

Etiquetado de imágenes: Un enfoque efectivo para entrenar un modelo de Visión Artificial para la detección de plagas y enfermedades de aguacate con datos limitados.

Ruth Clementina Barragán López
Laboratorio de Cómputo

José Eduardo Velazco Jiménez
Ingeniería en Sistemas Computacionales

Cynthia Alejandra Martínez Pinto
DEPI

Instituto Tecnológico de Cd Guzmán

Instituto Tecnológico de Cd Guzmán

Instituto Tecnológico de Cd Guzmán

ruth.bl@cdguzman.tecnm.mx

L20291027@cdguzman.tecnm.mx

cynthia.mp@cdguzman.tecnm.mx

Resumen. El cultivo de aguacate en la región sur de jalisco cobra cada vez mayor importancia para México como principal productor, el uso de visión artificial ayuda en la detección temprana de enfermedades y plagas para minimizar las posibles pérdidas.

Palabras Clave. Visión Artificial, etiquetado de imágenes, entrenamiento del modelo, convoluciones, aprendizaje automático.

I. ANTECEDENTES DEL PROBLEMA

El cultivo de aguacate es de gran importancia a nivel mundial por su valor nutricional y económico, se enfrenta a un desafío significativo: las plagas y enfermedades que amenazan su producción. En México, principal productor y exportador de aguacates, las pérdidas por estas condiciones pueden alcanzar hasta el 80% de la cosecha (Informe Anual de La CAMPAÑA CONTRA PLAGAS REGLAMENTADAS DEL AGUACATERO, Operada Con Recursos Del Componente de Sanidad Del Programa De, SAGARPA.). La detección temprana es crucial para el control efectivo y la minimización de estas pérdidas.

Los métodos tradicionales de detección, como la inspección visual por expertos, son lentos, costosos y propensos a errores humanos. Las técnicas de laboratorio, si bien son más precisas, también son costosas y lentas. La visión artificial surge como una alternativa prometedora, ofreciendo la posibilidad de automatizar la detección de plagas y enfermedades a través del análisis de imágenes de frutos del aguacate.

visión artificial requiere grandes cantidades de datos, lo que representa un desafío en el caso del aguacate, donde la disponibilidad de estos es limitada. El etiquetado de imágenes consiste en asignar etiquetas a estas, indicando la presencia o

ausencia de plagas y enfermedades, ya sea manualmente por expertos, automáticamente mediante algoritmos de aprendizaje automático o de forma colaborativa a través de plataformas en línea.

II. UBICACIÓN GEOGRÁFICA

El estudio se desarrolla en Ciudad Guzmán, Jalisco, México. Una región con una alta producción de aguacate y una gran incidencia de plagas y enfermedades. La elección de esta ubicación permite enfocar la investigación en un contexto específico y relevante para la problemática planteada.

III. RESULTADOS DE INVESTIGACIONES ANTERIORES

Diversos estudios han demostrado la efectividad del etiquetado de imágenes para mejorar el rendimiento de los modelos de visión artificial en la detección de plagas y enfermedades de aguacate. Algunos ejemplos relevantes incluyen:

- **Detección de plagas en aguacate utilizando un sistema de visión artificial basado en imágenes RGB y etiquetado colaborativo:** Este estudio alcanzó una

precisión del 88% en la detección de plagas de aguacate a través del etiquetado colaborativo. (Ortega-García et al., 2024)

Estos estudios proporcionan evidencia sólida del potencial del etiquetado de imágenes para mejorar la precisión de la detección de plagas y enfermedades de aguacate mediante la visión artificial.

IV. PLANTEAMIENTO DEL PROBLEMA.

Actualmente en el área de clasificación de imágenes, la cantidad de imágenes por clase necesaria para entrenar un modelo eficiente varía significativamente. Si bien las cifras de 100, 500 o 1000 imágenes por clase se utilizan a menudo como referencia inicial, este número puede variar en función de la complejidad del modelo, la variabilidad de las imágenes, la calidad de las imágenes y la disponibilidad de técnicas de aumento de datos o aprendizaje por transferencia, por lo que es importante buscar alternativas para hacer que el entrenamiento de un modelo para una visión artificial que no cuenta con un gran conjunto de imágenes sea igual de eficiente. Además de mencionar que esta situación limita la adopción de tecnologías en el monitoreo de cultivos de aguacate y, por lo tanto, afecta la capacidad para prevenir y controlar las plagas y enfermedades, ocasionando pérdidas económicas significativas a productores de aguacate.

