



Desarrollo de una red neuronal artificial para clasificar la severidad del Virus del Mosaico Dorado a nivel foliar en frijol (*Phaseolus vulgaris* L.)

Carlos David Ardón Muñoz
Bryan José Morales Calderón

Chiquimula, julio de 2020

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.

Contenido

1. Introducción	2
2. Marco teórico	3
2.1. Frijol	3
Mosaico dorado amarillo, mosaico dorado, mancha amarilla (VMDAF)	3
2.2. Visión artificial	4
2.2.1. Redes neuronales artificiales	4
2.2.2. GoogLeNet	6
2.2.3. ResNet	6
2.2.4. Aumento de datos	7
3. Objetivos	8
4. Hipótesis	8
5. Metodología	9
5.1. Localización	9
5.2. Delimitación temporal	9
5.3. Conjunto de datos	9
5.4. Preprocesamiento de imágenes	11
5.5. Entrenamiento de redes neuronales	12
5.6. Variables de respuesta	13
5.7. Análisis de la información	13
6. Resultados y Discusión	14
7. Conclusiones	19
8. Recomendaciones	20
9. Referencias	21
Anexos	23

Resumen

El Virus del Mosaico Dorado (BGMV) es una de las principales enfermedades que afectan al frijol (*Phaseolus vulgaris* L.) en el oriente de Guatemala. Para contrarrestarlo, los investigadores identifican la severidad en distintas variedades de frijol, lo que permite determinar su tolerancia y la efectividad del manejo del cultivo. Sin embargo, el proceso de identificación de la severidad se desarrolla normalmente de forma manual. Por lo que se desarrolló una red neuronal que le permite a cualquier persona, con apoyo de un teléfono inteligente Android, clasificar automáticamente la severidad. Para esto se recolectaron 3409 imágenes de hojas de frijol y se clasificaron en la escala del 1 al 9 propuesta por el CIAT. Se estudiaron dos arquitecturas de redes neuronales (GoogLeNet y ResNet50), dos formas de categorizar las muestras (alto-medio-bajo y alto-bajo) y dos niveles de extracción de características (todas las capas y capas intermedias). Las redes neuronales entrenadas sobre el conjunto de datos con tres categorías resultaron en modelos con *high bias*. Las redes neuronales entrenadas sobre los datos organizados en dos categorías y usando capas intermedias produjeron las mejores exactitudes sobre el conjunto de pruebas (ResNet50=96.68% y GoogLeNet=94.47%). Una prueba de McNemar con corrección de continuidad reveló que la diferencia es estadísticamente significativa (chi-squared=7, p-value=0.008) a un nivel de significación del 5%. Por lo que se desarrolló una aplicación para Android utilizando a ResNet50.

1. Introducción

El frijol es uno de los principales cultivos en Guatemala, sin embargo, es susceptible a enfermedades que pueden afectar su desempeño productivo. Algunas de estas enfermedades pueden detectarse con una simple inspección visual. Además, es posible que un experto pueda determinar la severidad de la infección con solo observar las hojas afectadas, como es el caso del Virus del Mosaico Dorado (BGMV).

Disponer de expertos que puedan determinar el nivel de tolerancia de una variedad de frijol a la infección por BGMV puede resultar en un costo elevado. Por lo que una aplicación informática que permita a un usuario sin formación en este ámbito realizar un diagnóstico, reduciría drásticamente los costos. Adicionalmente, una aplicación de esta naturaleza también es de beneficio para los expertos, pues proporcionaría un apoyo para confirmar su criterio, así como para ayudar en la formación de nuevos expertos.

Para desarrollar esta aplicación informática se utilizaron técnicas de aprendizaje automático supervisado, un área de la Inteligencia Artificial, que entrena modelos matemáticos a partir de una base de datos. En este caso la base de conocimientos consistió en un conjunto de fotografías de hojas de frijol; recolectadas en campo y clasificadas por un experto según la severidad de la infección por BGMV en una escala del 1 al 9.

Los modelos matemáticos utilizados fueron redes neuronales. Se entrenaron y compararon dos modelos de redes neuronales, GoogLeNet y ResNet50, para determinar cuál de estas dos topologías realiza la clasificación con una mayor exactitud. El conjunto de datos se agrupó en las severidades alta-baja y alta-medio-baja. Para cada agrupación se crearon conjuntos de entrenamiento (60%), validación (20%) y pruebas (20%). Únicamente los modelos entrenados sobre el conjunto alto-bajo dieron exactitudes altas. De donde se determinó que ResNet50 tiene una mejor exactitud (96.54%).

Finalmente, se creó una aplicación prototipo para dispositivos móviles Android, que permite a cualquier persona clasificar automáticamente nuevas fotografías. Además, se desarrolló una aplicación adicional que permite calcular la proporción de colores verde y amarillo en una imagen. Esta última aplicación puede servir como apoyo en el proceso de clasificación manual de la severidad.

2. Marco teórico

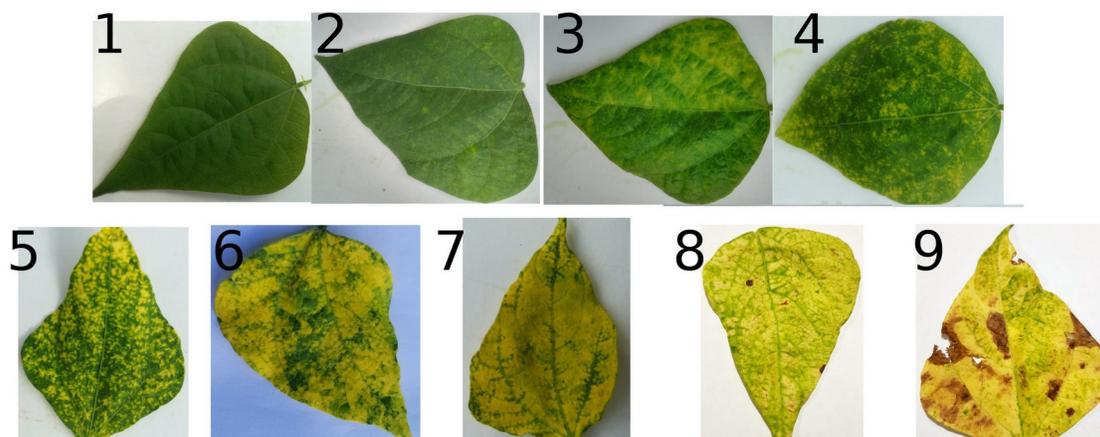
2.1. Frijol

El frijol es un cultivo de suma importancia para la población de Guatemala. Es una de sus principales fuentes de proteínas. “Son usados en todas las regiones por más del 80% de hogares, independientemente del nivel de pobreza” (Menchú & Méndez, 2011, p. 23). Para el departamento de Chiquimula, el frijol constituye uno de sus cultivos primordiales, aportando este departamento, según el Ministerio de Agricultura, Ganadería y Alimentación (2017) el 10% de la producción nacional. Por lo tanto, es trascendental el control de plagas y enfermedades que puedan reducir su producción.

