Detección imágenes o vídeos modificados mediante Redes Neuronales.

# Contexto

La desinformación y el crecimiento de noticias falsas en Internet es un problema que afecta cada vez más a los usuarios. Según un artículo del Instituto de Ciencias del Mar (ICM, 2020), esto se debe a varios factores, como el acceso sin control, el anonimato, la sobreabundancia de datos y la falta de regulación en Internet. La tecnología actual permite la manipulación y adulteración de contenido audiovisual de manera fácil, lo que causa que los usuarios estén mal informados. Además, la inteligencia artificial juega un papel crucial en el crecimiento de la generación de imágenes y videos adulterados, que afecta la integridad democrática y debe ser abordado en sus dimensiones política, social y cultural.

Por otro lado, la manipulación y el adulteramiento de imágenes y videos pueden tener consecuencias graves en diferentes ámbitos que afectan directamente a la sociedad, como la política, la justicia y los medios de comunicación. Según el informe de la Federación Internacional de Periodistas sobre noticias falsas (FIP, 2021), la creación de noticias falsas o contenido adulterado es una herramienta sutil en la manipulación de la opinión pública, lo que causa graves daños a la democracia y la libertad de expresión.

Además, la manipulación del contenido visual también puede ser utilizada para difamar y acosar a personas en línea. Según UNICEF (2021), el ciberacoso puede tener consecuencias psicológicas muy graves para las personas afectadas, además de causar daños en su reputación y tener un impacto negativo en la percepción de la sociedad.

Para concluir, la manipulación y el adulteramiento de imágenes y videos tienen consecuencias negativas en diferentes áreas y afectan directamente a la sociedad. Es importante tener en cuenta estas consecuencias y fomentar la ética y la integridad en la creación y difusión de contenido visual.

Las fuentes utilizadas son:

* Artículo del Instituto de Ciencias del Mar sobre la desinformación en redes sociales (<https://www.icm.es/2020/06/08/desinformacion-redes-sociales/>)
* Información de UNICEF sobre el ciberacoso (<https://www.unicef.org/es/end-violence/ciberacoso-que-es-y-como-detenerlo>)
* Informe de la Federación Internacional de Periodistas sobre noticias falsas (<https://www.ifj.org/fileadmin/user_upload/Fake_News_-_FIP_AmLat.pdf>)

# Propuesta

La propuesta de Trabajo Final de Grado consiste en desarrollar un sistema de detección de imágenes o videos modificados utilizando Redes Neuronales, con el objetivo de detectar imágenes o videos que han sido editados, manipulados o falsificados. El sistema utilizará técnicas de aprendizaje automático y análisis de imágenes para detectar anomalías y cambios en el contenido visual.

Para llevar a cabo este proyecto, se explorarán diferentes técnicas de procesamiento de imágenes y redes neuronales, y se compararán sus resultados en la detección de contenido falso o manipulado. Además, se investigará cómo el sistema puede ser utilizado en diferentes contextos, como la detección de noticias falsas, la protección de la privacidad y la seguridad en línea, y la prevención del ciberacoso.

El sistema se entrenará con una gran cantidad de datos para poder detectar y distinguir entre imágenes y videos genuinos y aquellos que han sido manipulados de alguna manera. Se utilizarán diferentes tipos de datos, como imágenes y videos reales y *generados por ordenador*, para mejorar la precisión y robustez del sistema.

Además, se estudiarán las *diferentes formas en que se pueden modificar imágenes y videos*, incluyendo el uso de técnicas de deepfake y la alteración de patrones, y se evaluará cómo el sistema puede detectar estas anomalías y cambios.

En resumen, la propuesta de Trabajo Final de Grado tiene como objetivo desarrollar un sistema de detección de imágenes o videos modificados utilizando técnicas de inteligencia artificial, y explorar sus aplicaciones prácticas en la detección de la manipulación de contenido visual en diferentes contextos. El proyecto combina técnicas de procesamiento de imágenes, aprendizaje automático y análisis de datos para proporcionar una solución robusta y precisa para la detección de contenido falso o manipulado.

