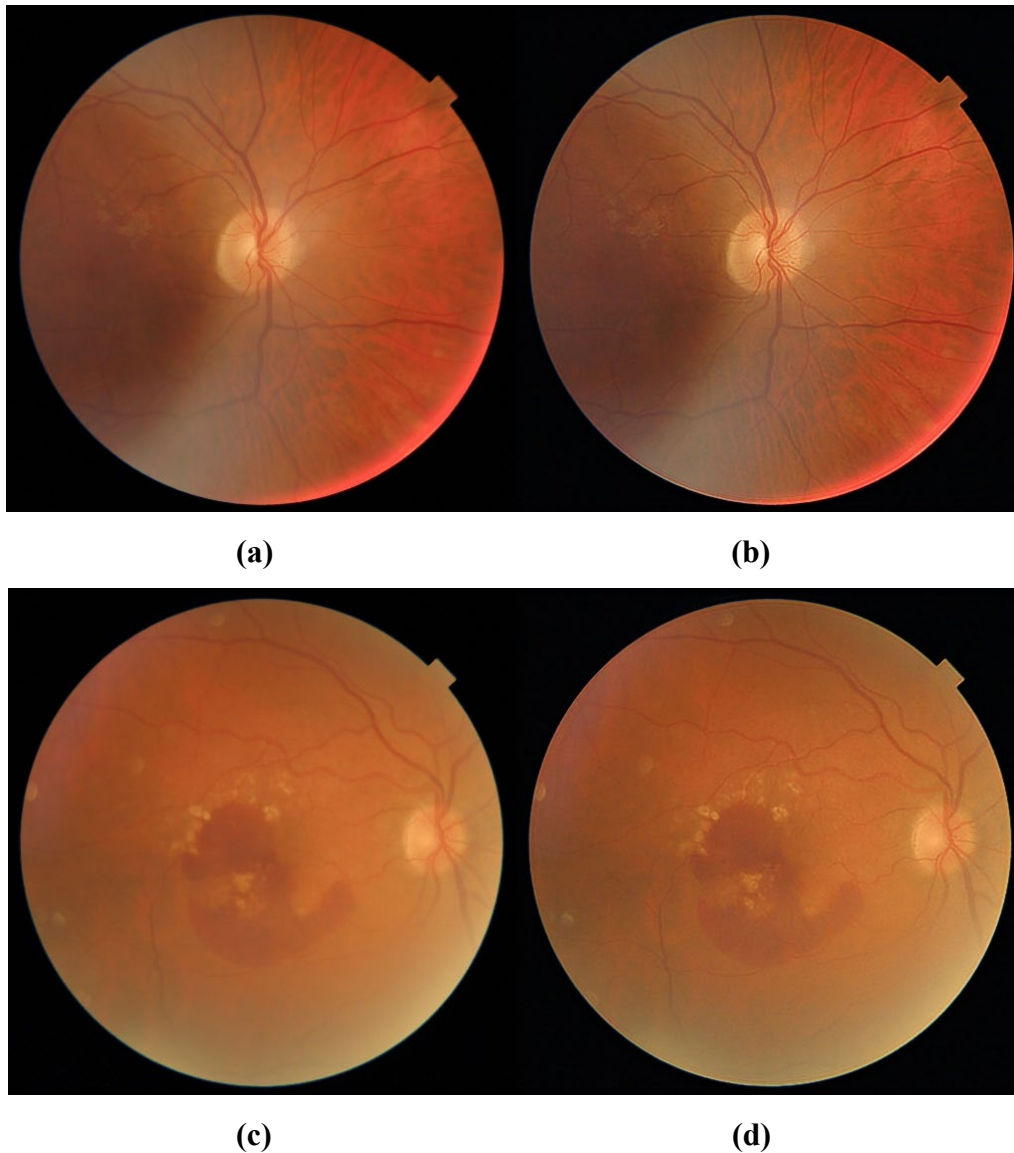


**Figura 4.7** Resultados visuales sobre las imágenes de test de mala calidad de la red que corrige la nitidez (a) Imagen original. (b) Imagen generada.

Para tratar de evaluar el funcionamiento de la red sobre este tipo de imágenes, se realizó una prueba adicional degradando este conjunto de imágenes con un filtro gaussiano de varianza pequeña ( $\sigma = 0.4$ ) obteniendo los resultados mostrados en la Tabla 4.4. Este valor de varianza se escogió tras observar los resultados de sucesivas pruebas con diferentes valores de este parámetro. De esta manera, las imágenes pierden nitidez y es posible percibir una mejora de la calidad tras la utilización del método propuesto. Como en el caso anterior, únicamente nos interesa la evaluación sin referencia, ya que comparamos la imagen original de mala calidad con la imagen generada de calidad mejorada. Sobre estas imágenes de mala calidad, los resultados muestran un 3.89%, 31.80% y 4.27% de mejora en términos de entropía, BRISQUE y NIQE respecto a la imagen original real. La entropía original es menor que el valor de entropía de la imagen generada, lo que indica la pérdida de detalles de la imagen y puede conducir a una detección falsa en sujetos con RD. El resultado obtenido sobre varias imágenes reales de ejemplo a las que se ha aplicado este procedimiento se puede observar en la Figura 4.8. En este caso, en la Figura 4.8 (b) se puede apreciar como la papila y los vasos sanguíneos en la parte izquierda de la imagen aparecen más definidos en comparación con la imagen original (Figura 4.8 (a)). Un caso similar se puede observar en la Figura 4.8 (c) y Figura 4.8 (d). Mientras que en la Figura 4.8 (c) se puede apreciar que algunas lesiones aparecen borrosas, en la Figura 4.8 (d) estas se muestran más nítidas. También se puede notar que tanto la papila como la mácula y los vasos sanguíneos aparecen más definidos en la Figura 4.8 (d).

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen original	31.03	12.98	3.64
Imagen generada	<b>21.16</b>	<b>12.43</b>	<b>3.78</b>

**Tabla 4. 4** Resultados de la evaluación sin referencia sobre las imágenes del conjunto de test de mala calidad de la red que corrige la nitidez añadiendo un filtro gaussiano con  $\sigma = 0.4$ .



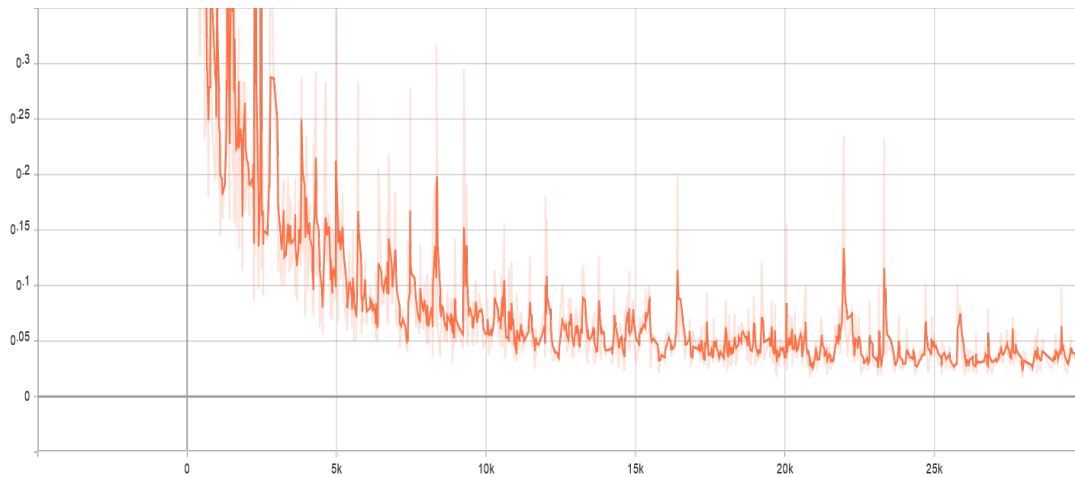
**Figura 4.8** Resultados visuales sobre las imágenes de test de mala calidad de la red que corrige la nitidez añadiendo un filtro gaussiano con varianza baja. Columna de la izquierda: imagen original. Columna de la derecha: imagen generada.

## 4.4 RESULTADOS PARA LA MEJORA DE CALIDAD DE IMÁGENES OSCURAS

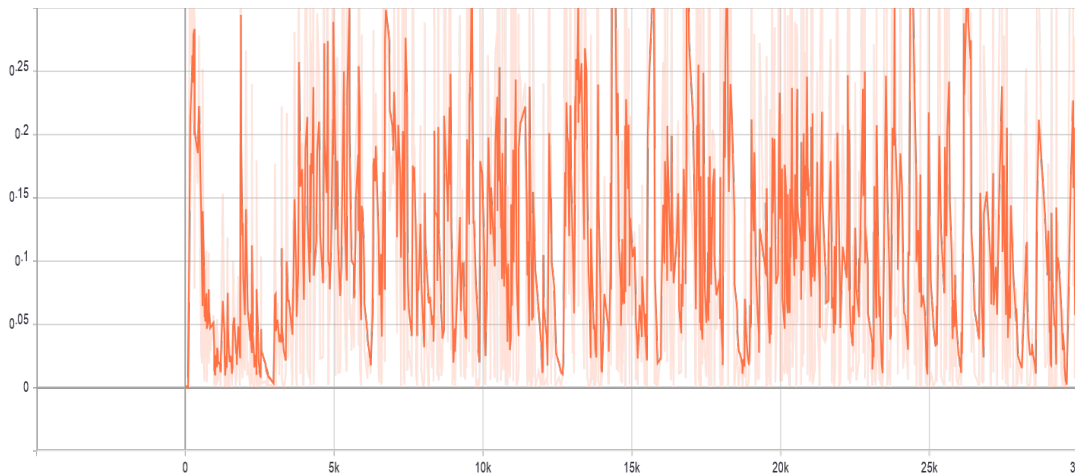
### 4.4.1 FASE DE ENTRENAMIENTO. RENDIMIENTO DE LA GAN

Para analizar el rendimiento de la GAN propuesta se muestran las pérdidas del generador y del discriminador en la fase de entrenamiento (Figura 4.9 y Figura 4.10 respectivamente). Analizar el rendimiento de la red en este sentido ha permitido determinar en qué momento podemos dejar de entrenar a la misma. Se puede observar (Figura 4.9) como la pérdida del generador disminuye rápidamente en las primeras 5000 épocas y se estabiliza en torno a un valor cercano a 0.05. Si nos centramos en la pérdida del discriminador (Figura 4.10), se puede ver como los valores oscilan y la red no llega a estabilizarse, lo que significa que tiende a detectar imágenes generadas como reales. Por este motivo, la red no responderá adecuadamente en todos los casos.





**Figura 4.9** Pérdida del generador de la red que corrige la iluminación.



**Figura 4.10** Pérdida del discriminador de la red que corrige la iluminación.

Visualmente, se puede observar cómo a medida que las épocas avanzan, las imágenes generadas son más similares a la original. En la Figura 4.11 se puede observar una imagen generada de ejemplo junto con las imágenes degradada y original después de 275 épocas de entrenamiento. En las Figuras 4.12 y 4.13 se muestran dos ejemplos adicionales obtenidos tras 5000 y 30000 épocas de entrenamiento, respectivamente.

Mientras que en la Figura 4.11 se aprecia un color diferente para las imágenes original (Figura 4.11 (b)) y generada (Figura 4.11 (c)), en las Figuras 4.12 y 4.13 los colores generados son muy similares al de la retinografía original. Es posible notar también que en la Figura 4.11 (c), la papila es de un tono similar al del fondo del ojo. En la última época del método este problema se soluciona como se puede observar en la Figura 4.13 (c), donde es posible distinguir el tono de la papila del tono del fondo del ojo.



(a)

(b)

(c)

**Figura 4.11** Resultados después de 275 épocas de la red que corrige la iluminación (a) Imagen degradada. (b) Imagen original. (c) Imagen generada



(a)

(b)

(c)

**Figura 4.12** Resultados después de 5000 épocas de la red que corrige la iluminación (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.



(a)

(b)

(c)

**Figura 4.13** Resultados después de 30000 épocas de la red que corrige la iluminación (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.

#### 4.4.2 FASE DE TEST

##### 4.4.2.1 Resultados sobre las imágenes de test de buena calidad

Los resultados de los parámetros calculados sobre el conjunto de test de la BD de buena calidad se muestran en la Tablas 4.5 y 4.6. Sobre este conjunto de imágenes de buena calidad, los resultados muestran un 3.12%, 2.28%, 14.59% y 156.81% de mejora en términos de BRISQUE, NIQE, SSIM y PSNR. Asimismo, varios resultados visuales se muestran en la Figura 4.14.

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen degradada	33.68	16.52	3.55
Imagen original	13.45	11.83	<b>3.90</b>
Imagen generada	<b>13.03</b>	<b>11.56</b>	3.89

**Tabla 4. 5** Resultados de la evaluación sin referencia sobre el conjunto de imágenes de test de buena calidad de la red que corrige la iluminación.

	EVALUACIÓN CON REFERENCIA	
	SSIM	PSNR
Imagen original - degradada	0.84	14.62
Imagen original - generada	<b>0.97</b>	<b>37.56</b>

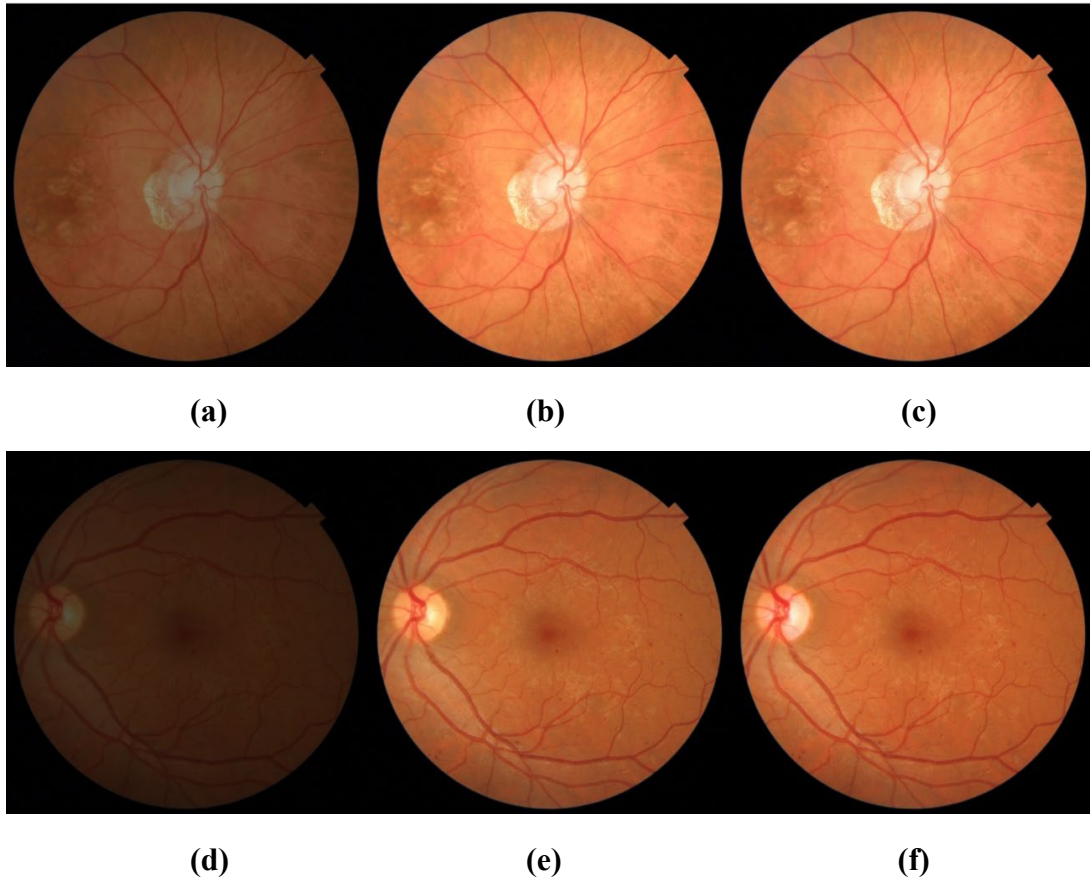
**Tabla 4. 6** Resultados de la evaluación con referencia sobre el conjunto de imágenes de test de buena calidad de la red que corrige la iluminación.

