

**Proyecto Final – Entrega 2**



**UNIVERSIDAD  
DE ANTIOQUIA**

1 8 0 3

Por: Elizabeth Ospina Rojas  
Sebastian López Gómez

Profesor: Dr. Raúl Ramos Pollán

Posgrado  
Departamento de Ingeniería  
Universidad de Antioquia  
Medellín, 2023-2

## Tabla de contenido

Contexto.....	3
Dataset .....	3
Desarrollo.....	4
Resultados .....	5
Conclusiones .....	8
Resultados Previos .....	8
Referencias.....	8

## **Contexto**

El maíz, un cultivo significativo por su alto consumo a nivel global, se ve afectado por diversos factores, incluida la cantidad de hojas en las plantas. Las hojas, necesarias para la fotosíntesis, desempeñan un papel crucial en la generación de biomasa, impactando directamente en la producción de granos. La implementación de algoritmos de Deep Learning para el conteo automático de hojas de maíz beneficia a los agricultores al mejorar la toma de decisiones, aumentando la eficiencia y potencialmente incrementando los ingresos. Además, facilita la identificación y mitigación de plantas estresadas o enfermas, generando datos útiles para comprender el desarrollo del maíz.

Los algoritmos destinados al análisis de condiciones de cultivo de maíz ofrecen un potencial para diferentes tipos de usuarios, desde agricultores individuales hasta empresas, organizaciones gubernamentales o investigadores. Su aplicación promete mejorar significativamente la eficiencia y la gestión en la producción de maíz, destacando su impacto beneficioso en la agricultura a nivel global.

## **Objetivo general**

Predecir la cantidad de hojas que se encuentran en imágenes provenientes de plantas de maíz, esta cantidad se encuentra entre 7 a 12 hojas.

## **Objetivos específicos**

- Usar modelos de clasificación para determinar el número de hojas de plantas de maíz por medio de etiquetas que van de 7 a 12 hojas.
- Usar modelos de regresión para determinar el número de hojas de plantas de maíz que van de 7 a 12 hojas.
- Comparar los modelos usados para predecir la cantidad de hojas de plantas de maíz.

## **Dataset**

El Dataset utilizado en esta implementación contiene de 12.229 imágenes de 568 x 351 píxeles. Además, se tiene un archivo .csv llamado 'corn\_leaf\_number.csv' que contiene las siguientes columnas:

1. Nombre del experimento a partir del cual se generó la imagen. 'corn\_exp1' se realizó en 2018 y 'corn\_exp23' se realizó en 2019.
2. Nombre del archivo de imagen.
3. Nombre del genotipo.
4. Número de puntas de hojas intactas
5. Número de puntas de hojas dañadas o cortadas
6. Número de puntas de hojas invisibles debido a oclusiones
7. Coordenadas de las puntas de las hojas intactas y dañadas.

En total de la base de datos se usaron 7344 imágenes, pesando 109.043 KB en total. Cada imagen está en escala de grises y formato .jpg, se escalan a un tamaño de 224 x 224 pixeles.

## Desarrollo

En este trabajo se desarrollan cuatro modelos de predicción para estimar la cantidad de hojas en plantas de maíz. Se aplicó tanto clasificación como regresión para evaluar distintos enfoques. Dos modelos se centraron en clasificar las plantas en rangos de hojas con etiquetas (7 a 12 hojas), mientras que los otros dos realizaron predicciones numéricas por medio de regresiones. Transfer Learning se empleó utilizando las arquitecturas preentrenadas Resnet50 y VGG16 para generar los modelos, permitiendo la adaptación de conocimientos previos a este contexto. Los notebooks utilizados para el desarrollo de este proyecto son los siguientes.

- 1\_resnet50\_classification\_DL
- 2\_resnet50\_regression\_DL
- 3\_vgg16\_classification\_DL
- 4\_vgg16\_regression\_DL

El Transfer Learning es una técnica de entrenamiento de modelos de Deep Learning donde un modelo preentrenado en una tarea específica se adapta a otra tarea relacionada. En este contexto, las redes neuronales preentrenadas, ResNet50 y VGG16, actúan como bloques de construcción fundamentales. Estas arquitecturas, ya entrenadas en grandes conjuntos de datos, han aprendido a reconocer patrones complejos en imágenes, lo cual es relevante para la detección y clasificación de hojas en el presente caso.

