






Generación de imágenes médicas en el área de la retinopatía diabética con ayuda de redes generativas antagónicas

Medical image generation in the diabetic retinopathy using generative adversarial network

C. Rioja-García ^a, M. Nakano ^a, O. U. Juárez-Sandoval ^{a,*}, K. Yanai ^b, G. Benítez-García ^b

^a Departamento de Ciencias de la Computación, Instituto Nacional Politécnico, Culhuacan, Ciudad de México, México.

^b Departamento de Ciencias de la Computación, Universidad de Electro-Comunicaciones, Chofu, Tokio, Japón.

Resumen

Una de las enfermedades que más afectan al sistema visual humano es la Retinopatía Diabética (RD), siendo esta una de las principales causas de ceguera a nivel mundial. Esta enfermedad se deriva de la Diabetes. Es importante que los oftalmólogos puedan detectar esta enfermedad a tiempo para poder darle un tratamiento adecuado. Se han propuesto trabajos de detección del grado de RD y detecciones de lesiones provocadas por la RD. Para mejorar la precisión de estos algoritmos es necesario entrenarlos de una manera correcta con una basta base de datos segmentada. A la fecha las bases de datos existentes en la RD contienen un número limitado de imágenes. Por lo cual se propone incrementar el número de imágenes de RD con ayuda de SinGAN (Aprendizaje de un modelo generativo a partir de una sola imagen natural) con esta red es posible crear nuevas imágenes a partir del entrenamiento de una sola imagen.

Palabras Clave: Aprendizaje Profundo, Retinopatía Diabética, Procesamiento de Imágenes, SinGAN.

Abstract

One of the diseases that most affect the human visual system is Diabetic Retinopathy (DR), being one of the main causes of blindness worldwide. This disease is derived from Diabetes. It is important for ophthalmologists to be able to detect this disease in time to be able to give it an adequate treatment. Several works have been proposed to detect the degree of DR and to detect lesions caused by DR. To improve the accuracy of these algorithms, it is necessary to train with a large database segmented in a correct way. To date, the existing DR databases contain a limited number of images. Therefore, it is proposed to increase the number of DR images with the help of SinGAN (Learning a Generative Model from a Single Natural Image). Using this network it is possible to create new images from a single image as the input.

Keywords: Deep Learning, Diabetic Retinopathy, Image processing, SinGAN.

1. Introducción

En el ámbito médico, el tratamiento de las enfermedades es más eficaz cuando se detecta en una fase temprana. La Retinopatía Diabética (RD) es una enfermedad ocular, siendo la principal causa de discapacidad visual en todo el mundo. Esta enfermedad se detecta mediante un examen sistemático del fondo del ojo (Li et al., 2019). La RD es una complicación de la diabetes que provoca que los vasos sanguíneos de la retina se hinchen, y por consecuencia se pierden líquidos y sangre (Boyd, 2020). La RD se detecta por la aparición de diferentes tipos de lesiones en una imagen de la retina. Las principales lesiones que se observan por la RD son las

siguientes: Microaneurismas (MA), Exudados blandos y duros (EX) y Hemorragias intraretinales (HI), (Alyoubi et al., 2020).

La pérdida de visión resultante de la RD puede prevenirse con un tratamiento oportuno (Chakraborty et al., 2013). Sin embargo, la detección oportuna de la RD sigue siendo un reto por varias circunstancias. La primera razón es la cantidad insuficiente de oftalmólogos para satisfacer la creciente demanda de pacientes con RD en todo el mundo, especialmente en los países en desarrollo (Resnikoff et al., 2020). Para poder cubrir la demanda de los pacientes con RD, se requiere otorgar a los oftalmólogos sistemas que puedan reducir tiempos y costos. Esto se podría realizar con algoritmos de detección de objetos que ayuden a localizar las lesiones causadas por la RD.

*Autor para la correspondencia: ojuarezs@ipn.mx

Correo electrónico: cesarrg@me.com (César Rioja García), mnakano@ipn.mx (Mariko Nakano), ojuarezs@ipn.mx (Oswaldo Ulises Juárez Sandoval), yanai@cs.uec.ac.jp (Keiji Yanai), gibrán@ieee.org (Gibrán de Jesús Benítez García).

Historial del manuscrito: recibido el 26/04/2023, última versión-revisada recibida el 05/05/2023, aceptado el 22/06/2023 en línea (postprint) desde el 03/07/2023, publicado el 05/01/2024. **DOI:** <https://doi.org/10.29057/icbi.v11i22.11022>



Es importante localizar las áreas afectadas y el tipo de lesión ocasionada por la RD, ya que como observamos en la Tabla 1 la presencia de hemorragias intraretinarias (HI) como lesión en la retina puede indicar una RD no proliferativo moderada, haciendo que el riesgo de la enfermedad del paciente aumente.

La detección de objetos en la RD es una tarea compleja.

- La técnica de aumento de datos no resuelve el problema de desbalance de datos. Por ejemplo, las imágenes del fondo del ojo con neovasos son sumamente limitadas, aunque la detección de este tipo de lesiones tiene una gran importancia para evitar la ceguera del paciente.

