

Redes Neuronales Convolucionales para la detección temprana de enfermedades en los cultivos de papa

Introducción.

La tecnología está inmersa en casi todas las áreas o ciencias en las cuales se tiene como objetivo resolver y facilitar un determinado problema o trabajo. En este mismo sentido, su aplicación se presencia en el sector agrícola, pues ha servido históricamente como herramienta entre el hombre y la naturaleza, ambos agentes marcan un proceso de cambio tecnológico y de innovación, las cuales impactan directamente en la producción (Olivares, Castro, Miranda, Morales, y Barrera, 2019).

A pesar de los avances tecnológicos en el campo de la agricultura, que han permitido a diversos países cubrir y satisfacer la demanda de alimentos de su población, existen factores como el cambio climático, declive de agentes polinizadores y enfermedades de cultivos, que presentan una amenaza real para la estabilidad y sustentabilidad de la agricultura en distintas localidades.

Actualmente la papa es una de las hortalizas mayormente extendidas en el mundo. En Bolivia es el tercer cultivo alimenticio más importante en términos de consumo humano. La producción de papa en Bolivia se ha incrementado en los últimos años, habiendo producido en el año agrícola 2019-2020 1,317,923 de toneladas métricas.

Sin embargo una de las grandes limitantes en la producción de papa son los problemas fitosanitarios o de enfermedades, que afectan plantas y tubérculos, generando pérdidas en los rendimientos y en la calidad del producto final, siendo los principales afectados los agricultores a pequeña escala. En este estudio se propone la alternativa del uso de aprendizaje profundo con modelos de redes neuronales convolucionales (CNN) como un enfoque viable para enfrentar esta problemática

Objetivo.

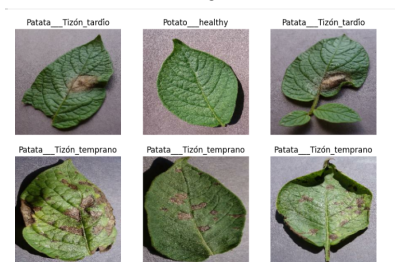
El objetivo de la presente trabajo es la detección temprana de las principales enfermedades como ser *Alternaria solani* (Tizón temprano) y *Phytophthora Infestans* (Tizón tardío) que afecta a los cultivos de papa, la detección se realizará a partir del estudio de sus hojas.

Conjunto de datos.

Para la obtención del conjunto de datos se recurre a PlantVillage un conjunto de datos disponible en kagge, el cual está compuesto por hojas de 14 tipos de cultivos distintos, correspondientes a 54.309 imágenes de hojas de plantas de manzana, arándano, cereza, maíz, uva, naranja, pera, pimiento morrón, papa, frambuesa, soja, calabaza, fresa y tomate. Para cada uno de estos cultivos se tienen la categoría sana y una o más categorías correspondientes a enfermedades causadas por hongos, bacterias, virus, ácaros y moho. Las imágenes se encuentran disponibles en su formato original (idéntico al tomado por la

cámara), en formato segmentado (substracción del fondo) y en escala de grises. Para este estudio, se seleccionaron las imágenes correspondientes a hojas de papa. El dataset seleccionado tiene un total de inicial de 2152 filas por cada fila existe una imagen y una etiqueta o clase a la que pertenece. Los campos del dataset seleccionado corresponden a image y label:

- Image son imágenes de hojas de papa de un tamaño de 256 pixeles por 256 píxeles, con 3 canales rgb, cada una perteneciente a cierta clase.
- Labels son las etiquetas o clases de cada imagen, tiene un total de 3 etiquetas, Patata___Tizón_tardío que hace referencia a hojas enfermas con tizon tardio, Patata___Tizón_temprano a hojas enfermas con tizon temprano y finalmente Potato___healthy hace referencia a hojas sanas.



Tizón tardío (*Phytophthora infestans*)

El tizón tardío es una enfermedad causada por el hongo *Phytophthora infestans* y es la más seria de las enfermedades fungosas que afectan a la papa. El daño que ocasiona puede devastar un cultivo en pocos días.