La arquitectura ResNet50 fue seleccionada debido a que fue la encontrada durante la revisión bibliográfica para la elaboración de este trabajo, esta red destaca por su estructura residual, que facilita el entrenamiento de redes profundas. En el modelo de clasificación, la capa de salida fue ajustada para generar la probabilidad de pertenencia a cada clase definida por el rango de hojas. Para regresión, se adaptó la capa de salida para producir un valor numérico directo. Por otra parte, debido a los resultados obtenidos con la primera red, se opta por hacer transfer learning con otra red, para esto se selecciona la VGG16 esta se caracteriza por su simplicidad y profundidad. Consiste en bloques de convoluciones seguidos por capas de pooling. En nuestro modelo de clasificación, la última capa densa fue adaptada para producir la salida en el rango de 7 a 12 hojas, utilizando una función de activación softmax. Para regresión, se modificó la última capa para generar una salida numérica directa utilizando la función de activación lineal.

Ambas arquitecturas aprendieron representaciones jerárquicas de las características de las imágenes durante las 30 épocas de entrenamiento como se muestra en la *Figura 1*. Este proceso de Transfer Learning permitió capitalizar conocimientos previos, acelerando el entrenamiento y mejorando el rendimiento en la tarea específica de predicción de hojas en plantas de maíz.

```
# Entrenar el modelo con los callbacks
history = model.fit(train_dataset.batch(64), validation_data=validation_dataset.batch(16), epochs=30, callbacks=callbacks)
```

*Figura 1 Entrenamiento del modelo*

Cabe destacar que se implementó un mecanismo de parada temprana en el entrenamiento, deteniendo el proceso si el error no variaba significativamente en diferentes épocas, lo que contribuyó a evitar sobreajustes innecesarios y mejorar la eficiencia del entrenamiento, un ejemplo de esto se muestra en la *Figura 2*.

```
Epoch 10: ReduceLROnPlateau reducing learning rate to 1e-06.
184/184 [=====] - 65s 356ms/step - loss: 0.4228 - accuracy: 0.7983 - recall: 0.7501 - val_loss: 1.8765 - val_accuracy: 0.4373 - val_recall: 0.3692 - lr: 4.0000e-06
Epoch 11/15
184/184 [=====] - ETA: 0s - loss: 0.3991 - accuracy: 0.8249 - recall: 0.7745Restoring model weights from the end of the best epoch: 6.
Epoch 11: val_loss did not improve from 1.57313
184/184 [=====] - 63s 344ms/step - loss: 0.3991 - accuracy: 0.8249 - recall: 0.7745 - val_loss: 1.8692 - val_accuracy: 0.4455 - val_recall: 0.3719 - lr: 1.0000e-06
Epoch 11: early stopping
```

*Figura 2. Early stopping*

## Resultados

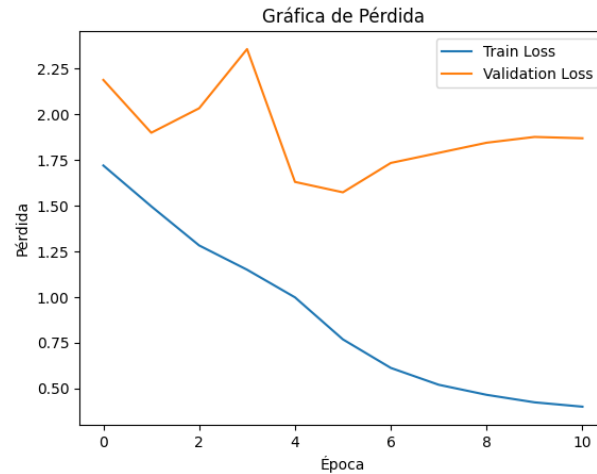
Los resultados de los indicadores se muestran en la *Tabla I*.

*Tabla I. Resultados indicadores*

ResNet50	Clasificación	Accuracy entrenamiento	0.8249
		Accuracy validación	0.4455
	Regresión	Mean Absolute Error entrenamiento	0.0178
		Mean Absolute Error validación	0.6953
VGG16	Clasificación	Accuracy entrenamiento	0.6317
		Accuracy validación	0.4401
	Regresión	Mean Absolute Error entrenamiento	0.2099
		Mean Absolute Error validación	0.6907