V. OBJETIVOS DE LA INVESTIGACIÓN

Objetivo General:

Desarrollar un sistema de detección de plagas y enfermedades de aguacate basado en visión artificial que utilice etiquetado de imágenes para optimizar el rendimiento del modelo y mejorar la precisión de la detección con datos limitados.

Objetivos específicos:

Demostrar la robustez del modelo de visión artificial para la detección de plagas y enfermedades de aguacate con datos limitados, utilizando gráficos para visualizar su rendimiento.

VI. FORMULACIÓN DE HIPÓTESIS

Hipótesis nula (H0):

Es posible entrenar un modelo de visión artificial con más del 80% de precisión para la detección de plagas y enfermedades de aguacate utilizando un conjunto de datos pequeño (menor a 100 imágenes).

Hipótesis alternativa (H1):

No es posible entrenar un modelo de visión artificial con más del 80% de precisión para la detección de plagas y enfermedades de aguacate utilizando un conjunto de datos pequeño (menor a 100 imágenes).

VII. JUSTIFICACIÓN

La detección temprana y precisa de plagas en los cultivos de aguacate es crucial para garantizar la salud y la productividad de los cultivos, así como para promover prácticas agrícolas sostenibles y reducir el uso de pesticidas. En este contexto, la visión artificial y el aprendizaje automático ofrecen herramientas poderosas para desarrollar sistemas automatizados de detección de plagas, que pueden ayudar a los agricultores a identificar y controlar las infestaciones de manera rápida y eficiente.

Sin embargo, el rendimiento de estos sistemas de detección de plagas puede verse afectado por varios factores, como el tamaño del conjunto de datos y el preprocesamiento de imágenes. Por lo tanto, es importante investigar cómo estos factores influyen en la eficiencia del modelo de visión artificial en la detección de plagas en cultivos de aguacate.

VIII. DISEÑO DEL MARCO TEÓRICO.

Para comprender el tema se deben tener nociones de los siguientes puntos:

- Aprendizaje automático
- Redes neuronales convolucionales (CNN)
- Convoluciones
- Clasificación de imágenes
- Técnicas de aumento de datos
- Etiquetado de imágenes
- Grounding Dino
- YOLO

IX. METODOLOGÍA

Recolección y preprocesamiento de datos:

- Se recolectó un conjunto de datos de imágenes de aguacate que contienen ejemplos de diferentes plagas y enfermedades.
- Las imágenes se preprocesaron para estandarizar su tamaño, formato y características relevantes para la detección de plagas y enfermedades.

Etiquetado de imágenes:

- Se aplicó una técnica de etiquetado de imágenes, como el etiquetado manual por expertos y el etiquetado automático mediante algoritmos de aprendizaje automático, para asignar etiquetas a las imágenes del conjunto de datos.

Aumento de Datos:

- Se implementaron técnicas de aumento de datos para ampliar artificialmente el tamaño del conjunto de datos y mejorar el rendimiento del modelo.

Entrenamiento del modelo:

- Se entrenó un modelo de visión artificial, utilizando una arquitectura CNN adecuada, con el conjunto de datos etiquetado.
- Se optimizaron los hiperparámetros del modelo para maximizar su precisión.

Evaluación del modelo:

- Se evaluó el rendimiento del modelo entrenado en un conjunto de datos de prueba independiente.
- Se utilizaron métricas de rendimiento como la precisión, la sensibilidad y la especificidad para evaluar la capacidad del modelo para detectar plagas y enfermedades de aguacate correctamente.

Análisis de resultados:

- Se analizaron los resultados obtenidos en la evaluación del modelo.

- Se determinó si el modelo alcanzó una precisión superior al 80% utilizando un conjunto de datos pequeño.
- Se discutieron las fortalezas y limitaciones del modelo y las técnicas utilizadas.

Conclusiones:

- Se formularon conclusiones sobre la viabilidad de entrenar un modelo de visión artificial con más del 80% de precisión para la detección de plagas y enfermedades de aguacate utilizando un conjunto de datos pequeño.
- Se proponen recomendaciones para futuras investigaciones y aplicaciones del modelo.