Mosaico dorado amarillo, mosaico dorado, mancha amarilla (VMDAF)

Es una enfermedad viral que puede causar importantes pérdidas en el cultivo de frijol. Se puede identificar visualmente debido a que “en las hojas se observa un moteado de tonos amarillos hasta amarillo fuerte con venas más blancas de lo normal. La hoja puede 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). Por lo que es factible detectarlo por medio de imágenes de las hojas y clasificar la severidad de la infección, para lo cual es posible utilizar la figura 1 como referencia.

Figura 1. Escala diagramática para evaluación de la severidad del Virus del Mosaico Dorado, en el oriente de Guatemala, 2019.



2.2. Visión artificial

Esta disciplina estudia el análisis de imágenes por medio de computador para la identificación y clasificación de objetos en las mismas, tal como lo haría un humano. Esta tarea se desarrolla rutinariamente en el cerebro, sin embargo, ha demostrado ser especialmente complicada de realizar por medio de programas informáticos, al igual que otras aplicaciones, como el análisis del lenguaje natural. Debido a esta complejidad, las investigaciones han dejado de tratar de crear programas imperativos basados en reglas que simulen el comportamiento del cerebro humano en este aspecto; y se han centrado en crear programas que aprenden a partir de un conjunto de datos (Alpaydin, 2016).

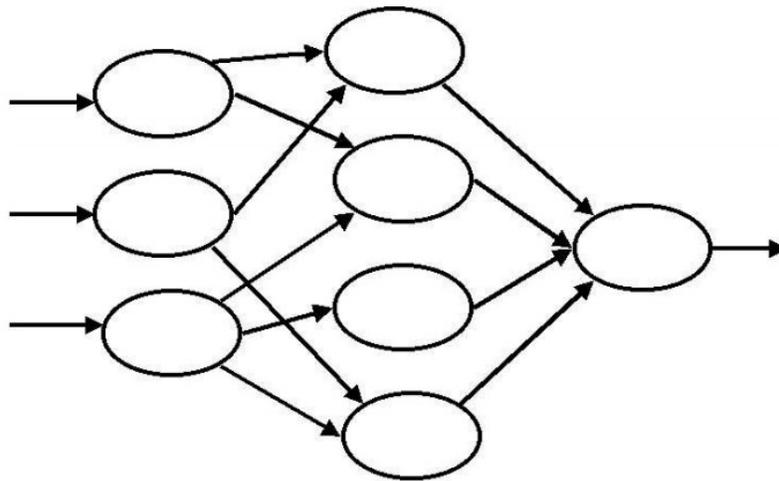
Para los programas de visión artificial, el conjunto de datos de entrenamiento son imágenes. Estas deberán recolectarse en grandes cantidades y ser procesadas antes de utilizarse por los programas. Según Sucar y Gómez (2011) algunas modificaciones comunes son remover defectos como problemas por movimiento o desenfoco, mejorar ciertas propiedades como color, contraste o estructura y agregar ‘colores falsos’ a imágenes monocromáticas (p. 1). Este procesamiento del conjunto de datos es fundamental para el éxito de los algoritmos de aprendizaje automático informado.

Los algoritmos de aprendizaje automático informado son programas de inteligencia artificial que requieren un conjunto de datos clasificados por expertos humanos. Estos programas toman los datos de muestra y construyen un modelo matemático que les permite, posteriormente, identificar y clasificar imágenes que no se encontraban originalmente en el conjunto de datos de entrenamiento. Existe diversidad de algoritmos aplicables al aprendizaje automático informado, tales como la regresión lineal y logística. Sin embargo, en cuanto al análisis de imágenes, las redes neuronales han demostrado ser especialmente útiles (Krizhevsky, Sutskever, & Hinton, 2012; Szegedy et al., 2015; He, Zhang, Ren, & Sun, 2016).

2.2.1. Redes neuronales artificiales

Las redes neuronales artificiales (ANN por sus siglas en inglés) son modelos matemáticos inspirados en el funcionamiento de sus contrapartes biológicas. De acuerdo con Flores-López y Fernández-Fernández (2008), las ANN “imitan la estructura *hardware* del sistema nervioso, centrándose en el funcionamiento del cerebro humano, basado en el aprendizaje a través de la experiencia” (p. 16). Las ANN están formadas por nodos, llamados neuronas, cada una posee una función de activación y se organizan en capas (*layers*), tal como se muestra en la figura 2.

Figura 2. Red neuronal de tres capas de profundidad, tres características de entrada y un nodo de salida.



Fuente: L. E. Sucar y G. Gómez, 2011, p. 164.

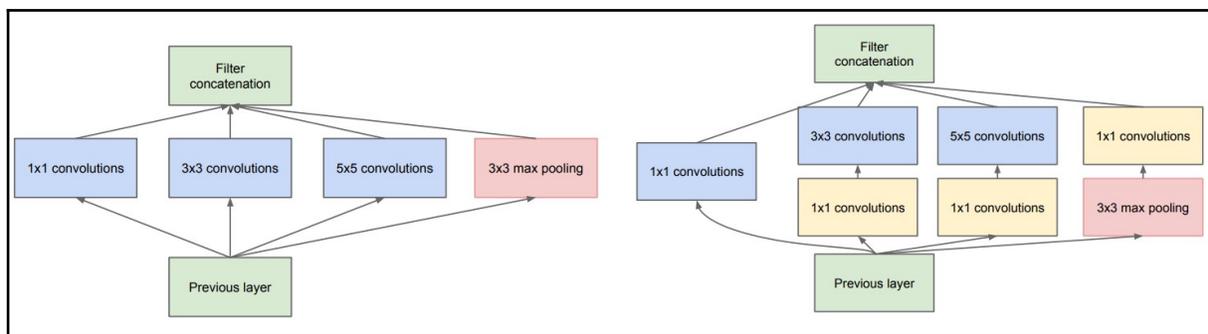
En las ANN es común que cada neurona tenga múltiples entradas y una sola salida. La capa de la izquierda es la entrada de datos, mientras que la capa del extremo derecho es la de salida. Todas las capas internas son llamadas ocultas. La eficacia y eficiencia de las ANN depende en gran medida de la calidad del conjunto de datos de entrenamiento, sin embargo, también es importante considerar el tipo de función de activación de cada neurona, la cantidad de neuronas por capa y la cantidad de capas ocultas. Todas estas características son llamadas la arquitectura o topología de la red neuronal (Gnana-Sheela & Deepa, 2013).

