

# Aprendizaje No Supervisado para la Mejora de Imágenes en Condiciones de Baja Iluminación

## Resumen

Este proyecto tuvo como objetivo comparar dos enfoques complementarios para la mejora de imágenes en condiciones de baja iluminación: uno basado en técnicas clásicas de procesamiento digital y otro soportado por redes neuronales profundas. Se implementaron procesos de ajuste manual—incluyendo corrección gamma, ecualización local del histograma y reducción de ruido—junto con el modelo Zero-DCE++ (Li, Guo, & Loy, 2021), una red neuronal convolucional entrenada para mejorar imágenes sin supervisión directa. Ambos métodos se aplicaron sobre una muestra representativa de la base de datos *The Dark Face* (Yang et al., 2020) y los resultados fueron evaluados mediante métricas objetivas como BRISQUE, NIQE, entropía y contraste RMS. Los hallazgos evidencian mejoras significativas en la calidad perceptual respecto al estado original, con fortalezas diferenciadas en cada enfoque. Esta experiencia permitió comprender con mayor profundidad las capacidades y limitaciones de cada método, así como su aplicabilidad en contextos reales de preprocesamiento de imágenes.

## 1. Introducción

La mejora de imágenes es una etapa crítica dentro del procesamiento digital, ya que condiciona directamente la calidad, utilidad e interpretabilidad de una imagen en diversas aplicaciones. En campos como la medicina, la visión por computador, la seguridad o la teledetección, una imagen clara puede ser la diferencia entre un análisis exitoso y un resultado erróneo (Gonzalez & Woods, 2018).

Cuando se trabaja con imágenes capturadas en condiciones de iluminación deficiente, como escenas nocturnas o ambientes con exposición limitada, mejorar la imagen se vuelve esencial. El preprocesamiento—entendido como el conjunto de operaciones previas que buscan acondicionar la imagen antes de aplicar técnicas más avanzadas—cumple aquí un rol fundamental: permite rescatar información visual, corregir deficiencias y facilitar tanto la interpretación humana como el análisis automatizado.

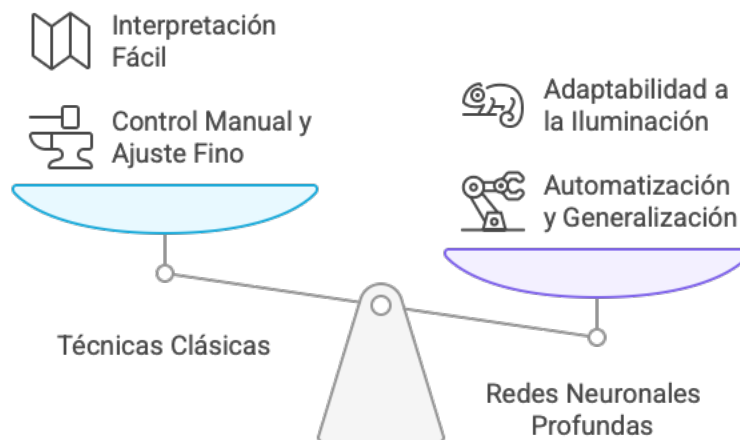
En este contexto, el presente laboratorio se orientó a estudiar y comparar distintas estrategias de mejora aplicadas a imágenes oscuras. A continuación, se detallan los métodos utilizados, las condiciones experimentales y las métricas de evaluación empleadas para analizar sus resultados.

## 2. Material y Métodos

### 2.1. Enfoque y contextualización técnica

Para abordar la mejora de imágenes en condiciones de baja iluminación, se optó por comparar dos enfoques contrastantes pero complementarios: uno basado en técnicas clásicas de procesamiento digital y otro apoyado en redes neuronales profundas.

El primer enfoque, manual y controlado, emplea transformaciones bien conocidas como ajustes gamma, ecualización del histograma y reducción de ruido. Estas técnicas permiten intervenir de manera directa en el comportamiento de la intensidad y el contraste de la imagen, siendo altamente configurables y fácilmente interpretables. Resultan especialmente útiles en contextos donde no se dispone de modelos entrenados o donde se requiere un ajuste fino y personalizado.