X. PROCESO METODOLÓGICO

Recolección de datos

Se ha llevado a cabo la recolección imágenes de manera manual de dos diferentes malestares más comunes en el cultivo de aguacate, como lo son: Quemadura y Roña. Donde se han logrado recabar la cantidad de 55 imágenes para las quemaduras en el aguacate y 64 elementos para la enfermedad “Roña”. Además, se han estandarizado a (640 x 640) pixeles debido a que las herramientas utilizadas (YOLOv8) requieren de estas proporciones para su correcto funcionamiento.

Etiquetado de imágenes:

El proceso en la etiquetación de imágenes de buena calidad es lo más delicado de este proceso de acuerdo con el instituto de ingeniería del conocimiento (Conocimiento, 2022) donde mencionan que “se puede producir una discriminación intencional o explícita (disparate treatment) y, por otro, una discriminación no intencional (disparate impact).” Así mismo mencionan el error humano como un origen en los sesgos de la IA, “La definición o etiquetado de los datos, por error humano o criterios subjetivos”, por lo que se eligió dividir esta etapa en dos fases:

1. Clasificación de las imágenes del conjunto de datos recabado por un especialista en el campo de enfermedades de aguacate.
2. Etiquetado automático mediante la herramienta Grounding Dino.

Aumento de datos

Utilizar técnicas de “Aumento de imágenes” o “Data Augmentation”, específicamente “Voltar: Horizontal, Vertical”, “Rotación: Entre -15° y +15°”, “Difuminar: Aumento a 1.1px” fue un factor importante en el proceso, pues debido a este se ha aumentado el conjunto de datos en un 30%.

Entrenamiento del modelo

Debido al etiquetado o “labeling” que nos ha proporcionado la herramienta Grounding Dino, se procede a entrenar un modelo de visión artificial utilizando una arquitectura CNN como lo es YOLO en su versión 8, con el conjunto de datos etiquetado. Pasando como parámetro un conjunto de datos especificado por el archivo YAML. Donde: clase 0 = Roña y clase 1 = Quemadura

```

1 train: ../train/images
2 val: ../valid/images
3 test: ../test/images
4
5 nc: 2
6 names: ['0', '1']
    
```

Ilustración 1 Archivo Yalm

Así mismo se optimizan los hiper parámetros esperando un rendimiento por mayor al 80%, configurado el número de épocas (epochs = 100), el tamaño del lote (batch = 64), el tamaño de la imagen (imgsz = 640), y el dispositivo de entrenamiento (device = 'cuda').



Ilustración 2 Conjunto de pruebas de entrenamiento (YOLOv8)

Evaluación del modelo

Se han recopilado los informes y estadísticas de rendimiento que YOLOv8 ofrece al entrenar un modelo, en la cual se puede observar algunas imágenes del conjunto de pruebas siendo etiquetadas y prediciendo su posible causa de enfermedad.



Ilustración 3 Conjunto de pruebas de predicciones de validación

Además, con un conjunto de pruebas de 10 imágenes con diferentes tipos de luminosidad y fondo se ha puesto a prueba el modelo, en el cual como se muestra en la tabla de resultados se tiene un promedio de 0.85 (85%) en detección de Roña y 0.816 en pruebas de quemaduras. Así mismo, como se muestra en la *ilustración 4* hay un pico de

degradación en la precisión, esto se debe a las condiciones en la captura de la *ilustración 5* donde se observa que la luminosidad y el contraste con las hojas es baja, aún así logra detectar el aguacate, pero no su enfermedad como si lo hace en la *ilustración 6* donde las condiciones son adecuadas.