En el caso de las imágenes de test de buena calidad, los resultados visuales muestran una reconstrucción satisfactoria. Se puede observar que en el caso de la Figura 4.14 (a), el porcentaje de oscuridad que se ha añadido es leve. La corrección de la iluminación de esta imagen da lugar a la Figura 4.14 (c), en la cual tanto el tono del fondo de ojo como de la papila y los vasos sanguíneos son idénticos a los de la imagen original (Figura 4.14 (b)). En la Figura 4.14 (d) se muestra otro ejemplo de corrección del color de una retinografía a la que se ha aplicado un porcentaje mayor de oscuridad. Del mismo modo se aprecia que el color del fondo de ojo, la papila, la mácula y los vasos sanguíneos son similares en la imagen generada por el método (Figura 4.14 (f)) y la imagen original (Figura 4.14 (e)).

##### 4.4.2.2 Resultados sobre las imágenes de test de mala calidad

Los resultados sobre el conjunto de test de la BD de mala calidad se muestran en la Tabla 4.7. Para obtener los resultados de este apartado, se han escogido únicamente las imágenes que presentaban algún tipo de oscuridad.

Aunque los resultados obtenidos muestran una mejora de la imagen generada respecto de la original, en algunos casos (como en el ejemplo de la Figura 4.15), la red no corrige completamente las zonas más claras de la imagen. En la parte inferior izquierda de la imagen mostrada en la Figura 4.15 (b) es posible observar una gran saturación del color que impide la visión completa de los vasos sanguíneos de esa zona. Por este motivo,



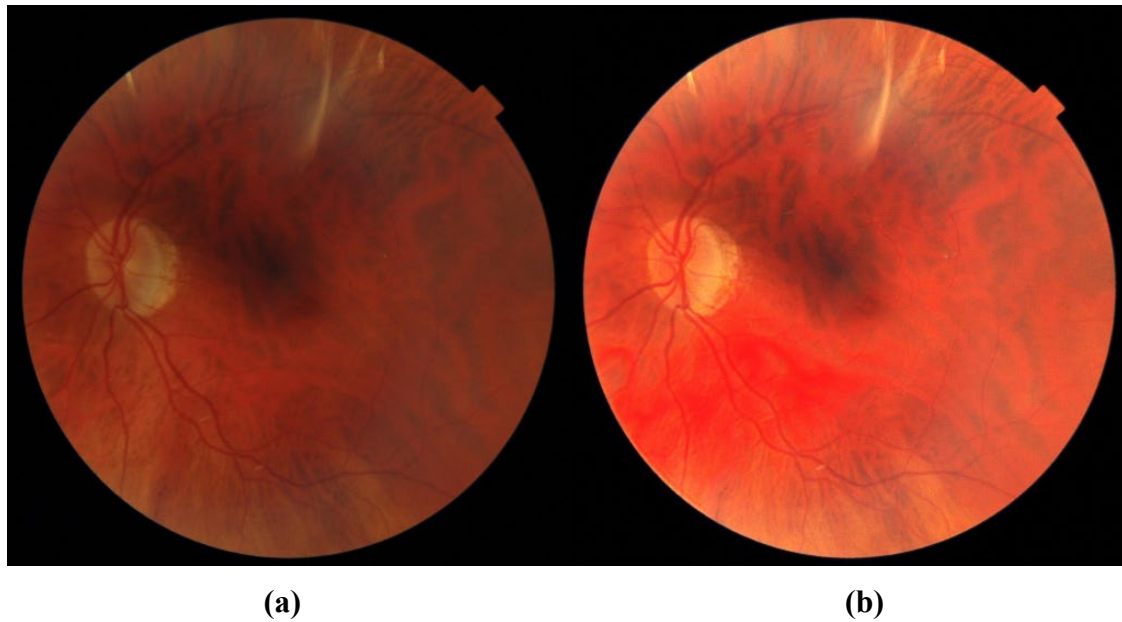
**Figura 4.14** Resultados sobre imágenes del conjunto de test de buena calidad de la red que corrige la iluminación. Columna de la izquierda: imagen degradada. Columna central: imagen original. Columna de la derecha: imagen generada

para tratar de evaluar el funcionamiento de la red sobre este tipo de imágenes, se realizó una prueba adicional oscureciendo levemente este tipo de imágenes. En este caso, se oscurece la imagen aplicando la máscara negra indicando un peso de 0.7 para la imagen original y un peso de 1 para la máscara. Se puede observar que, al añadir un porcentaje de oscuridad leve a la imagen original, los resultados obtenidos mejoran sensiblemente (ver Tabla 4.8).

Como en el caso anterior, únicamente nos interesa la evaluación sin referencia, ya que comparamos la imagen original de mala calidad con la imagen generada de calidad mejorada. Sobre estas imágenes de mala calidad, los resultados muestran un 50.62% y un 8.33% de mejora en términos de BRISQUE y entropía. En la Figura 4.16 se muestra el resultado sobre tres imágenes de este conjunto.

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen original	30.76	<b>11.88</b>	3.60
Imagen generada	<b>18.82</b>	16.04	<b>4.16</b>

**Tabla 4. 7** Resultados de la evaluación sin referencia sobre las imágenes del conjunto de test de mala calidad de la red que corrige la iluminación.



**Figura 4.15** Resultados visuales sobre las imágenes de test de mala calidad de la red que corrige la iluminación (a) Imagen original. (b) Imagen generada.

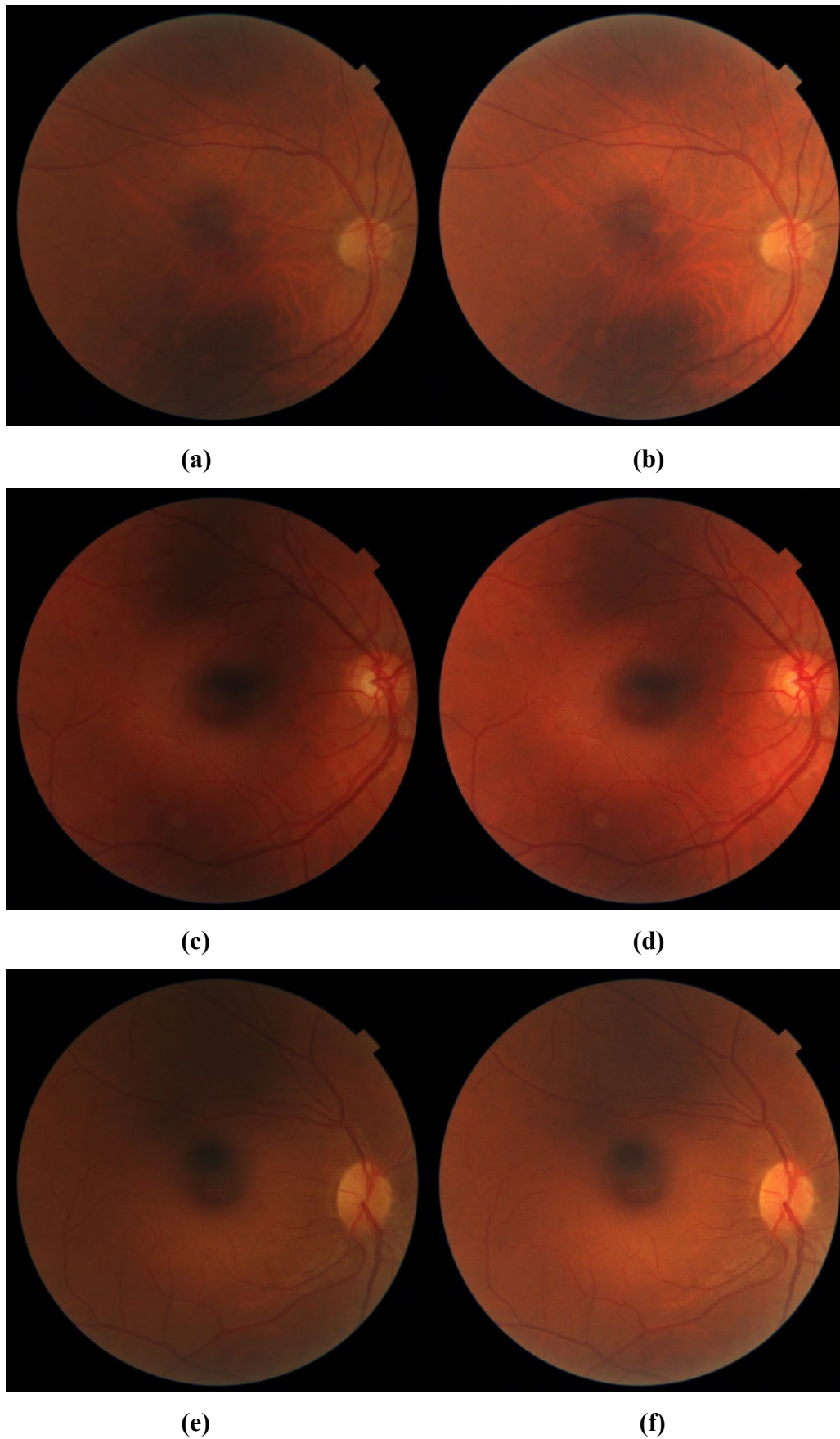
	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen original	30.76	<b>11.88</b>	3.60
Imagen generada	<b>15.19</b>	12.36	<b>3.90</b>

**Tabla 4. 8** Resultados de la evaluación sin referencia sobre las imágenes del conjunto de test de mala calidad añadiendo un porcentaje leve de oscuridad.

En la Figura 4.16 (a) es posible ver ciertas zonas oscurecidas en la parte superior e inferior de la imagen que impiden la correcta visión de algunos vasos sanguíneos. No obstante, tras su corrección (Figura 4.16 (b)) estas zonas aparecen con un tono más uniforme lo que permite una mejor apreciación de los vasos sanguíneos de la zona. Un ejemplo similar es el mostrado en la Figura 4.16 (c) donde también se aprecia una zona oscura en la parte superior de la imagen, y en cuya corrección (Figura 4.16 (d)) se puede observar la mácula más definida. Finalmente, en la Figura 4.16 (e) se muestra el caso de una retinografía con una zona demasiado oscura en la parte superior de la imagen y en la mácula. De manera análoga a los casos anteriores, tras la corrección de la iluminación de esta imagen (que se muestra en la Figura 4.16 (f)), los vasos sanguíneos de la zona superior de la imagen se pueden identificar más fácilmente.

#### 4.5 RESULTADOS PARA AMBAS REDES EN SERIE

Para comprobar la efectividad del método propuesto se han realizado pruebas utilizando ambas redes. En primer lugar, las imágenes se introducen en la red que corrige la iluminación de la imagen y posteriormente se corrige la nitidez con la segunda red.



**Figura 4.16** Resultados visuales sobre las imágenes de test de mala calidad de la red que corrige la iluminación añadiendo un porcentaje leve de oscuridad. Columna de la izquierda: imagen original. Columna de la derecha: imagen generada.

#### 4.5.1 RESULTADOS SOBRE LAS IMÁGENES DE TEST DE BUENA CALIDAD

Los resultados obtenidos a la salida de la red que corrige la oscuridad de las imágenes del conjunto de test de buena calidad son los mostrados en el apartado 4.4.2.2. A continuación, esas imágenes se han introducido en la red que corrige la borrosidad de las imágenes obteniendo los resultados recogidos en las Tablas 4.9 y 4.10. Los resultados obtenidos muestran un 3.38%, 2.57%, 5.68% y 4.84% de mejora respecto de la imagen original en términos de NIQE, entropía, SSIM y PSNR. En las Figuras 4.17 y 4.18 se muestra el resultado para una imagen de ejemplo que ha sido corregida por ambas redes.

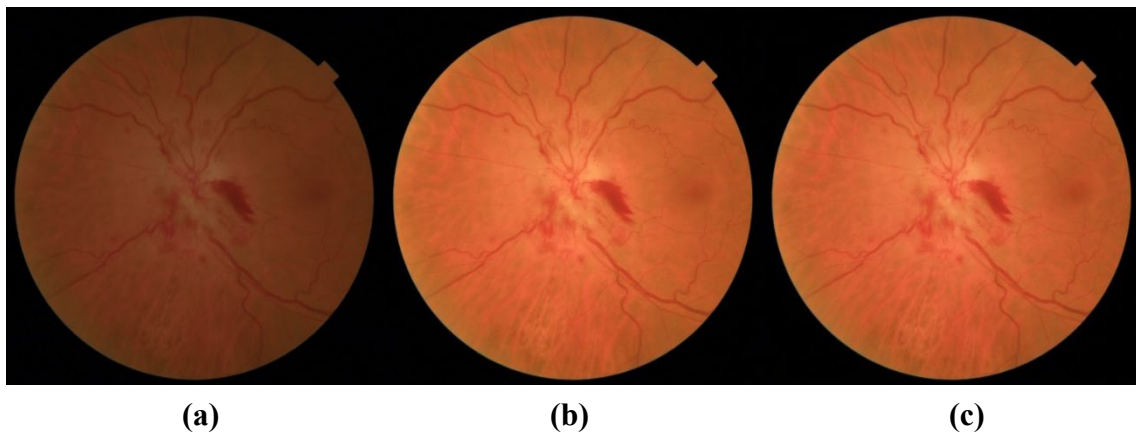
En la Figura 4.17(a) se observa la imagen inicial que va a atravesar las dos redes. Tras la corrección de la iluminación de la imagen se obtiene la imagen mostrada en la Figura 4.17 (c). Tras esto, se ha degradado la nitidez de la imagen obteniendo la Figura 4.18 (a) y se ha abordado la corrección de la borrosidad a través de la segunda red. La Figura 4.18(c) se ha obtenido tras atravesar las dos redes. En comparación con la imagen original real (Figura 4.18(b)) la imagen generada presenta una calidad similar.

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen degradada	68.71	14.84	3.81
Imagen original	<b>13.45</b>	11.83	3.90
Imagen generada	26.86	<b>11.43</b>	<b>4.01</b>

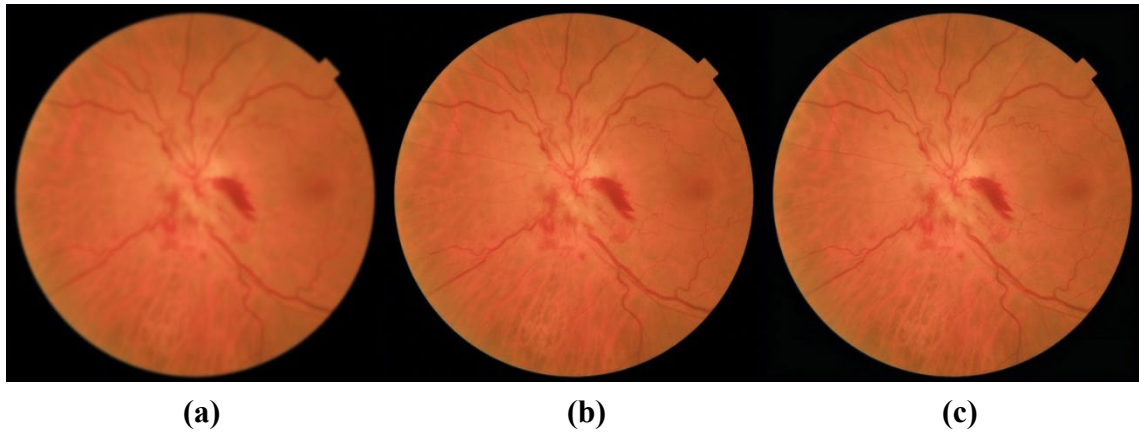
**Tabla 4. 9** Resultados de la evaluación sin referencia sobre las imágenes de test de buena calidad para ambas redes en serie.

	EVALUACIÓN CON REFERENCIA	
	SSIM	PSNR
Imagen original - degradada	0.88	36.35
Imagen original - generada	<b>0.93</b>	<b>38.11</b>

**Tabla 4. 10** Resultados de la evaluación con referencia sobre las imágenes de test de buena calidad para ambas redes en serie.



**Figura 4.17** Resultado de una imagen del conjunto de test de buena calidad tras la corrección de su iluminación. (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.



**Figura 4.18** Resultado de una imagen del conjunto de test de buena calidad tras la corrección de su nitidez. (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.

