**Informe sobre la Selección de Arquitecturas para la Clasificación de Enfermedades Oculares del Ciclo 2**

# **1. Planteamiento del Problema**

Las discapacidades visuales son una preocupación global significativa. En Colombia, alrededor de 4000 personas por millón requieren servicios de baja visión, con el glaucoma y la retinopatía diabética siendo las principales causas. La detección temprana es crucial, y este proyecto busca mejorar la precisión del diagnóstico mediante el uso de redes neuronales convolucionales pre-entrenadas.

# **2. Proceso de Evaluación y Selección de Arquitecturas**

## 2.1. Metodología

Se evaluaron tres arquitecturas (VGG19, ResNet50, y EfficientNet) utilizando un dataset con imágenes de fondo de ojo de cuatro clases, utilizando la tasa de aprendizaje que permitió obtener los mejores resultados para cada arquitectura y repitiendo los experimentos cinco veces.

## 2.2. Hiperparámetros y Configuraciones

* **Learning Rates:** 0.001 para Efficientnet y Resnet50, 0.0001 para VGG19
* **Batch Size:** 16
* **Número de Épocas:** 60
* **Repeticiones:** 5 por arquitectura
* **Función de Pérdida:** Entropía cruzada
* **Función de Activación:** Softmax
* **Optimización:** Adam
* **Validación Cruzada:** Implementada en este ciclo 2

## 2.3 Dataset utilizado

El conjunto de datos utilizado en este estudio es el "Clasificación de las enfermedades oculares", compilado por Guna Venkat Doddi en 2022. Este dataset contiene alrededor de 1,000 imágenes diagnósticas de retina para cada clase, incluyendo retinopatía diabética, cataratas, glaucoma y ojos normales. Con el fin de evaluar la capacidad del modelo para distinguir entre cada una de estas patologías, se decidió realizar la clasificación en cuatro categorías: normal, retinopatía diabética, glaucoma, catarata.

## 2.4 Ajuste de paciencia

En este estudio, se utilizó una configuración de "paciencia" (patience) que permite reducir la tasa de aprendizaje si la precisión de validación no mejora después de 5 épocas. Si la precisión de validación no mejora después de 10 épocas, el entrenamiento se detiene. Esta técnica ayudó a prevenir el sobreajuste y a mantener la eficiencia del modelo.

# **3. Análisis de Resultados**

## 3.1. VGG19

* **Varianza General:**

| learning\_rate | varianza\_val\_loss | varianza\_val\_accuracy |
| --- | --- | --- |
| 0.0001 | 0.005846542959 | 8.65E-06 |

* **Resumen General:**

| learning\_rate | repeticion | mejor\_val\_accuracy | mejor\_val\_loss |
| --- | --- | --- | --- |
| 0.0001 | 3 | 0,937 | 0.3347 |

* **Resumen resultados 0.0001:**

| repeticion | max\_val\_accuracy | epoch\_max\_val\_accuracy | min\_val\_loss | epoch\_min\_val\_loss |
| --- | --- | --- | --- | --- |
| 1 | 0.9705882072 | 12 | 0.4250106812 | 59 |
| 2 | 0.9705882072 | 12 | 0.4011312723 | 58 |
| 3 | 0.9705882072 | 9 | 0.3346978128 | 59 |
| 4 | 0.9632353187 | 8 | 0.5517255664 | 58 |
| 5 | 0.9705882072 | 13 | 0.353236258 | 41 |

