



Revista Ciencia Multidisciplinaria CUNORI

Clasificación automática de la severidad del Virus del Mosaico Dorado en frijol

Automatic classification of severity of Golden Mosaic Virus in beans

Carlos David Ardón Muñoz

Universidad de San Carlos de Guatemala
<http://orcid.org/0000-0003-4359-4665>
carlos_david@cunori.edu.gt

Bryan José Morales Calderón

Universidad de San Carlos de Guatemala
bryanjosemoralescalderon@gmail.com

Recibido: 25/10/2020

Publicado: 19/01/2021

Referencia del artículo

Ardón Muñoz, C. D. y Morales Calderón, B. J. (2021). Clasificación automática de la severidad del Virus del Mosaico Dorado en frijol. Revista Ciencia Multidisciplinaria CUNORI. 5(1). 41-49. DOI: <https://doi.org/10.36314/cunori.v5i1.151>

Resumen



Se realizó un estudio comparativo entre dos arquitecturas de redes neuronales, InceptionV3 y ResNet50, para determinar cuál es la mejor en la clasificación de la severidad del Virus del Mosaico Dorado en frijol. Para esto, se recolectaron y clasificaron manualmente 3409 imágenes de hojas de frijol. Las muestras se organizaron en dos formas de agrupación, tres categorías (alto, medio y bajo) y dos categorías (alto y bajo). Además, los entrenamientos se realizaron con dos niveles de extracción de características: todas las capas y desde capas intermedias. Los resultados de los entrenamientos sobre los datos distribuidos en tres categorías produjeron modelos con high-bias. Por otro lado, los modelos entrenados sobre el conjunto de datos con dos categorías produjeron las exactitudes más altas sobre el conjunto de pruebas al extraer características desde capas intermedias (ResNet50=96.68% e InceptionV3=94.47%). Con una prueba de McNemar se determinó que la diferencia es estadísticamente significativa a un nivel de significancia del 5%. Por tanto, ResNet50 con extracción de características desde capas intermedias posee la exactitud más alta en la tarea estudiada.

Palabras clave

inteligencia artificial, redes neuronales, visión artificial

Abstract



We conducted a comparative study between two neural network architectures, InceptionV3 and ResNet50, to determine which one is better at classifying the severity of the Golden Mosaic Virus in beans. For this, we collected and manually classified 3409 bean leaf images. We organized the samples in two ways: three categories (high, medium, and low) and two categories (high and low). Furthermore, we trained the models with two levels of feature extraction: all layers and intermediate layers. The results of the training on the data distributed in three categories produced models with high-bias. On the other hand, models trained on the data set with two categories produced the highest accuracy on the test set when extracting features from intermediate layers (ResNet50=96.68% and InceptionV3=94.47%). With a McNemar test, we determine that the difference is statistically significant at a 5% significance level. Therefore, ResNet50 with feature extraction from intermediate layers has the highest accuracy in the studied task.

Keywords

artificial intelligence, neural networks, computer vision





Introducción

El frijol es uno de los cultivos esenciales para la población guatemalteca. Se consume independientemente del nivel de pobreza en por lo menos el 80% de los hogares del país (Menchú & Méndez, 2011, p. 23). En el caso de Chiquimula, constituye uno de los principales cultivos, según el Ministerio de Agricultura, Ganadería y Alimentación (2017), este departamento aporta el 10% de la producción nacional. Sin embargo, existen múltiples enfermedades y plagas que pueden afectar su rendimiento. Por tal motivo, es de suma importancia el control de las mismas. En el oriente de Guatemala, una de las principales enfermedades que afecta al frijol es el Virus del Mosaico Dorado (VMDAF). Esta es una enfermedad viral que puede llegar a causar la pérdida total del cultivo. Algunas de las características que presentan las plantas infectadas son tonos amarillos en las hojas, venas más blancas de lo normal y las hojas tienden a enrollarse hacia la parte inferior (Instituto Interamericano de Cooperación para la Agricultura, Proyecto Red Sistema de Integración Centroamericano de Tecnología Agrícola, & Cooperación Suiza para el Desarrollo, 2008). Debido a estos rasgos, es posible detectarlo por simple inspección visual.