Tabla 1: Grados de RD

Grado	Severidad	Síntomas
DR0	No hay RD aparente	No hay presencia de ninguna lesión de RD
DR1	RD no proliferativo leve	Presencia solamente de MA, no hay HI
DR2	RD no proliferativo moderada	Presencia de cuando menos una HI
DR3	RD no proliferativo severa	Presencia de más de 20 HI en cada cuadrante
DR4	RD proliferativo	Además de 3, presencia de neovasos o hemorrágica vítrea

Para crear un modelo de detección de lesiones con la ayuda del aprendizaje profundo se requiere un previo entrenamiento. Este entrenamiento se recomienda que sea con el mayor número de imágenes posibles, con esto podemos asegurar excelentes resultados al momento de evaluar la precisión del modelo. Hay que solucionar varios retos que se presentan para crear un algoritmo que tenga una excelente precisión.

Las bases de datos que existen sobre la RD carecen de un balance entre los datos, esto quiere decir que se tiene una desproporción de imágenes como se muestra en la Figura 1 (He et al., 2020, Wei et al. 2020), de los diferentes grados de la RD que se muestran en la Tabla 1. La segmentación de lesiones que se otorgan en estas bases de datos, son reducidas, y como hemos dicho anteriormente, la localización de las lesiones generadas por la RD es de gran importancia para la asignación del grado de enfermedad y determinar el riesgo en el paciente.

La Red Generativa Antagónica es creada por Goodfellow et al., (2014) para sintetizar señales o imágenes, la cual, presenta una alternativa para incrementar el número de datos de entrenamiento. La DCGAN (Radford et al. 2016), se han usado para aumentar datos de entrenamiento, sintetizando señales de Electromiografía (Aicent et al. 2020) y Electroencefalografía (Wang et al. 2018). Ambos trabajos sintetizan señales unidimensionales, ya que, en los dos campos no presentan el problema de desbalance de datos y la variación de formas de señales es relativamente pequeña.

Como se había mencionado anteriormente, el número de imágenes con algunas lesiones de RD son sumamente limitadas, lo que dificulta el uso de la mayoría de las GANs, tales como DCGAN (Radford et al. 2016), WGAN (Arjovsky et al. 2017), CycleGAN (Zhu et al. 2017) y SRGAN (Ledig et al. 2017), que requieren una gran cantidad de datos de

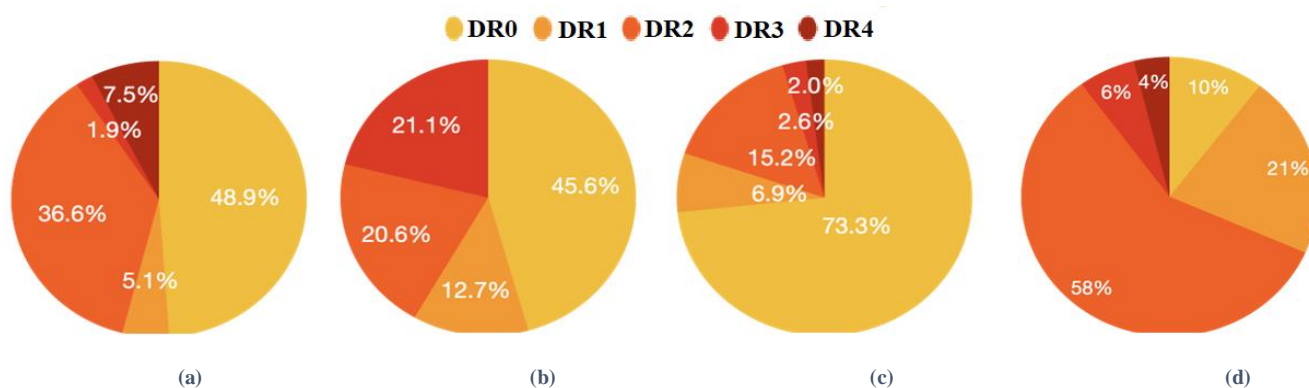


Figura 1: Desbalance de las bases de datos. (a) DDR Dataset, (b) Messidor, (c) EyePACS y (d) Retinal Lesions

La técnica de aumento de datos en aprendizaje profundo es un método establecido para incrementar artificialmente el número de datos de entrenamiento (Shorten et al. 2019). En el caso de imágenes, se aplican rotación, traslación, inversión, escalamiento entre otros procesamientos de imágenes. Sin embargo, esta técnica no es suficiente para la tarea de detección de lesiones y asignación del grado de la RD por las siguientes dos razones:

- Diferencia de objetos comunes, como perro, bicicleta, coche, etc., las lesiones de la RD no tienen una forma determinada. La técnica de aumento de datos no genera lesiones con diferente forma, las cuales son requeridas para un entrenamiento más adecuado.

entrenamiento para sintetizar imágenes con ciertas características. Tomando en cuenta lo anterior, este trabajo tiene como intención explorar la generación de nuevas imágenes en el área de la RD, para aumentar el número de imágenes, con las cuales se pueda ayudar a entrenar los modelos de detección de lesiones de RD. Para resolver los obstáculos antes mencionados, en este artículo se propone crear imágenes sintéticas del fondo del ojo y de lesiones de RD mediante el uso de SinGAN (Shaham et al, 2019). ya que SinGAN requiere solo una imagen para el entrenamiento y tiene diferentes usos, tales como armonización y edición, que son útiles para generar imágenes sintéticas con lesiones de RD. Las lesiones que se seleccionaron para este artículo únicamente son las HI, como se menciona anteriormente con

base a la Tabla 1, la presencia de estas lesiones incrementa el riesgo en los pacientes.