**Discusión:**

La arquitectura **VGG19** mostró resultados consistentes, alcanzando una **precisión máxima de validación del 93.7%** en la tercera repetición, con una **pérdida mínima de validación de 0.3347**. Los valores de varianza, tanto en **precisión de validación** como en **pérdida de validación**, fueron bajos: 8.65E-06 y 0.0058, respectivamente; estos valores lo posicionan en el segundo puesto en términos de varianza de pérdida de validación, último puesto en términos de mejor pérdida de validación, y segundo puesto en términos de precisión, comparado con las otras dos arquitecturas. La tasa de aprendizaje utilizada fue de 0.0001. A pesar de su alto rendimiento, los valores de pérdida en algunas repeticiones fueron más elevados, lo que sugiere que la arquitectura podría estar más sensible a ciertos patrones en los datos (Simonyan & Zisserman, 2014). Cabe destacar un punto a favor de la arquitectura, y es que la simplicidad del diseño de VGG19, basada en capas convolucionales apiladas, permite una fácil adaptación y ajuste del modelo (Simonyan & Zisserman, 2014); sin embargo, las variaciones en la pérdida de validación sugieren que existe una menor estabilidad comparada con las otras arquitecturas evaluadas.

## 3.2. ResNet50

* **Varianza General:**

| learning\_rate | varianza\_val\_loss | varianza\_val\_accuracy |
| --- | --- | --- |
| 0.001 | 0.0009785415549 | 6.07E-06 |

* **Resumen General:**

| learning\_rate | repeticion | mejor\_val\_accuracy | mejor\_val\_loss |
| --- | --- | --- | --- |
| 0.001 | 1 | 0.952 | 0.2899 |

* **Resumen resultados 0.001:**

| repeticion | max\_val\_accuracy | epoch\_max\_val\_accuracy | min\_val\_loss | epoch\_min\_val\_loss |
| --- | --- | --- | --- | --- |
| 1 | 0.9541062713 | 23 | 0.2899044752 | 29 |
| 2 | 0.951690793 | 21 | 0.337315768 | 21 |
| 3 | 0.9468598962 | 15 | 0.3696849346 | 15 |
| 4 | 0.9492753744 | 14 | 0.3782716393 | 14 |
| 5 | 0.9492753744 | 13 | 0.3542591035 | 21 |

**Discusión:**

**ResNet50** alcanzó una **precisión de validación máxima del 95.2%** con una **pérdida de validación mínima de 0.2899** en la primera repetición. La varianza en la precisión de validación fue extremadamente baja (6.07E-06), lo que indica que el modelo mantuvo un rendimiento constante a través de las diferentes repeticiones estos valores lo posicionan en el primer puesto en términos de varianza de pérdida de validación, segundo puesto en términos de mejor pérdida de validación, y último puesto en términos de precisión, comparado con las otras dos arquitecturas. Algunos valores de pérdida de validación fueron más altos, esto sugiere que podría haber margen para una mayor optimización en los parámetros del modelo. La tasa de aprendizaje óptima para ResNet50 fue 0.001, similar a EfficientNet. Cabe destacar un punto a favor de la arquitectura; su diseño está caracterizado por sus bloques residuales, esto facilita la transmisión de gradientes a través de sus capas más profundas, generalmente dando como resultado una excelente precisión sin problemas de sobreajuste (Vipas.Ai, 2024); adicionalmente, este modelo demuestra, en la mayoría de casos, una capacidad estable para manejar características complejas en las imágenes de retina, con una baja varianza en sus resultados y un buen equilibrio entre precisión y pérdida de validación (S. Ortiz & M. A. Goenaga Jimenez, 2023).

## 3.3. EfficientNet (Versión B3)

* **Varianza General:**

| learning\_rate | varianza\_val\_loss | varianza\_val\_accuracy |
| --- | --- | --- |
| 0.001 | 0.001006149902 | 3.59E-05 |

* **Resumen General:**

| learning\_rate | repeticion | mejor\_val\_accuracy | mejor\_val\_loss |
| --- | --- | --- | --- |
| 0.001 | 3 | 0.959 | 0.26 |

* **Resumen resultados 0.001:**

| repeticion | max\_val\_accuracy | epoch\_max\_val\_accuracy | min\_val\_loss | epoch\_min\_val\_loss |
| --- | --- | --- | --- | --- |
| 1 | 0.9541062713 | 13 | 0.3016205132 | 18 |
| 2 | 0.9541062713 | 17 | 0.3219943941 | 25 |
| 3 | 0.9613526464 | 25 | 0.2600095868 | 26 |
| 4 | 0.944444418 | 16 | 0.3570893109 | 16 |
| 5 | 0.9468598962 | 11 | 0.2989785075 | 21 |