Hojas.

En el campo, los primeros síntomas de la enfermedad se presentan con frecuencia en las hojas inferiores. Los síntomas consisten en pequeñas manchas de color entre verde claro y verde oscuro que se convierten en lesiones pardas o negras según la condición del ambiente. Las lesiones se inician frecuentemente en las puntas y los bordes de las hojas. Una aureola verde clara o amarilla de algunos milímetros de ancho suele separar el tejido muerto del sano.

Tizón temprano (*Alternaria solani*)

El tizón temprano es un problema serio en muchas áreas del mundo donde no sólo afecta a la papa, sino también al tomate y a otras solanáceas. El tizón temprano ha sido menos estudiado que el tizón tardío, pero en los últimos años se ha observado que es una enfermedad importante en muchas áreas templadas donde se cultiva la papa. La enfermedad ataca al follaje y algunas veces también a los tubérculos.

Hojas.

En las hojas se desarrollan lesiones más o menos circulares, de color marrón oscuro, con anillos dispuestos concéntricamente semejan un tablero de tiro al blanco; primero se desarrollan en las hojas inferiores, más viejas. Según las condiciones ambientales y la variedad de papa, las lesiones se agrandan de 0,5-2,0 cm de diámetro, y a éstas se asocian áreas cloróticas alrededor y entre las lesiones. En condiciones de sequedad, las lesiones pueden perforarse dejando huecos (que se asemejan a orificios de bala). Las hojas pueden volverse completamente cloróticas, secar y morir. La enfermedad generalmente ocasiona la defoliación, pero las hojas secas algunas veces quedan colgando de la planta.

Pre procesamiento y balanceo de datos.

Balanceo de datos.

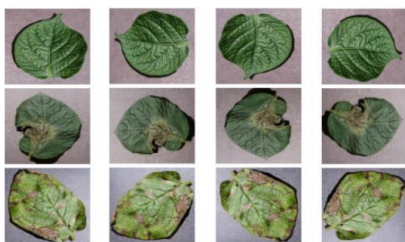
Inicialmente se tiene un total de 1000 imágenes de hojas con tizón tardío, 1000 imágenes de hojas de tizón temprano y 152 imágenes de hojas sanas, siendo 152 la cantidad significativamente más pequeña en comparación a las otras 2 clases, razón por la cual se considera pertinente un balanceo de datos. Se realiza el balanceo de datos reduciendo a 300 imágenes hojas con tizón tardío y 300 imágenes hojas de tizón temprano teniendo un total finalmente de 752 muestras.

Preprocesamiento de imágenes

Las operaciones de preprocesamiento en las imágenes antes de alimentarlas al modelo de red neuronal tienen como objetivo mejorar la calidad de los datos de entrada y aumentar la capacidad del modelo para generalizar y aprender patrones relevantes.

Se consideraron las siguientes operaciones de preprocesamiento:

- Resizing para redimensionar las imágenes a un tamaño de 256 x 256 píxeles. Esta operación garantiza que todas las imágenes tengan la misma dimensión antes de ser procesadas por el modelo. Al estandarizar el tamaño de las imágenes, se evita cualquier disparidad de escala que pueda afectar el rendimiento del modelo.
- Rescaling para normalizar los valores de los píxeles de las imágenes. Lo que hace esta operación es dividir todos los valores de píxeles por 255 para que estén en el rango de 0 a 1. Esta normalización es importante porque los modelos de redes neuronales suelen funcionar mejor con valores de entrada en un rango más pequeño y uniforme.
- RandomFlip para un aumento de datos de volteo aleatorio a las imágenes. El volteo aleatorio se aplica de manera horizontal (mode="horizontal") y vertical (mode="vertical"). Esto permite aumentar la cantidad de datos de entrenamiento al generar versiones adicionales de las imágenes con cambios de orientación. Además, el volteo aleatorio ayuda al modelo a ser más robusto al aprender a reconocer patrones independientemente de su orientación.
- RandomRotation para una transformación de rotación aleatoria a las imágenes con un factor de rotación de 0.5. Esto significa que las imágenes pueden rotar hasta un ángulo máximo de 50 grados en sentido horario o antihorario. Al igual que el aumento de volteo, esto aumenta la variabilidad en los datos de entrenamiento y ayuda al modelo a aprender características más robustas.