*Figura 1. Técnicas clásicas vs. Redes neuronales*

Por otro lado, el segundo enfoque se basa en el modelo Zero-DCE++, una red neuronal convolucional profunda diseñada para realizar mejoras de imagen directamente desde el espacio de píxeles, sin necesidad de supervisión (Li, Guo, & Loy, 2021). Este tipo de

modelos aprovecha el aprendizaje profundo para identificar patrones complejos de iluminación y aplicar transformaciones adaptativas que mejoran la percepción visual general de la escena. Su ventaja principal radica en la automatización y en la capacidad de generalizar sobre un conjunto amplio de condiciones de iluminación deficiente.

La elección de estos dos puntos de vista responde a la intención de explorar tanto el control detallado y comprensible del procesamiento manual como el potencial adaptativo y automatizado del aprendizaje profundo. Al compararlos, se puede valorar no solo la eficacia de cada técnica, sino también su aplicabilidad práctica, su flexibilidad y su impacto visual.

## **2.2. Imágenes utilizadas**

Las imágenes utilizadas en este experimento fueron seleccionadas de la base de datos pública The Dark Face (Rakshit, 2022), desarrollada por el grupo STAR (Spatial and Temporal Restoration, Understanding and Compression Team) del Wangxuan Institute of Computer Technology, Universidad de Pekín. Este conjunto de datos proporciona más de 6.000 imágenes reales capturadas en condiciones de baja iluminación durante la noche, en escenarios como calles, puentes, parques y edificios educativos.

Para evaluar los métodos empleados en este laboratorio, se utilizó una muestra de 4 imágenes seleccionadas aleatoriamente, que reflejan condiciones reales de visibilidad limitada. Estas imágenes representan un reto significativo para las técnicas de mejora debido al alto nivel de oscuridad y a la diversidad de escenas urbanas.

## **2.3. Método 1: Ajuste manual**

Como se observa en la Figura 2, el primer enfoque aplicado consiste en una serie de operaciones de mejora de imagen diseñadas de forma manual. El objetivo es incrementar el brillo, mejorar el contraste y reducir el ruido perceptual, mediante técnicas bien establecidas del procesamiento digital de imágenes. La secuencia de operaciones es la siguiente:



Figura 2. Método 1, ajuste manual con técnicas clásicas

1. Ajuste de gamma: Se aplica una corrección gamma con un valor de 0.5, lo que amplifica las intensidades bajas y aclara la imagen general. Esto mejora la visibilidad en áreas oscuras sin saturar las zonas brillantes.
2. Ecualización del histograma (CLAHE): Para mejorar el contraste se convierte la imagen al espacio de color LAB, donde se aísla el canal de luminancia ( $L$ ) y se aplica la técnica CLAHE (Contrast Limited Adaptive Histogram Equalization). Esta técnica divide la imagen en pequeñas regiones ( $8 \times 8$ ) y realiza una ecualización local, evitando la sobreexposición en áreas homogéneas. Se utilizó un límite de recorte de 2.0.
3. Reducción de ruido: Para suavizar artefactos y ruido de alta frecuencia, se emplea un *filtro de mediana* con un kernel de  $3 \times 3$ . Este filtro es eficaz para preservar los bordes mientras elimina puntos ruidosos aislados.
4. Reconversión a espacio RGB: Tras el procesamiento en LAB, la imagen es reconvertida a RGB para su visualización final, asegurando la fidelidad del color.

Este flujo de procesamiento fue implementado en Python utilizando bibliotecas como OpenCV, PIL y NumPy para el procesamiento de las imágenes, pero el procedimiento se ha mantenido independiente del lenguaje de programación para facilitar su comprensión y reproducción. La configuración específica se puede resumir así:

*gamma: 0.5*

*hist\_eq: "clahe"*

*clahe\_clip\_limit: 2.0*

*clahe\_tile\_grid\_size: (8, 8)*

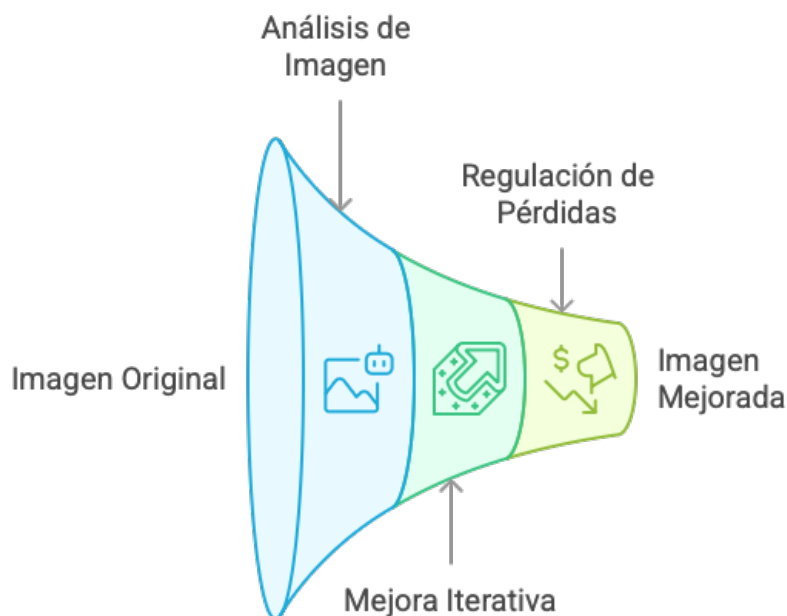
*denoise: "median"*

*median\_kernel\_size: 3*

Este método se caracteriza por su transparencia y facilidad de ajuste, permitiendo un control preciso sobre el resultado final de la imagen.

## 2.4. Método 2: Red neuronal Zero-DCE++

El segundo enfoque implementado corresponde a una red neuronal convolucional profunda llamada Zero-DCE++, entrenada específicamente para la mejora de imágenes en condiciones de baja iluminación sin necesidad de supervisión directa.



*Figura 3. Método 2, ajuste con Red neuronal Zero-DCE++*

Zero-DCE++ emplea una arquitectura basada en convoluciones separables en profundidad, sin operaciones de pooling, lo que permite preservar los detalles espaciales

de la imagen. Como se ilustra en la Figura 3, el proceso se inicia con la imagen original (extremo izquierdo), la cual es sometida a un análisis automático de imagen (bloque azul), donde la red genera un mapa de realce. Este mapa se aplica de forma iterativa (bloque verde) sobre la imagen de entrada, mejorando progresivamente su luminancia. Finalmente, mediante una regulación de pérdidas (bloque amarillo), el modelo ajusta sus parámetros para producir una imagen mejorada (extremo derecho).

Las funciones de pérdida utilizadas para entrenar el modelo fueron:

1. **L\_TV** (*Total Variation*): promueve la suavidad del mapa de realce.
2. **L\_spa** (*Spatial Consistency*): fuerza la coherencia espacial entre la imagen original y la mejorada.
3. **L\_color** (*Color Constancy*): minimiza desviaciones en el equilibrio de color.
4. **L\_exp** (*Exposure Control*): promueve niveles de exposición adecuados.

Durante el entrenamiento, se utilizaron 5.000 imágenes de entrenamiento y 1.000 de validación, todas extraídas del conjunto Dark Face. El modelo se entrenó durante 5 épocas, con un tamaño de lote de 12, utilizando el optimizador Adam y una tasa de aprendizaje de  $1e-4$ .

La implementación fue realizada en Python utilizando la biblioteca PyTorch, ampliamente empleada para el desarrollo y entrenamiento de redes neuronales profundas. Además, se apoyó en librerías como OpenCV para el manejo de imágenes, NumPy para cálculos matriciales y Matplotlib para la visualización de métricas de entrenamiento. Las funciones de pérdida fueron implementadas como módulos personalizados en `torch.nn.Module`, mientras que la arquitectura de la red está definida en la clase `enhance_net_nopool`.

Tanto la definición de la red como el proceso de entrenamiento pueden encontrarse en el siguiente repositorio de [GitHub \(Camilo, 2024\)](#), y estuvieron motivados por el trabajo de ejemplo disponible en Kaggle (Nguyen, 2023).

Este enfoque ofrece una solución robusta y escalable, especialmente adecuada para automatizar el procesamiento de grandes volúmenes de imágenes sin necesidad de

intervención manual. A diferencia del método anterior, Zero-DCE++ no requiere ajustes empíricos para cada imagen, ya que la red ha aprendido directamente a realzar imágenes oscuras mediante un entrenamiento orientado a múltiples criterios de calidad visual.