#### 4.5.2 RESULTADOS SOBRE LAS IMÁGENES DE TEST DE MALA CALIDAD

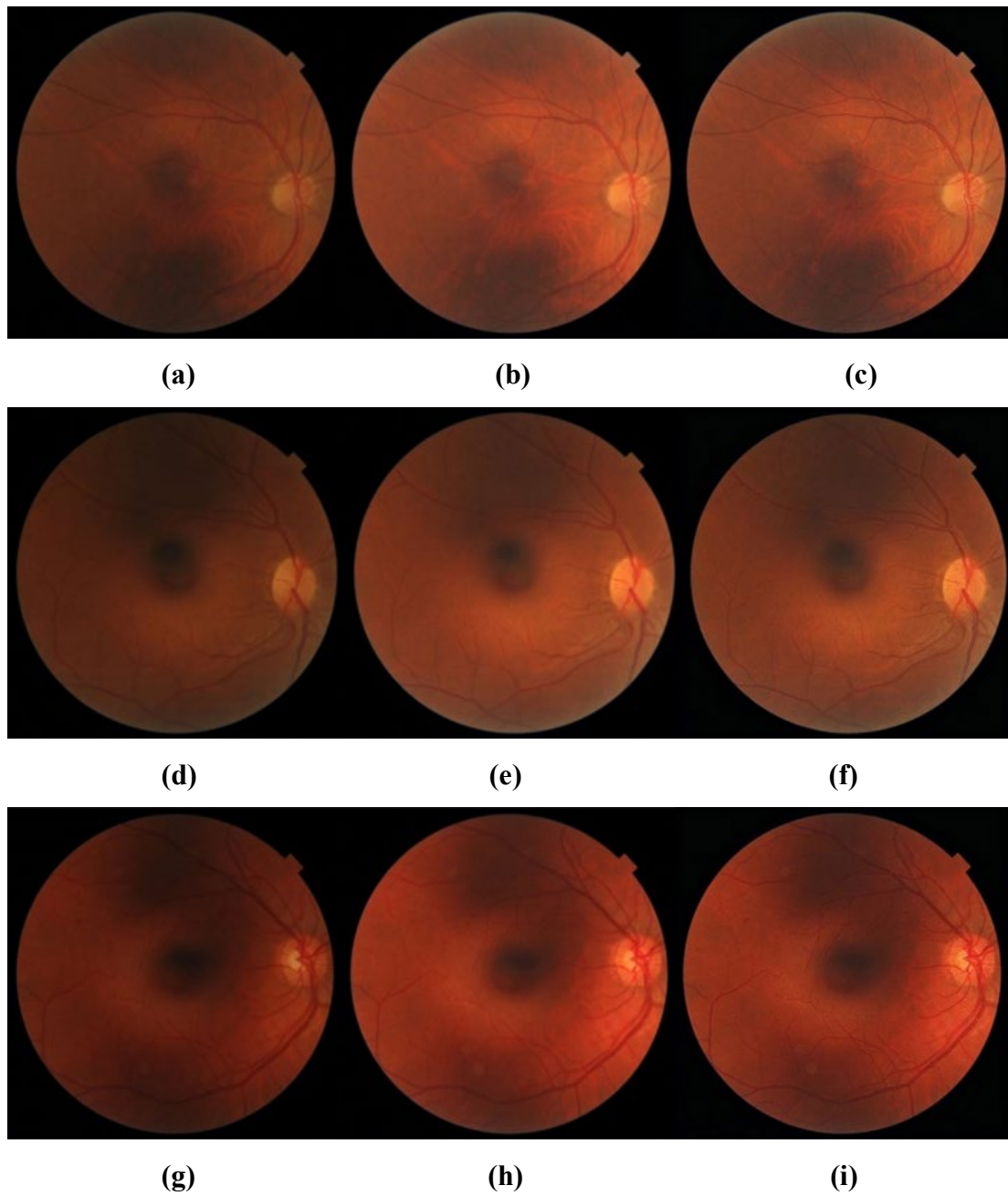
Los resultados obtenidos a la salida de la red que corrige la oscuridad de las imágenes son los mostrados en el apartado 4.4.2.3. A continuación, esas imágenes se han introducido en la red que corrige la borrosidad de las imágenes obteniendo los resultados recogidos en la Tabla 4.11. Los resultados obtenidos muestran un 88.95%, 21.17% y 2.46% de mejora en términos de BRISQUE, NIQE y entropía.

	EVALUACIÓN SIN REFERENCIA		
	BRISQUE	NIQE	ENTROPÍA
Imagen original	23.61	14.07	4.07
Imagen generada	<b>3.02</b>	<b>11.09</b>	<b>4.17</b>

**Tabla 4. 11** Resultados de la evaluación sin referencia sobre las imágenes de test de mala calidad para ambas redes en serie.

En la Figura 4.19 se muestran los resultados obtenidos tras pasar distintas imágenes de ejemplo por ambas redes. La figura 4.19 (a) muestra una imagen con falta de iluminación en la parte inferior. Esto provoca que los vasos sanguíneos de esa zona no se aprecien con claridad. Tras la corrección de la iluminación (Figura 4.19 (b)) esta zona aparece más clara permitiendo una mayor distinción de los vasos sanguíneos. A continuación, tras la corrección de la nitidez de la imagen (Figura 4.19 (c)), es posible percibir mejor los vasos sanguíneos tanto en la zona inferior de la imagen como en la papila. Un segundo ejemplo similar se muestra en la Figura 4.19 (d), la cual presenta una zona oscurecida en la parte superior de la imagen. En la Figura 4.19 (e) se muestra la imagen tras corregir la oscuridad de la misma. En esta figura se observan de manera más clara los vasos sanguíneos de la parte superior. Un último ejemplo se puede observar en la Figura 4.19 (g) donde aparece una imagen con una zona oscura en la parte superior y otra en la parte inferior. Se puede observar que tras la corrección de la iluminación y de la nitidez de la imagen tanto la mácula como los vasos sanguíneos aparecen mejor definidos. Asimismo, las zonas superior e inferior presentan un tono más uniforme.





**Figura 4.19** Resultados visuales sobre imágenes del conjunto de test de mala calidad tras pasar por ambas redes. Columna de la izquierda: imagen original. Columna central: imagen tras la corrección de iluminación. Columna de la derecha: imagen tras la corrección de nitidez.



# CAPÍTULO 5

---

## DISCUSIÓN

5.1	INTRODUCCIÓN .....	81
5.2	CORRECCIÓN DE LA NITIDEZ DE IMÁGENES DE RETINA .....	81
5.3	CORRECCIÓN DE LA ILUMINACIÓN DE IMÁGENES DE RETINA.....	84
5.4	COMPARACIÓN CON ESTUDIOS PREVIOS .....	86



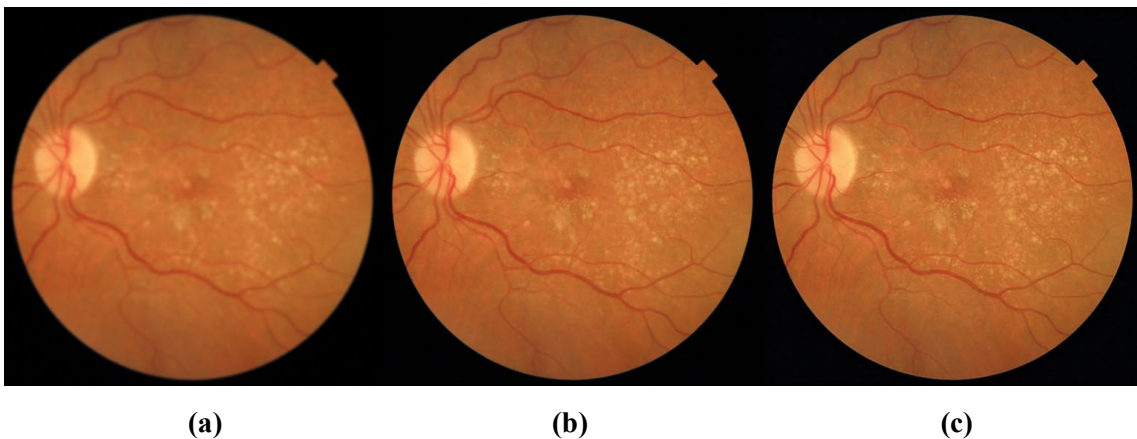
## 5.1 INTRODUCCIÓN

En la oftalmología moderna, se ha observado que la aplicación de la fotografía del fondo de ojo es fundamental en la detección temprana de enfermedades como el glaucoma, las cataratas y la RD. Estas imágenes de fondo a menudo sufren de falta de calidad debido a factores como ruido, bajo contraste e iluminación no uniforme. Por ello, en este TFG se ha desarrollado y evaluado un sistema de mejora de calidad de retinografías para que puedan ser posteriormente evaluadas por un experto o un sistema automático. El método ha sido desarrollado y evaluado utilizando una colección de imágenes facilitadas por el IOBA de la Universidad de Valladolid y el Hospital Clínico Universitario de Valladolid.

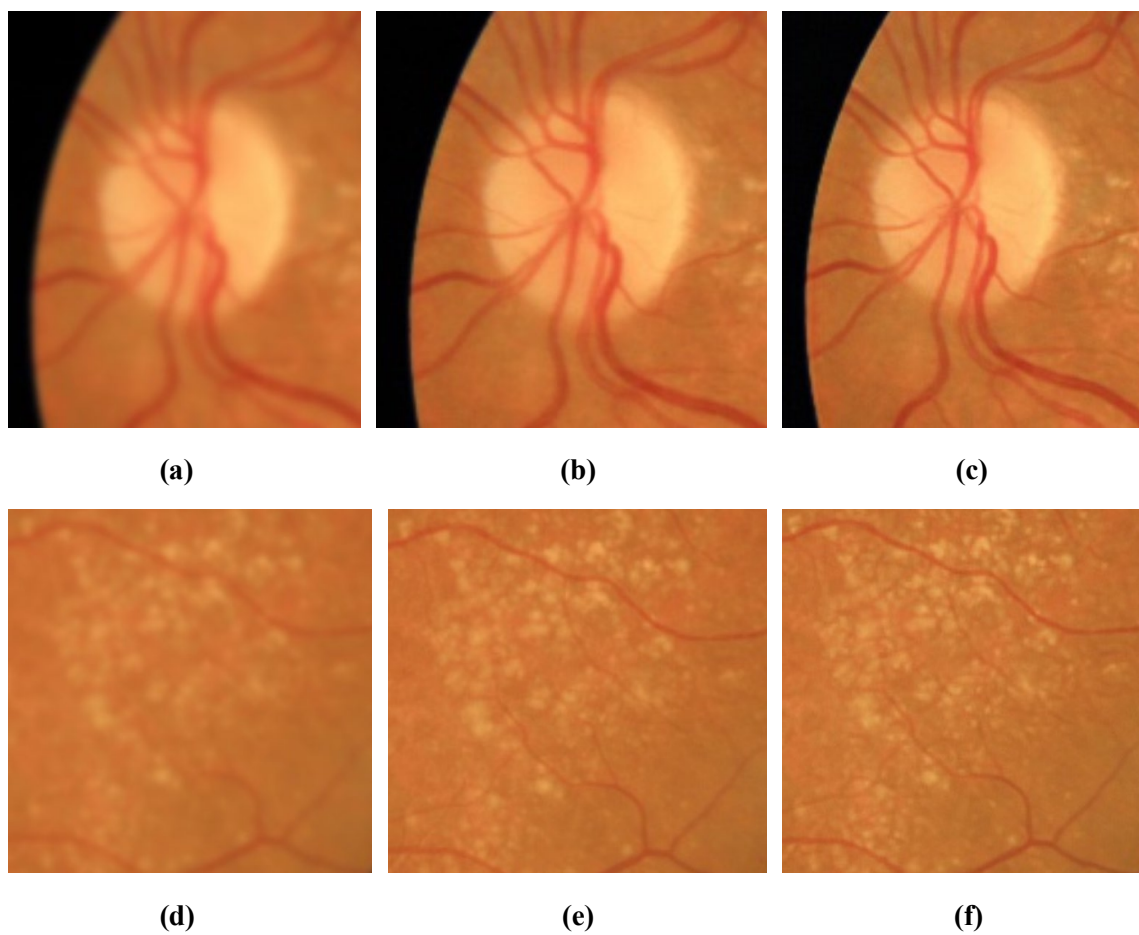
En este capítulo se realiza una valoración crítica de los resultados obtenidos con el método desarrollado. Se van a discutir los resultados obtenidos a partir de la BD creada específicamente para este trabajo. Para ello, se han analizado por separado los resultados del método de corrección de iluminación y del método de corrección de nitidez. Finalmente, los resultados se contrastan con los obtenidos en estudios previos y se comprueba la utilidad del estudio.

## 5.2 CORRECCIÓN DE LA NITIDEZ DE IMÁGENES DE RETINA

El método propuesto se basa en una red de tipo GAN que tiene como objetivo proporcionar imágenes con nitidez corregida. Se han degradado las imágenes de buena calidad tanto del conjunto de entrenamiento como del conjunto de test mediante la aplicación de un filtro gaussiano y se ha abordado la mejora de dichas imágenes borrosas y la comparación con las imágenes originales sin degradar. En la Figura 5.1 se muestra un ejemplo de la reconstrucción de imágenes de retina procedentes de las imágenes buena calidad del conjunto de test. En la imagen generada es posible observar una mayor definición tanto de los exudados como de los vasos sanguíneos en comparación con la imagen original. Esta mayor definición se puede apreciar mejor en la Figura 5.2, donde se han ampliado ciertas zonas de la imagen en las que se aprecia una mejora de la nitidez.



**Figura 5. 1** Comparación de la calidad de las imágenes sobre una imagen del conjunto de test de buena calidad. (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.

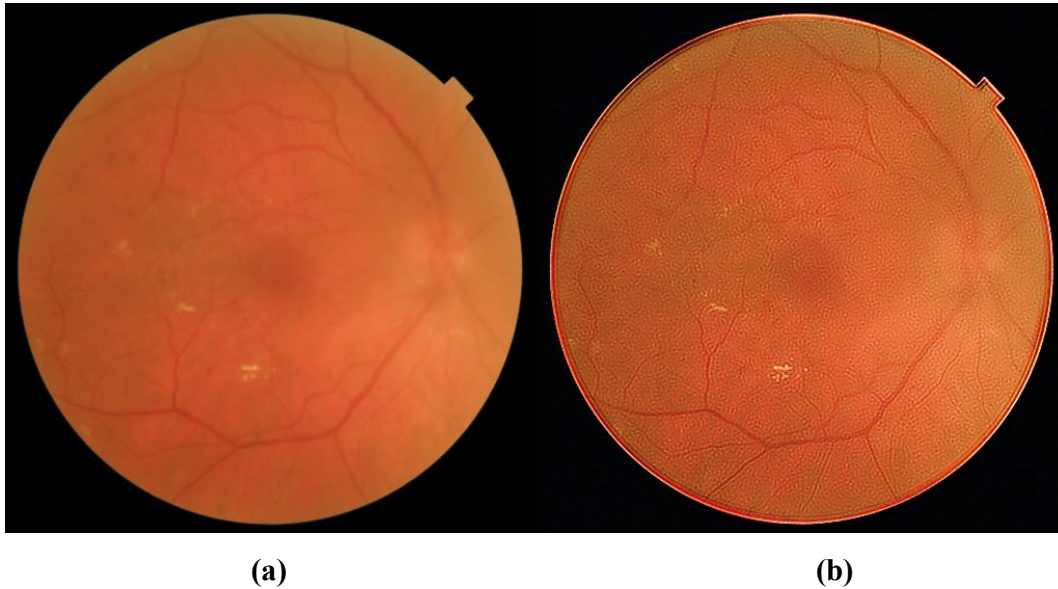


**Figura 5. 2** Comparación con zoom de la nitidez de una retinografía perteneciente al conjunto de test de buena calidad. (a) Imagen degradada con zoom en la papila. (b) Imagen original con zoom en la papila. (c) Imagen generada con zoom en la papila. (d) Imagen degradada con zoom en los exudados. (e) Imagen original con zoom en los exudados. (f) Imagen generada con zoom en los exudados.