Un experto humano es capaz de no solo determinar la existencia de la infección, sino también puede establecer su severidad por medio de la observación. Para esto, el Centro Integral de Agricultura Tropical (CIAT) desarrolló una escala general para evaluación de enfermedades virales (van-Schoonhoven & Pastor-Corrales, 1987). La escala se puede utilizar para clasificar la severidad de la infección por medio de la observación de hojas de los cultivos. No obstante, este método implica la formación de los profesionales humanos en este ámbito, que puede ser costoso y tardado, además, podría ser afectado por la subjetividad de la persona.

Para evitar los problemas asociados a la dependencia de expertos humanos o para apoyar a los mismos, se pueden utilizar herramientas de análisis de imágenes por medio de computadoras. Esta es una tarea que se realiza rutinariamente en el cerebro humano, pero que resulta complicada para programas estándar de computador. Debido a esta complejidad, las investigaciones han dejado de tratar de simular el comportamiento del cerebro por medio de programas imperativos basados en reglas y se han centrado en crear programas que aprenden a partir de un conjunto de datos (Alpaydin, 2016). Por este motivo, los sistemas de aprendizaje automático son utilizados ampliamente en tareas de análisis de imágenes.

Uno de los tipos de algoritmos de aprendizaje automático que se utiliza actualmente son las redes neuronales convolucionales. Estos son algoritmos que requieren un conjunto de datos previamente clasificados por expertos humanos. A partir de estos datos se construye un modelo matemático para identificar y clasificar imágenes que no se encontraban originalmente en el conjunto de entrenamiento. Para la clasificación de imágenes, las redes neuronales han demostrado ser especialmente útiles y capaces de alcanzar una alta exactitud (Krizhevsky, Sutskever, & Hinton, 2012). En consecuencia, las redes neuronales son el principal método utilizado actualmente para clasificar imágenes de forma automática.

Para desarrollar una red neuronal para una aplicación específica deben establecerse algunos parámetros antes de entrenarla con un conjunto de datos. El conjunto de características tales como la función de activación de cada neurona, la cantidad de neuronas por capa y la cantidad de capas son llamadas la arquitectura o topología de la red neuronal (Gnana-Sheela & Deepa, 2013). No obstante, no existe un consenso sobre cuáles son los valores óptimos para estos parámetros y pueden ser distintos para cada aplicación. Según Stathakis (2009) el método tradicional para encontrar estos valores es por prueba y error. Esto puede ser un proceso tardado y que no encuentre los valores óptimos.

Por lo anterior, en este estudio se entrenaron y compararon dos arquitecturas que han demostrado una gran exactitud en la clasificación de imágenes, GoogLeNet (InceptionV3) y ResNet50, las cuales han ganado el reto de reconocimiento visual a gran escala ImageNet (Szegedy et al., 2015; He, Zhang, Ren, & Sun, 2016). Esto requirió la recolección de muestras fotográficas de hojas infectadas en distintas escalas, para posteriormente, entrenar a las redes neuronales y determinar cuál de las dos clasifica la severidad del Virus del Mosaico Dorado a nivel foliar con mayor exactitud. Para de esta forma, proveer de una aplicación informática que le permita a cualquier persona clasificar la severidad de la infección de manera automática.

Materiales y métodos

Para el conjunto de datos de entrenamiento se recolectaron un total de 3409 fotografías de hojas de frijol con Virus del Mosaico Dorado en diferentes severidades, utilizando cámaras de dispositivos móviles de al menos 12 mega-píxeles. Las imágenes fueron clasificadas manualmente por



un humano utilizando la escala del uno al nueve propuesta por el CIAT. Durante el proceso de entrenamiento las imágenes se sometieron a un proceso de aumento de datos usando cortes, acercamientos, rotaciones, desplazamientos, volteo y modificación del brillo. Además, las imágenes se organizaron en dos formas de agrupación: tres grupos (alto, medio y bajo) y dos grupos (alto y bajo). Finalmente, para cada forma de agrupación, se distribuyeron de manera aleatoria y estratificada en conjuntos de entrenamiento (60%), validación (20%) y pruebas (20%).