# Objetivos

**Objetivos Generales:**

1. Desarrollar un sistema de detección de imágenes y videos modificados utilizando técnicas de redes neuronales.
2. Evaluar el rendimiento del sistema en términos de precisión, velocidad y eficiencia y otras métricas de interés.
3. Comparar los resultados obtenidos con otros sistemas de detección de imágenes y videos modificados existentes.

**Objetivos Específicos:**

1. Realizar una revisión de la literatura existente en detección de imágenes y videos modificados, incluyendo las técnicas de redes neuronales.
2. Diseñar e implementar un conjunto de algoritmos de redes neuronales para la detección de imágenes y videos modificados.
3. Seleccionar y preparar un conjunto de datos adecuado para el entrenamiento y evaluación del sistema.
4. Evaluar el rendimiento del sistema en términos de precisión, velocidad y eficiencia mediante pruebas rigurosas.
5. Comparar los resultados obtenidos con otros sistemas de detección de imágenes y videos modificados existentes en la literatura y realizar un análisis crítico de las fortalezas y debilidades de cada enfoque.

# Metodología

La metodología Kanban es un enfoque ágil utilizado en la gestión de proyectos que se centra en la visualización del trabajo y la eliminación de cuellos de botella.

La metodología Kanban se basa en el uso de tableros Kanban, que son tableros visuales que muestran el trabajo en diferentes estados. Los tableros Kanban suelen tener tres columnas principales: "Pendiente", "En proceso" y "Terminado".

La metodología Kanban se enfoca en la mejora continua, lo que significa que se enfoca en identificar y eliminar los cuellos de botella en el proceso de trabajo para mejorar el flujo de trabajo. Para lograr esto, Kanban utiliza una serie de prácticas, incluyendo:

1. Limitar el trabajo en progreso para evitar la sobrecarga del equipo y mantener un flujo de trabajo constante.
2. Identificar y medir los cuellos de botella en el proceso de trabajo para encontrar oportunidades de mejora.
3. Implementar cambios en el proceso de trabajo de forma incremental para evitar interrupciones significativas y minimizar el riesgo de fracaso.
4. Fomentar la colaboración y la comunicación entre los miembros del equipo para mejorar la eficiencia y la calidad del trabajo.

# Metodología para la implementación del Proyecto

1. **Recopilación y preparación de datos (dataset):**

* Identificar fuentes de datos adecuadas para el conjunto de entrenamiento y el conjunto de prueba.
* Descargar y organizar las imágenes o videos en carpetas de acuerdo con su clasificación (modificado o no modificado).
* Verificar la calidad y la resolución de las imágenes o videos.
* Realizar el preprocesamiento de los datos.
* Dividir los datos en conjuntos de entrenamiento, validación y prueba.

1. **Selección y entrenamiento de la red neuronal:**

* Seleccionar una arquitectura adecuada de red neuronal (**A determinar**)
* Inicializar los pesos de la red neuronal y establecer una función de pérdida adecuada para el entrenamiento.
* Entrenar la red neuronal utilizando el conjunto de entrenamiento y ajustar los hiperparámetros.
* Evaluar el rendimiento de la red neuronal en el conjunto de validación.
* Realizar pruebas de sobreajuste para asegurarse de que la red.

1. **Validación de la red neuronal:**

* Evaluar el rendimiento de la red neuronal en el conjunto de prueba para determinar la precisión, la sensibilidad y la especificidad.
* Realizar un análisis de errores para determinar las limitaciones y áreas de mejora de la red neuronal.
* Comparar el rendimiento de la red neuronal con otras técnicas de detección de imágenes o videos modificados existentes.

1. **Optimización y mejora de la red neuronal:**

* Ajustar los hiperparámetros de la red neuronal según los resultados de la validación.
* Realizar el aumento de datos para aumentar la cantidad y la variabilidad del conjunto de entrenamiento. (Según la necesidad y tiempo es opcional)
* Realizar la transferencia de aprendizaje para aprovechar el conocimiento previo de una red pre-entrenada en tareas similares.