Actualmente no existe un consenso sobre cuál es la mejor arquitectura de red neuronal. Además, una arquitectura que funciona bien para cierta aplicación en particular, puede que no tenga un buen desempeño en otras aplicaciones. Por lo tanto, es necesario encontrar para cada aplicación la estructura adecuada. Según Stathakis (2009) el método tradicional para determinar la mejor estructura es por prueba y error. Lo cual puede ser un proceso tardado y que no alcance la respuesta óptima. No obstante, existen algunas arquitecturas que han demostrado gran exactitud en la clasificación de imágenes, algunas de estas topologías son GoogLeNet y ResNet, las cuales han ganado el “Reto de reconocimiento visual a gran escala ImageNet”.

2.2.2. GoogLeNet

Se trata de una red neuronal convolucional profunda, con una arquitectura compleja. Tal como lo indican Szegedy et al. (2015), la red tiene una profundidad de 22 capas con parámetros y 27 si se cuentan las capas *max-pooling*, las cuales se usan para reducción de muestreo. Las capas están agrupadas en *inception modules*, que a su vez se encuentran apilados un módulo después del otro. Debido al tamaño de esta red neuronal, en la figura 3 se muestran únicamente la topología de las dos clases de módulos desarrollados en esta arquitectura.

Figura 3. Módulos de GoogLeNet utilizados en una de las dos redes neuronales comparadas para clasificar la severidad del Virus del Mosaico Dorado en frijol, en el oriente de Guatemala, 2019.



Fuente: adaptado de Szegedy et al., 2015, p. 4.

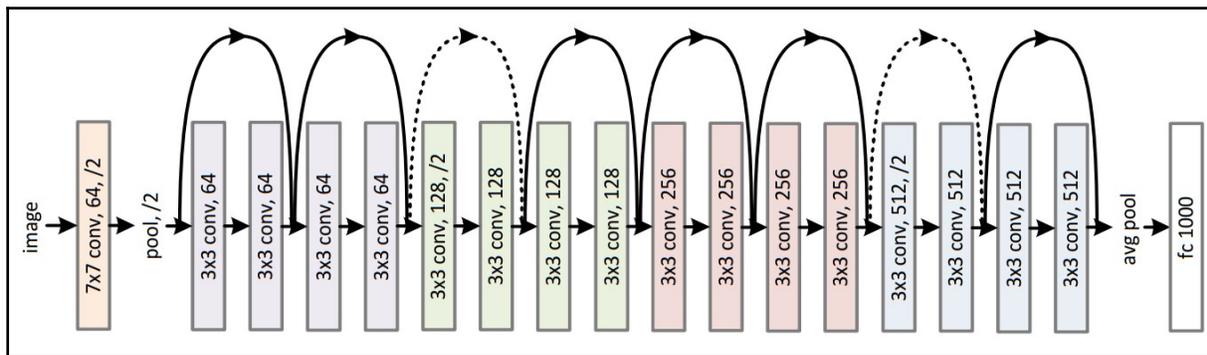
Este tipo de ANN recibe el nombre de “convolucional profunda”, debido a que algunas capas utilizan la operación matemática convolución y que contiene múltiples capas ocultas, convirtiéndola en profunda. En contraposición al estándar en donde solamente existe una capa oculta.

2.2.3. ResNet

Es un tipo de red neuronal convolucional profunda, que utiliza saltos entre las conexiones para reutilizar los valores de activación de capas previas. Esto se realiza para evitar el problema del desvanecimiento del gradiente. Es común que las redes neuronales usen cálculos de gradientes para encontrar el costo mínimo de la función objetivo, sin embargo, a medida que el gradiente se acerca al óptimo, los pasos tienden a hacerse más pequeños, lo que provoca que el entrenamiento tarde más. Con la reutilización de valores anteriores, se espera acelerar la convergencia del gradiente y por tanto, el tiempo de entrenamiento (He, Zhang, Ren, & Sun, 2016).

Esta topología tiene múltiples variantes en cuanto al número de capas. He, Zhang, Ren y Sun (2016) propusieron y compararon dos versión, la primera con 18 capas y la segunda con 34, llamándolas ResNet-18 y ResNet-34, respectivamente. En la figura 4 se detalla la arquitectura de ResNet-18. Para este estudio se utilizó ResNet-50 que es una variante de las propuestas originalmente. Se diferencia de las anteriores, en que tiene un mayor número de capas ocultas.

Figura 4. Topología de ResNet-18 utilizada como base para la arquitectura de una de las dos redes neuronales comparadas para clasificar la severidad del Virus del Mosaico Dorado en frijol, en el oriente de Guatemala, 2019.



Fuente: K. He, X. Zhang, S. Ren y J. Sun, 2016.

2.2.4. Aumento de datos

Otro concepto relevante para el entrenamiento de cualquier tipo de algoritmo de aprendizaje automático informado es el aumento de datos (*data augmentation*). Consiste en crear nuevas versiones modificadas de las muestras. Por ejemplo, en el caso de fotografías, podrían crearse nuevas imágenes rotando, desplazando, volteando o recortando la fotografía original. Esto se realiza con el fin de incrementar el número de muestras, así como, minimizar el problema del sobre-ajuste.

El sobre-ajuste (*overfitting*) es un problema en el aprendizaje automático, caracterizado por una alta precisión para clasificar cuando se evalúa sobre el conjunto de datos de entrenamiento, pero una baja precisión cuando se evalúa en un conjunto de pruebas separado (Subramanian & Simon, 2013, p. 636). Es posible combatir este problema con métodos como el abandono de neuronas (*dropout*), lo cual se implementa en ResNet, pero no en GoogLeNet. Adicionalmente, el aumento de datos se puede utilizar en las dos arquitecturas mencionadas para reducir el sobre-ajuste.