Para el entrenamiento por transferencia de aprendizaje se utilizaron implementaciones de las redes neuronales en Tensorflow y pre-entrenadas en el conjunto de datos de ImageNet. Así mismo, a partir de pruebas piloto se determinaron los valores para los hiperparámetros mostrados en la tabla 1. Se estudiaron dos formas de extracción de características de las redes neuronales: con todas las capas y a partir de capas intermedias. En el caso de InceptionV3 se extrajeron las características de la capa mixed, mientras que para ResNet50, fue a partir de la capa conv5_block1_pre-act_relu. Durante el entrenamiento, la capa de clasificación se entrenó durante 35 epochs y cinco epochs adicionales de fine-tuning.

Tabla No. 1 Hiperparámetros usados en los entrenamientos de redes neuronales

Hiperparámetro	Valor
Función de optimización	RMSprop Stochastic gradient descent (fine-tuning)
Función de pérdida	Categorical cross-entropy
Tasa de aprendizaje	0.001
Momentum	0.9 (fine-tuning)
Batch size	20

Fuente: propia.

Resultados

Las neuronales entrenadas sobre los datos distribuidos en tres grupos (alto, medio y bajo) produjeron exactitudes muy bajas en la clasificación,

siendo la más alta ResNet50 a partir de capas intermedias con una exactitud del 83.4% en el conjunto de pruebas. La matriz de confusión revela que las redes neuronales tiene una alta inexactitud para diferenciar la clase del medio. Mientras que los resultados obtenidos por las redes neuronales entrenadas sobre los datos distribuidos en dos grupos (alto y bajo) fueron mejores (tabla 2), llegando a alcanzar 96.68% en el conjunto de pruebas.

Tabla No. 2 Exactitudes de las redes neuronales en la clasificación de la severidad del Virus del Mosaico Dorado en frijol en dos niveles

		Exactitud al epoch 40		
Arquitectura	Nivel	Conjunto de entrenamiento (%)	Conjunto de validación (%)	Conjunto de pruebas (%)
InceptionV3	Todas las capas	96.28	90.57	92.12
	Capa intermedia	95.98	94.59	94.47
ResNet50	Todas las capas	96.28	90.57	96.27
	Capa intermedia	96.03	95.98	96.68

Fuente: propia.

Debido a que las redes neuronales entrenadas sobre los datos distribuidos en tres grupos tienen exactitudes bajas en el conjunto de entrenamiento, tanto como en el conjunto de pruebas, esto indica un problema de high-bias. Por tal motivo, se descartan estas redes neuronales pues no tienen utilidad práctica en la tarea de clasificación. Por otra parte, cuando las redes neuronales se entrenan sobre los datos distribuidos en dos categorías se obtienen mejores resultados. Sin embargo, las versiones entrenadas utilizando todas las capas de la red presentan una diferencia en la exactitud entre los conjuntos de validación y pruebas, lo que indica la existencia de high-variance. En consecuencia, las únicas redes neuronales con resultados aceptables son las entrenadas sobre los datos distribuidos en dos categorías y usando capas intermedias.

Las dos redes neuronales seleccionadas obtuvieron una alta exactitud en el conjunto de pruebas. En el caso de InceptionV3 alcanzó el 94.47% de exactitud, mientras ResNet50 el 96.68%. A estos resultados se les aplicó una prueba de McNemar con corrección de continuidad, la cual produjo los valores de: $\chi^2=7.0312$, $df=1$ y $p\text{-value}=0.00801$. En consecuencia, a un nivel de significancia del 5% la diferencia es estadísticamente significativa.



Discusión

La arquitectura de red neuronal que obtuvo los mejores resultados fue ResNet50 entrenada sobre el conjunto de datos distribuido en dos categorías y a partir de capas intermedias. La exactitud alcanzada en esta tarea (96.68%) es lo suficientemente alta como para utilizarse en aplicaciones prácticas. Además, debido a que las características se extraen desde capas intermedias, su implementación en dispositivos móviles es factible por requerir menos poder de cómputo.