Clasificación usando una red neuronal convolucional CNN.

El dataset utilizado corresponde a un conjunto de datos supervisados, en este tipo de datos cada muestra está compuesta por un conjunto de características (variables independientes) y una etiqueta o valor objetivo (variable dependiente). En el contexto de la clasificación de imágenes, los datos supervisados consistirán en imágenes (características) junto con las etiquetas que indican la clase a la que pertenece cada imagen.

Para la división del dataset de trabajo se considera una división tradicional de los datos en entrenamiento, validación y prueba es conocida como "validación cruzada" o "cross-validation". En este enfoque, los datos se dividen en conjuntos de entrenamiento y prueba, y luego se reserva una porción adicional para la validación.

Para una primera ejecución se toman en cuenta las siguientes consideraciones:

- Datos de Entrenamiento. Se selecciona el 70%, este conjunto de datos se utiliza para ajustar parámetros(pesos,bias) .
- Datos de Validación. Se selecciona el 10% este conjunto de datos se utiliza para ajustar y seleccionar los hiperparametros (taza de aprendizaje, nro de capas , neuronas por capa, tamaño del mini lote)
- Datos de Entrenamiento.-Se selecciona el 20% este conjunto de datos se utiliza para evaluar la generalización del modelo, es decir ayuda a evaluar el comportamiento que tendrá con datos de la vida real.

Para una segunda ejecución:

- Datos de Entrenamiento. Se selecciona el 40%
- Datos de Validación .Se selecciona el 10%
- Datos de Entrenamiento.-Se selecciona el 50%.

Para el modelo de clasificación se considera una red neuronal convolucional CNN simple con varias capas convolucionales y de agrupación para procesar las imágenes de hojas de papa.

Dado que se trata de un conjunto de datos pequeño y poco variado se ve pertinente configurar este tipo de red, red neuronal convolucional CNN simple. Tras la experimentación con diferentes números de capas y profundidad se logró el siguiente CNN.

```
## input_shape es un parámetro que se usa para definir la forma de los datos de entrada a una red neuronal
input_shape=(batch_size,image_size,image_size,channels)

modelo_cnn=models.Sequential([
    rescalar_normalizar,
    aumento_datos_r_v,
    tf.keras.layers.Conv2D(filters = 16, kernel_size = (3,3), activation = 'relu', input_shape = input_shape),
    tf.keras.layers.MaxPool2D((2,2)),
    tf.keras.layers.Conv2D(64, (3,3), activation = 'relu'),
    tf.keras.layers.MaxPool2D((2,2)),
    tf.keras.layers.Conv2D(128, (3,3), activation = 'relu'),
    tf.keras.layers.MaxPool2D((2,2)),
    tf.keras.layers.Conv2D(64, (3,3), activation = 'relu'),
    tf.keras.layers.MaxPool2D((2,2)),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(128, activation = 'relu'),
    tf.keras.layers.Dense(64, activation = 'softmax'),
])
modelo_cnn.build(input_shape=input_shape)
```

Las capas convolucionales en la CNN se encargan de extraer características relevantes de las imágenes mediante el uso de filtros convolucionales. Luego, las capas de agrupación (max pooling) reducen la dimensionalidad y preservan las características más importantes. Se han definido varias capas convolucionales (Conv2D) seguidas de capas de agrupación (MaxPool2D) para extraer y reducir las características de las imágenes. Luego, se utiliza una capa Flatten para aplanar los datos y una capa densa (Dense) con activación ReLU para aprender patrones más complejos. Finalmente, se aplica una capa densa de salida con activación softmax para clasificar las enfermedades de la papa.

