Detección imágenes o vídeos modificados mediante Redes Neuronales.

# Indicación del nivel de seguimiento de la planificación prevista para el desarrollo del TFG, y de los ajustes efectuados y/o previstos.

Para el desarrollo del trabajo de “*Detección imágenes o vídeos modificados mediante Redes Neuronales"*, se tiene un seguimiento de la planificación prevista para el desarrollo del proyecto, así como de los ajustes que se han ido efectuado, con el objetivo de alcanzar los objetivos propuestos.

En este sentido, en la primera entrega del proyecto se cumplió con la planificación prevista, la cual consistía en la definición del planteamiento de los objetivos, la revisión de la bibliografía relevante y la definición del marco teórico y metodológico. Para ello, se realizó una búsqueda exhaustiva de las principales bibliografías que existen sobre el tema.

Sin embargo, durante el desarrollo del proyecto se presentaron *algunas dificultades, principalmente relacionadas con el enfoque a seguir*. Al principio se creó un clasificador a partir del dataset seleccionado, pero los resultados obtenidos no fueron los deseados. Por lo tanto, se tuvo que replantear la idea y enfocarse más en el **preprocesamiento de la imagen antes de pasarla al modelo**.

En este sentido, los ajustes que se han efectuado consisten en darle más atención a la generación del modelo, resaltando las características principales de las imágenes para poder obtener un modelo eficiente.

Por lo tanto, se espera que, en la evaluación del rendimiento del sistema, en términos de precisión, velocidad y eficiencia y otras métricas de interés, se logren alcanzar los objetivos propuestos, y se espera comparar los resultados obtenidos con otros sistemas de detección de imágenes y videos modificados existentes.

Por otro lado, la idea con el desarrollo del proyecto es poder crear no solo un modelo de detección de manipulación en imágenes, sino también otro modelo capaz de identificar el tipo de modificación que se ha realizado sobre la imagen en cuestión.

# Indicación de cambios introducidos en los objetivos fijados y/o en la metodología de trabajo prevista, ambos justificados adecuadamente.

En relación con los objetivos, se ha considerado importante enfocarse en el estudio de resaltar las características principales de la imagen, ya que esto es fundamental para el correcto funcionamiento del sistema de detección de imágenes y videos modificados. Para lograr este objetivo, se ha dedicado un tiempo considerable a investigar diferentes técnicas de extracción de características de imágenes, las cuales han sido analizadas y evaluadas en función de su eficacia y eficiencia en el procesamiento de imágenes.

Se ha estudiado la técnica de *ELA* (Error Level Analysis), la cual se basa en analizar las diferencias en los niveles de compresión de una imagen después de ser editada. También se ha investigado el uso de la *transformada de Wavelet*, que se basa en el análisis de la imagen en diferentes escalas y orientaciones para resaltar las características de interés. La *transformada BDCT* (Block Discrete Cosine Transform) se ha empleado para resaltar las características de la imagen a nivel de bloques. La técnica de *Sobel*, que permite detectar bordes en las imágenes, se ha utilizado para identificar posibles áreas modificadas. Además, se ha estudiado la detección de la *ROI* (Region of Interest), que permite identificar las regiones de la imagen que son más relevantes para la detección de modificaciones. Por último, se ha analizado la *decorrelación* de imágenes para identificar las regiones con mayor contraste.

Por otra parte, se ha planteado un nuevo objetivo que consiste en detectar el tipo de modificación realizada en la imagen, para lo cual se ha considerado importante realizar una segmentación del objeto en cuestión. De esta manera, se logrará identificar con mayor precisión el área o zona modificada, lo que permitirá una mejor evaluación del rendimiento del sistema.

Todos estos cambios han sido adecuadamente justificados en función de los resultados esperados y de la experiencia adquirida en el desarrollo del proyecto. Además, se ha considerado importante la revisión constante de los objetivos y la metodología de trabajo, con el fin de garantizar que se están cumpliendo los objetivos planteados al comienzo del proyecto.

# Explicación general de la metodología que se está siguiendo para alcanzar los objetivos.

La metodología seguida para alcanzar los objetivos de este TFG se ha basado en una investigación exhaustiva de las técnicas y métodos previamente desarrollados en el ámbito de la detección de manipulación de imágenes. Esta investigación ha permitido la identificación de las principales tendencias y desafíos en esta área, así como las soluciones y enfoques más prometedores.

Una vez recopilada la suficiente información, se procedió a la formulación de la idea principal para el desarrollo del modelo de detección de manipulación de imágenes. Posteriormente, se trabajó en la construcción y recopilación de un dataset completo que permitiera la correcta formación y validación del modelo.