1. **Implementación y pruebas en tiempo real:**

* Implementar la red neuronal en un sistema en tiempo real (Aplicación Movil o Pagina Web).
* Realizar pruebas en un conjunto de datos de prueba en tiempo real para asegurarse de que la red funciona correctamente en diferentes escenarios y condiciones.
* Analizar los resultados de las pruebas y realizar ajustes en la implementación.

# Planificación

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TAREA** | **DESCRIPCIÓN** | **PROGRESO** | **INICIO** | **FIN** |
| **Fases de Presentación** |  |  |  |  |
| Tarea 1 | Reunión inicial con el Tutor del proyecto. | 100% | 19-2-23 | 19-2-23 |
| Tarea 2 | Realizar planificación del proyecto. | 90% | 19-2-23 | 6-3-23 |
| **Fases de Planificación** |  |  |  |  |
| Tarea 1 | Definición y planteamiento de objetivos. | 50% | 1-3-23 | 10-3-23 |
| Tarea 2 | Revisión de bibliografía relevante | 0% | 1-3-23 | 10-3-23 |
| Tarea 3 | Establecimiento de metodología y herramientas a utilizar. | 0% | 1-3-23 | 3-3-23 |
| Tarea 4 | Seguir revisando bibliografía y recopilando información relevante. | 0% | 4-3-23 | 15-3-23 |
| Tarea 5 | Definir el marco teórico y metodológico. | 0% | 11-3-23 | 16-3-23 |
| **Fases de Implementación** |  |  |  |  |
| Tarea 5 | **Realizar la primera sesión de seguimiento.** | 0% | **6-3-23** | **12-3-23** |
| Tarea 1 | Iniciar la implementación del proyecto. | 0% | 12-3-23 | 27-3-23 |
| Tarea 2 | Identificar las variables a medir y los indicadores de análisis. | 0% | 12-3-23 | 14-3-23 |
| Tarea 3 | Continuar con la implementación del proyecto | 0% | 14-3-23 | 27-3-23 |
| Tarea 4 | Realizar el análisis preliminar de los resultados. | 0% | 19-3-23 | 27-3-23 |
| Tarea 6 | Corregir posibles desviaciones. | 0% | 27-3-23 | 6-4-23 |
| Tarea 7 | Continuar con la implementación del proyecto. | 0% | 6-4-23 | 23-4-23 |
| Tarea 8 | **Realizar la segunda sesión de seguimiento.** | 0% | **17-4-23** | **23-4-23** |
| Tarea 9 | Evaluar los avances y corregir posibles desviaciones. | 0% | 23-4-23 | 22-5-23 |
| Tarea 10 | Consolidar los resultados y comenzar a redactar el informe. | 0% | 23-4-23 | 22-5-23 |
| Tarea 11 | **Realizar la tercera sesión de seguimiento.** | 0% | **22-5-23** | **28-5-23** |
| Tarea 12 | Evaluar los avances y corregir posibles desviaciones. | 0% | 22-5-23 | 12-6-23 |
| Tarea 13 | Consolidar los resultados y continuar redactando el informe. | 0% | 22-5-23 | 12-6-23 |
| Tarea 14 | Preparar el informe final para la revisión en la cuarta sesión de seguimiento. | 0% | 13-6-23 | 18-6-23 |
| Tarea 15 | **Comenzar a preparar la presentación.** | 0% | 13-6-23 | 18-6-23 |
| **Fases de Evaluación y Presentación** | |  |  |  |
| Tarea 1 | **Realizar la cuarta sesión de seguimiento.** | 0% | **12-6-23** | **18-6-23** |
| Tarea 2 | Corregir el informe final y ajustar la presentación. | 0% | 12-6-23 | 17-6-23 |
| Tarea 3 | Preparar la propuesta de presentación para la revisión en la quinta sesión de seguimiento. | 0% | 18-6-23 | 24-6-23 |
| Tarea 4 | Realizar los ajustes necesarios en el informe final. | 0% | 18-6-23 | 24-6-23 |
| Tarea 5 | **Realizar la quinta sesión de seguimiento.** | 0% | **26-6-23** | **30-6-23** |
| Tarea 5 | Ajustar la presentación y finalizar el informe. | 0% | 26-6-23 | 30-6-23 |
| Tarea 6 | Preparar el dossier y el póster. | 0% | 26-6-23 | 30-6-23 |
| Tarea 7 | Entrega del Dossier. | 0% | **26-6-23** | **2-7-23** |
| Tarea 8 | Entrega del Poster. | 0% | **26-6-23** | **6-7-23** |
| Tarea 9 | **Presentar y defender el TFG ante el tribunal.** | 0% | **21-7-23** |  |