Por otra parte, la baja exactitud en la clasificación de la severidad con más de dos grupos concuerda con los resultados obtenidos por Esgario, Krohling y Ventura (2020), quienes estudiaron la clasificación de la severidad de enfermedades en café. Lo que indica que los modelos actuales de redes neuronales convolucionales tienen altas exactitudes para determinar la existencia de una enfermedad, pero bajas para clasificar su severidad en más de dos categorías. Por tal motivo, futuras investigaciones podrían explorar la utilización de métodos alternativos como agrupación (clustering), regresión, segmentación y árboles de decisiones para mejorar los resultados en la clasificación de severidad con más de dos categorías.

Agradecimientos

Este proyecto fue ejecutado gracias al apoyo financiero del Departamento de Agricultura de los Estados Unidos (USDA, por sus siglas en inglés). Las opiniones expresadas en esta publicación son las de su(s) autor(es) o institución(es) a las que pertenecen. La mención de empresas o productos comerciales no implica la aprobación o preferencia sobre otros de naturaleza similar que no se mencionan.

Referencias

Alpaydin, E. (2016). Machine Learning: the new AI. MIT Press.

Esgario, J. G. M., Krohling, R. A., & Ventura, J. A. (2020). Deep learning for classification and severity estimation of coffee leaf biotic stress. En *Computers and Electronics in Agriculture*, 169. <https://doi.org/10.1016/j.compag.2019.105162>

- Gnana-Sheela, K., & Deepa S. N. (2013). Review on Methods to Fix Number of Hidden Neurons in Neural Networks. En *Mathematical Problems in Engineering*, 2013. <https://doi.org/10.1155/2013/425740>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. En *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- Instituto Interamericano de Cooperación para la Agricultura, Proyecto Red Sistema de Integración Centroamericano de Tecnología Agrícola, & Cooperación Suiza para el Desarrollo. (2008). Guía de identificación y manejo integrado de las enfermedades del frijol en América Central. IICA. Recuperado de <http://repiica.iica.int/docs/B0891E/B0891E.pdf>
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. En *Advances in Neural Information Processing Systems*, 25(2). <https://doi.org/10.1145/3065386>
- Menchú, M. T., & Méndez, H. (2011). Análisis de la situación alimentaria en Guatemala. Guatemala: Instituto de Nutrición de Centro América y Panamá. Recuperado de http://www.incap.int/index.php/es/publicaciones/publicaciones-incap/doc_view/663-guatemala-informe-analisis-de-situacion-alimentaria
- Ministerio de Agricultura, Ganadería y Alimentación. (2017). Informe situacional del frijol. DIPLAN/MAGA.
- Stathakis, D. (2009). How many hidden layers and nodes?. *International Journal of Remote Sensing*, 30(8), 2133-2147. <https://doi.org/10.1080/01431160802549278>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. En *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>

Van-Schoonhoven, A., & Pastor-Corrales, M. A. (1987). Sistema estándar para la evaluación de germoplasma de frijol. Centro Integral de Agricultura Tropical.

Sobre los autores

Carlos David Ardón Muñoz

Es Ingeniero en Ciencias y Sistemas egresado de la Universidad de San Carlos de Guatemala. Se desempeña como profesor de las Carreras de Ingeniería en el Centro Universitario de Oriente, donde imparte cursos relacionados con ciencia de la computación y metodología de la investigación.

Bryan José Morales Calderón

Es Ingeniero Agrónomo en Sistemas de Producción, egresado de la Universidad de San Carlos de Guatemala. Ha realizado investigaciones en maíz como investigador principal y en frijol como investigador adjunto. Actualmente labora como Director Municipal de Planificación en la Municipalidad de San Jacinto, Chiquimula.

Copyright (c) Carlos David Ardón Muñoz y Bryan José Morales Calderón



Este texto está protegido por una licencia [Creative Commons 4.0](https://creativecommons.org/licenses/by/4.0/).

Usted es libre para compartir, copiar y redistribuir el material en cualquier medio o formato y adaptar el documento, remezclar, transformar y crear a partir del material para cualquier propósito, incluso comercialmente, siempre que cumpla la condición de **atribución**: usted debe reconocer el crédito de una obra de manera adecuada, proporcionar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que tiene el apoyo del licenciante o lo recibe por el uso que hace.