El resto de este artículo está organizado de la siguiente manera. En la sección 2 se describen las características y estructura de la red SinGAN. El desarrollo de este artículo junto con los resultados obtenidos se detalla en la sección 3. Finalmente, en la sección 4 se describe la conclusión de este trabajo.

2. SinGAN

Las Redes Generativas Antagónicas son modelos de aprendizaje profundo que tienen dos redes neuronales, una de ellas es la red neuronal generadora. La red neuronal generadora tiene como objetivo crear nuevas imágenes a partir de los datos de entrenamiento.

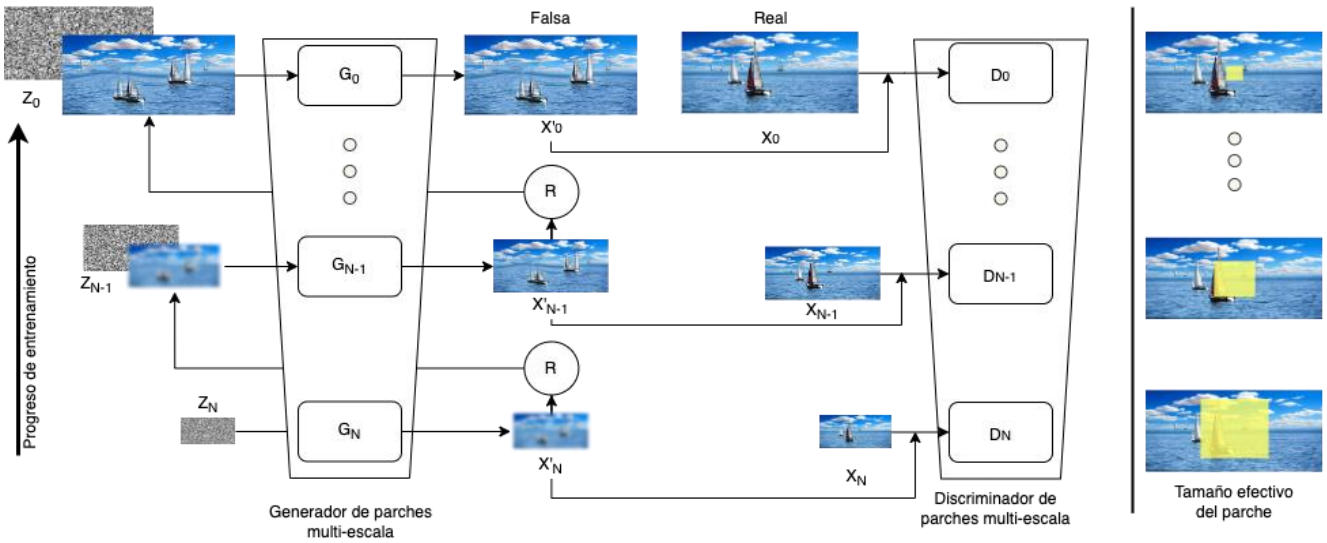


Figura 3: Esquema en pirámide del algoritmo SinGAN

Por otro lado, tenemos la red neuronal discriminadora. Esta red neuronal tiene como tarea identificar las imágenes creadas por la red neuronal generadora y el conjunto de datos con el cual se haya entrenado. Los resultados de las imágenes generadas por las GANs han sido muy satisfactorios, dando imágenes casi realistas.

Uno de los tipos de redes GANs es el modelo de SinGAN el cual es un modelo que a partir de una sola imagen es capaz de generar nuevas imágenes aleatorias con estructuras y texturas complejas, relacionadas a la imagen de entrada como se ilustra en la Figura 2.



Figura 2: a) Imagen original. b) Imagen generada por el modelo.

SinGAN, va más allá de la detección de bordes y texturas, para ello proponen capturar la información de la imagen, trabajando con diferentes escalas en la imagen, como se muestra en la Figura 3. El modelo de SinGAN consiste en una pirámide de generadores $\{G_0, \dots, G_N\}$ entrenada contra otra pirámide de discriminadores $\{D_0, \dots, D_N\}$. El entrenamiento empieza con el generador de baja resolución G_N , introduciendo

una imagen de ruido Gaussiano Z_N como entrada única. El generador G_N y el discriminador D_N en combinación con la imagen original X_N generan una imagen similar X'_N y del mismo tamaño que Z_N . Después la imagen generada X'_N se sobre muestra con un factor R para obtener una imagen de entrada que se introduce junto con la imagen de ruido Gaussiano Z_{N-1} en la etapa superior de la pirámide, con la cual se genera una imagen X'_{N-1} . Así sucesivamente, la imagen de mayor escala y resolución se genera con la imagen sobre muestreada de la salida de la etapa anterior junto con la imagen de ruido Gaussiano.

SinGAN se puede usar en diversas aplicaciones, tales como edición, super-resolución, animación, etc. El número de etapas N es un factor que debe ajustarse dependiendo de la aplicación.