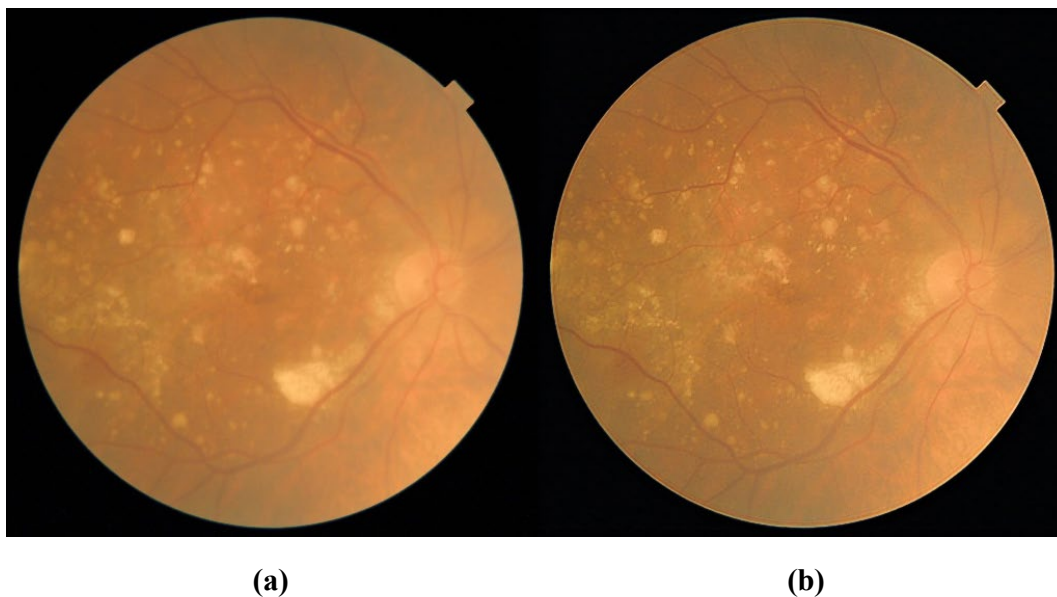
En una segunda parte se abordó la mejora de imágenes que, de partida, tenían una calidad insuficiente debido a que presentaban algún grado de borrosidad. Estas imágenes están contenidas en el conjunto de test de mala calidad. Esta etapa es necesaria para comprobar la efectividad del método sobre imágenes reales extraídas de la práctica clínica. Si estas imágenes presentan un grado de borrosidad similar al utilizado en el entrenamiento el método funciona correctamente. Sin embargo, la red no responde adecuadamente sobre imágenes que ya presentan una nitidez adecuada o una borrosidad leve, como se muestra en la Figura 5.3. La imagen original (Figura 5.3 (a)) presenta la zona de la papila difuminada, mientras que en el resto de la imagen se pueden apreciar los vasos sanguíneos. Tras la utilización del método (Figura 5.3 (b)) los vasos sanguíneos cercanos a la papila se logran distinguir mejor. Sin embargo, en esta imagen se observa ruido granular sobre todo en aquellas zonas que aparecían mejor definidas en la imagen original.

En este contexto, cabe destacar que, en las imágenes de mala calidad, muchas de ellas ya son nítidas en gran parte de la imagen o no responden a una degradación puramente gaussiana. En la fase de entrenamiento únicamente se ha entrenado con imágenes a las que se ha aplicado un filtro gaussiano para degradar las mismas. Por ello, la red sabe cómo corregir este tipo de imágenes, pero no sabe cómo procesar aquellas que

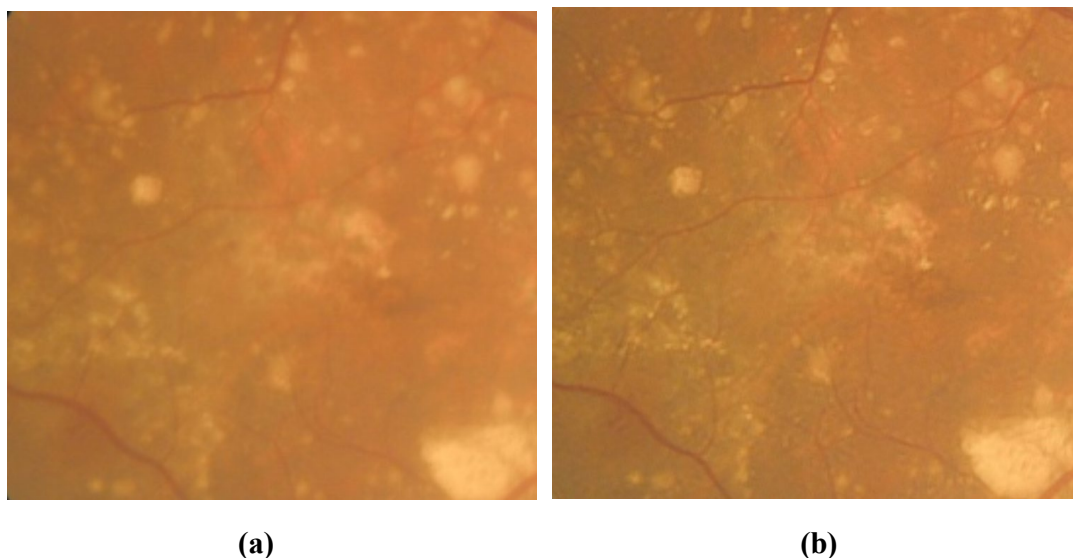
ya tienen una nitidez correcta o que no responden a una degradación provocada por un filtro gaussiano. Para tratar de evaluar el funcionamiento de la red sobre este tipo de imágenes, se realizó una prueba adicional degradando este conjunto de imágenes con un filtro gaussiano de varianza pequeña ( $\sigma = 0.4$ ) obteniendo los resultados mostrados en la Figura 5.4. En la Figura 5.5 se puede observar cómo la calidad de la imagen generada es mayor que la de la imagen original, pudiéndose apreciar mejor tanto los vasos sanguíneos como los exudados duros, lo que puede permitir que los sistemas de detección de RD automáticos existentes sean más efectivos en el diagnóstico de la enfermedad.



**Figura 5. 3** Efecto de la corrección de la nitidez de una imagen de mala calidad del conjunto de test. (a) Imagen original. (b) Imagen generada.



**Figura 5. 4** Efecto de la corrección de la nitidez en una imagen de mala calidad del conjunto de test a la que se ha aplicado un filtro gaussiano de varianza baja. (a) Imagen original. (b) Imagen generada.



**Figura 5. 5** Detalle del efecto de la corrección de la nitidez en una imagen de mala calidad del conjunto de test. (a) Imagen original. (b) Imagen generada.

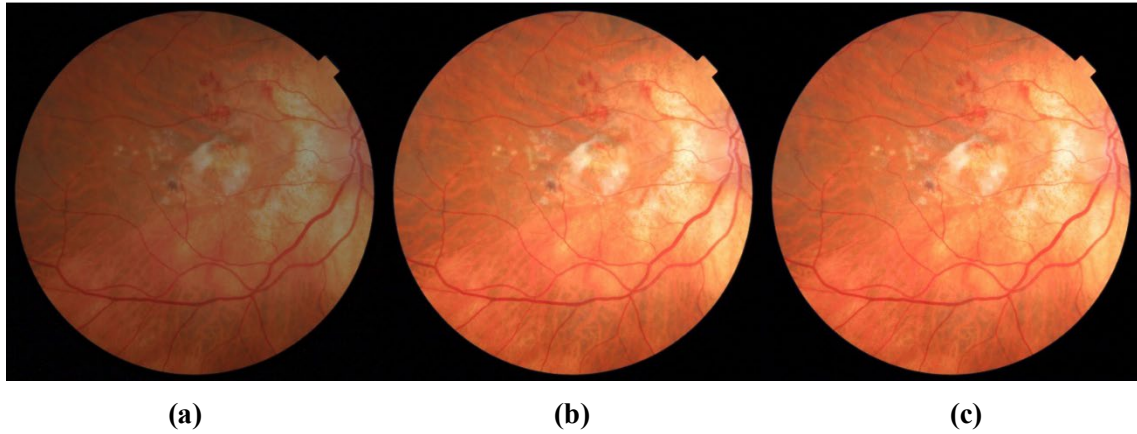
Sobre el conjunto de imágenes de test de mala calidad, los resultados muestran un 3.89%, 31.80% y 4.27% de mejora en términos de entropía, BRISQUE y NIQE respecto a la imagen original real, lo que indica que la pérdida de detalles de la imagen es menor y puede mejorar la detección de la RD en sistemas automáticos.

### 5.3 CORRECCIÓN DE LA ILUMINACIÓN DE IMÁGENES DE RETINA

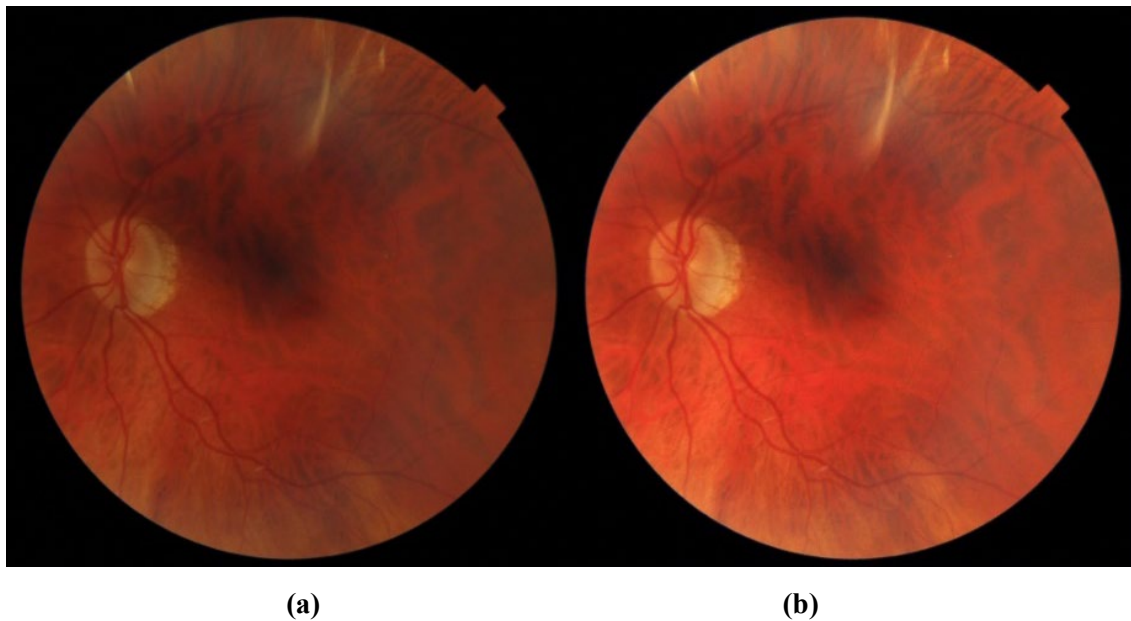
Para la corrección de la iluminación en retinografías, en este TFG se empleó también una red de tipo GAN. En el conjunto de test, se han degradado las imágenes de buena calidad mediante la aplicación de una máscara negra a la imagen y se ha abordado la mejora de dichas imágenes oscuras y la comparación con las imágenes originales sin degradar. En la Figura 5.6 se muestra un ejemplo de la reconstrucción de imágenes de retina procedentes de las imágenes del conjunto de test de buena calidad. Se puede observar como la imagen degradada (Figura 5.6 (a)) se encuentra ligeramente oscurecida. Si nos centramos en las imágenes original (Figura 5.6 (b)) y generada por el método (Figura 5.6 (c)) se puede distinguir en la parte superior izquierda una mejor visión de los vasos sanguíneos en comparación con la imagen oscurecida.

Una vez evaluado el método sobre las imágenes de test que presentaban una calidad suficiente, es necesario comprobar el funcionamiento sobre el conjunto de imágenes de test de mala calidad. Estas imágenes extraídas de la práctica clínica muestran una calidad insuficiente debido a una incorrecta iluminación. Un ejemplo de los resultados obtenidos se puede observar en la Figura 5.7. Tanto la papila como la zona superior de la imagen original (Figura 5.7 (a)) son más oscuras que el resto de la imagen. Tras la utilización del método (Figura 5.7 (b)) estas zonas aparecen más iluminadas pudiéndose apreciar mejor los vasos sanguíneos en dichas áreas. No obstante, si las imágenes tienen una iluminación suficiente como se muestra en la Figura 5.8, la red no funciona de forma adecuada. En la Figura 5.8 (b) se observa que el área de la papila aparece de un color rojizo, no corrigiendo correctamente la iluminación.



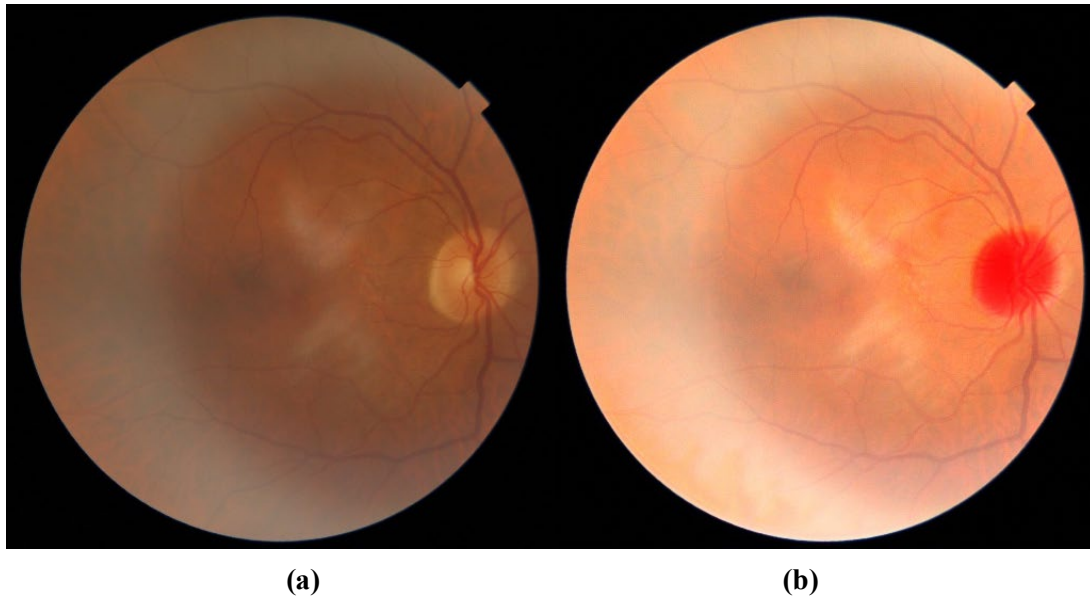


**Figura 5. 6** Efecto de la corrección de la iluminación de las imágenes. (a) Imagen degradada. (b) Imagen original. (c) Imagen generada.

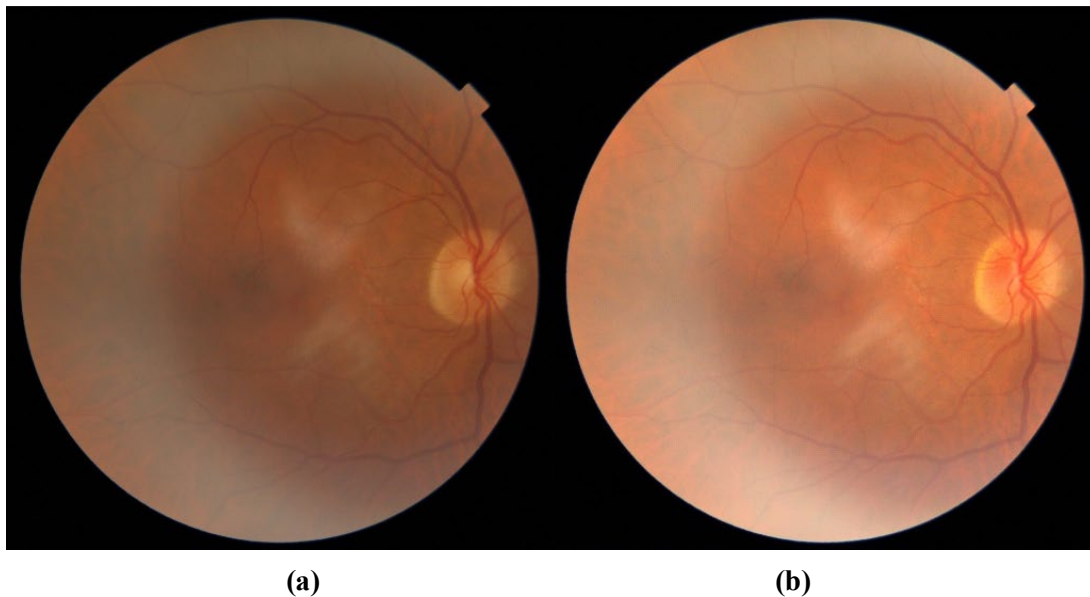


**Figura 5. 7** Efecto de la corrección de la iluminación sobre una imagen del conjunto de test de mala calidad. (a) Imagen original. (b) Imagen generada.