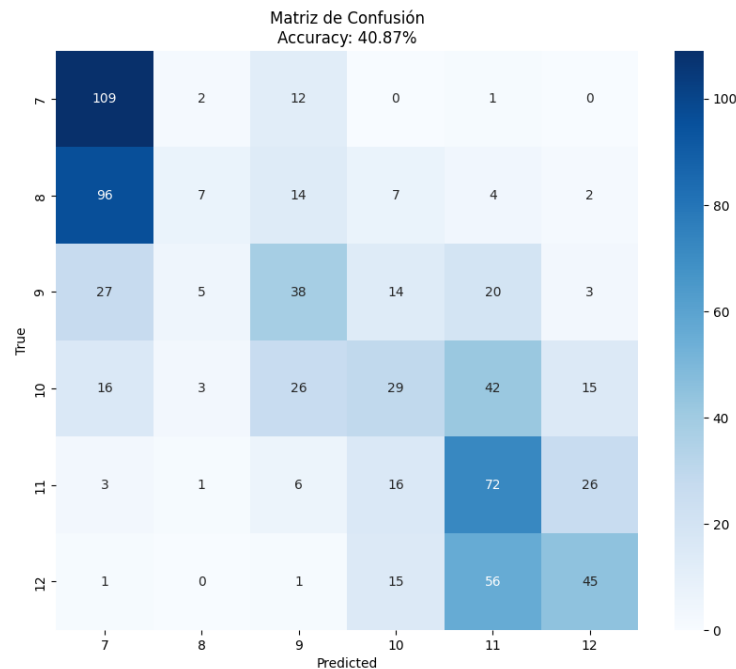
En la exploración de modelos de predicción aplicando transfer learning con las redes ResNet50 y VGG16, el indicador accuracy tuvo un valor mayor en entrenamiento que en validación, esto permite observar una posible tendencia al overfitting. El fenómeno de overfitting se manifestó cuando los modelos mostraron un ajuste excesivo a los datos de entrenamiento, resultando en un rendimiento subóptimo en conjuntos de validación o datos nuevos.

Para esta aplicación, las arquitecturas preentrenadas ResNet50 y VGG16, aunque potentes, mostraron cierta susceptibilidad al sobreajuste al ser utilizadas en el conjunto de datos propuesto. Este comportamiento se tradujo en un desempeño sobresaliente en los datos de entrenamiento, pero una capacidad limitada para generalizar patrones a nuevos datos, como se evidenció en los conjuntos de validación. Esto se puede observar en la gráfica de la *Figura 3*, donde se evidencia que la pérdida disminuye en el transcurso de las épocas para el entrenamiento, pero esa misma pérdida tiene un comportamiento aleatorio terminando incluso en crecimiento en el transcurso de las épocas para el conjunto de validación.



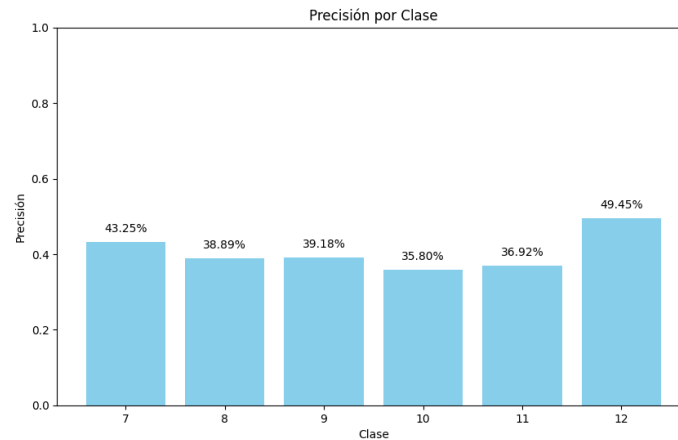
*Figura 3 Pérdida*

En cuanto al modelo de clasificación con la ResNet50, en la *Figura 4* se observa una matriz de confusión, esta muestra en el eje “x” las etiquetas predichas y en el eje “y” las etiquetas reales, permitiendo así ver los resultados de manera global, y mostrando que existen etiquetas como 7 hojas que tuvieron una cantidad total de aciertos de 109, siendo la mayor cantidad de aciertos del modelo, pero también tuvo una cantidad de errores en cuanto a la predicción de 7 hojas siendo el valor correcto 8 hojas de 96 imágenes.