La Figura 4 muestra un generador de etapa n , G_n , de SinGAN, en donde se introduce la imagen sobre muestreada de X'_{n+1} y la imagen de ruido Gaussiano Z_n como entradas. El generador tiene una arquitectura residual de 5 bloques que consisten en una capa de convolución de 3×3 , normalización de lotes (BN) y activación "Reaky ReLU". La salida del generador es una nueva imagen X'_n con mayor escala que X'_{n+1} , la cual será utilizada junto con X_n en el discriminador D_n (Shaham et al, 2019).

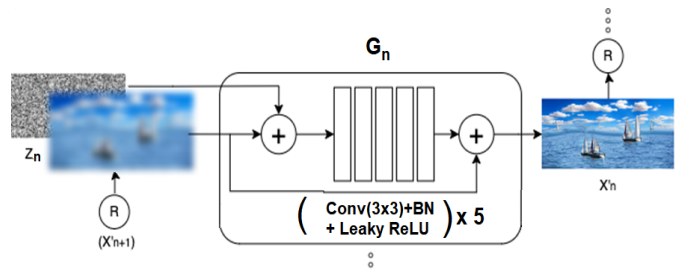


Figura 4: Generación de escala sencilla

La función de pérdida de la etapa n de SinGAN es la suma de la pérdida antagónica \mathcal{L}_{adv} y la pérdida de reconstrucción \mathcal{L}_{rec} , como se muestra en (1).

$$\min_{G_n} \max_{D_n} \mathcal{L}_{adv}(G_n, D_n) + \alpha \mathcal{L}_{rec}(G_n) \quad (1)$$

La función de pérdida antagonista $\mathcal{L}_{adv}(G_n, D_n)$ es una función común de GAN y la función de pérdida de reconstrucción está dada por (2).

$$\mathcal{L}_{rec} = \|G_n(0, R(X'_{n+1})) - X_n\|^2 \quad (2)$$

donde $\|\cdot\|^2$ es la norma L2, y $R(X'_{n+1})$ es el resultado del sobre muestreo de X'_{n+1} . La función de pérdida de la primera etapa se expresa como: $\mathcal{L}_{rec} = \|G_N(Z_N) - X_N\|^2$, ya que en esta etapa se genera de una imagen de ruido una imagen parecida a X_N .

Cuando el entrenamiento de SinGAN ha finalizado, este modelo es capaz de generar imágenes aleatorias con las características de la imagen con la que se ha entrenado el modelo (Shaham et al, 2019).

3. Resultados experimentales

Como hemos mencionado anteriormente uno de los grandes problemas en el campo de la RD y del aprendizaje profundo, es la limitada cantidad de imágenes de muestras del fondo del ojo. La intención de estos resultados es aumentar las imágenes que se tienen de la RD, esto ayudará a tener más muestras para su entrenamiento y conseguir mejorar los resultados de los modelos de aprendizaje profundo aplicados a la detección del grado de la RD.

Las imágenes que se han seleccionado para trabajar provienen de la base de datos Retinal Lesions (Wei et al., 2021). En esta base de datos han colaborado 45 oftalmólogos segmentado 12K imágenes con 8 lesiones provocadas por la RD, dando un total de 290k segmentaciones manuales en la base de datos.

Se han realizado diferentes combinaciones de imágenes, como generar imágenes nuevas, editarlas y armonizarlas. Las imágenes de cada grado de la RD son seleccionadas aleatoriamente, tomando como criterio la asignación que otorga la base de datos Retinal Lesions y la asignación de grado que otorga Kaggle.

Este trabajo tiene dos resultados de síntesis usando dos técnicas diferentes de SinGAN: edición y armonización. Edición es explorar la generación de una nueva imagen a partir de la imagen real de la base de datos con un grado de RD. La imagen sintetizada debe pertenecer al mismo grado de RD que la imagen de entrada. Por otro lado, en el proceso de armonización, las lesiones obtenidas de una imagen real se superponen en otra imagen sana para generar artificialmente una imagen nueva con lesiones de RD. Como se había mencionado anteriormente la lesión seleccionada es HI.

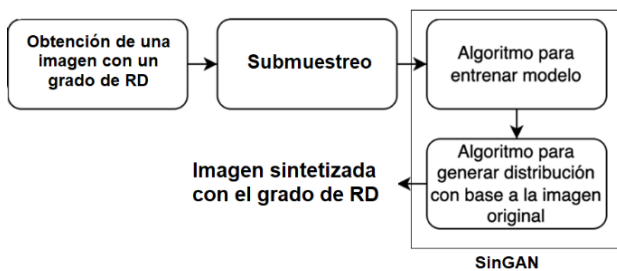


Figura 5: Proceso de Edición

3.1. Síntesis de imágenes en diferentes grados de RD

Para sintetizar imágenes con diferentes grados de RD, se utilizó la técnica de edición de SinGAN, la cual está mostrada en la Figura 5. en donde se selecciona aleatoriamente una imagen del conjunto de imágenes etiquetadas con un grado de RD (DR0, DR1, ..., DR4). La imagen seleccionada es la imagen de entrada X_0 de SinGAN. A esta imagen se le aplica el proceso de submuestreo para obtener las imágenes de baja escala $X_1, X_2, \dots, X_{N-1}, X_N$. Después de entrenar el modelo de SinGAN desde la etapa N hasta etapa 0, como se muestra en la Figura 3, se obtiene una imagen sintetizada del mismo grado de RD.