En las imágenes del conjunto de test de mala calidad se puede observar que muchas de ellas presentan una iluminación adecuada en gran parte de la imagen, y por tanto la red no proporciona resultados satisfactorios. Esto puede deberse a que la red GAN, únicamente se ha entrenado con imágenes que eran oscuras en su totalidad, por lo que la red no consigue resultados satisfactorios sobre aquellas que ya tienen una iluminación correcta, al menos en una parte de la imagen. Por este motivo, para poder evaluar el método sobre estas imágenes ha sido necesario oscurecerlas levemente. En la Figura 5.9 se muestra la corrección de la iluminación de la misma imagen que se ha mostrado en la Figura 5.8, pero a la que se ha oscurecido previamente (con un peso de la imagen original de 0.7 y un peso de la máscara negra de 1). En la Figura 5.9 (b) el color de la papila se ha generado correctamente, y la totalidad de la imagen aparece mejor iluminada.



**Figura 5. 8** Efecto de la corrección de la iluminación sobre una imagen del conjunto de test de mala calidad con iluminación adecuada. (a) Imagen original. (b) Imagen generada.



**Figura 5. 9** Efecto de la corrección de la iluminación sobre una imagen con iluminación adecuada a la que se ha añadido un porcentaje leve de oscuridad. (a) Imagen original. (b) Imagen generada.

#### 5.4 COMPARACIÓN CON ESTUDIOS PREVIOS

Se ha realizado una comparativa de los resultados obtenidos en este TFG con los obtenidos en estudios previos (Tabla 5.1). La mayoría de los estudios previos miden sus resultados en base a las medidas BRISQUE, NIQE, entropía, PSNR y SSIM. Por ello, en este TFG se decidió incluir todas estas medidas para poder analizar los resultados. Se observan resultados similares a los logrados por otros autores en estudios previos, si bien las técnicas aplicadas son diversas.

Para obtener los resultados globales del método desarrollado en este TFG se ha calculado la media de los resultados obtenidos para cada conjunto de la BD. En las imágenes de mala calidad, los resultados escogidos para el cálculo de esta media han sido

Referencia	EVALUACIÓN SIN REFERENCIA			EVALUACIÓN CON REFERENCIA	
	BRISQUE	NIQE	Entropía	PSNR	SSIM
Mitra et al. (2018)	45.16	3.32	6.69	16.38	0.78
Zhou et al., (2018)	46.13	4.30	6.74	17.73	0.73
Xiong et al., (2017)	43.61	3.87	6.67	17.26	0.87
Zhao et al., (2019)	40.62	2.74	6.89	19.24	0.89
Qureshi et al., (2019)			4.60	23.78	0.98
Método propuesto para corregir nitidez	<b>22.96</b>	12.70	3.95	<b>38.04</b>	0.93
Método propuesto para corregir iluminación	<b>13.49</b>	11.73	3.89	<b>37.36</b>	0.97
Método propuesto	<b>18.23</b>	12.21	3.92	<b>37.69</b>	0.95

**Tabla 5. 1** Comparación cuantitativa entre el método propuesto y otros métodos.

los obtenidos al degradar levemente este grupo de imágenes. Estos resultados son ligeramente inferiores a los obtenidos en trabajos previos sobre la mejora de la calidad en retinografías. No obstante, la comparación debe hacerse con precaución ya que en trabajos previos utilizan distintas BDs y la forma de obtener los resultados varía entre unos y otros.

Como se puede observar en esta tabla, en otros estudios se han obtenido valores más bajos de NIQE y valores más altos de entropía. El hecho de que, en la mayoría de ellos, estos valores sean aproximadamente iguales puede ser debido a que sus autores hicieron uso de las mismas BDs. Otro aspecto interesante es que las BDs utilizadas en los estudios previos suelen estar formadas por un número mayor de imágenes que la BD empleada para validar este trabajo. Sin embargo, aunque estos estudios tienen BDs mayores, las imágenes empleadas en este TFG son muy variables y más aproximadas a la práctica clínica real.

Como se ha mencionado anteriormente, los valores más bajos de BRISQUE y NIQE indican una mejor calidad de imagen, mientras que los valores de entropía más altos se refieren a una mejor calidad. El método global desarrollado logra la mejor calidad de imagen en términos de BRISQUE con un valor de 18.23. En cuanto a la evaluación con referencia, el enfoque presentado en este trabajo obtiene el valor más alto de PSNR con 37.69. Además, obtiene el segundo valor más alto de SSIM con 0.95. Estos resultados están en línea con los más altos de los obtenidos en estudios previos siendo superiores a ellos en algunos casos.

Cabe destacar que ninguno de los estudios previos mencionados incluye ejemplos sobre imágenes de mala calidad extraídas de la práctica clínica. Esto quiere decir que todas las imágenes presentadas en dichos estudios han sido previamente degradadas mediante operaciones de diezmado o la aplicación de algún tipo de filtro. Además, a diferencia de otros métodos, el modelo presentado mejoró el contraste reduciendo tanto

la borrosidad como unificando el color de las imágenes de fondo de ojo sin necesidad de realizar un preprocesado costoso y ningún paso de postprocesado.

# CAPÍTULO 6

---

## CONCLUSIONES Y LÍNEAS FUTURAS

6.1	INTRODUCCIÓN .....	91
6.2	CONTRIBUCIONES ORIGINALES .....	91
6.3	CONCLUSIONES .....	92
6.4	LIMITACIONES Y LÍNEAS FUTURAS .....	93



## 6.1 INTRODUCCIÓN

En este capítulo se exponen las principales contribuciones y conclusiones de este TFG. Tras la exposición en los anteriores capítulos del método propuesto en este trabajo, los resultados obtenidos y la discusión de los mismos, en la sección 6.2 se han recogido las aportaciones originales del trabajo. A continuación, se presentan las principales conclusiones del estudio. Finalmente, la sección 6.4 recoge las posibles líneas futuras de investigación.

## 6.2 CONTRIBUCIONES ORIGINALES

La principal aportación de este TFG radica en la propuesta de un método automático para la mejora de calidad de retinografías basado en *Deep Learning*. El método empleado se basa en redes de tipo GAN, ampliamente utilizadas para mejorar la resolución de todo tipo de imágenes. Además, se han tratado de mejorar los resultados de otros métodos centrados en la mejora de calidad de retinografías previamente publicados. Las principales contribuciones del trabajo realizado son:

- 1) Implementación de una etapa previa que modifica las imágenes de fondo de ojo degradando su calidad en términos de nitidez e iluminación. Esta etapa es necesaria puesto que, a partir de estas imágenes de menor calidad, se extraen las características de entrada a la red.
- 2) Implementación de una red de tipo GAN basada en el método propuesto en (Ledig et al., 2017) pero con las siguientes modificaciones novedosas. En primer lugar, se han eliminado los bloques de interpolación presentes en la red del generador de imágenes y se ha adaptado para imágenes de tamaño 512 x 512 en vez de para imágenes de 256 x 256. En segundo lugar, se ha degradado la calidad de las imágenes a través de un filtro gaussiano y de una máscara negra en lugar de a través de una operación de diezmado. El método desarrollado proporciona resultados similares a los obtenidos en otros estudios previos.
- 3) Verificación de la idoneidad del método en imágenes extraídas de la práctica clínica y con diferentes niveles de calidad.
- 4) Evaluación de los resultados obtenidos empleando dos criterios diferentes: una evaluación de la imagen con referencia a la original y una evaluación de las imágenes sin referencia a otra.
- 5) Obtención de medidas globales de la calidad de las retinografías. Para ello, se aplicaron, entre otros, los métodos BRISQUE y NIQE. Estas medidas se han desarrollado fuera del ámbito médico, con un propósito más general y con la finalidad de evaluar la calidad de diversos tipos de imágenes. En este trabajo se ha podido comprobar su aplicabilidad sobre imágenes de retina.

### 6.3 CONCLUSIONES

En este TFG se ha abordado una tarea importante en el ámbito del procesado automático de retinografías para el diagnóstico de enfermedades como la RD: la mejora de calidad de las mismas. Es necesario tener en cuenta que la correcta interpretación de las imágenes de fondo de ojo en un contexto clínico, ya sea por un experto o por un sistema automático, sólo puede hacerse si las imágenes tienen calidad suficiente. En este contexto, el método propuesto cumple el objetivo de mejorar la calidad de las retinografías para aumentar el número de imágenes interpretables en sistemas de cribado o en un entorno clínico, evitándose así descartar muchas de las imágenes capturadas.

Cabe destacar que se han cumplido todos los objetivos planteados para este TFG:

- 1) Se ha llevado a cabo una exhaustiva revisión bibliográfica de los diferentes estudios que se han empleado para la mejora de calidad de retinografías. En la revisión se incluyeron las investigaciones de referencia en el campo, así como las investigaciones más novedosas y recientes. La revisión realizada se ha empleado como punto de partida para el desarrollo del algoritmo propuesto en este trabajo.
- 2) Se ha creado una BD de retinografías en colaboración con el IOBA y el Hospital Clínico Universitario de Valladolid. Esta BD contiene 1000 imágenes con diferentes niveles de calidad.
- 3) Se ha empleado el lenguaje de programación Python y sus librerías de *Deep Learning: Keras y TensorFlow*, aprovechando las enormes posibilidades que ofrecen en el procesado de imágenes digitales.
- 4) Se ha diseñado y desarrollado un método completamente automático para la mejora de calidad de imágenes de fondo de ojo. Para ello, se han empleado dos redes de tipo GAN.
- 5) Se ha evaluado el funcionamiento del método propuesto empleando medidas estadísticas. Posteriormente, se ha realizado el análisis de dichos resultados.

Teniendo en cuenta el trabajo realizado y el grado de cumplimiento de los objetivos propuestos, las conclusiones más relevantes que se pueden extraer de este TFG son:

- 1) La fase de preprocesado afecta considerablemente a los resultados del estudio. Dependiendo del grado de oscuridad o borrosidad que se aplique a la imagen de partida, se consigue un mayor o menor grado de mejora de la calidad de la imagen.
- 2) El uso de redes de tipo GAN es adecuado para la mejora de calidad de retinografías. El método puede llegar a ser muy preciso si la cantidad de imágenes disponibles durante el entrenamiento es elevada.
- 3) Los métodos desarrollados son totalmente automáticos una vez que se han encontrado los parámetros óptimos a utilizar durante la fase de entrenamiento. Esto significa que no se requiere la intervención del usuario en ningún momento del procesado de la imagen.



- 4) Los resultados obtenidos con el método propuesto son similares a los alcanzados en estudios previos e indican que el método desarrollado puede ser útil en el contexto de la mejora de calidad de retinografías.
- 5) El método propuesto podría ser empleado en un entorno clínico o de cribaje de enfermedades oculares ya que, si se aplica como paso previo a la evaluación clínica de las imágenes, permitiría reducir el número de imágenes no graduables.

## 6.4 LIMITACIONES Y LÍNEAS FUTURAS

A pesar de que los resultados obtenidos son satisfactorios y similares a los de estudios previos, este estudio no está exento de limitaciones. Algunas de ellas se exponen a continuación:

- 1) Una de las mayores limitaciones del método propuesto es que corrige correctamente las imágenes borrosas y las oscuras. Sin embargo, no responde correctamente a imágenes que ya tienen nitidez e iluminación adecuadas. En las imágenes de mala calidad sucede que muchas de ellas ya son nítidas en gran parte de la imagen o que no responden a un ruido puramente gaussiano, puesto que la calidad de la imagen se evalúa también en base a otros factores diferentes. La red responde muy bien ante imágenes borrosas, pero si tiene que procesar una imagen que ya es nítida lo hace incorrectamente. Lo mismo ocurre con las imágenes oscuras. La red consigue corregir las imágenes oscuras, pero no consigue procesar aquellas que tienen una iluminación correcta.
- 2) Otra de las limitaciones del presente trabajo es el tamaño de la BD con la que se ha trabajado. Para poder extender las conclusiones del estudio realizado sería necesario completar la BD con un mayor número de imágenes. Además, sería necesario que estas nuevas retinografías que formasen parte de la BD fueran más representativas de la variabilidad existente en la práctica clínica, es decir, utilizando imágenes de distintos retinógrafos o lugares de captura y tomadas por distintos especialistas. Esta heterogeneidad dotaría de un carácter más universal al método y permitiría comprobar la efectividad del mismo con mayor precisión.
- 3) Otra de las limitaciones del método propuesto tiene que ver con el uso de redes neuronales. En los métodos que utilizan *Deep Learning* resulta imposible determinar cómo llegan las redes neuronales a una conclusión. Además, el coste computacional de estos modelos es enorme debido a la gran cantidad de iteraciones y de datos que es necesario almacenar.

Teniendo en cuenta estas limitaciones, se proponen algunas líneas futuras de investigación centradas en realizar ciertas modificaciones sobre el método propuesto para intentar mejorar los resultados obtenidos. En este contexto, se plantean dos opciones:

- 1) Incluir una fase previa para comprobar qué imágenes necesitan corrección de nitidez e iluminación.

- 2) Entrenar con imágenes de buena calidad para que la red aprenda cómo tratarlas. Además, sería recomendable degradar solo determinadas partes de las imágenes en lugar de la totalidad de la misma.

Además, en este trabajo no se ha planteado la posibilidad de que el método pueda utilizarse para mejorar la calidad de retinografías de diferentes resoluciones y grados de apertura del FOV. Por ello, se plantea explorar las posibilidades de adaptar el método desarrollado a una mayor variabilidad de imágenes ampliando la BD disponible durante el entrenamiento. En este contexto, sería deseable utilizar las imágenes de resolución completa (1956x1934 píxeles), aunque para ello se necesitará disponer de un ordenador con un gran número de recursos.

Los resultados obtenidos sugieren que el método desarrollado puede resultar útil para la ayuda al diagnóstico de la RD, pudiéndose emplear como un paso previo a tareas de segmentación, algoritmos de detección de lesiones o sistemas de graduación de la severidad de enfermedades oculares como la RD. El método implementado puede conducir a la reducción de costes en los programas de detección, lo que facilitaría el diagnóstico y cribado de poblaciones extensas.

## BIBLIOGRAFÍA

- Abiru, N., Kawasaki, E., & Eguch, K. (2002). Current knowledge of Japanese type 1 diabetic syndrome. *Diabetes/Metabolism Research and Reviews*, 18(5), 357–366. <https://doi.org/10.1002/dmrr.323>
- Abràmoff, M. D., Garvin, M. K., & Sonka, M. (2010). Retinal Imaging and Image Analysis. *IEEE Rev Biomed Eng*, 169–208. <https://doi.org/10.1109/RBME.2010.2084567>.Retinal
- Achille, A., & Soatto, S. (2018). Information Dropout: Learning Optimal Representations Through Noisy Computation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12), 2897–2905. <https://doi.org/10.1109/TPAMI.2017.2784440>
- Aggarwal, C. C. . (2018). *Neural Networks and Deep Learning*. Springer International Publishing.
- Ahirwar, K. (2019). *Generative Adversarial Networks Projects. Build next-generation generative models using TensorFlow and Keras*. Packt Publishing.
- Akhtar, P., & Azhar, F. (2010). A Single Image Interpolation Scheme for Enhanced Super Resolution in Bio-Medical Imaging. *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, 1–5. <https://doi.org/10.1109/ICBBE.2010.5518164>
- Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging*, 30(4), 449–459. <https://doi.org/10.1007/s10278-017-9983-4>
- Alemañy Martorell, Jaime; Villar Valdés, R. (2003). *Oftalmología*. La Habana: Editorial Ciencias Médicas.
- Aliseda, D., & Berastegui, L. (2008). Retinopatía diabética. *Anales Del Sistema Sanitario de Navarra*, 31, 23–34. [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=19169292%5Cnhttp://recyt.fecyt.es/index.php/ASSN/article/download/5210/4402](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19169292%5Cnhttp://recyt.fecyt.es/index.php/ASSN/article/download/5210/4402)
- American Diabetes Association. (1997). Report of the expert committee on the diagnosis and classification of Diabetes Mellitus. *Diabetes Care*, 20, 1183–1197.
- Armanious, K., Jiang, C., Fischer, M., Küstner, T., Nikolaou, K., Gatidis, S., & Yang, B. (2018). MedGAN: Medical Image Translation using GANs. In *arXiv e-prints* (p. arXiv:1806.06397). <https://ui.adsabs.harvard.edu/abs/2018arXiv180606397A>
- Avendi, M. R., Kheradvar, A., & Jafarkhani, H. (2016). A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 30, 108–119. <https://doi.org/https://doi.org/10.1016/j.media.2016.01.005>