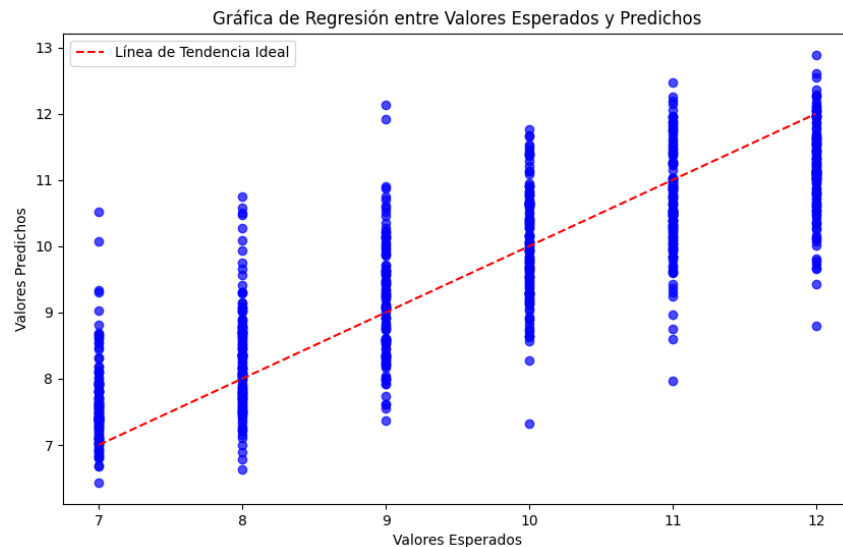
*Figura 4 Matriz de confusión*

Este mismo modelo de clasificación permite visualizar el accuracy o precisión por cada una de las clases o etiquetas, mostrando la mayor precisión en la predicción de imágenes con 12 hojas, con un total de 49.45% y la menor en 10 hojas con 35.80%.



*Figura 5 Precisión por clase*

Por último, para los modelos de regresión se analizó la gráfica que se muestra en la *Figura 5*, que muestra como es la distancia de los puntos con respecto al modelo planteado



*Figura 6. Regresión entre valores esperados y predichos*

Por otra parte, al evaluar la VGG16, se realizaron análisis de la matriz de confusión y la representación gráfica de la curva de regresión. Se observó que, en este modelo, la precisión para las clases asociadas a las etiquetas 7, 8, 9, 10, 11 y 12 fue del 66.7%, 31.15%, 27.15%, 35.71%, 0% y 43.50%, respectivamente. Se destaca una particularidad en la incapacidad de la VGG16 para reconocer las imágenes de plantas que poseen 11 hojas.

En términos generales, se evidencia que tanto la ResNet50 como la VGG16, utilizadas tanto para clasificación como para regresión, presentan deficiencias en el conteo preciso de hojas. Esto sugiere que los datos utilizados podrían ser insuficientes, indicando la necesidad de aumentar la cantidad de datos para mejorar el rendimiento de los modelos.

## **Conclusiones**

Los modelos ResNet50 y VGG16 mostraron una tendencia al sobreajuste, con un rendimiento superior en el conjunto de entrenamiento que en el de validación. Este resultado sugiere que, a pesar de la capacidad de Transfer Learning para aprovechar los conocimientos previos, las arquitecturas preentrenadas pueden ser susceptibles al sobreajuste en el contexto específico del conjunto de datos de hojas de maíz. Es importante abordar este problema para mejorar la generalización de los modelos.

Los resultados obtenidos sugieren que los modelos ResNet50 y VGG16 podrían mejorar su rendimiento con un conjunto de datos más extenso y representativo. La insuficiencia de datos actual puede ser una limitación, ya que dificulta a los modelos aprender a identificar patrones de hojas de maíz en imágenes nuevas y diferentes. Aumentar la cantidad y diversidad de datos podría ayudar a los modelos a generalizar mejor y proporcionar una representación más precisa de la variabilidad en las imágenes de hojas de maíz.

## **Resultados Previos**

Miao et al (2021) en su artículo “Automation of leaf counting in maize and sorghum using deep learning” presenta dos enfoques de aprendizaje automático para el conteo de hojas de maíz y sorgo: conteo por regresión mediante CNN, y conteo por detección.

Tras su estudio obtuvo que el conteo por regresión tuvo un rendimiento ligeramente mejor que el enfoque de conteo por detección, sin embargo, el enfoque por detección tuvo un rendimiento más consistente en plantas con diferentes números de hojas. Además, se obtuvo que el conteo por regresión era más susceptible a errores cuando las imágenes estaban tomadas en ángulos oblicuos, mientras que el enfoque por detección era más susceptible cuando las imágenes contenían hojas superpuestas. De acuerdo con los autores, ambos enfoques tienen el potencial de cumplir con el objetivo de automatización, pero redirigirse a diferentes enfoques podría presentar resultados con mejor desempeño.

## **Referencias**

- [1] Miao, C., Guo, A., Thompson, A. M., Yang, J., Ge, Y., & Schnable, J. C. (2021). Automation of leaf counting in maize and sorghum using deep learning. *Plant Phenome Journal*, 4(1). <https://doi.org/10.1002/ppj2.20022>
- [2] Miao, Chenyong (2020). corn leaf tip pixel annotations and images. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.13056524.v1>