# 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. Exposing Photo Manipulation with Inconsistent Shadows Eric Kee, Dartmouth College James O’Brien, University of California, Berkeley and Hany Farid, Dartmouth College.
5. How to Detect Faked Photos - A reprint from American Scientist the magazine of Sigma Xi, The Scientific Research Society.
6. Image Forgery Identification using Convolution Neural Network N. Hema Rajini.
7. PROTECTING PRESIDENT ZELENSKYY AGAINST DEEP FAKES A PREPRINT Matyáš Bohácek ,Hany Farid.
8. Detecting Deep-Fake Videos from Aural and Oral Dynamics Shruti Agarwal and Hany Farid University of California, Berkeley Berkeley, CA USA.
9. Detecting Deep-Fake Videos from Phoneme-Viseme Mismatches Shruti Agarwal and Hany Farid University of California, Berkeley Berkeley, CA USA.
10. Detecting Deep-Fake Videos from Appearance and Behavior Shruti Agarwal1 , Tarek El-Gaaly2 , Hany Farid1 , and Ser-Nam Lim2.
11. Exposing Digital Forgeries in Interlaced and De-Interlaced Video Weihong Wang, Student Member, IEEE, and Hany Farid, Member, IEEE.
12. Exposing Digital Forgeries in Video by Detecting Duplication Weihong Wang, Hany Farid.
13. Detecting Real-Time Deep-Fake Videos Using Active Illumination -Candice R. Gerstner Hany Farid.
14. Advances in Computer Vision and Pattern Recognition - Husrev Taha Sencar Luisa Verdoliva Nasir Memon.
15. A Coarse-to-fine Deep Convolutional Neural Network Framework for Frame Duplication Detection and Localization in Forged Videos Chengjiang Long Arslan Basharat Anthony Hoogs.
16. DEEPFAKE DETECTION: CURRENT CHALLENGES AND NEXT STEPS Siwei Lyu
17. Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics Yuezun Li1 , Xin Yang1 , Pu Sun2 , Honggang Qi2 and Siwei Lyu1 1 University at Albany, State University of New York, USA 2 University of Chinese Academy of Sciences, China.
18. DETECTION OF REAL-TIME DEEPFAKES IN VIDEO CONFERENCING WITH ACTIVE PROBING AND CORNEAL REFLECTION Hui Guo, Xin Wang, Siwei Lyu.
19. Celeb-DF: A Large-scale Challenging Dataset for DeepFake Forensics Yuezun Li1 , Xin Yang1 , Pu Sun2 , Honggang Qi2 and Siwei Lyu1 1 University at Albany, State University of New York, USA 2 University of Chinese Academy of Sciences, China.
20. IMPERCEPTIBLE ADVERSARIAL EXAMPLES FOR FAKE IMAGE DETECTION Quanyu Liao1 , Yuezun Li3 , Xin Wang2† , Bin Kong2 , Bin Zhu4 , Siwei Lyu3 , Youbing Yin2 , Qi Song2 , Xi Wu1†.
21. In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking Yuezun Li, Ming-Ching Chang and Siwei Lyu University at Albany, State University of New York, USA.
22. EXPOSING DEEP FAKES USING INCONSISTENT HEAD POSES Xin Yang? , Yuezun Li? and Siwei Lyu.