- Bing, X., Zhang, W., Zheng, L., & Zhang, Y. (2019). Medical Image Super Resolution Using Improved Generative Adversarial Networks. *IEEE Access*, 7, 145030–145038. <https://doi.org/10.1109/ACCESS.2019.2944862>
- Bishop, C. M. (1995). Neural networks for pattern recognition. In *Oxford University Press*.
- Boucher, M. C. ., Gresset, J. A. ., Angioi, K. ., & Olivier, S. (2003). Effectiveness and safety of screening for diabetic retinopathy with two nonmydriatic digital images compared with the seven standard stereoscopic photographic fields. *Canadian Journal of Ophthalmology*, 38(7), 557–568.
- Boyd, S. (2010). *La retina normal. Enfermedades y cirugía vitreoretiniana*. Panamá: Jaypee-Highlights Medical Publishers.
- Bruna, J., Sprechmann, P., & LeCun, Y. (2016). *Super-resolution with deep convolutional sufficient statistics*. 2009, 1–17.
- Chuang, Y., Chen, Y., Wang, Y., & Kao, M. (2018). *Deep Photo Enhancer : Unpaired Learning for Image Enhancement from Photographs with GANs*.
- Dabelea, D., Mayer-Davis, E. J., Saydah, S., Imperatore, G., Linder, B., Divers, J., Bell, R., Badaru, A., Talton, J. W., Crume, T., Liese, A. D., Merchant, A. T., Lawrence, J. M., Reynolds, K., Dolan, L., Liu, L. L., Hamman, R. F., & Study, for the S. for D. in Y. (2014). Prevalence of Type 1 and Type 2 Diabetes Among Children and Adolescents From 2001 to 2009. *JAMA*, 311(17), 1778–1786. <https://doi.org/10.1001/jama.2014.3201>
- Denton, E., Szlam, A., & Fergus, R. (2015). *Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks*. 1–9. <http://papers.nips.cc/paper/5773-deep-generative-image-models-using-a-laplacian-pyramid-of-adversarial-networks.pdf>
- Do, C. B., & Ng, A. Y. (2006). Transfer learning for text classification. In Y. Weiss, B. Schölkopf, & J. C. Platt (Eds.), *Advances in Neural Information Processing Systems 18* (pp. 299–306). MIT Press. <http://papers.nips.cc/paper/2843-transfer-learning-for-text-classification.pdf>
- Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>
- Druet, C., Tubiana-Rufi, N., Chevenne, D., Rigal, O., Polak, M., & Levy-Marchal, C. (2006). Characterization of insulin secretion and resistance in type 2 diabetes of adolescents. *Journal of Clinical Endocrinology and Metabolism*, 91(2), 401–404. <https://doi.org/10.1210/jc.2005-1672>
- Duchon, C. E. (1979). Lanczos Filtering in One and Two Dimensions. In *Journal of applied meteorology* (Vol. 18, Issue 8, pp. 1016–1022). [https://doi.org/10.1175/1520-0450\(1979\)018<1016:LFIOAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<1016:LFIOAT>2.0.CO;2)
- Early Treatment Diabetic Retinopathy Study Research Group. (1991). Early photocoagulation for diabetic retinopathy. ETDRS report number 9. *Ophthalmology*, 98(5 Suppl), 766–785.
- Elliott, A. (2005). Medical imaging. *Nuclear Instruments and Methods in Physics*

- Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 546(1), 1–13.  
<https://doi.org/https://doi.org/10.1016/j.nima.2005.03.127>
- Enderle, John D. & Bronzino, J. D. (2012). Introduction to Biomedical Engineering. In *Journal of Chemical Information and Modeling* (Third Edit, Vol. 53, Issue 9). Academic Press.
- Federación Internacional de Diabetes. (2019). Atlas de la Diabetes de la FID. Actualización. In *Atlas de la Diabetes de la FID*. [http://www.idf.org/sites/default/files/Atlas-poster-2014\\_ES.pdf](http://www.idf.org/sites/default/files/Atlas-poster-2014_ES.pdf)
- Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321, 321–331. <https://doi.org/https://doi.org/10.1016/j.neucom.2018.09.013>
- Fu, Y., & Wang, S. (2016). A no reference image quality assessment metric based on visual perception. *Algorithms*, 9(4). <https://doi.org/10.3390/a9040087>
- García Gadañón, M. (2008). *Procesado de Retinografías Basado en Redes Neuronales para la Detección Automática de Lesiones Asociadas a la retinopatía Diabética*. Universidad de Valladolid.
- Gedraite, E. S., & Hadad, M. (2011). Investigation on the effect of a Gaussian Blur in image filtering and segmentation. *Proceedings Elmar - International Symposium Electronics in Marine, September*, 393–396.
- Gonzalez, Rafael C. & Woods, R. E. (2018). *Digital Image Processing* (4th Editio). Pearson Education.
- Gonzalez, R. C. ., & Woods, R. E. . (2000). *Processamento de imagens digitais* (E. Blücher (ed.)).
- Goodfellow, I. ., Pouget-Abadie, J. ., Mirza, M. ., Xu, B. ., Warde-Farley, D. ., Ozair, S. ., Courville, A. ., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2672–2680.
- Grisan, E. (2005). *Automatic Analysis of Retinal Images: Retinopathy Detection and Grading*. PhD dissertation.
- Gupta, B., & Tiwari, M. (2019). Color retinal image enhancement using luminosity and quantile based contrast enhancement. *Multidimensional Systems and Signal Processing*, 30(4), 1829–1837. <https://doi.org/10.1007/s11045-019-00630-1>
- Guy, Chris & Ffytche, D. (2005). *An Introduction to the Principles of Medical Imaging*. Imperial College Press.
- Hadjikhani, N., & Tootell, R. B. H. (2000). Projection of rods and cones within human visual cortex. *Human Brain Mapping*, 9(1), 55–63. [https://doi.org/10.1002/\(SICI\)1097-0193\(2000\)9:1<55::AID-HBM6>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-0193(2000)9:1<55::AID-HBM6>3.0.CO;2-U)
- Halban, P. A., Polonsky, K. S., Bowden, D. W., Hawkins, M. A., Ling, C., Mather, K. J., Powers, A. C., Rhodes, C. J., Sussel, L., & Weir, G. C. (2014).  $\beta$ -cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *The Journal of Clinical Endocrinology and Metabolism*, 99(6), 1983–1992.

<https://doi.org/10.1210/jc.2014-1425>

- Han, B. G., Yang, H. S., & Moon, Y. S. (2018). Locally adaptive contrast enhancement using convolutional neural network. *2018 IEEE International Conference on Consumer Electronics, ICCE 2018, 2018-Janua*, 1–2. <https://doi.org/10.1109/ICCE.2018.8326096>
- Haykin, S. (1999). *Neural Networks: a comprehensive foundation*. Prentice Hall International, Upper Saddle River.
- He, K., Sun, J., & Tang, X. (2011). *Single Image Haze Removal Using Dark Channel Prior*. 33(12), 2341–2353.
- Health, D. of. (2002). *National Service Framework for Diabetes: Delivery Strategy*.
- Holz, Frank G.; Spaide, R. (2010). *Medical Retina. Focus on Retinal Imaging*. Springer.
- Horé, A., & Ziou, D. (2010). Image Quality Metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition*, 2366–2369. <https://doi.org/10.1109/ICPR.2010.579>
- Hush, D. R. ., & Horne, B. G. (1992). An overview of neural networks. Part I: Static Networks. *Informática y Automática*, 25(1), 19–36.
- Ilangovan, S. S., Mahanty, B., & Sen, S. (2016). Biomedical imaging techniques. *Medical Imaging: Concepts, Methodologies, Tools, and Applications, April 2018*, 413–435. <https://doi.org/10.4018/978-1-5225-0571-6.ch016>
- Javitt, J. C., Canner, J. K., Frank, R. G., Steinwachs, D. M., & Sommer, A. (1990). Detecting and treating retinopathy in patients with type I diabetes mellitus. A health policy model. *Ophthalmology*, 97(4), 483–485. [https://doi.org/10.1016/s0161-6420\(90\)32573-3](https://doi.org/10.1016/s0161-6420(90)32573-3)
- Jebadurai, J., & Peter, J. D. (2018). Super-resolution of retinal images using multi-kernel SVR for IoT healthcare applications. *Future Generation Computer Systems*, 83, 338–346. <https://doi.org/10.1016/j.future.2018.01.058>
- Jebaseeli, T. J., Deva Durai, C. A., & Peter, J. D. (2019). Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM. *Optik*, 199, 163328. <https://doi.org/https://doi.org/10.1016/j.ijleo.2019.163328>
- Jennings, D., Flint, A., Turton, B. C. H., & Nokes, L. D. M. (1995). *Introduction to Medical Electronics Applications* (1<sup>a</sup> edición). Edward Arnold, Hodder Headline PLC.
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual Losses for Style Transfer and SR. *Eccv*, 1–5. [https://doi.org/10.1007/978-3-319-46475-6\\_43](https://doi.org/10.1007/978-3-319-46475-6_43)
- Kajan, S.; Goga, J.; Lacko, K. and Pavlovičová, J. (2020). Detection of Diabetic Retinopathy Using Pretrained Deep Neural Networks. *2020 Cybernetics & Informatics (K&I)*.
- Kaur, A., & Kaur, G. (2018). A review on image enhancement with deep learning approach. *ACCENTS Transactions on Image Processing and Computer Vision*, 4(11), 16–20. <https://doi.org/10.19101/tipcv.2018.411002>

- Kels, B. D., Grzybowski, A., & Grant-Kels, J. M. (2015). Human ocular anatomy. *Clinics in Dermatology*, 33(2), 140–146. <https://doi.org/10.1016/j.clindermatol.2014.10.006>
- Kharroubi, A. T. (2015). Diabetes mellitus: The epidemic of the century. *World Journal of Diabetes*, 6(6), 850. <https://doi.org/10.4239/wjd.v6.i6.850>
- Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 1646–1654. <https://doi.org/10.1109/CVPR.2016.182>
- Kim, M., Yun, J., Cho, Y., Shin, K., Jang, R., Bae, H.-J., & Kim, N. (2019). Deep Learning in Medical Imaging. *Neurospine*, 16(4), 657–668. <https://doi.org/10.14245/ns.1938396.198>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Lasmar, N.-E., Stitou, Y., & Berthoumieu, Y. (2009). Multiscale Skewed Heavy Tailed Model for Texture Analysis. *Proceedings of the 16th IEEE International Conference on Image Processing*, 2257–2260.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 105–114. <https://doi.org/10.1109/CVPR.2017.19>
- Lee, P. P., Feldman, Z. W., Ostermann, J., Brown, D. S., & Sloan, F. A. (2003). Longitudinal rates of annual eye examinations of persons with diabetes and chronic eye diseases. *Ophthalmology*, 110(10), 1952–1959. [https://doi.org/10.1016/S0161-6420\(03\)00817-0](https://doi.org/10.1016/S0161-6420(03)00817-0)
- Li, C., & Wand, M. (2016). *Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis*.
- Li, Y., & Shen, L. (2018). Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. In *Sensors* (Vol. 18, Issue 2). <https://doi.org/10.3390/s18020556>
- Lin, D. Y., Blumenkranz, M. S., Brothers, R. J., & Grosvenor, D. M. (2002). The sensitivity and specificity of single-field nonmydriatic monochromatic digital fundus photography with remote image interpretation for diabetic retinopathy screening: a comparison with ophthalmoscopy and standardized mydriatic color photography. *American Journal of Ophthalmology*, 134(2), 204–213. [https://doi.org/10.1016/s0002-9394\(02\)01522-2](https://doi.org/10.1016/s0002-9394(02)01522-2)
- Lippmann, R. P. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2), 4–22.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42(1995), 60–88.

- <https://doi.org/10.1016/j.media.2017.07.005>
- Liu, H., Xu, J., Wu, Y., Guo, Q., Ibragimov, B., & Xing, L. (2018a). Learning deconvolutional deep neural network for high resolution medical image reconstruction. *Information Sciences*, 468(August), 142–154. <https://doi.org/10.1016/j.ins.2018.08.022>
- Liu, H., Xu, J., Wu, Y., Guo, Q., Ibragimov, B., & Xing, L. (2018b). Learning deconvolutional deep neural network for high resolution medical image reconstruction. *Information Sciences*, 468, 142–154. <https://doi.org/https://doi.org/10.1016/j.ins.2018.08.022>
- López Gálvez, M. I. (2004). Escala internacional de severidad de la retinopatía y del edema macular diabético . In *Archivos de la Sociedad Española de Oftalmología* (Vol. 79, pp. 149–150). scielo.es .
- Lu, L., Zheng, Y., Carneiro, G., & Yang, L. (2017). *Deep Learning and Convolutional Neural Networks for Medical Image Computing: Precision Medicine, High Performance and Large-Scale Datasets* (1st ed.). Springer Publishing Company, Incorporated.
- Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift Fur Medizinische Physik*, 29(2), 102–127. <https://doi.org/10.1016/j.zemedi.2018.11.002>
- Mahapatra, D., & Bozorgtabar, B. (2019). *Progressive Generative Adversarial Networks for Medical Image Super resolution*. 1–21.
- Mahapatra, D., Bozorgtabar, B., & Garnavi, R. (2019). Image super-resolution using progressive generative adversarial networks for medical image analysis. *Computerized Medical Imaging and Graphics*, 71, 30–39. <https://doi.org/10.1016/j.compmedimag.2018.10.005>
- Martín, B. ., & Sanz, A. (2006). *Redes neuronales y Sistemas Borrrosos*. RA-MA editorial.
- Martínez Rubio, M., Moya Moya, M., Bellot Bernabé, A., & Belmonte Martínez, J. (2012). Cribado de retinopatía diabética y teleoftalmología. *Archivos de La Sociedad Espanola de Oftalmologia*, 87(12), 392–395. <https://doi.org/10.1016/j.ofal.2012.04.004>
- Mathieu, M., Couprie, C., & LeCun, Y. (2016). Deep multi-scale video prediction beyond mean square error. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, 2015*, 1–14.
- Maurer, Calvin R.; Fitzpatrick, J. M. (1993). A Review of Medical Image Registration. *Interactive Image-Guided Neurosurgery*, 17–44.
- Mitra, A., Roy, S., Roy, S., & Kumar, S. (2018). Computer Methods and Programs in Biomedicine: Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract. *Computer Methods and Programs in Biomedicine*, 156, 169–178. <https://doi.org/10.1016/j.cmpb.2018.01.001>
- Mitra, A., Roy, S., Roy, S., & Setua, S. K. (2018). Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract.