Justificación del clasificador.

Los seres vivos detectamos y reconocemos formas, objetos, colores, contornos en imágenes con facilidad. Esta tarea no es complicada en nuestro entorno, pero si se presenta una gran dificultad para los sistemas de visión por computadora.

Uno de los principales elementos a considerar para la elección del clasificador es la capacidad de extracción de características por parte del clasificador. La extracción de características juega un papel importante para la identificación de un objeto. Particularmente en el procesamiento de imágenes se utiliza la extracción de color, la textura, la morfología, los bordes, entre otros.

Las características más relevantes que se pueden utilizar en la detección de enfermedades de las plantas son textura significa cómo se distribuye el color en la imagen, la aspereza, la dureza de la imagen, características las cuales son utilizadas para la detección de áreas de plantas infectadas.(Yandún Velasteguí, M. A, 2020).

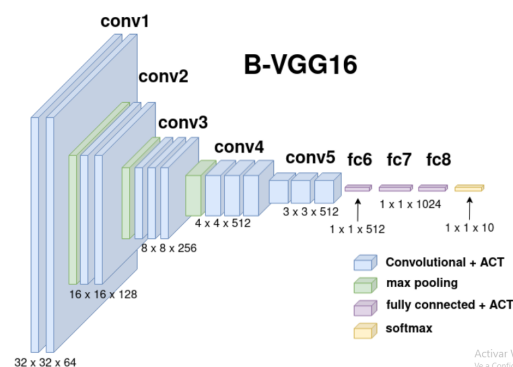
En el presente trabajo se eligió una red neuronal convolucional para la detección de enfermedades en plantas.

Las redes neuronales convolucionales (CNN) son efectivas para procesar imágenes y extraer automáticamente características relevantes (LeCun et al., 1998). Las capas convolucionales en una CNN analizan las imágenes a diferentes niveles de abstracción, permitiendo que el modelo aprenda representaciones internas altamente discriminatorias para la clasificación de imágenes (Krizhevsky et al., 2012).

La estructura jerárquica de las capas en una red neuronal permite que las primeras capas aprendan características de bajo nivel, como bordes y texturas, mientras que las capas posteriores aprenden características de nivel superior basadas en las características de las capas anteriores (Zeiler & Fergus, 2014). Esto permite que el modelo capture información contextual y relaciones espaciales en las imágenes, lo que es esencial para una clasificación precisa.

Las capas de pooling y las capas convolucionales con tamaños de filtro pequeños permiten que las redes neuronales sean sensibles a características locales en las imágenes, independientemente de su escala.

Para la arquitectura de una red neuronal convolución existen varios enfoques, en entre ellos algunos que resultaron efectivos para un contexto similar fue trabajos para diagnóstico de enfermedades del tomate es presentado por Rangarajan y por Brahimi . En ambos trabajos se utilizaron arquitecturas de RNC pre entrenadas. En Rangarajan utilizaron las arquitecturas AlexNet y VGG16 para reconocer 6 enfermedades del tomate, alcanzando porcentajes de eficacia en clasificación del 97.49% y 97.29% respectivamente. Brahimi en su trabajo presenta resultados de clasificación de 9 tipos de enfermedades del tomate alcanzando una eficacia del 99.18%.



En este caso hay que tener en cuenta que se trabaja con grandes conjuntos de datos así también se considera una gran variabilidad en sus datos, este tipo de modelos son los que mejor se adaptan a datos de la vida real y mejores resultados arrojan.

Sin embargo para este presente trabajo es suficiente una red neuronal simple ya que se maneja un conjunto de datos pequeños y poca variabilidad de datos ya que el conjunto de datos de imágenes no considera los efectos de luz o la calidad de la imagen los datos, pero así también existe la probabilidad de que no funcione adecuadamente con datos reales.