3.1.1. Síntesis de imagen de grado 0 de RD

En esta prueba se selecciona una imagen sana mostrada por la Figura 6a de grado 0 de RD para entrenar el modelo. El resultado generado por el modelo es la figura 6b. Las Figuras 7a y 7b son ampliaciones de una sección de las Figuras 6a y 6b. En la Figura 7a podemos observar la calidad y algunas características de la imagen original con la que se está entrenando, y en la Figura 7b podemos ver con más detalle la calidad y las características de la imagen generada por el modelo.

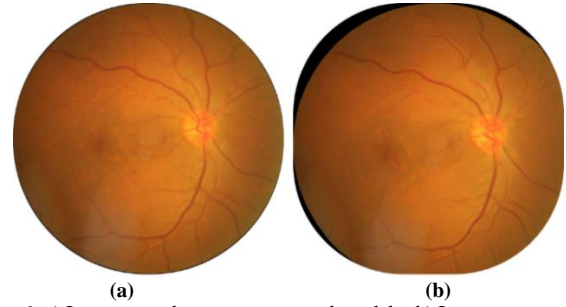


Figura 6: a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

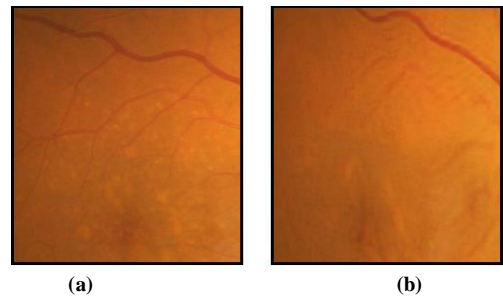


Figura 7: Ampliación de las imágenes de la figura 6. a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

La imagen sintetizada es parecida a la imagen original, sin embargo, existen diferencias entre ambas imágenes, por ejemplo, la forma de los vasos sanguíneos y la textura del fondo del ojo. Esta diferencia no se puede generar por la técnica de aumento de datos.

3.1.2. Síntesis de imagen de grado 1 de RD

En esta prueba se selecciona una imagen del grado 1 de RD (Figura 8a) para entrenar el modelo. El resultado generado por el modelo es la Figura 8b. Las Figuras 9a y 9b son ampliaciones de una sección de las Figuras 8a y 8b. En la Figura 9a podemos observar la calidad y algunas

características de la imagen con la que se está entrenando, y en la Figura 9b podemos ver con más detalle la calidad y las características de la imagen generada por el modelo.

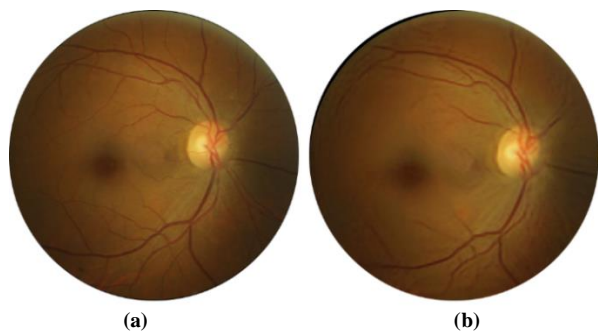


Figura 8: a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

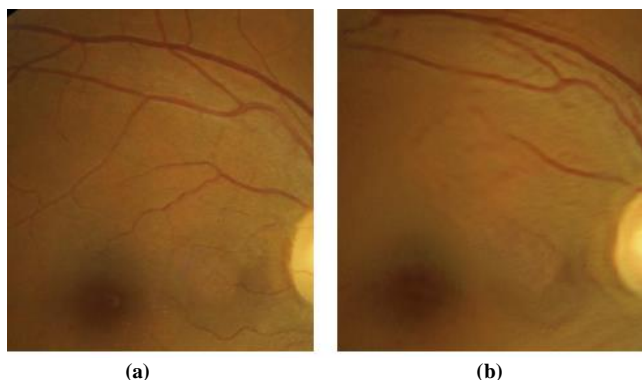


Figura 9: Ampliación de las imágenes de la Figura 8. a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

Nuevamente, la imagen sintetizada es muy parecida a la imagen original, sin embargo, tiene una ligera diferencia que no se puede generar por la técnica de aumento de datos.

3.1.3. Síntesis de imagen de grado 2 de RD

En esta prueba se selecciona una imagen del grado 2 de RD (Figura 10a) para entrenar el modelo. El resultado generado por el modelo es la Figura 10b. La Figura 11a es la ampliación de una sección de la Figura 10a, en la cual podemos observar la calidad y algunas características de la imagen con la que se está entrenando el modelo, y las Figuras 11a y 11b son una ampliación de una sección de la Figura 10a y 10b, respectivamente.

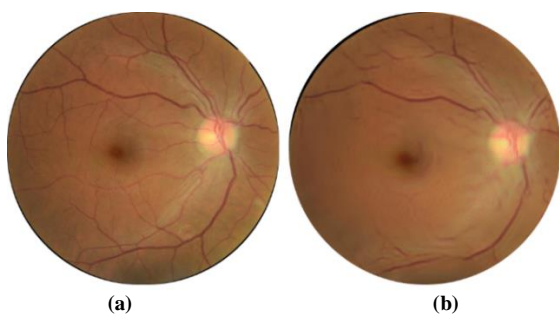


Figura 10: a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.