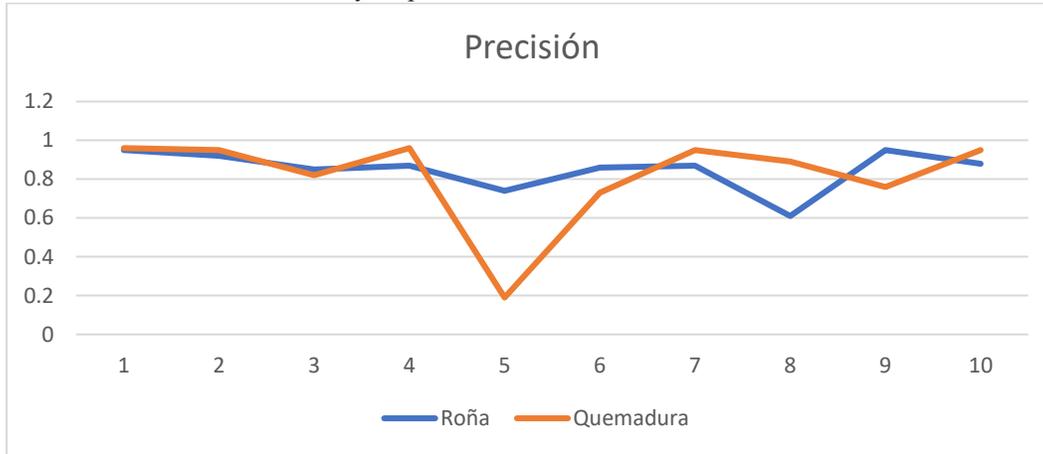


Ilustración 4 Precisión en predicciones manuales (Pruebas implementando el modelo y pasando una fotografía como parámetro)

Prueba	Roña	Quemadura
1	0.95	0.96
2	0.92	0.95
3	0.85	0.82
4	0.87	0.96
5	0.74	0.19
6	0.86	0.73
7	0.87	0.95
8	0.61	0.89
9	0.95	0.76
10	0.88	0.95
Promedios	0.85	0.816

Tabla 1 Precisiones en pruebas manuales (Pruebas implementando el modelo y pasando una fotografía como parámetro)



Ilustración 6 Fotografía con malas condiciones (Bajo contraste con el fondo, objeto de detección desenfocado y poco visible)



Ilustración 5 Fotografía en buenas condiciones (El objeto está enfocado, no hay objetos obstruyendo y el objeto es muy visible)

XI. ANÁLISIS DE RESULTADOS:

Por otro lado, en las pruebas de rendimiento que ofrece ultralytics del propio YOLO muestra una

precisión inconsistente en las 100 épocas de entrenamiento, donde se obtuvo un 0.7545 (75%) en promedio de precisión.

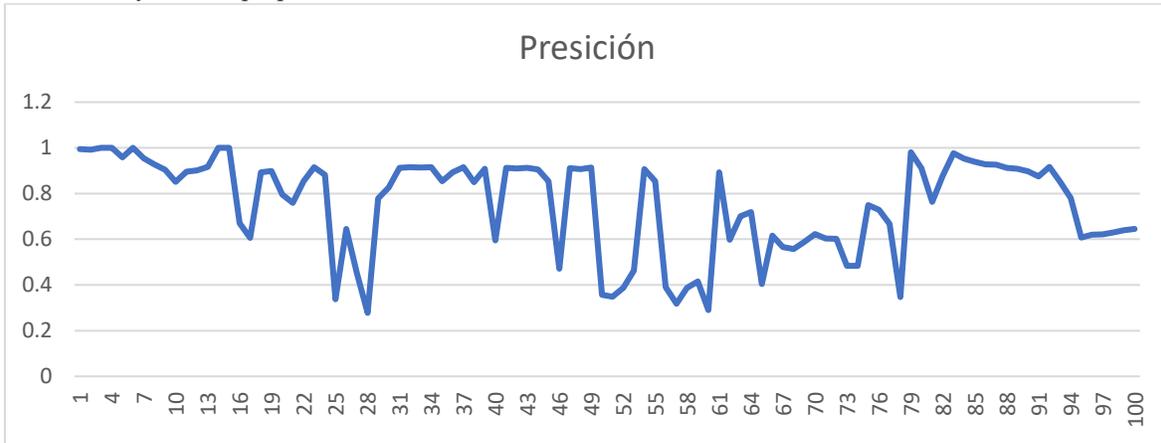


Ilustración 7 Precisión por épocas en pruebas de YOLOv8

En la ilustración 8 se muestran las gráficas que muestran el proceso de entrenamiento y validación del modelo. En general, las pérdidas (tanto para el entrenamiento como la validación) muestran una tendencia descendente, lo que indica que el modelo mejora a lo largo del tiempo. Sin embargo, las métricas como la precisión y el recall presentan fluctuaciones significativas, lo que sugiere inestabilidad o variabilidad en el desempeño del modelo, especialmente en los datos de validación

debido a la cantidad limitada de datos. Sin embargo, las métricas de precisión media (mAP@50 y mAP@50-95) muestran una mejora constante, lo que indica que el modelo se está ajustando mejor a lo largo del entrenamiento, aunque la generalización en el conjunto de validación parece ser un desafío. Lo cual supone que el modelo está progresando en su capacidad de aprendizaje, pero hay espacio para mejorar la estabilidad y generalización en validación.