Las capas indispensables para una red neuronal simple son las siguientes mismas con las que se experimentó y se alcanzó una configuración aceptable para este conjunto de datos

- Capa de entrada: Esta capa toma las imágenes como entrada. utilizando una capa de convolución inicial para extraer características de bajo nivel de las imágenes.
- Capas convolucionales: Esta capa se agrega varias capas convolucionales para aprender características más abstractas y complejas a medida que se profundiza en la red. utilizando diferentes tamaños de filtros y número de filtros en cada capa para capturar diferentes patrones.
- Capas de pooling: Después de cada capa convolucional, se agrega una capa de pooling para reducir la dimensionalidad de las características y permitir una mejor generalización.
- Capas totalmente conectadas: Después de las capas convolucionales y de pooling, aplanan las características y se agrega una o varias capas totalmente conectadas. Estas capas aprenden las relaciones entre las características extraídas y realizan la clasificación final.
- Capa de salida: La capa de salida tiene tantos nodos como clases diferentes de enfermedades desees detectar en los cultivos de papa. se utiliza una función de activación adecuada, como softmax, para obtener una distribución de probabilidades sobre las clases.

En resumen la elección de una red neuronal convolucional se justifica por su capacidad para capturar información contextual y relaciones espaciales en las imágenes, así como por su capacidad para aprender características de bajo nivel, como bordes y texturas, y características de nivel superior basadas en las capas anteriores. También se han citado trabajos previos en diagnóstico de enfermedades de plantas que demuestran la eficacia de las CNN en contextos similares.

Resultados.

Los resultados obtenidos son los siguientes:

Para una primera ejecución con 100 épocas, y una división de datos 80(train)/20(test) se obtuvo:



Para la matriz de confusión se realizó la predicción correspondiente con el conjunto de datos de test de 20% y se obtuvo una predictibilidad 94% bastante bueno, con la matriz :

	Tizón_tardío	Tizón_temprano	hojas sanas
Tizón_tardío	171	22	4
Tizón_temprano	0	199	0
hojas sanas	0	0	100

Para una segunda ejecución con 100 épocas, y una división de datos 50(train)/50(test) se obtuvo una predictibilidad 91% inferior al caso anterior ya que se tuvo un conjunto de datos más pequeño para entrenamiento, con una matriz de:

	Tizón_tardío	Tizón_temprano	hojas sanas
Tizón_tardío	85	17	6
Tizón_temprano	0	110	0
hojas sanas	0	0	54

Los resultados obtenidos son bastante buenos y aceptables para un conjunto de datos como este. Es importante destacar que este proyecto de inteligencia artificial es un primer paso para abordar el problema de las enfermedades en los cultivos de papa. Se espera que los resultados obtenidos impulsen futuras investigaciones y mejoras en la detección y control de enfermedades en este importante cultivo.

Conclusiones.

La detección de los diferentes tipos de enfermedades que se presentan en los cultivos de papa durante los procesos de desarrollo, es de gran relevancia ya que la detección oportuna de una plaga en etapa temprana ahorra tiempo y costos al momento de aplicar el tratamiento necesario, generando productos sanos para el consumo y reducción de la contaminación por el uso excesivo de fungicidas

El modelo implementado a partir de redes neuronales convolucionales, adaptado para clasificar enfermedades en cultivos de papa, presentan grandes resultados con precisiones superiores a 90% en todas las configuraciones experimentales. Y los mapas de características muestran que el modelo logra detectar correctamente en la imagen las zonas donde existe una patología.

Referencias:

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097-1105. DOI: 10.1145/3065386

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324. DOI: 10.1109/5.726791

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. *European Conference on Computer Vision*, 818-833. DOI: 10.1007/978-3-319-10590-1_53

Yandún Velasteguí, M. A, 2020. Detección de enfermedades en cultivos. DOI <https://doi.org/10.48190/cumbres.v6n1a4>

<https://tizon.inia.cl/assets/boletines/inia%20carillanca%20-%20principales%20enfermedades%20que%20afectan%20el%20cultivo%20de%20la%20papa.pdf>