3. Objetivos

General

Desarrollar una red neuronal artificial para clasificar la severidad de la infección provocada por el Virus del Mosaico Dorado (BGMV) a nivel foliar en frijol (*Phaseolus vulgaris* L.) de la cosecha 2019, en el oriente de Guatemala.

Específicos

1. Construir una base de datos de fotografías clasificadas en las diferentes severidades de la escala de evaluación de enfermedades virales propuesta por el CIAT, para su utilización en algoritmos de aprendizaje automático supervisado.
2. Determinar la mejor arquitectura de red neuronal en la clasificación de imágenes de la base de datos, determinada por la exactitud en la clasificación sobre un conjunto de pruebas.
3. Crear una aplicación informática prototipo para la clasificación automática de la severidad del daño causado por el Virus del Mosaico Dorado, mediante la toma de fotografías en el cultivo de frijol.

4. Hipótesis

H₀: Las arquitecturas de redes neuronales GoogLeNet y ResNet50 presentan una proporción de errores similar en la clasificación de la severidad de la infección provocada por el Virus del Mosaico Dorado a nivel foliar en frijol.

5. Metodología

5.1. Localización

Las muestras fotografías de hojas de frijol se tomaron en plantaciones pertenecientes a agricultores de los municipios de Camotán, Jocotán, Ipala, San José la Arada y San Juan Ermita del departamento de Chiquimula y del municipio de San Diego, Zacapa.

5.2. Delimitación temporal

Las muestras se recolectaron durante los meses de octubre y noviembre de 2019. Las cuales fueron clasificadas durante el mes de diciembre del mismo año.

Luego, las muestras fueron utilizadas para realizar el entrenamiento de redes neuronales durante enero de 2020 y los resultados fueron analizados en febrero, a partir de los que se desarrollaron las aplicaciones prototipo para dispositivos móviles en el mismo mes.

Finalmente, la divulgación de resultados, demostración y recolección de datos de retroalimentación de uso de las aplicaciones se llevó a cabo en el mes de mayo de 2020 de manera virtual por medio de una video conferencia, debido a las restricciones sanitarias impuestas por el Gobierno de Guatemala en respuesta a la pandemia de COVID-19.

5.3. Conjunto de datos

El conjunto de datos consiste en 3409 fotografías de hojas de frijol tomadas sobre un fondo blanco. Para lo que se utilizaron cámaras digitales estándar con color RGB de dispositivos móviles como *smartphones*, de al menos 12 megapíxeles.

Las imágenes fueron posteriormente descargadas a un computador personal, donde un experto humano las clasificó manualmente de acuerdo a la escala general para evaluación de enfermedades virales propuesta por van-Schoonhoven y Pastor-Corrales (1987), la distribución de las muestras se encuentra en la tabla 1.

Tabla 1. Cantidad de muestras recolectadas por cada nivel de severidad del Virus del Mosaico Dorado en frijol, según la escala CIAT, en el oriente de Guatemala, 2019.

Severidad	Número de muestras
1	442
2	186
3	199
4	767
5	962
6	409
7	355
8	59
9	30
Total:	3409

Para realizar el entrenamiento, se crearon dos conjuntos de datos, el primero agrupando las muestras en 3 categorías: alto, medio y bajo. Esta distribución se muestra en la tabla 2. Además, el segundo conjunto de datos se creó agrupando en 2 categorías: alto y bajo; cuya distribución se muestra en la tabla 3.

Tabla 2. Distribución de muestras en tres categorías de severidad del Virus del Mosaico Dorado en frijol, en el oriente de Guatemala, 2019.

Severidad	Categoría	Número de muestras
1	Baja	628
2		
3	Media	966
4		
5	Alta	1815
6		
7		
8		
9		

Tabla 3. Distribución de muestras en dos categorías de severidad del Virus del Mosaico Dorado en frijol, en el oriente de Guatemala, 2019.

Severidad	Categoría	Número de muestras
1		
2	Bajo	827
3		
4		
5		
6	Alto	2582
7		
8		
9		

Los dos conjuntos creados se dividieron aleatoriamente y de manera estratificada en tres sub-conjuntos: entrenamiento (60%), validación (20%) y pruebas (20%). El conjunto de entrenamiento se utilizó para alimentar las redes neuronales. Mientras que el conjunto de validación se utilizó durante el proceso optimización de hiperparámetros para verificar el sobre-ajuste. Finalmente el conjunto de pruebas se utilizó para determinar la exactitud del modelo en la clasificación.

5.4. Preprocesamiento de imágenes

Todas las imágenes recolectadas fueron sometidas a un proceso de modificación previo a utilizarlas en el entrenamiento de las redes neuronales:

- Recortado: las imágenes fueron recortadas manualmente utilizando un rectángulo circunscrito a la hoja, para minimizar la cantidad de pixeles que no representan a la muestra.
- Escala: con el objetivo de reducir el tamaño de mega bytes del conjunto de imágenes. Todas las fotografías fueron reducidas a la mitad de su tamaño original utilizando un algoritmo con suavizado (*antialiasing*) y permitiendo una pérdida del 5% (calidad=95%).

5.5. Entrenamiento de redes neuronales

Se evaluaron dos topologías de redes neuronales, GoogLeNet (InceptionV3) y ResNet50 implementadas en TensorFlow. Las cuales se seleccionaron debido a su alta exactitud en la clasificación de ImageNet. En cuanto a ResNet se seleccionó la variante de 50 capas, en contraposición a 150 capas, debido a que finalmente se ejecutaría sobre un dispositivo móvil. Por lo que se requiere que la red neuronal no sea demasiado profunda para que las inferencias no demoren mucho tiempo.

Se utilizó la técnica de transferencia de aprendizaje, utilizando modelos previamente entrenados sobre ImageNet y sustituyendo la capa de clasificación final por una nueva para el problema específico. A partir de pruebas piloto de entrenamiento sobre el conjunto de datos se determinó que cuando se utilizan todas las capas, la exactitud y la pérdida no tienden a converger en un valor. Debido a esto, se incluyeron en los entrenamientos versiones modificadas de las redes neuronales, donde las características se extrajeron de capas intermedias. En el caso de GoogLeNet las características intermedias se extrajeron a partir de la capa *mixed*, mientras que para ResNet50 fue a partir de la capa *conv5_block1_preact_relu*.