## 2.5. Métricas de evaluación

Debido a que el conjunto de datos empleado no incluye imágenes con referencia o etiquetas de calidad, se optó por métricas de evaluación sin referencia, lo que permite realizar comparaciones objetivas entre los métodos propuestos. Las métricas utilizadas fueron:

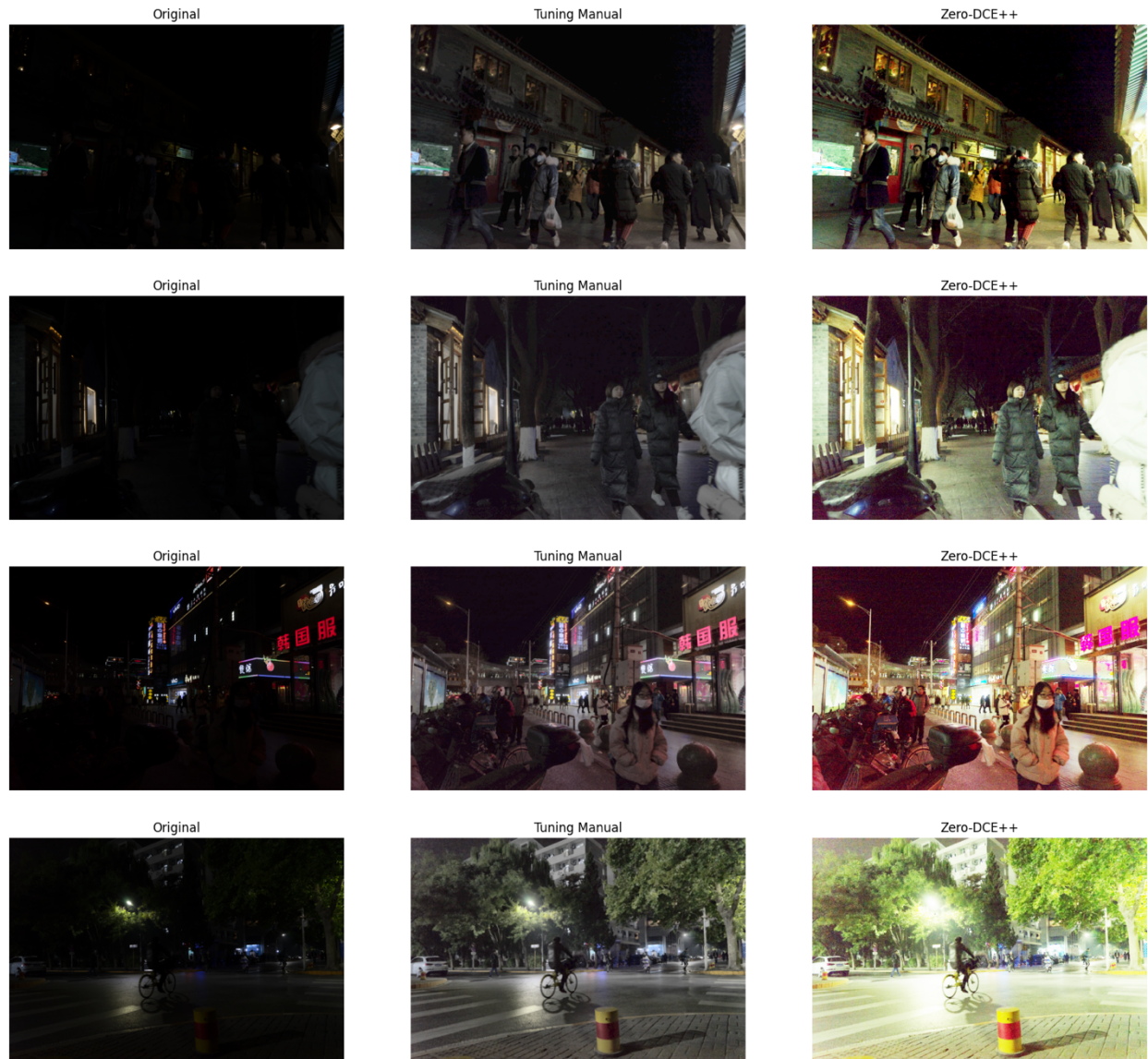
- **BRISQUE** (*Blind/Referenceless Image Spatial Quality Evaluator*): cuantifica la calidad percibida de la imagen sin necesidad de una imagen de referencia.
- **NIQE** (*Natural Image Quality Evaluator*): estima la naturalidad estadística de la imagen mejorada.
- **Entropía**: mide la cantidad de información o nivel de detalle presente en la imagen.
- **Contraste RMS**: cuantifica la variación de intensidades, representando el contraste global de la imagen.