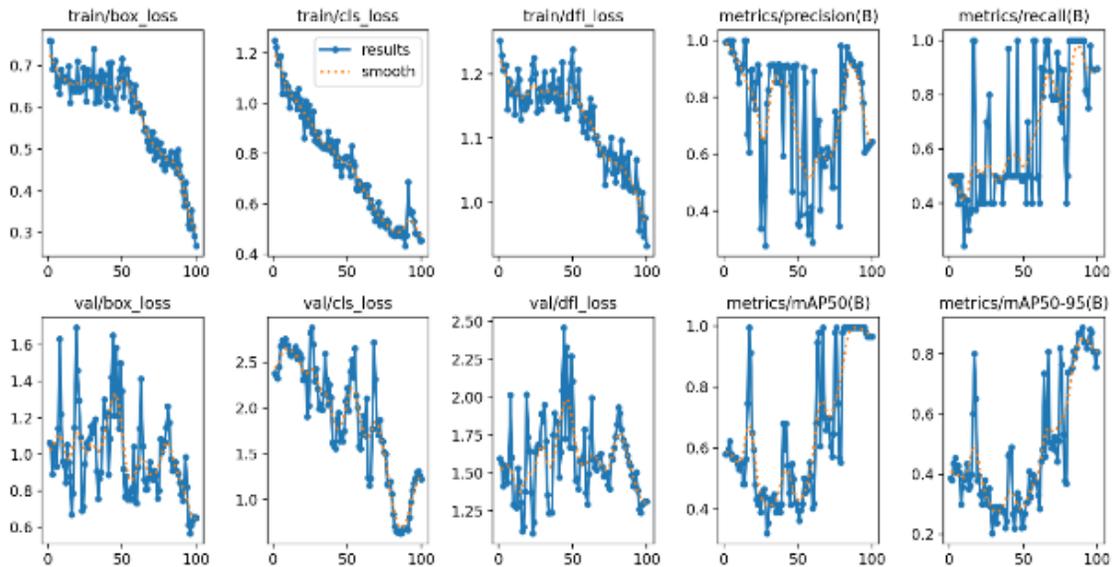


Ilustración 8 Métricas de ultralytics del modelo.

De igual manera, en la siguiente gráfica se determina en que niveles de confianza el modelo tiene un buen rendimiento. Para el modelo representado, el nivel de confianza óptimo está en 0.663, donde el F1-score global es máximo (0.94).

Esto indica que, en ese punto, el modelo tiene un buen equilibrio entre precisión y recall, superando con creces el nivel estable de 80% que se esperaba en la investigación.