Las redes neuronales fueron entrenadas durante 40 *ephocs*. La capa de clasificación se entrenó durante 35 *ephocs*, congelando todas las demás. Mientras que en los 5 *ephocs* restantes se utilizó *fine-tuning*, entrenando las capas cercanas a la de clasificación. Los hiperparámetros utilizados durante el entrenamiento se detallan en la tabla 4.

Tabla 4. Hiperparámetros usados en los entrenamientos de las redes neuronales usadas para la clasificación de la severidad del Virus del Mosaico Dorado en frijol, en el oriente de Guatemala, 2019.

Parámetro	Valor
Función de optimización	RMSprop <i>Stochastic gradient descent (fine-tuning)</i>
Función de pérdida	<i>Categorical cross-entropy</i>
Tasa de aprendizaje	0.001
Momentum	0.9 (<i>fine-tuning</i>)
<i>Ephocs</i>	40
<i>Batch size</i>	20

Para minimizar el problema del sobre-ajuste se utilizaron técnicas de aumento de datos. La implementación se realizó haciendo uso de los métodos declarados en el *framework* Keras. Se aplicó

únicamente al conjunto de entrenamiento, en el momento de creación de cada batch (en tiempo de ejecución). Cuando cada imagen se carga, aleatoriamente puede ser modificada. Los procedimientos de aumento de datos utilizados fueron: corte (*shear*), acercamientos, rotaciones, desplazamientos, brillo y volteo.

5.6. Variables de respuesta

Exactitud: magnitud adimensional que representa el porcentaje de imágenes clasificadas correctamente en el conjunto de datos, definida como:

$$Exactitud = \frac{VP + VN}{VP + FP + VN + FN}$$

Dónde:

VP = cantidad de verdaderos positivos

VN = cantidad de verdaderos negativos

FP = cantidad de falsos positivos

FN = cantidad de falsos negativos

Tiempo de entrenamiento: cantidad de horas, minutos y segundos que la red neuronal requiere para optimizar los parámetros sobre el conjunto de entrenamiento en determinado número de *epochs*.

5.7. Análisis de la información

El criterio utilizado para determinar la mejor arquitectura de red neuronal fue la comparación de las exactitudes en la clasificación sobre el conjunto de pruebas. Para determinar si la diferencia entre las exactitudes de las redes neuronales no se debe a factores aleatorios del muestreo se aplicó la prueba de McNemar.

No se utilizaron pruebas como t de Student o análisis de varianza debido a que estas suponen que las observaciones son independientes y se calculan sobre varias repeticiones. Condiciones que no se cumple en este diseño, puesto que las observaciones de los dos tratamientos se toman sobre el mismo conjunto de datos y las redes neuronales se entrenan una sola vez. Este es un problema común en los experimentos con algoritmos de aprendizaje automático. Problema abordado por Dietterich (1998) que encontró que la mejor prueba para determinar si un algoritmo de aprendizaje automático tiene mejor desempeño que otro en una tarea en particular es la prueba de McNemar, debido a que minimiza la posibilidad de errores de tipo I en las pruebas de hipótesis.

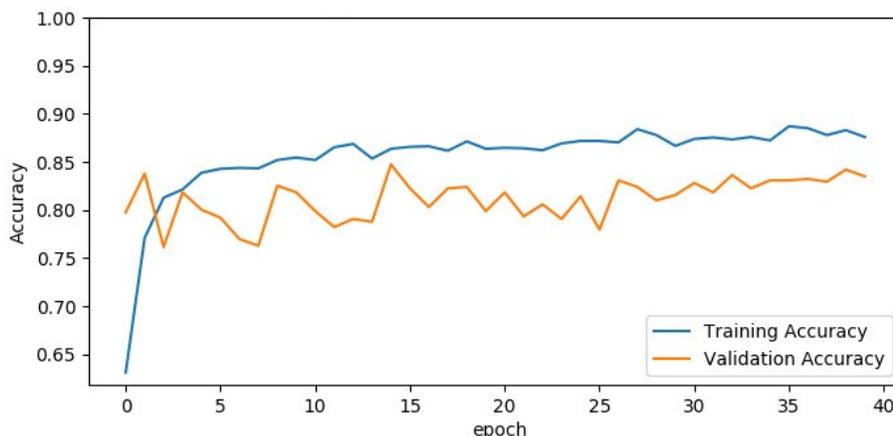
6. Resultados y Discusión

Las exactitudes en la clasificación sobre el conjunto de datos organizado en 3 categorías puede observarse en la tabla 5. Estas exactitudes son al *epoch* 40 de entrenamiento. En todos los conjuntos las exactitudes en la clasificación son bajas, lo que indica un problema de *high bias*. La figura 5 muestra la tendencia en las exactitudes sobre los conjuntos de entrenamiento y validación para ResNet50 (que alcanzó el valor más alto sobre el conjunto de pruebas), de donde se concluye que más *epochs* no mejorarían la exactitud, pues esta ha empezado a converger.

Tabla 5. Resultados de las redes neuronales entrenadas para clasificar la severidad del Virus del Mosaico Dorado en frijol en tres niveles, en el oriente de Guatemala, 2019.