Figura 11: Ampliación de las imágenes de la Figura 10. a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

En esta prueba, se produjo una imagen borrosa, ya que SinGAN introduce una imagen de ruido Gaussiano, y no se puede controlar la calidad de la imagen sintetizada.

3.1.4. Síntesis de imagen de grado 3 de RD

En esta prueba se selecciona una imagen del grado 3 de RD (Figura 12a) para entrenar el modelo. El resultado generado por el modelo es la Figura 12b. Las Figuras 13a y 13b son ampliaciones de una sección de las Figuras 12a y 12b, en las cuales podemos observar la calidad y algunas características de las imágenes con la que se está entrenando y su respectivo resultado.

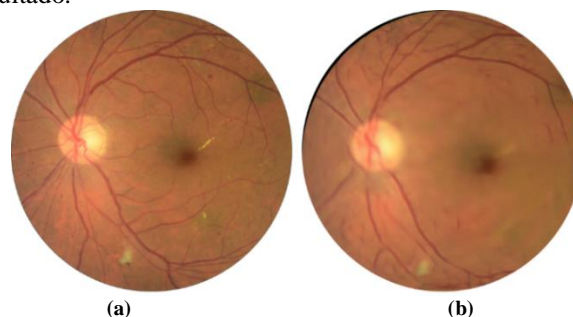


Figura 12: a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

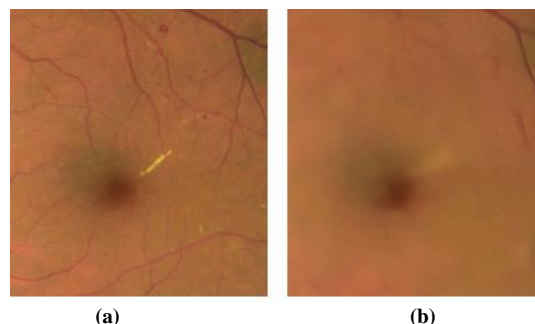


Figura 13: Ampliación de las imágenes de la Figura 12. a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

3.1.5. Síntesis de imagen de grado 4 de RD

En esta prueba se selecciona una imagen de grado 4 de RD (Figura 14a) para entrenar el modelo. El resultado generado por el modelo es la Figura 14b. Las Figuras 15a y 15b son ampliaciones de una sección de las Figuras 14a y 14b, respectivamente, en las cuales podemos observar la calidad y algunas características de ambas imágenes.

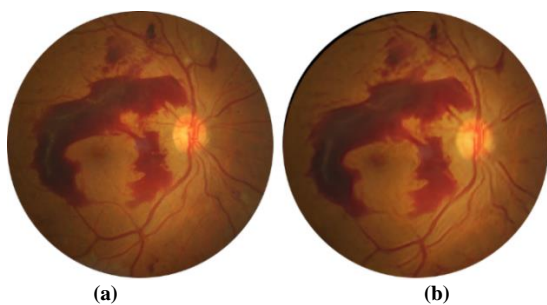


Figura 14: a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

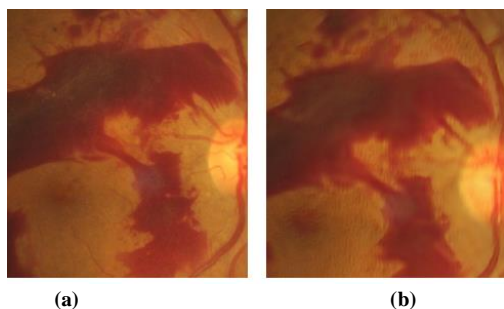


Figura 15: Ampliación de las imágenes de la Figura 14. a) Imagen con la que se entrena el modelo. b) Imagen generada por el modelo.

En este caso, SinGAN produjo una imagen muy similar a la imagen de entrenamiento, pero con un ligero cambio en la forma de la Hemorragia, como se puede observar en la Figura 15.

3.2. Armonización de imágenes sanas con imágenes insanas

Una imagen sintética del fondo del ojo con las lesiones de HIs se puede generar usando la técnica de armonización de SinGAN, introduciendo una imagen sana junto con las lesiones HIs de la imagen insana. El proceso de armonización se puede observar en la Figura 16.

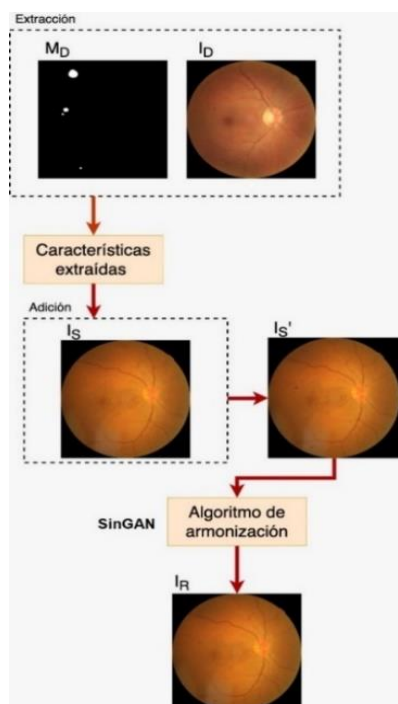


Figura 16: Algoritmo de Armonización.