En una primera aproximación, se aplicó una clasificación de las imágenes. Sin embargo, se observó que esto no era suficiente para obtener un modelo preciso y efectivo en la detección de manipulaciones. Por lo tanto, se decidió profundizar en el preprocesamiento de las imágenes para resaltar las características clave y relevantes de las mismas.

A partir de ahí, se ha trabajado en la exploración y prueba de diferentes técnicas de preprocesamiento de imágenes con el objetivo de optimizar la precisión del modelo y obtener los mejores resultados posibles. Esta evaluación se ha llevado a cabo mediante la comparación de los resultados obtenidos con cada técnica aplicada y seleccionando aquellas que mejoran la calidad de la detección.

A continuación, mostraré algunos de los preprocesamientos que se han aplicado a las imágenes:

## Ejemplo: Imagen 1

Un río al lado de un cuerpo de agua

Descripción generada automáticamente

**Transformada wavelet bidimensional discreta**

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

**Canales Y, Cb y Cr**

Imagen que contiene reloj

Descripción generada automáticamente

**Métodos utilizados para el preprocesamiento de la imagen**Interfaz de usuario gráfica

Descripción generada automáticamente

# Desarrollo del trabajo

El modelo que se ha desarrollado es una red neuronal convolucional (CNN) ya que se utiliza para clasificar imágenes, en este caso específicamente para detectar si una imagen ha sido modificada o alterada. El uso de CNN se debe a su capacidad para aprender patrones en imágenes y su eficacia en la clasificación de imágenes,las capas de convolución y *max pooling* se utilizan para extraer características importantes de la imagen y reducir la dimensionalidad, mientras que las capas totalmente conectadas se utilizan para la clasificación final.

El uso de la función de activación *ReLU* utilizado en la primera capa totalmente conectada ayuda a evitar la desaparición del gradiente y a acelerar el proceso de entrenamiento,la función de activación *softmax* se utiliza en la segunda capa para producir una salida de dos dimensiones, lo que permite la clasificación en dos clases: *"modificada"* o *"no modificada".*

El optimizador *Adam* se utiliza para ajustar los pesos del modelo y mejorar su precisión y se utiliza una tasa de abandono para evitar el sobreajuste del modelo y mejorar su capacidad de generalización.

Al guardar el modelo con la precisión de validación más alta y evaluar su precisión en el conjunto de pruebas, se asegura que el modelo sea lo suficientemente preciso y generalizable para su uso en la detección de modificaciones o alteraciones de imágenes en situaciones del mundo real.

El modelo que ha obtenido los mejores resultados es aquel en el que se ha aplicado ELA a la imagen antes de ser introducida en el modelo. Se ha utilizado un tamaño de imagen de 128x128 píxeles para el procesamiento.

A continuación, muestro los resultados del entrenamiento de los diferentes modelos.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| técnica de preprocesamiento aplicada | Training Accuracy | Validation  Accuracy | Training Loss | Validation  Loss |
| ELA | **0,96** | 0,93 | 0,05 | 0,17 |
| BDCT | 0,845 | 0,78 | 0,35 | 0,475 |
| Canny | 0,782 | 0,78 | 0,48 | 0,515 |
| Enhance Features | 0,782 | 0,78 | 0,495 | 0,51 |
| Sobel | 0,78 | 0,78 | 0,49 | 0,52 |

## MODELO 1 ELA\_128

### Resultados

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamente Gráfico, Gráfico de rectángulos

Descripción generada automáticamente

## MODELO 2 BDCT\_128

### Resultados

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamente Gráfico

Descripción generada automáticamente

## MODELO 3 CANNY\_128

### Resultados

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamente

## MODELO 4 ENCH\_128

### Resultados

Gráfico

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamente

## MODELO 5 SOBEL\_128

### Resultados

Gráfico

Descripción generada automáticamente

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Gráfico, Gráfico de barras

Descripción generada automáticamente

Bibliografía

1. "Digital Image Forensics” - Hany Farid.
2. Exposing Photo Manipulation From User-Guided 3-D Lighting Analysis Tiago Carvalhoa , Hany Faridb and Eric Keec aCemaden, S˜ao Jos´e dos Campos, SP, Brazil bDartmouth College, Hanover, NH, USA cColumbia University, New York, NY, USA.
3. Exposing Photo Manipulation with Inconsistent Reflections James F. O’Brien University of California, Berkeley and Hany Farid Dartmouth College.
4. Image Forgery Identification using Convolution Neural Network N. Hema Rajini.
5. ELA (Error Level Analysis) - <https://fotoforensics.com/tutorial-ela.php>
6. Sobel Filter - <https://spec.oneapi.io/oneipl/0.5/filtering/sobel.html>
7. An efficient post-processing adaptive filtering technique to rectifying the flickering effects - <https://www.researchgate.net/publication/351462926_An_efficient_post-processing_adaptive_filtering_technique_to_rectifying_the_flickering_effects/figures?lo=1>