Exactitud al epoch 40					
Arquitectura	Nivel	Conjunto de entrenamiento (%)	Conjunto de validación (%)	Conjunto de pruebas (%)	Tiempo de entrenamiento medio por <i>epoch</i> (s)
InceptionV3	Todas las capas	87.28	75.17	78.56	167.18
	Capa intermedia	85.85	80.86	80.91	154.08
ResNet50	Todas las capas	88.50	75.03	75.93	136.28
	Capa intermedia	87.58	83.50	83.40	123.05

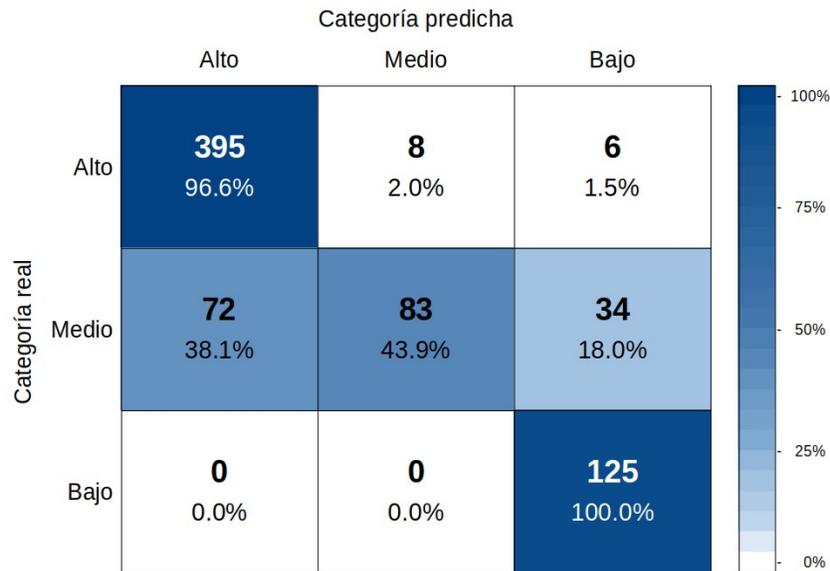
Figura 5. Exactitudes de ResNet50 usando capas intermedias sobre los conjuntos de entrenamiento y validación, al clasificar la severidad del Virus del Mosaico Dorado en frijol en tres niveles, en el oriente de Guatemala, 2019.



La matriz de confusión de la figura 6, muestra que el principal problema que causa la baja exactitud es la delimitación de la clasificación del medio. Esto concuerda con los resultados de Esgario, Krohling y Ventura (2020) que analizaron el problema de clasificación de severidad de enfermedades

en café y quienes también encontraron una baja exactitud al clasificar severidad usando redes neuronales convolucionales.

Figura 6. Matriz de confusión de ResNet50 usando capas intermedias sobre el conjunto de validación, al clasificar la severidad del Virus del Mosaico Dorado en frijol en tres niveles, en el oriente de Guatemala, 2019.



Por otra parte, cuando los datos se organizan únicamente en dos categorías (lo que reduce el número de fronteras entre categorías) la exactitud mejora, tal como se observa en la tabla 6. En donde para ambas topologías la exactitud más alta sobre el conjunto de pruebas se alcanza al extraer las características desde capas intermedias.

Tabla 6. Resultados de las redes neuronales entrenadas para clasificar la severidad del Virus del Mosaico Dorado en frijol en dos niveles, en el oriente de Guatemala, 2019.

Exactitud al ephoc 40					
Arquitectura	Nivel	Conjunto de entrenamiento (%)	Conjunto de validación (%)	Conjunto de pruebas (%)	Tiempo de entrenamiento medio por ephoc (s)
InceptionV3	Todas las capas	96.28	90.57	92.12	168.05
	Capa intermedia	95.98	94.59	94.47	153.75
ResNet50	Todas las capas	96.28	90.57	96.27	136.78
	Capa intermedia	96.03	95.98	96.68	124.00

En las figuras 7 y 8 se muestran las exactitudes sobre los conjuntos de entrenamiento y validación para las dos topologías extrayendo características desde capas intermedias para el conjunto de datos con dos categorías. En ambas se observa que han alcanzado a converger. En el caso de InceptionV3, existe un leve sobre-ajuste expresado en la distancia entre las curvas de entrenamiento y validación.

Figura 7. Exactitudes de InceptionV3 usando capas intermedias sobre los conjuntos de entrenamiento y validación, al clasificar la severidad del Virus del Mosaico Dorado en frijol en dos niveles, en el oriente de Guatemala, 2019.

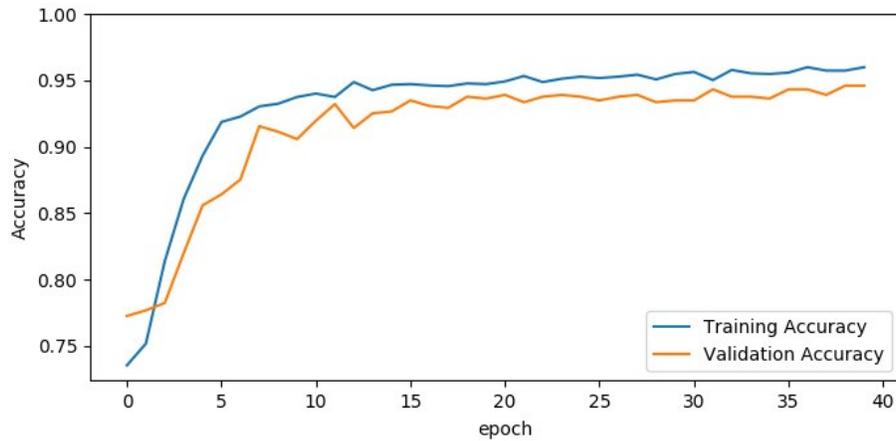
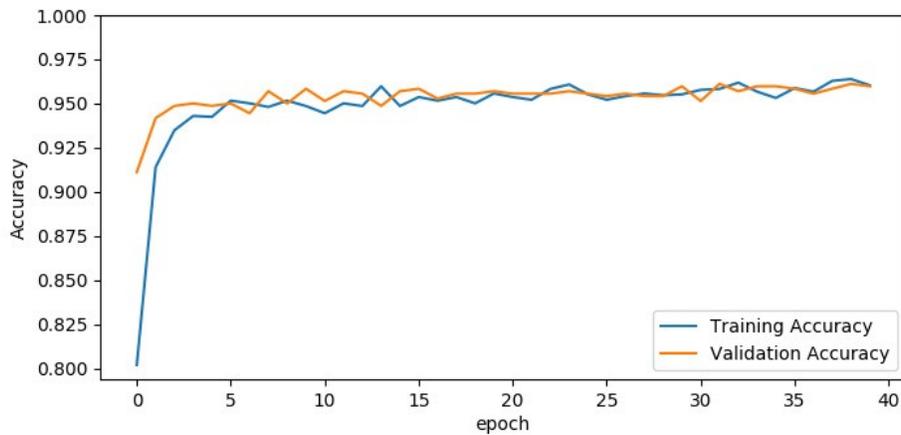


Figura 8. Exactitudes de ResNet50 usando capas intermedias sobre los conjuntos de entrenamiento y validación, al clasificar la severidad del Virus del Mosaico Dorado en frijol en dos niveles, en el oriente de Guatemala, 2019.