Estas métricas fueron aplicadas tanto a las imágenes originales como a los resultados obtenidos mediante los métodos de ajuste manual y Zero-DCE++, permitiendo un análisis comparativo riguroso y coherente

## 3. Resultados

### 3.1. Comparación visual

A continuación, se muestran 4 ejemplos representativos del conjunto de imágenes evaluadas. Cada grupo incluye la imagen original (izquierda), la versión con ajustes manuales (centro) y la imagen mejorada con Zero-DCE++ (derecha).



*Figura 4. Resultado procesamiento de las imagenes 1259.png, 2454.png, 2796.png y 2835.png.*

En general, las imágenes mejoradas con el enfoque manual muestran un aumento notable en la visibilidad sin sacrificar la naturalidad. El contraste es más marcado y la imagen conserva los detalles clave.

Las imágenes procesadas con Zero-DCE++ tienden a tener una iluminación más agresiva, lo que mejora la visibilidad global, aunque en ciertos casos puede llevar a la pérdida de detalles en zonas brillantes y la generación de artefactos visuales.



### 3.2. Observaciones específicas por imagen:

- **1259.png:** Zero-DCE++ aporta una imagen más viva, con colores cálidos y alto contraste, pero pierde algunos detalles en áreas muy iluminadas. El ajuste manual mantiene una representación más sobria y equilibrada.
- **2454.png:** Zero-DCE++ destaca al iluminar uniformemente la escena sin sobresaturación. El ajuste manual también mejora, aunque conserva un tono más oscuro.
- **2796.png:** El ajuste manual resulta superior al mantener los detalles finos (letras y bordes), mientras que Zero-DCE++ ofrece mayor colorido pero menos nitidez.
- **2835.png:** Zero-DCE++ sobreexpone la escena, generando una imagen artificial. En contraste, el ajuste manual ofrece un equilibrio visual convincente, sin perder texturas.

### 3.3. Comparación cuantitativa

Se presentan los valores obtenidos de cada métrica para las cuatro imágenes utilizadas:

<i>Imagen</i>	<i>Método</i>	<i>BRISQUE</i>	<i>NIQE</i>	<i>Entropía</i>	<i>Contraste RMS</i>
1259.png	Original	44.75	5.15	2.84	15.04
	Tuning Manual	41.92	4.78	4.45	40.20
	Zero-DCE++	38.93	3.77	4.17	80.87
2454.png	Original	67.16	5.75	2.09	11.86
	Tuning Manual	48.05	5.10	3.84	32.56
	Zero-DCE++	48.40	4.57	3.59	67.82
2796.png	Original	33.89	4.42	2.87	23.79
	Tuning Manual	39.56	4.88	4.61	43.83
	Zero-DCE++	39.98	3.81	4.48	74.05
2835.png	Original	16.22	4.02	3.43	15.94
	Tuning Manual	35.48	4.37	4.91	37.83
	Zero-DCE++	11.37	2.99	5.00	52.27

*Tabla 1. Evaluación de métricas*

Zero-DCE++ mejora de forma consistente el **contraste RMS** (hasta 400% respecto a la imagen original y 100% respecto al método manual) y reduce significativamente los valores de **NIQE**, lo cual indica mayor naturalidad percibida.

**BRISQUE** presenta cierta variabilidad: en algunas imágenes mejora (e.g., 2835.png), mientras que en otras se mantiene similar o ligeramente superior al ajuste manual.

El ajuste manual tiende a obtener mejores valores de **entropía** en imágenes con elementos finos, lo cual evidencia su capacidad para preservar detalles.