El procedimiento del algoritmo de armonización es el siguiente:

1. La etapa de extracción consiste en obtener las áreas de las lesiones de una imagen (I_D) con daño causado por la RD. Esta información la obtenemos a partir de una imagen de las lesiones segmentadas (M_D) proporcionada por la base de datos Retinal Lesions (Wei et al. 2020), de esta manera podemos extraer el área dañada. Esta etapa da como resultado el área con las lesiones de una imagen del fondo del ojo dañado por la RD.
2. Las áreas de las lesiones que fueron extraídas de la imagen I_D se adicionan a las lesiones de la imagen sana I_S , como resultado se obtiene una imagen nueva I_S' , misma que representa a una imagen sana con lesiones añadidas; en la Figura 17a podemos observar que las lesiones añadidas se distinguen fácilmente del resto de la imagen.
3. Finalmente, la técnica de armonización de SinGAN se aplica a la imagen I_S' . Este proceso de armonización provee como resultado una imagen totalmente nueva. Como se puede observar las lesiones que se han añadido a la imagen sana I_S tienen un aspecto similar al resto de la imagen insana, dando como resultado una imagen insana que se puede clasificar como una imagen con lesiones HIs producidas por la RD.

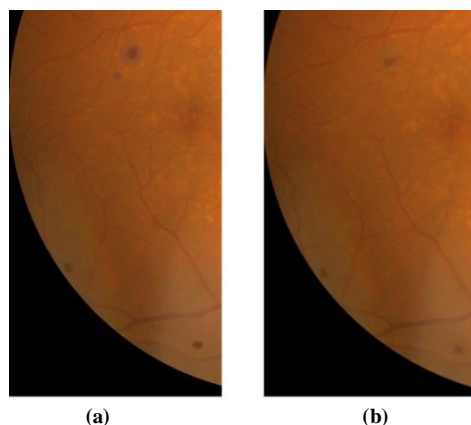


Figura 17: (a) Resultado de adición de lesión I_S' , (b) Resultado de Armonización I_R .

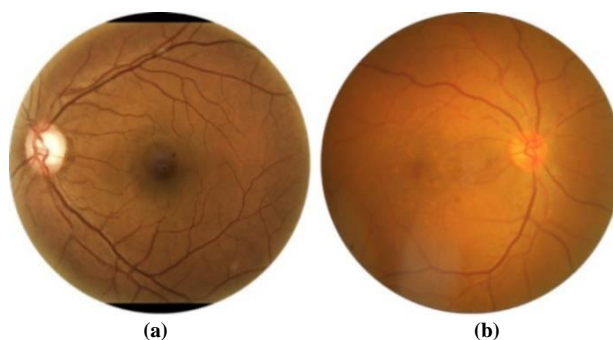


Figura 18: (a) Imagen con hemorragia (b) Imagen sana.

Las Figuras 18 y 19 muestran otro ejemplo de armonización, la Figura 18a es una imagen con HI y la Figura 18b es una imagen sana seleccionada aleatoriamente de la base de imágenes.

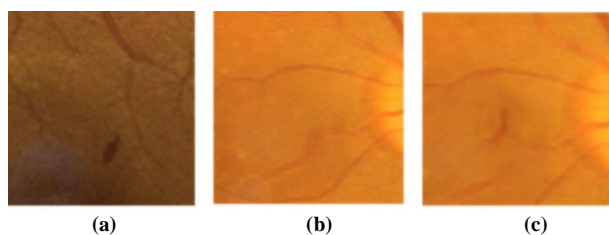


Figura 19: (a) Imagen con hemorragia (b) Imagen sana (c) Imagen nueva con hemorragia

Las Figuras 19a y 19b son ampliaciones de una sección de las Figuras 18a y 18b. En la Figura 19a se observa la lesión de HI y está lesión se sobrepone en la región cerca del disco óptico de la Figura 19b. La Figura 19c muestra el resultado de la armonización, donde se observa la lesión HI en la imagen sana. Se puede observar que en la Figura 19c, la lesión HI está armonizada con la textura alrededor de la lesión, produciendo un aspecto más natural. Usando esta técnica de armonización de SinGAN, podemos incrementar el número de imágenes con ciertas lesiones de RD, contribuyendo a mejorar el proceso de entrenamiento.

4. Conclusiones

Hoy en día el aprendizaje profundo juega un papel muy importante en la vida diaria, facilitando algunas tareas médicas de rutina, como puede ser el caso de RD. Dado el gran número de pacientes diagnosticados con RD y el limitado número de oftalmólogos, se deben proponer métodos automatizados de diagnóstico de la RD.

Este trabajo se realizó para ayudar a mejorar el funcionamiento de los algoritmos basados en el aprendizaje profundo para la asignación del grado de la RD y detección de lesiones. El principal obstáculo para obtener un buen funcionamiento del algoritmo es la limitada cantidad de datos de entrenamiento y el desbalance de datos entre los diferentes grados de RD y los tipos de lesiones; aunado a esto, la forma irregular de las lesiones impide que la técnica de aumento de datos sea eficiente. Para solucionar este obstáculo, en este trabajo se presentó el uso de un modelo GAN, especialmente SinGAN; el cual, se diferencia de otros GANs ya que solamente requiere una imagen de entrenamiento para obtener imágenes similares. Esta cualidad de SinGAN es favorable en casos donde la base de datos es limitada como el de la RD. Para obtener imágenes sintéticas del fondo de los ojos se emplearon dos técnicas del SinGAN, “edición” y “armonización”. Usando la “edición”, se generaron imágenes sintéticas a partir de las imágenes del fondo de los ojos con diferentes grados de RD. Mientras que empleando la técnica de “armonización”, se armonizaron lesiones HIs en imágenes sanas, segmentando las lesiones de la imagen insana y sobreponiéndolas en las imágenes sanas.