Para comparar las exactitudes más altas entre las dos topologías y poder determinar si la diferencia es significativa se necesita calcular la tabla de contingencia que se muestra en la tabla 7. Con estos datos se realiza la prueba de McNemar.

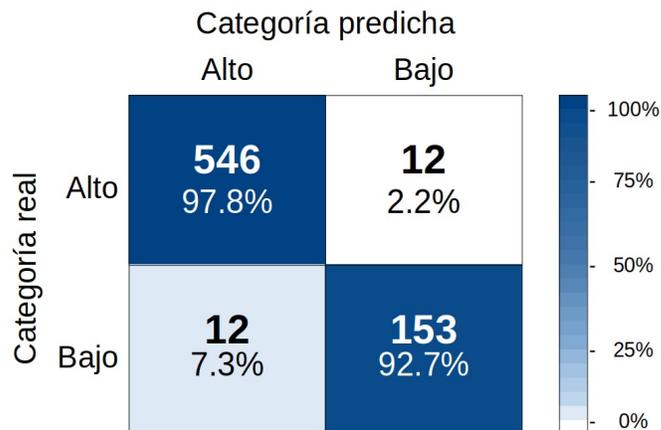
Tabla 7. Tabla de contingencia de InceptionV3 y ResNet50 usando capas intermedias sobre el conjunto de pruebas, al clasificar la severidad del Virus del Mosaico Dorado en frijol en dos niveles, en el oriente de Guatemala, 2019.

		ResNet50		
InceptionV3		Correcto	Incorrecto	Totales
Correcto		675	8	683
Incorrecto		24	16	40
Totales		699	24	723

La prueba de McNemar se realizó utilizando la implementación existente en el lenguaje de programación estadístico R, usando corrección de continuidad. Lo que produjo como resultado a $\chi^2 = 7.0312$, $df = 1$ y $p\text{-value} = 0.00801$. Debido a que $p\text{-value}$ es menor que 0.05, se rechaza la hipótesis nula. Existe una diferencia significativa en la proporción de errores en la clasificación entre las dos topologías a un nivel de significación del 0.05. La red neuronal que produce la exactitud más alta sobre el conjunto de datos recolectados es ResNet50 con extracción de características de capas intermedias.

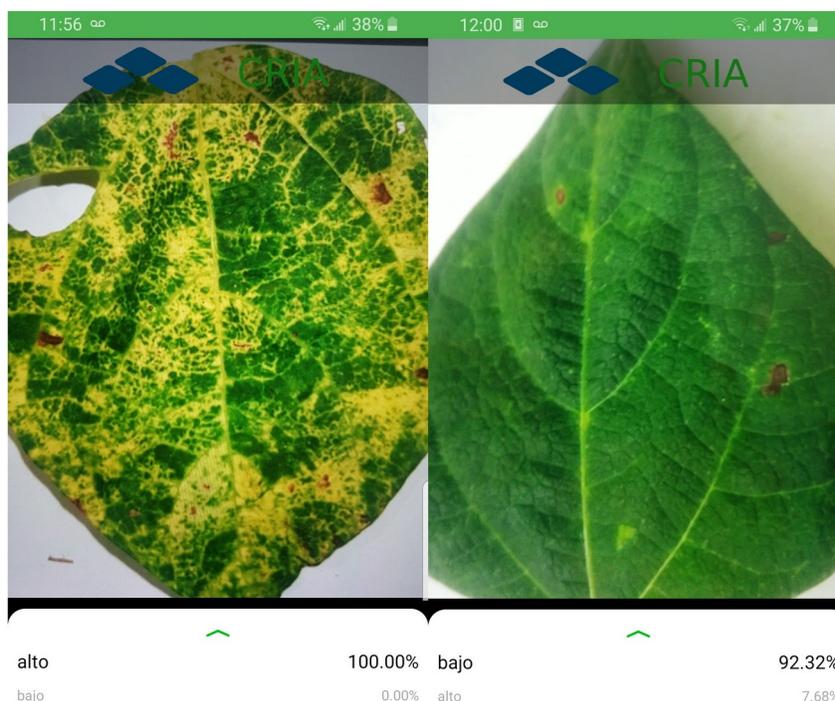
En la figura 9 se muestra la matriz de confusión de la mejor red neuronal. De donde se observa que la exactitud para clasificar muestras con severidad alta es del 97.8% mientras que para clasificar muestras de severidad baja es 92.7%.

Figura 9. Matriz de confusión de ResNet50 usando capas intermedias sobre el conjunto de validación, al clasificar la severidad del Virus del Mosaico Dorado en frijol en dos niveles, en el oriente de Guatemala, 2019.



A partir de la red neuronal seleccionada, se desarrolló una aplicación prototipo para el sistema operativo Android, haciendo uso de TensorFlow lite. La aplicación está basada en el ejemplo de la documentación del mencionado *framework*. La red neuronal originalmente en formato HDF5 (‘.h5’) fue convertida a un archivo FlatBuffer (‘.tflite’) de coma flotante, así como al formato cuantizado. La figura 10 muestra la interfaz gráfica de la aplicación.

Figura 10. Interfaz gráfica de la aplicación para Android con TensorFlow lite para clasificar la severidad del Virus del Mosaico Dorado en frijol en dos niveles, en el oriente de Guatemala, 2019.



7. Conclusiones

1. Se recolectaron 3409 imágenes de muestras de hojas de frijol, las cuales fueron recortadas, reducidas en tamaño y clasificadas manualmente según la escala propuesta por el CIAT para enfermedades virales. Lo que permitió crear un conjunto de datos de acceso público para futuras investigaciones con sistemas de aprendizaje supervisado que involucren al Virus del Mosaico Dorado.
2. Se encontró que, para el conjunto de datos recolectados, ResNet50 alcanzó una mejor exactitud (96.68%) en comparación a GoogLeNet (94.47%), cuando los datos se encuentran organizados en dos categorías (alta y baja) y las características se extraen desde capas intermedias. Una prueba de McNemar con corrección de continuidad reveló que la diferencia es estadísticamente significativa ($\chi^2=7$, $p\text{-value}=0.008$) a un nivel de significación del 5%.
3. Con el desarrollo de la aplicación para dispositivos móviles que implementa redes neuronales, se provee una herramienta para clasificar fácilmente la severidad (alta o baja) del Virus del Mosaico Dorado a nivel foliar en frijol. La clasificación se realiza en el orden de centenas de mili-segundos cuando se ejecuta en un solo hilo en un procesador (CPU), mientras que en el caso de ejecutarse sobre un procesador de gráficos (GPU) es capaz de realizar las inferencias en el orden de decenas de mili-segundos.