### **3.4. Análisis crítico**

- El ajuste manual resulta especialmente útil en escenas con fuentes de luz puntuales o cuando se busca mantener detalles sutiles. Además, proporciona resultados más sobrios, con menor riesgo de saturación excesiva.
- Zero-DCE++ ofrece un rendimiento general superior en condiciones de oscuridad extrema, elevando la exposición global de manera agresiva. No obstante, esto puede conllevar la pérdida de textura y generación de áreas artificialmente planas o brillantes.

Ambos enfoques tienen méritos: uno destaca por su control fino y bajo costo computacional, el otro por su capacidad de automatización y respuesta en tiempo real a diversas condiciones visuales.

## **4. Conclusiones**

### **4.1. Resumen de hallazgos**

Del análisis realizado se concluye lo siguiente:

- Ambos métodos mejoran significativamente la visibilidad y calidad perceptual de imágenes en condiciones de baja iluminación, aunque con enfoques y ventajas distintas.
- Zero-DCE++ presenta un desempeño superior en la mayoría de las métricas, con resultados consistentes y automatizados, lo que lo hace ideal para sistemas de visión por computadora en escenarios de baja exposición.
- El ajuste manual, aunque limitado por la intervención humana, ofrece mayor control y preservación de detalles en escenas específicas, especialmente cuando se requiere evitar saturaciones o mantener texturas finas.

- En resumen, el ajuste manual es preferible en entornos controlados o con recursos limitados, mientras que Zero-DCE++ es más adecuado para procesos automatizados en condiciones visuales variables.

#### **4.2. Limitaciones y reflexiones**

- El ajuste manual requiere una configuración personalizada por imagen, lo que limita su escalabilidad en escenarios con grandes volúmenes de datos.
- Zero-DCE++ requiere entrenamiento previo y acceso a hardware con capacidad de cómputo (como GPU), lo cual puede no ser viable en entornos con recursos limitados.
- Algunas métricas objetivas como BRISQUE pueden no captar completamente la percepción visual subjetiva del observador humano. Por esta razón, se recomienda complementar el análisis cuantitativo con evaluaciones cualitativas y funcionales, especialmente cuando el objetivo es la interpretación visual o el uso posterior de las imágenes en tareas de análisis

#### **4.3. Propuestas futuras**

- Explorar combinaciones híbridas: usar Zero-DCE++ como preprocesamiento seguido de un ajuste manual fino.
- Evaluar redes alternativas como RetinexNet (Wei et al., 2018), EnlightenGAN (Jiang et al., 2019) o modelos supervisados con imágenes de referencia.

## 5. Bibliografia

Yang, W., Yuan, Y., Ren, W., Liu, J., Scheirer, W. J., Wang, Z., Zhang, et al. (2020). *Advancing image understanding in poor visibility environments: A collective benchmark study*. *IEEE Transactions on Image Processing*, 29, 5737–5752. <https://doi.org/10.1109/TIP.2020.2981922>

Li, C., Guo, C., & Loy, C. C. (2021). *Learning to enhance low-light image via zero-reference deep curve estimation*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/TPAMI.2021.3063604>

Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing* (4th ed.). Pearson.

Rakshit, S. (2022). *Dark Face Dataset* [Data set]. Kaggle. <https://www.kaggle.com/datasets/soumikrakshit/dark-face-dataset/data>

Camilo, D. (2024). *Low-light image enhancer: Training pipeline* [Jupyter notebook]. GitHub. [https://github.com/davidcamilo0710/low\\_light\\_enhancer/blob/main/pipeline.ipynb](https://github.com/davidcamilo0710/low_light_enhancer/blob/main/pipeline.ipynb)

Nguyen, U. (2023). *Zero-DCE vehicle* [Code notebook]. Kaggle. <https://www.kaggle.com/code/uiter22521498/zero-dce-vehicle>

Wei, C., Wang, W., Yang, W., & Liu, J. (2018). *Deep Retinex Decomposition for Low-Light Enhancement*. In *Proceedings of the British Machine Vision Conference (BMVC)*. <https://arxiv.org/abs/1808.04560>

Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., & Wang, Z. (2019). *EnlightenGAN: Deep Light Enhancement without Paired Supervision*. *arXiv preprint arXiv:1906.06972*. <https://arxiv.org/abs/1906.06972>