Las imágenes generadas por el modelo propuesto mantienen la estructura del fondo de los ojos, sin embargo, se puede observar que la calidad de estas imágenes generadas por la técnica de “edición” podría ser insuficiente para la asignación de un grado de RD. Mientras que la técnica de “armonización” puede ser un enfoque prometedor, ya que los resultados obtenidos son reales. Usando esta técnica, se pueden sintetizar imágenes del fondo de los ojos con lesiones de RD. Para determinar el uso real del sistema propuesto, tanto la

técnica de “edición” como la de “armonización”, requieren evaluaciones por oftalmólogos expertos.

Como trabajo futuro, estas imágenes sintéticas serán sometidas a la evaluación de oftalmólogos para dar una calificación numérica del sistema y así poder evaluar la veracidad y precisión que este tipo de imágenes pudiera llegar a tener. Así mismo, se evaluará el sistema automatizado basado en el aprendizaje profundo. La evaluación se realizará haciendo un entrenamiento con imágenes reales e imágenes sintéticas, mismas que son obtenidas de diversas imágenes reales con diferentes grados, diferentes posiciones y formas de lesiones de la RD.

Recientemente, se han propuesto otros GANs, tales como ConSinGAN (Hinz et al. 2021) y One-Shot GAN (Sushko et al. 2021) que se entrenan con solo una imagen y están mostrando buen desempeño en la tarea de “edición”. El uso de estos modelos de GANs para generar imágenes sintéticas con la modalidad de “edición” pueden ser una alternativa adicional al objetivo de este trabajo.

Referencias

- Alyoubi, W. L., Shalash, W. M., Abulkhair, M. F. (2020). Diabetic retinopathy detection through deep learning techniques: A review. *Informatics in Medicine Unlocked*, 20, 100377.
- Anicet Zanini, Rafael; Luna Colombini, Esther (2020). Parkinson's Disease EMG Data Augmentation and Simulation with DCGANs and Style Transfer, *Sensor* 20(9)
DOI:10.3390/s20092605.
- Arjovsky, M., Chintala S., Bottou, L. (2017) Wasserstein generative adversarial networks 34th International Conference on Machine Learning, ICML 2017, 1, pp. 298-321
- Boyd, K. (2020). American Academy of Ophthalmology-What is Diabetic Retinopathy. Accessed: Sep, 10, 2021.
- Chakraborty, C., Gupta, B., Ghosh, S. K. (2013). A review on telemedicine-based WBAN framework for patient monitoring. *Telemedicine and e-Health*, 19(8), 619-626.
DOI: 10.1089/tmj.2012.0215,
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- He, A., Li, T., Li, N., Wang, K., Fu, H. (2020). CABNet: category attention block for imbalanced diabetic retinopathy grading. *IEEE Transactions on Medical Imaging*, 40(1), 143-153.
DOI: 10.1109/TMI.2020.3023463
- Hinz, T., Fisher, M., Wang, O., Wermter, S. (2021) Improved Techniques for Training Single-Image GANs, *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1299-1308,
DOI: 10.1109/WACV48630.2021.00134.
- Ledig C., et al. (2017), Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 105-114
DOI: 10.1109/CVPR.2017.19.
- Li, T., Gao, Y., Wang, K., Guo, S., Liu, H., Kang, H. (2019). Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. *Information Sciences*, 501, 511-522.
DOI: 10.1016/j.ins.2019.06.011
- Redford, A., Metz L., Chintala S. (2016), Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, 4th International Conference on Learning Representations, ICLR 2016
- Resnikoff, S., Lansingh, V. C., Washburn, L., Felch, W., Gauthier, T. M., Taylor, H. R., Wiedemann, P. (2020). Estimated number of ophthalmologists worldwide will we meet the needs?. *British Journal of Ophthalmology*, 104(4), 588-592.
DOI: 0.1136/bjophthalmol-2019-314336
- Shaham, T. R., Dekel, T., Michaeli, T. (2019). SinGAN: Learning a generative model from a single natural image. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4570-4580.
- Shorten, Connor; Khoshgoftaar, Taghi M. (2019). A survey on Image Data Augmentation for Deep Learning. *Mathematics and Computers in Simulation*. 6(60).
DOI:10.1186/s40537-019-0197-0

- Sushko, V. Gall, J., Khoreva, A., (2021) One-Shot GAN: Learning to Generate Samples from Single Images and Videos," IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 2596-2600. DOI: 10.1109/CVPRW53098.2021.00293.
- Wang, F., Zhong, S., Peng, J., Jiang, J., Liu, Y., (2018). Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks. *Multimedia Modeling. Lecture Notes in Computer Science*. 10705. pp.82–93
DOI:10.1007/978-3-319-73600-6_8
- Wei, Q., Li, X., Yu, W., Zhang, X., Zhang, Y., Hu, B., Chen, Y. (2021). Learn to segment retinal lesions and beyond. In 2020 25th International Conference on Pattern Recognition, pp. 7403-7410. DOI: 10.1109/CVPRW53098.2021.00