- Computer Methods and Programs in Biomedicine*, 156, 169–178. <https://doi.org/10.1016/j.cmpb.2018.01.001>
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Transactions on Image Processing*, 21(12), 4695–4708. <https://doi.org/10.1109/TIP.2012.2214050>
- Mittal, Anish, Soundararajan, R., & Bovik, A. C. (2013). Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters*, 20(3), 209–212. <https://doi.org/10.1109/LSP.2012.2227726>
- Moini, J. (2019). *Epidemiology of Diabetes*. Elsevier.
- Mompín Poblet, J. (1988). *Introducción a la bioingeniería*. Marcombo.
- Ong, G. L., Ripley, L. G., Newsom, R. S., Cooper, M., & Casswell, A. G. (2004). Screening for sight-threatening diabetic retinopathy: comparison of fundus photography with automated color contrast threshold test. *American Journal of Ophthalmology*, 137(3), 445–452. <https://doi.org/10.1016/j.ajo.2003.10.021>
- Ophthalmology, A. A. of. (2012). *Curso de Ciencias Básicas y Clínicas en Oftalmología (Parte II)*.
- Oppenheim, A. V. ., Willsky, A. S. ., & Nawab, S. H. (1997). *Signals and Systems* (2nd ed.). Prentice Hall.
- Ortiz Pérez, S. (2015). *Manual de enfermería oftalmológica*. Lab. 2ª ed. Thea (ed).
- Patton, N., Aslam, T. M., MacGillivray, T., Deary, I. J., Dhillon, B., Eikelboom, R. H., Yogesana, K., & Constable, I. J. (2006). Retinal image analysis: concepts, applications and potential. *Progress in Retinal and Eye Research*, 25(1), 99–127. <https://doi.org/10.1016/j.preteyeres.2005.07.001>
- Planck, M., Kim, K. I., & Kwon, Y. (2008). Example-based Learning for and JPEG Artifact Removal Example-based Learning for Single-Image Super-Resolution and JPEG Artifact Removal. *Biological Cybernetics*, August.
- Polesel, A., Ramponi, G., & Mathews, V. J. (2000). *Image Enhancement via Adaptive Unsharp Masking*. 9(3), 505–510.
- Qidwai, U., & Chen, C. H. . (2009). *Digital Image Processing: An Algorithmic Approach with MATLAB*. Taylor & Francis Ltd.
- Qureshi, I., Ma, J., & Shaheed, K. (2019). A Hybrid Proposed Fundus Image Enhancement Framework for Diabetic Retinopathy. *Algorithms*, 12(1), 1–16. <https://doi.org/10.3390/a12010014>
- Radford, A., Metz, L., & Chintala, S. (2016). *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*. 1–16.
- Riordan-Eva, Paul; Cunningham, E. T. (2012). *Vaughan y Asbury. Oftalmología general*. (McGRAW-HILL (ed.); 18 Edición).
- Rousseau, F. (2014). *Brain Hallucination*. March. <https://doi.org/10.1007/978-3-540-88682-2>
- Ruderman, D. L. (1994). The statistics of natural images. *Network: Computation in*

- Neural Systems*, 5(4), 517–548. [https://doi.org/10.1088/0954-898X\\_5\\_4\\_006](https://doi.org/10.1088/0954-898X_5_4_006)
- Saha, S. K., Fernando, B., Cuadros, J., Xiao, D., & Kanagasingham, Y. (2018). *Automated Quality Assessment of Colour Fundus Images for Diabetic Retinopathy Screening in Telemedicine*. 869–878.
- Sajjadi, M. S. M., Scholkopf, B., & Hirsch, M. (2017). EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis. *Proceedings of the IEEE International Conference on Computer Vision, 2017-Octob*, 4501–4510. <https://doi.org/10.1109/ICCV.2017.481>
- Sawhney, G. S. (2007). *Fundamentals of Biomedical Engineering* (N. A. I. P. L. Publishers (ed.)).
- Setiawan, A. W., Mengko, T. R., Santoso, O. S., & Suksmono, A. B. (2013). Color retinal image enhancement using CLAHE. *Proceedings - International Conference on ICT for Smart Society 2013: "Think Ecosystem Act Convergence", ICISS 2013, March 2015*, 215–217. <https://doi.org/10.1109/ICTSS.2013.6588092>
- Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., & Ma, J. (2017). *MSR-net: Low-light Image Enhancement Using Deep Convolutional Network*. <http://arxiv.org/abs/1711.02488>
- Simonyan, K. ., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *CoRR, Abs/1409.1556*.
- Srivastava, N. ., Hinton, G. ., Krizhevsky, A. ., Sutskever, I. ., & Salakhutdinov, R. . (2014). Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>
- Starck, J., Murtagh, F., Candès, E. J., & Donoho, D. L. (2003). *Gray and Color Image Contrast Enhancement by the Curvelet Transform*. 12(6), 706–717.
- Stewart, M. W. (2017). *Diabetic Retinopathy: Current Pharmacologic Treatment and Emerging Strategies*. Springer.
- Sun, J., Sun, J., Xu, Z., & Shum, H. Y. (2008). Image super-resolution using gradient profile prior. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*. <https://doi.org/10.1109/CVPR.2008.4587659>
- Takano, N., & Alaghband, G. (2019). *SRGAN: Training Dataset Matters*. 1–7. <http://arxiv.org/abs/1903.09922>
- Tao, L., Zhu, C., Song, J., Lu, T., Jia, H., & Xie, X. (2017). Low-light image enhancement using CNN and bright channel prior. *IEEE International Conference on Image Processing (ICIP), 0*. <http://arxiv.org/abs/1711.02488>
- Thapa, D., Raahemifar, K., Bobier, W. R., & Lakshminarayanan, V. (2014). Comparison of super-resolution algorithms applied to retinal images. *Journal of Biomedical Optics*, 19(5), 056002. <https://doi.org/10.1117/1.jbo.19.5.056002>
- Vernon, D. (1991). *Machine Vision. Automated Visual Inspection and Robot Vision*. Prentice Hall.
- Vieira, A., & Ribeiro, B. (2018). *Introduction to Deep Learning Business Applications for Developers*. Springer Science+Business Media New York.

- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (2019). ESRGAN: Enhanced super-resolution generative adversarial networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11133 LNCS, 63–79. [https://doi.org/10.1007/978-3-030-11021-5\\_5](https://doi.org/10.1007/978-3-030-11021-5_5)
- Watson, S., & Lowe, G. (2019). Ocular anatomy and physiology relevant to anaesthesia. *Anaesthesia and Intensive Care Medicine*, 20(12), 710–715. <https://doi.org/10.1016/j.mpaic.2019.10.004>
- Wilkinson, C. P., Ferris, F. L., Klein, R. E., Lee, P. P., Agardh, C. D., Davis, M., Dills, D., Kambik, A., Pararajasegaram, R., & Verdager, J. T. (2003). Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, 110(9), 1677–1682. [https://doi.org/https://doi.org/10.1016/S0161-6420\(03\)00475-5](https://doi.org/https://doi.org/10.1016/S0161-6420(03)00475-5)
- Wilkinson CP, Ferris FL 3RD, Klein RE, Lee PP, Agardh CD, D. M. et al. (2003). *Global Diabetic Retinopathy Project Group. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales.*
- World Health Organization. (2005). *Prevention of blindness from diabetes mellitus : report of a WHO consultation in Geneva, Switzerland, 9-11 November 2005.* World Health Organization. <https://apps.who.int/iris/handle/10665/43576>
- Xiong, L., Li, H., & Xu, L. (2017a). An enhancement method for color retinal images based on image formation model. *Computer Methods and Programs in Biomedicine*, 143, 137–150. <https://doi.org/https://doi.org/10.1016/j.cmpb.2017.02.026>
- Xiong, L., Li, H., & Xu, L. (2017b). Computer Methods and Programs in Biomedicine: An enhancement method for color retinal images based on image formation model. *Computer Methods and Programs in Biomedicine*, 143, 137–150. <https://doi.org/10.1016/j.cmpb.2017.02.026>
- Yanoff, Myron; Sassani, J. (2019). *Ocular Pathology.* (8th Editio). Elsevier.
- Yeh, R. A., Chen, C., Lim, T. Y., Schwing, A. G., Hasegawa-johnson, M., & Do, M. N. (2017). *Semantic Image Inpainting with Deep Generative Models.*
- Zhao, H., Li, H., Maurer-stroh, S., & Cheng, L. (2018). Synthesizing retinal and neuronal images with generative adversarial nets. *Medical Image Analysis*, 49, 14–26. <https://doi.org/10.1016/j.media.2018.07.001>
- Zhao, H., Yang, B., Cao, L., & Li, H. (2019). Data-Driven Enhancement of Blurry Retinal Images via Generative Adversarial Networks. In D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, & A. Khan (Eds.), *Medical Image Computing and Computer Assisted Intervention -- MICCAI 2019* (pp. 75–83). Springer International Publishing.
- Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, G. (2019). *Handbook of Medical Image Computing and Computer Assisted Intervention* (1st Editio). Academic Press.
- Zhou, M., Jin, K., Wang, S., Ye, J., & Qian, D. (2018). *Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment.* 65(3), 521–527.
- Abiru, N., Kawasaki, E., & Eguch, K. (2002). Current knowledge of Japanese type 1 diabetic syndrome. *Diabetes/Metabolism Research and Reviews*, 18(5), 357–366.

- <https://doi.org/10.1002/dmrr.323>
- Abràmoff, M. D., Garvin, M. K., & Sonka, M. (2010). Retinal Imaging and Image Analysis. *IEEE Rev Biomed Eng*, 169–208. <https://doi.org/10.1109/RBME.2010.2084567>.Retinal
- Achille, A., & Soatto, S. (2018). Information Dropout: Learning Optimal Representations Through Noisy Computation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12), 2897–2905. <https://doi.org/10.1109/TPAMI.2017.2784440>
- Aggarwal, C. C. . (2018). *Neural Networks and Deep Learning*. Springer International Publishing.
- Ahirwar, K. (2019). *Generative Adversarial Networks Projects. Build next-generation generative models using TensorFlow and Keras*. Packt Publishing.
- Akhtar, P., & Azhar, F. (2010). A Single Image Interpolation Scheme for Enhanced Super Resolution in Bio-Medical Imaging. *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, 1–5. <https://doi.org/10.1109/ICBBE.2010.5518164>
- Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. *Journal of Digital Imaging*, 30(4), 449–459. <https://doi.org/10.1007/s10278-017-9983-4>
- Alemañy Martorell, Jaime; Villar Valdés, R. (2003). *Oftalmología*. La Habana: Editorial Ciencias Médicas.
- Aliseda, D., & Berastegui, L. (2008). Retinopatía diabética. *Anales Del Sistema Sanitario de Navarra*, 31, 23–34. [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=19169292%5Cnhttp://recyt.fecyt.es/index.php/ASSN/article/download/5210/4402](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19169292%5Cnhttp://recyt.fecyt.es/index.php/ASSN/article/download/5210/4402)
- American Diabetes Association. (1997). Report of the expert committee on the diagnosis and classification of Diabetes Mellitus. *Diabetes Care*, 20, 1183–1197.
- Armanious, K., Jiang, C., Fischer, M., Küstner, T., Nikolaou, K., Gatidis, S., & Yang, B. (2018). MedGAN: Medical Image Translation using GANs. In *arXiv e-prints* (p. arXiv:1806.06397). <https://ui.adsabs.harvard.edu/abs/2018arXiv180606397A>
- Avendi, M. R., Kheradvar, A., & Jafarkhani, H. (2016). A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 30, 108–119. <https://doi.org/https://doi.org/10.1016/j.media.2016.01.005>
- Bing, X., Zhang, W., Zheng, L., & Zhang, Y. (2019). Medical Image Super Resolution Using Improved Generative Adversarial Networks. *IEEE Access*, 7, 145030–145038. <https://doi.org/10.1109/ACCESS.2019.2944862>
- Bishop, C. M. (1995). Neural networks for pattern recognition. In *Oxford University Press*.
- Boucher, M. C. ., Gresset, J. A. ., Angioi, K. ., & Olivier, S. (2003). Effectiveness and safety of screening for diabetic retinopathy with two nonmydriatic digital images

- compared with the seven standard stereoscopic photographic fields. *Canadian Journal of Ophthalmology*, 38(7), 557–568.
- Boyd, S. (2010). *La retina normal. Enfermedades y cirugía vitreorretiniana*. Panamá: Jaypee-Highlights Medical Publishers.
- Bruna, J., Sprechmann, P., & LeCun, Y. (2016). *Super-resolution with deep convolutional sufficient statistics*. 2009, 1–17.
- Chuang, Y., Chen, Y., Wang, Y., & Kao, M. (2018). *Deep Photo Enhancer : Unpaired Learning for Image Enhancement from Photographs with GANs*.
- Dabelea, D., Mayer-Davis, E. J., Saydah, S., Imperatore, G., Linder, B., Divers, J., Bell, R., Badaru, A., Talton, J. W., Crume, T., Liese, A. D., Merchant, A. T., Lawrence, J. M., Reynolds, K., Dolan, L., Liu, L. L., Hamman, R. F., & Study, for the S. for D. in Y. (2014). Prevalence of Type 1 and Type 2 Diabetes Among Children and Adolescents From 2001 to 2009. *JAMA*, 311(17), 1778–1786. <https://doi.org/10.1001/jama.2014.3201>
- Denton, E., Szlam, A., & Fergus, R. (2015). *Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks*. 1–9. <http://papers.nips.cc/paper/5773-deep-generative-image-models-using-a-laplacian-pyramid-of-adversarial-networks.pdf>
- Do, C. B., & Ng, A. Y. (2006). Transfer learning for text classification. In Y. Weiss, B. Schölkopf, & J. C. Platt (Eds.), *Advances in Neural Information Processing Systems 18* (pp. 299–306). MIT Press. <http://papers.nips.cc/paper/2843-transfer-learning-for-text-classification.pdf>
- Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>
- Druet, C., Tubiana-Rufi, N., Chevenne, D., Rigal, O., Polak, M., & Levy-Marchal, C. (2006). Characterization of insulin secretion and resistance in type 2 diabetes of adolescents. *Journal of Clinical Endocrinology and Metabolism*, 91(2), 401–404. <https://doi.org/10.1210/jc.2005-1672>
- Duchon, C. E. (1979). Lanczos Filtering in One and Two Dimensions. In *Journal of applied meteorology* (Vol. 18, Issue 8, pp. 1016–1022). [https://doi.org/10.1175/1520-0450\(1979\)018<1016:LFIOAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<1016:LFIOAT>2.0.CO;2)
- Early Treatment Diabetic Retinopathy Study Research Group. (1991). Early photocoagulation for diabetic retinopathy. ETDRS report number 9. *Ophthalmology*, 98(5 Suppl), 766–785.
- Elliott, A. (2005). Medical imaging. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 546(1), 1–13. <https://doi.org/https://doi.org/10.1016/j.nima.2005.03.127>
- Enderle, John D. & Bronzino, J. D. (2012). Introduction to Biomedical Engineering. In *Journal of Chemical Information and Modeling* (Third Edit, Vol. 53, Issue 9). Academic Press.
- Federación Internacional de Diabetes. (2019). Atlas de la Diabetes de la FID.

- Actualización. In *Atlas de la Diabetes de la FID*. [http://www.idf.org/sites/default/files/Atlas-poster-2014\\_ES.pdf](http://www.idf.org/sites/default/files/Atlas-poster-2014_ES.pdf)
- Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321, 321–331. <https://doi.org/https://doi.org/10.1016/j.neucom.2018.09.013>
- Fu, Y., & Wang, S. (2016). A no reference image quality assessment metric based on visual perception. *Algorithms*, 9(4). <https://doi.org/10.3390/a9040087>
- García Gadañón, M. (2008). *Procesado de Retinografías Basado en Redes Neuronales para la Detección Automática de Lesiones Asociadas a la retinopatía Diabética*. Universidad de Valladolid.
- Gedraite, E. S., & Hadad, M. (2011). Investigation on the effect of a Gaussian Blur in image filtering and segmentation. *Proceedings Elmar - International Symposium Electronics in Marine, September*, 393–396.
- Gonzalez, Rafael C. & Woods, R. E. (2018). *Digital Image Processing* (4th Editio). Pearson Education.
- Gonzalez, R. C. ., & Woods, R. E. . (2000). *Processamento de imagens digitais* (E. Blücher (ed.)).
- Goodfellow, I. ., Pouget-Abadie, J. ., Mirza, M. ., Xu, B. ., Warde-Farley, D. ., Ozair, S. ., Courville, A. ., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 2672–2680.
- Grisan, E. (2005). *Automatic Analysis of Retinal Images: Retinopathy Detection and Grading*. PhD dissertation.
- Gupta, B., & Tiwari, M. (2019). Color retinal image enhancement using luminosity and quantile based contrast enhancement. *Multidimensional Systems and Signal Processing*, 30(4), 1829–1837. <https://doi.org/10.1007/s11045-019-00630-1>
- Guy, Chris & Ffytche, D. (2005). *An Introduction to the Principles of Medical Imaging*. Imperial College Press.
- Hadjikhani, N., & Tootell, R. B. H. (2000). Projection of rods and cones within human visual cortex. *Human Brain Mapping*, 9(1), 55–63. [https://doi.org/10.1002/\(SICI\)1097-0193\(2000\)9:1<55::AID-HBM6>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-0193(2000)9:1<55::AID-HBM6>3.0.CO;2-U)
- Halban, P. A., Polonsky, K. S., Bowden, D. W., Hawkins, M. A., Ling, C., Mather, K. J., Powers, A. C., Rhodes, C. J., Sussel, L., & Weir, G. C. (2014).  $\beta$ -cell failure in type 2 diabetes: postulated mechanisms and prospects for prevention and treatment. *The Journal of Clinical Endocrinology and Metabolism*, 99(6), 1983–1992. <https://doi.org/10.1210/jc.2014-1425>
- Han, B. G., Yang, H. S., & Moon, Y. S. (2018). Locally adaptive contrast enhancement using convolutional neural network. *2018 IEEE International Conference on Consumer Electronics, ICCE 2018, 2018-Janua*, 1–2. <https://doi.org/10.1109/ICCE.2018.8326096>
- Haykin, S. (1999). *Neural Networks: a comprehensive foundation*. Prentice Hall International, Upper Saddle River.