**Discusión:**

**EfficientNet** se destacó como la mejor arquitectura evaluada, presentando los mejores resultados en cuanto a precisión y pérdida de validación. Con una tasa de aprendizaje de 0.001, este modelo mostró una precisión constante y una pérdida de validación más baja que las demás arquitecturas, incluso si resultó obtener el último puesto en términos de varianza de pérdida de validación. Debido a la manera en la que EfficientNet ajusta de manera eficiente la resolución de entrada, la profundidad y el número de filtros, esta arquitectura es particularmente eficiente y adecuada para su implementación en sistemas con limitaciones computacionales (Elmoufidi & Amoun, 2021).

.

# **4. Conclusiones y Justificación de Selección**

## 4.1 Conclusiones con respecto a las tablas de varianza, tablas generales y tablas por learning rate

A partir lo anterior concluimos que:

* **EfficientNet** fue seleccionada por su excelente rendimiento general y su eficiencia en la utilización de recursos. Con los mejores resultados tanto en precisión como en pérdida de validación, es ideal para la clasificación de enfermedades oculares, especialmente en entornos con limitaciones de recursos computacionales.
* **ResNet50** presenta un equilibrio entre precisión y pérdida de validación. Su arquitectura residual permite un entrenamiento más eficiente y estable, especialmente en tareas que requieren la extracción de características complejas, como la clasificación de enfermedades oculares. Sin embargo, en algunas repeticiones mostró una pérdida de validación ligeramente más alta que EfficientNet, esta inestabilidad fue un factor clave para no seleccionarla como la arquitectura final para el ciclo 3.
* **VGG19**, a pesar de que destaca por su simplicidad y capacidad de ofrecer resultados de alto rendimiento, mostró algunas variaciones en la pérdida, esto y su sensibilidad a patrones de datos no lo hacen un candidato final para modelo final en el ciclo 3.

## 4.2 Conclusiones con respecto a las tablas de métricas de cada arquitectura y matrices de confusión:

Adicionalmente y a partir de las matrices de confusión y las tablas de métricas para los modelos ResNet50, VGG19 y EfficientNet (las cuales se encontrarán en la hoja de cálculo adjunta en Anexos) , se puede realizar el siguiente análisis comparativo:

### **1. Exactitud General (Accuracy)**

* **VGG19**: 94%
* **EfficientNet**: 95%
* **ResNet50**: 93%

El modelo EfficientNet logra la mayor exactitud global en comparación con VGG19 y ResNet50, aunque las diferencias no son grandes. El desempeño de VGG19 es muy cercano al de EfficientNet, mientras que ResNet50 tiene un rendimiento algo menor en términos de exactitud global.

### **2. Precisión por Clase**

* **Cataract**:
  + VGG19: 0.97
  + EfficientNet: 0.99
  + ResNet50: 0.93

EfficientNet tiene la mejor precisión en la clase Catarata, mientras que VGG19 también muestra un buen desempeño. ResNet50 tiene la precisión más baja en esta clase.

* **Diabetic Retinopathy**:
  + VGG19: 0.99
  + EfficientNet: 1.00
  + ResNet50: 0.99

Todos los modelos tienen un excelente desempeño en la clase de Retinopatía Diabética, siendo EfficientNet el único que alcanza una precisión perfecta de 1.00.

* **Glaucoma**:
  + VGG19: 0.88
  + EfficientNet: 0.92
  + ResNet50: 0.89

EfficientNet también supera a los otros dos modelos en la clase de Glaucoma, con una precisión de 0.92. VGG19 tiene la menor precisión para esta clase, con 0.88.

* **Normal**:
  + VGG19: 0.94
  + EfficientNet: 0.90
  + ResNet50: 0.90

En la clase Normal, VGG19 sobresale con una precisión de 0.94, superando tanto a EfficientNet como a ResNet50, que comparten una precisión de 0.90.