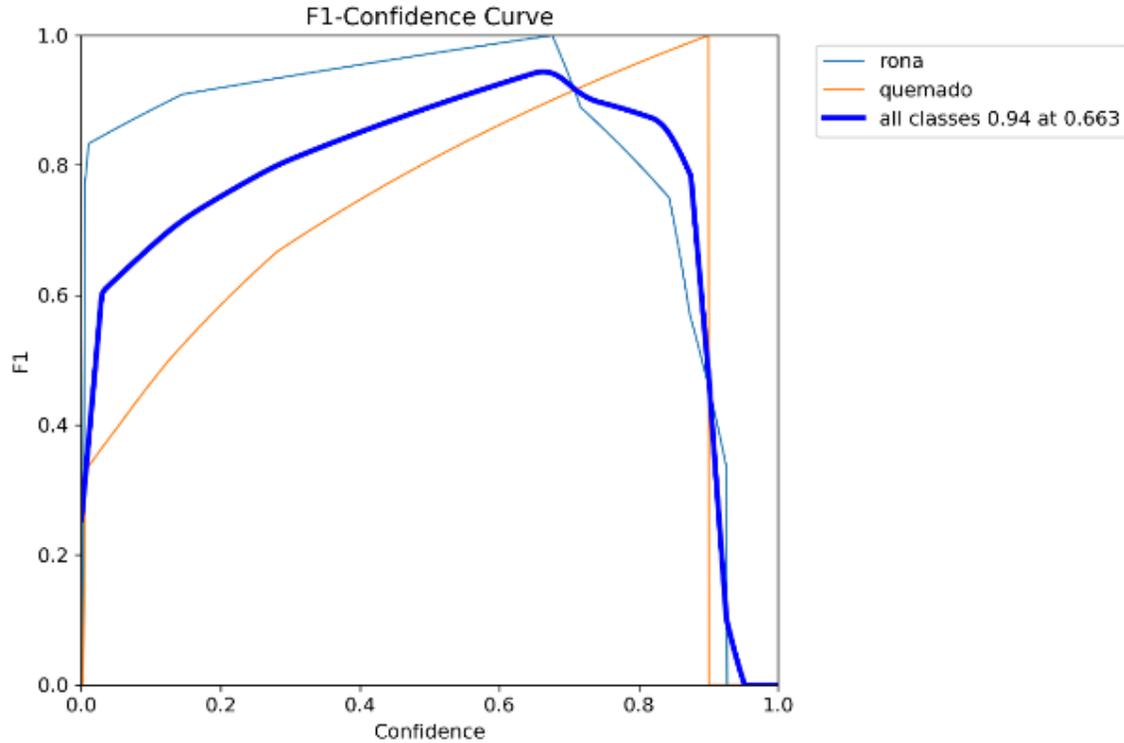


Ilustración 9 Nivel de confianza según ultralytics de YOLOv8

XII. CONCLUSIONES:

Finalmente, se ha notado en las diversas gráficas mostradas anteriormente que en algunos casos el modelo es capaz de superar el 80% en precisión y confianza en las pruebas realizadas manualmente, pero es inviable confirmar que puede lograr una consistencia con el conjunto de datos presentado. El hecho de etiquetar imágenes muestra que es posible tomar fotografías en campo sin necesidad de un estudio que controle la luminosidad. Por lo tanto y como recomendaciones para

investigaciones futuras; variar las épocas de entrenamiento y agregar más imágenes al conjunto de datos, pues si bien no se logró demostrar que con menos de 100 imágenes se puede lograr una consistencia precisión requerida en esta investigación, si se pudo demostrar que entre más imágenes se tengan de una clase, menos variaciones habrá en la confianza y (Ilustración 4, Tabla 1)

XIII. BIBLIOGRAFÍA.

1. Informe anual de la CAMPAÑA CONTRA PLAGAS REGLAMENTADAS DEL AGUACATERO, operada con recursos del Componente de Sanidad del Programa de. (n.d.). Retrieved September 26, 2024, from https://www.gob.mx/cms/uploads/attachment/file/157266/Plagas_reglamentadas_de_l_aguacatero.pdf
2. Campos, E., & Manuel, J. (2021). CLASIFICADOR DE RED NEURONAL CONVOLUCIONAL PARA IDENTIFICAR ENFERMEDADES DEL FRUTO DE AGUACATE (*Persea americana* Mill.) A PARTIR DE IMÁGENES DIGITALES. *Agrociencia*, 55(8), 695–709. <https://doi.org/10.47163/agrociencia.v55i8.2662>
3. Kosaraju, Raj Prateek. (2021). *Deep CNNs for large scale species classification*. ArXiv.org. <https://arxiv.org/abs/2102.01863>
4. *Classification of dog skin diseases using deep learning with images captured from multispectral imaging device (pdf) | Paperity*. (2022). Paperity.org. <https://paperity.org/p/288276052/classification-of-dog-skin-diseases-using-deep-learning-with-images-captured-from>
5. Y. Asnawi, A.V.H. Simanjuntak, U. Muksin, Okubo, M., S.I. Putri, S. Rizal, & M. Syukri. (2022). Enhancement of convolutional neural network for urban environment parking space classification. *DOAJ (DOAJ: Directory of Open Access Journals)*. <https://doi.org/10.22034/gjesm.2022.03.02>
6. majianjia. (2021, November 29). *GitHub - majianjia/nnom: A higher-level Neural Network library for microcontrollers*. GitHub. <https://github.com/majianjia/nnom>
7. Training AI for Self-Driving Vehicles: the Challenge of Scale | NVIDIA Technical Blog. (2017, October 10). NVIDIA Technical Blog. <https://developer.nvidia.com/blog/training-self-driving-vehicles-challenge-scale/>
8. Yang, Y., & Perdikaris, P. (n.d.). *Physics-informed deep generative models*. Retrieved April 18, 2024, from <http://bayesiandeeplearning.org/2018/papers/16.pdf>

9. Conocimiento, del. (2022, March 9). *Inteligencia Artificial responsable: sesgos y explicabilidad* - IIC. Instituto de Ingeniería Del Conocimiento. <https://www.iic.uam.es/innovacion/inteligencia-artificial-responsable-sesgos-y-explicabilidad>