- He, K., Sun, J., & Tang, X. (2011). *Single Image Haze Removal Using Dark Channel Prior*. 33(12), 2341–2353.
- Health, D. of. (2002). *National Service Framework for Diabetes: Delivery Strategy*.
- Holz, Frank G.; Spaide, R. (2010). *Medical Retina. Focus on Retinal Imaging*. Springer.
- Horé, A., & Ziou, D. (2010). Image Quality Metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition*, 2366–2369. <https://doi.org/10.1109/ICPR.2010.579>
- Hush, D. R. ., & Horne, B. G. (1992). An overview of neural networks. Part I: Static Networks. *Informática y Automática*, 25(1), 19–36.
- Ilangovan, S. S., Mahanty, B., & Sen, S. (2016). Biomedical imaging techniques. *Medical Imaging: Concepts, Methodologies, Tools, and Applications, April 2018*, 413–435. <https://doi.org/10.4018/978-1-5225-0571-6.ch016>
- Javitt, J. C., Canner, J. K., Frank, R. G., Steinwachs, D. M., & Sommer, A. (1990). Detecting and treating retinopathy in patients with type I diabetes mellitus. A health policy model. *Ophthalmology*, 97(4), 483–485. [https://doi.org/10.1016/s0161-6420\(90\)32573-3](https://doi.org/10.1016/s0161-6420(90)32573-3)
- Jebadurai, J., & Peter, J. D. (2018). Super-resolution of retinal images using multi-kernel SVR for IoT healthcare applications. *Future Generation Computer Systems*, 83, 338–346. <https://doi.org/10.1016/j.future.2018.01.058>
- Jebaseeli, T. J., Deva Durai, C. A., & Peter, J. D. (2019). Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM. *Optik*, 199, 163328. <https://doi.org/https://doi.org/10.1016/j.ijleo.2019.163328>
- Jennings, D., Flint, A., Turton, B. C. H., & Nokes, L. D. M. (1995). *Introduction to Medical Electronics Applications* (1<sup>a</sup> edición). Edward Arnold, Hodder Headline PLC.
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual Losses for Style Transfer and SR. *Eccv*, 1–5. [https://doi.org/10.1007/978-3-319-46475-6\\_43](https://doi.org/10.1007/978-3-319-46475-6_43)
- Kajan, S.; Goga, J.; Lacko, K. and Pavlovičová, J. (2020). Detection of Diabetic Retinopathy Using Pretrained Deep Neural Networks. *2020 Cybernetics & Informatics (K&I)*.
- Kaur, A., & Kaur, G. (2018). A review on image enhancement with deep learning approach. *ACCENTS Transactions on Image Processing and Computer Vision*, 4(11), 16–20. <https://doi.org/10.19101/tpcv.2018.411002>
- Kels, B. D., Grzybowski, A., & Grant-Kels, J. M. (2015). Human ocular anatomy. *Clinics in Dermatology*, 33(2), 140–146. <https://doi.org/10.1016/j.clindermatol.2014.10.006>
- Kharroubi, A. T. (2015). Diabetes mellitus: The epidemic of the century. *World Journal of Diabetes*, 6(6), 850. <https://doi.org/10.4239/wjd.v6.i6.850>
- Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 1646–1654.

---

<https://doi.org/10.1109/CVPR.2016.182>

- Kim, M., Yun, J., Cho, Y., Shin, K., Jang, R., Bae, H.-J., & Kim, N. (2019). Deep Learning in Medical Imaging. *Neurospine*, *16*(4), 657–668. <https://doi.org/10.14245/ns.1938396.198>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Lasmar, N.-E., Stitou, Y., & Berthoumieu, Y. (2009). Multiscale Skewed Heavy Tailed Model for Texture Analysis. *Proceedings of the 16th IEEE International Conference on Image Processing*, 2257–2260.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 105–114. <https://doi.org/10.1109/CVPR.2017.19>
- Lee, P. P., Feldman, Z. W., Ostermann, J., Brown, D. S., & Sloan, F. A. (2003). Longitudinal rates of annual eye examinations of persons with diabetes and chronic eye diseases. *Ophthalmology*, *110*(10), 1952–1959. [https://doi.org/10.1016/S0161-6420\(03\)00817-0](https://doi.org/10.1016/S0161-6420(03)00817-0)
- Li, C., & Wand, M. (2016). *Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis*.
- Li, Y., & Shen, L. (2018). Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. In *Sensors* (Vol. 18, Issue 2). <https://doi.org/10.3390/s18020556>
- Lin, D. Y., Blumenkranz, M. S., Brothers, R. J., & Grosvenor, D. M. (2002). The sensitivity and specificity of single-field nonmydriatic monochromatic digital fundus photography with remote image interpretation for diabetic retinopathy screening: a comparison with ophthalmoscopy and standardized mydriatic color photography. *American Journal of Ophthalmology*, *134*(2), 204–213. [https://doi.org/10.1016/s0002-9394\(02\)01522-2](https://doi.org/10.1016/s0002-9394(02)01522-2)
- Lippmann, R. P. (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine*, *4*(2), 4–22.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*(1995), 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Liu, H., Xu, J., Wu, Y., Guo, Q., Ibragimov, B., & Xing, L. (2018a). Learning deconvolutional deep neural network for high resolution medical image reconstruction. *Information Sciences*, *468*(August), 142–154. <https://doi.org/10.1016/j.ins.2018.08.022>
- Liu, H., Xu, J., Wu, Y., Guo, Q., Ibragimov, B., & Xing, L. (2018b). Learning deconvolutional deep neural network for high resolution medical image reconstruction. *Information Sciences*, *468*, 142–154.



<https://doi.org/https://doi.org/10.1016/j.ins.2018.08.022>

- López Gálvez, M. I. (2004). Escala internacional de severidad de la retinopatía y del edema macular diabético . In *Archivos de la Sociedad Española de Oftalmología* (Vol. 79, pp. 149–150). scieloes .
- Lu, L., Zheng, Y., Carneiro, G., & Yang, L. (2017). *Deep Learning and Convolutional Neural Networks for Medical Image Computing: Precision Medicine, High Performance and Large-Scale Datasets* (1st ed.). Springer Publishing Company, Incorporated.
- Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift Fur Medizinische Physik*, 29(2), 102–127. <https://doi.org/10.1016/j.zemedi.2018.11.002>
- Mahapatra, D., & Bozorgtabar, B. (2019). *Progressive Generative Adversarial Networks for Medical Image Super resolution*. 1–21.
- Mahapatra, D., Bozorgtabar, B., & Garnavi, R. (2019). Image super-resolution using progressive generative adversarial networks for medical image analysis. *Computerized Medical Imaging and Graphics*, 71, 30–39. <https://doi.org/10.1016/j.compmedimag.2018.10.005>
- Martín, B. ., & Sanz, A. (2006). *Redes neuronales y Sistemas Borrrosos*. RA-MA editorial.
- Martínez Rubio, M., Moya Moya, M., Bellot Bernabé, A., & Belmonte Martínez, J. (2012). Cribado de retinopatía diabética y teleoftalmología. *Archivos de La Sociedad Espanola de Oftalmologia*, 87(12), 392–395. <https://doi.org/10.1016/j.ofal.2012.04.004>
- Mathieu, M., Couprie, C., & LeCun, Y. (2016). Deep multi-scale video prediction beyond mean square error. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, 2015*, 1–14.
- Maurer, Calvin R.; Fitzpatrick, J. M. (1993). A Review of Medical Image Registration. *Interactive Image-Guided Neurosurgery*, 17–44.
- Mitra, A., Roy, S., Roy, S., & Kumar, S. (2018). Computer Methods and Programs in Biomedicine: Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract. *Computer Methods and Programs in Biomedicine*, 156, 169–178. <https://doi.org/10.1016/j.cmpb.2018.01.001>
- Mitra, A., Roy, S., Roy, S., & Setua, S. K. (2018). Enhancement and restoration of non-uniform illuminated Fundus Image of Retina obtained through thin layer of cataract. *Computer Methods and Programs in Biomedicine*, 156, 169–178. <https://doi.org/10.1016/j.cmpb.2018.01.001>
- Mittal, A, Moorthy, A. K., & Bovik, A. C. (2012). No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Transactions on Image Processing*, 21(12), 4695–4708. <https://doi.org/10.1109/TIP.2012.2214050>
- Mittal, Anish, Soundararajan, R., & Bovik, A. C. (2013). Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters*, 20(3), 209–212. <https://doi.org/10.1109/LSP.2012.2227726>

- Moini, J. (2019). *Epidemiology of Diabetes*. Elsevier.
- Mompín Poblet, J. (1988). *Introducción a la bioingeniería*. Marcombo.
- Ong, G. L., Ripley, L. G., Newsom, R. S., Cooper, M., & Casswell, A. G. (2004). Screening for sight-threatening diabetic retinopathy: comparison of fundus photography with automated color contrast threshold test. *American Journal of Ophthalmology*, 137(3), 445–452. <https://doi.org/10.1016/j.ajo.2003.10.021>
- Ophthalmology, A. A. of. (2012). *Curso de Ciencias Básicas y Clínicas en Oftalmología (Parte II)*.
- Oppenheim, A. V. ., Willsky, A. S. ., & Nawab, S. H. (1997). *Signals and Systems* (2nd ed.). Prentice Hall.
- Ortiz Pérez, S. (2015). *Manual de enfermería oftalmológica*. Lab. 2ª ed. Thea (ed).
- Patton, N., Aslam, T. M., MacGillivray, T., Deary, I. J., Dhillon, B., Eikelboom, R. H., Yogesán, K., & Constable, I. J. (2006). Retinal image analysis: concepts, applications and potential. *Progress in Retinal and Eye Research*, 25(1), 99–127. <https://doi.org/10.1016/j.preteyeres.2005.07.001>
- Planck, M., Kim, K. I., & Kwon, Y. (2008). Example-based Learning for and JPEG Artifact Removal Example-based Learning for Single-Image Super-Resolution and JPEG Artifact Removal. *Biological Cybernetics*, August.
- Polesel, A., Ramponi, G., & Mathews, V. J. (2000). *Image Enhancement via Adaptive Unsharp Masking*. 9(3), 505–510.
- Qidwai, U., & Chen, C. H. . (2009). *Digital Image Processing: An Algorithmic Approach with MATLAB*. Taylor & Francis Ltd.
- Qureshi, I., Ma, J., & Shaheed, K. (2019). A Hybrid Proposed Fundus Image Enhancement Framework for Diabetic Retinopathy. *Algorithms*, 12(1), 1–16. <https://doi.org/10.3390/a12010014>
- Radford, A., Metz, L., & Chintala, S. (2016). *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*. 1–16.
- Riordan-Eva, Paul; Cunningham, E. T. (2012). *Vaughan y Asbury. Oftalmología general*. (McGRAW-HILL (ed.); 18 Edición).
- Rousseau, F. (2014). *Brain Hallucination. March*. <https://doi.org/10.1007/978-3-540-88682-2>
- Ruderman, D. L. (1994). The statistics of natural images. *Network: Computation in Neural Systems*, 5(4), 517–548. [https://doi.org/10.1088/0954-898X\\_5\\_4\\_006](https://doi.org/10.1088/0954-898X_5_4_006)
- Saha, S. K., Fernando, B., Cuadros, J., Xiao, D., & Kanagasíngam, Y. (2018). *Automated Quality Assessment of Colour Fundus Images for Diabetic Retinopathy Screening in Telemedicine*. 869–878.
- Sajjadi, M. S. M., Scholkopf, B., & Hirsch, M. (2017). EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis. *Proceedings of the IEEE International Conference on Computer Vision, 2017-Octob*, 4501–4510. <https://doi.org/10.1109/ICCV.2017.481>

- Sawhney, G. S. (2007). *Fundamentals of Biomedical Engineering* (N. A. I. P. L. Publishers (ed.)).
- Setiawan, A. W., Mengko, T. R., Santoso, O. S., & Suksmono, A. B. (2013). Color retinal image enhancement using CLAHE. *Proceedings - International Conference on ICT for Smart Society 2013: "Think Ecosystem Act Convergence", ICISS 2013, March 2015*, 215–217. <https://doi.org/10.1109/ICTSS.2013.6588092>
- Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., & Ma, J. (2017). *MSR-net: Low-light Image Enhancement Using Deep Convolutional Network*. <http://arxiv.org/abs/1711.02488>
- Simonyan, K. ., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *CoRR, Abs/1409.1556*.
- Srivastava, N. ., Hinton, G. ., Krizhevsky, A. ., Sutskever, I. ., & Salakhutdinov, R. . (2014). Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>
- Starck, J., Murtagh, F., Candès, E. J., & Donoho, D. L. (2003). *Gray and Color Image Contrast Enhancement by the Curvelet Transform*. 12(6), 706–717.
- Stewart, M. W. (2017). *Diabetic Retinopathy: Current Pharmacologic Treatment and Emerging Strategies*. Springer.
- Sun, J., Sun, J., Xu, Z., & Shum, H. Y. (2008). Image super-resolution using gradient profile prior. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*. <https://doi.org/10.1109/CVPR.2008.4587659>
- Takano, N., & Alagband, G. (2019). *SRGAN: Training Dataset Matters*. 1–7. <http://arxiv.org/abs/1903.09922>
- Tao, L., Zhu, C., Song, J., Lu, T., Jia, H., & Xie, X. (2017). Low-light image enhancement using CNN and bright channel prior. *IEEE International Conference on Image Processing (ICIP), 0*. <http://arxiv.org/abs/1711.02488>
- Thapa, D., Raahemifar, K., Bobier, W. R., & Lakshminarayanan, V. (2014). Comparison of super-resolution algorithms applied to retinal images. *Journal of Biomedical Optics*, 19(5), 056002. <https://doi.org/10.1117/1.jbo.19.5.056002>
- Vernon, D. (1991). *Machine Vision. Automated Visual Inspection and Robot Vision*. Prentice Hall.
- Vieira, A., & Ribeiro, B. (2018). *Introduction to Deep Learning Business Applications for Developers*. Springer Science+Business Media New York.
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (2019). ESRGAN: Enhanced super-resolution generative adversarial networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11133 LNCS, 63–79. [https://doi.org/10.1007/978-3-030-11021-5\\_5](https://doi.org/10.1007/978-3-030-11021-5_5)
- Watson, S., & Lowe, G. (2019). Ocular anatomy and physiology relevant to anaesthesia. *Anaesthesia and Intensive Care Medicine*, 20(12), 710–715. <https://doi.org/10.1016/j.mpaic.2019.10.004>
- Wilkinson, C. P., Ferris, F. L., Klein, R. E., Lee, P. P., Agardh, C. D., Davis, M., Dills,

- D., Kampik, A., Pararajasegaram, R., & Verdager, J. T. (2003). Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, *110*(9), 1677–1682. [https://doi.org/https://doi.org/10.1016/S0161-6420\(03\)00475-5](https://doi.org/https://doi.org/10.1016/S0161-6420(03)00475-5)
- Wilkinson CP, Ferris FL 3RD, Klein RE, Lee PP, Agardh CD, D. M. et al. (2003). *Global Diabetic Retinopathy Project Group. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales.*
- World Health Organization. (2005). *Prevention of blindness from diabetes mellitus : report of a WHO consultation in Geneva, Switzerland, 9-11 November 2005.* World Health Organization. <https://apps.who.int/iris/handle/10665/43576>
- Xiong, L., Li, H., & Xu, L. (2017a). An enhancement method for color retinal images based on image formation model. *Computer Methods and Programs in Biomedicine*, *143*, 137–150. <https://doi.org/https://doi.org/10.1016/j.cmpb.2017.02.026>
- Xiong, L., Li, H., & Xu, L. (2017b). Computer Methods and Programs in Biomedicine: An enhancement method for color retinal images based on image formation model. *Computer Methods and Programs in Biomedicine*, *143*, 137–150. <https://doi.org/10.1016/j.cmpb.2017.02.026>
- Yanoff, Myron; Sassani, J. (2019). *Ocular Pathology.* (8th Editio). Elsevier.
- Yeh, R. A., Chen, C., Lim, T. Y., Schwing, A. G., Hasegawa-johnson, M., & Do, M. N. (2017). *Semantic Image Inpainting with Deep Generative Models.*
- Zhao, H., Li, H., Maurer-stroh, S., & Cheng, L. (2018). Synthesizing retinal and neuronal images with generative adversarial nets. *Medical Image Analysis*, *49*, 14–26. <https://doi.org/10.1016/j.media.2018.07.001>
- Zhao, H., Yang, B., Cao, L., & Li, H. (2019). Data-Driven Enhancement of Blurry Retinal Images via Generative Adversarial Networks. In D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, & A. Khan (Eds.), *Medical Image Computing and Computer Assisted Intervention -- MICCAI 2019* (pp. 75–83). Springer International Publishing.
- Zhou, S. Kevin; Rueckert, Daniel; Fichtinger, G. (2019). *Handbook of Medical Image Computing and Computer Assisted Intervention* (1st Editio). Academic Press.
- Zhou, M., Jin, K., Wang, S., Ye, J., & Qian, D. (2018). *Color Retinal Image Enhancement Based on Luminosity and Contrast Adjustment.* *65*(3), 521–527.