### **3. Recall por Clase**

* **Cataract**:
  + VGG19: 0.97
  + EfficientNet: 0.96
  + ResNet50: 0.94

VGG19 tiene el mejor desempeño en términos de recall para la clase de Catarata, mientras que EfficientNet y ResNet50 tienen valores ligeramente más bajos.

* **Diabetic Retinopathy**:
  + Todos los modelos (VGG19, EfficientNet, ResNet50) tienen un recall perfecto de 1.00.
* **Glaucoma**:
  + VGG19: 0.96
  + EfficientNet: 0.92
  + ResNet50: 0.86

VGG19 destaca con un recall muy alto en la clase Glaucoma, con 0.96, mientras que ResNet50 tiene el valor más bajo con 0.86. EfficientNet se encuentra en un término medio con 0.92.

* **Normal**:
  + VGG19: 0.85
  + EfficientNet: 0.93
  + ResNet50: 0.91

Aquí, EfficientNet tiene el mejor recall con 0.93, superando tanto a ResNet50 como a VGG19, que queda con el valor más bajo (0.85).

### **4. F1-Score por Clase**

* **Cataract**:
  + VGG19: 0.97
  + EfficientNet: 0.97
  + ResNet50: 0.93

EfficientNet y VGG19 empataron con un F1-score de 0.97 para la clase de Catarata, mientras que ResNet50 tiene un valor inferior (0.93).

* **Diabetic Retinopathy**:
  + Todos los modelos obtienen un F1-score de 1.00.
* **Glaucoma**:
  + VGG19: 0.92
  + EfficientNet: 0.92
  + ResNet50: 0.87

Tanto VGG19 como EfficientNet tienen un F1-score de 0.92 para Glaucoma, mientras que ResNet50 queda por detrás con 0.87.

* **Normal**:
  + VGG19: 0.89
  + EfficientNet: 0.91
  + ResNet50: 0.90

EfficientNet supera a los otros dos modelos con un F1-score de 0.91 en la clase Normal. VGG19 tiene el valor más bajo (0.89), mientras que ResNet50 se queda en 0.90.

### **Conclusiones Comparativas**

* **EfficientNet** parece ser el mejor modelo en términos de precisión, recall y F1-score en las clases más problemáticas como Catarata y Glaucoma, lo que lo hace ideal para su selección como modelo final.
* **VGG19** tiene un desempeño muy sólido, especialmente en Glaucoma y Normal, pero su exactitud general es inferior a la de EfficientNet.
* **ResNet50**, aunque tiene un buen desempeño, queda rezagado en varias clases, especialmente en Glaucoma, lo que reduce su fiabilidad en comparación con EfficientNet y VGG19.

# 5. Anexos

En el siguiente enlace de hojas de cálculo se encontrará todos los resultados (Tablas generales, específicas y de varianza) de todas las arquitecturas, así como las tablas de métricas y mapas de calor (matrices de confusión) de VGG19, Efficientnet y Resnet50: https://docs.google.com/spreadsheets/d/1\_gel0anltmhzDncGMzjnL1fWJpdfIdtOnRYPSwJ3BlA/edit?usp=sharing

# 6. Referencias:

Elmoufidi, Abdelali & Amoun, Hind. (2021). EfficientNetB3 Architecture for Diabetic Retinopathy Assessment using Fundus Images. 10.21203/rs.3.rs-609899/v1.

S. Ortiz & M. A. Goenaga Jimenez. (2023, 2 de noviembre). Deep Learning-Based Ocular Disease Classification in Fundus Images. IEEE Colombian Caribbean Conference (C3), Barranquilla, Colombia, 2023. pp. 1-6. doi: 10.1109/C358072.2023.10436234.  
Simonyan, K., & Zisserman, A. (2014, 4 septiembre). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv.org. <https://arxiv.org/abs/1409.1556>

Vipas.Ai. (2024, 8 agosto). RESNet-50: A robust Image Classification Model - VIPAS.AI - Medium. Medium. <https://medium.com/@vipas.ai/resnet-50-a-robust-image-classification-model-9bd3fffbc3e3>