8. Recomendaciones

1. Debido a que las redes neuronales alcanzaron una baja exactitud en la clasificación de la severidad al utilizar tres categorías. Es útil realizar nuevas investigaciones con el objetivo de mejorar la exactitud. Una posibilidad es verificar si la utilización de sistemas de aprendizaje automático no supervisado como la agrupación (*clustering*) son beneficiosas para las redes neuronales al eliminar posibles inconsistencias en la clasificación realizada por humanos. Adicionalmente, pueden probarse otras tecnologías como la segmentación, sistemas de regresión y árboles de decisiones.
2. Las redes neuronales han probado ser eficientes en la identificación de enfermedades a nivel foliar en plantas. Sin embargo, existen muy pocas investigaciones relacionadas con el frijol. Esto significa que existe una importante área de investigación para nuevos sistemas de identificación de estrés biótico y abiótico detectable a nivel foliar por medio de imágenes en cultivos de frijol.
3. Es importante que los usuarios de las aplicaciones para dispositivos móviles consideren la calidad de la imagen para que la clasificación sea correcta. Uno de los principales aspectos a considerar es enfocar correctamente la imagen evitando movimientos. Además, se deben evitar las condiciones luminosas extremas, mucha luz o muy poca, pues esto puede evitar apreciar correctamente los colores que es una parte fundamental en la clasificación. Finalmente, es preferible colocar la muestra sobre un fondo blanco para minimizar la posibilidad de equivocaciones en la clasificación.
4. La visión por computadora, utilizando modelos de aprendizaje automático, es un área relativamente nueva que actualmente se encuentra en auge. Por lo que sería conveniente que los institutos de investigación, universidades e investigadores puedan continuar explorando aplicaciones de la inteligencia artificial en distintas disciplinas (industria, medicina, educación, economía y otros) que lleven al desarrollo de nuevos sistemas de clasificación automática o de apoyo a expertos.

9. Referencias

- Alpaydin, E. (2016). *Machine Learning: the new AI*. Cambridge: MIT Press.
- Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural computation*, 10(7), 1895-1923. Recuperado de <https://sci2s.ugr.es/keel/pdf/algorithm/articulo/dietterich1998.pdf>
- 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. doi:10.1016/j.compag.2019.105162
- Flores-López, R., & Fernández-Fernández, J. M. (2008). *Las redes neuronales artificiales: Fundamentos teóricos y aplicaciones prácticas*. Netbiblo: La Coruña.
- 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. doi: 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. Recuperado de <https://arxiv.org/pdf/1512.03385.pdf>
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. En *Advances in Neural Information Processing Systems*, 25(2). Recuperado de <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
- 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*. Managua, Nicaragua: IICA. Recuperado de <http://repiica.iica.int/docs/B0891E/B0891E.pdf>
- 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*. Guatemala: DIPLAN/MAGA.

- Stathakis, D. (2009). How many hidden layers and nodes?. *International Journal of Remote Sensing*, 30(8), 2133-2147. Recuperado de http://dstath.users.uth.gr/papers/IJRS2009_Stathakis.pdf
- Subramanian, J., & Simon, R. (2013). Overfitting in prediction models—is it a problem only in high dimensions?. *Contemporary clinical trials*, 36(2), 636-641. Recuperado de <https://brb.nci.nih.gov/techreport/Subramanian-Overfitting.pdf>
- Sucar, L. E., & Gómez, G. (2011). *Visión computacional*. Puebla, México: Instituto Nacional de Astrofísica, Óptica y Electrónica. Recuperado de <https://ccc.inaoep.mx/~esucar/Libros/vision-sucar-gomez.pdf>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. En *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9. Recuperado de <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/43022.pdf>
- van-Schoonhoven, A., & Pastor-Corrales, M. A. (1987). *Sistema estándar para la evaluación de germoplasma de frijol*. Colombia: Centro Integral de Agricultura Tropical.

Anexos

Anexo 1: repositorios de código fuente, instaladores y conjunto de datos

Repositorio	Dirección web
Aplicación Android para clasificar severidad por medio de ReNet50	https://github.com/carlosdavidardon/nn-golden-mosaic-app https://play.google.com/store/apps/details?id=org.iica.lite.examples.classification (instalador en Google Play)
Aplicación Android para calcular la proporción de coloración verde y amarilla	https://github.com/carlosdavidardon/hue-analysis-app https://play.google.com/store/apps/details?id=com.iica.hueanalysis (instalador en Google Play)
Conjunto de datos de imágenes de hojas clasificadas según severidad del Virus del Mosaico Dorado	https://github.com/carlosdavidardon/golden-mosaic-virus-dataset
Programas para realizar entrenamiento de redes neuronales y operaciones auxiliares	https://github.com/carlosdavidardon/golden-mosaic-virus-classifier

Anexo 2: aplicación Android para clasificación de severidad con redes neuronales

Funcionamiento

Iniciar la aplicación y enfocar la hoja de muestra durante unos segundos hasta que la red neuronal estabilice su predicción. Es preferible que la hoja se encuentre sobre un fondo blanco y que el entorno se encuentre bien iluminado.

Instalación

Descargar la aplicación desde la tienda oficial de Android en Google Play:

<https://play.google.com/store/apps/details?id=org.iica.lite.examples.classification>

Presionar el botón “Instalar”.

Requerimientos

Para ejecutar la aplicación se requiere como mínimo tener Android 5.0 (LOLLIPOP), equivalente al API nivel 21.

Atribución

Esta aplicación se basa en el ejemplo de la documentación oficial de TensorFlow lite, el cual se encuentra publicado bajo la licencia Apache 2.0.

Anexo 3: aplicación Android para determinar la proporción de coloración verde y amarilla

Funcionamiento

Iniciar la aplicación se presiona el botón 'Seleccionar imagen' que ofrece dos opciones:

- Tomar fotografía: utiliza la cámara del dispositivo para capturar una imagen.
- Seleccionar de la biblioteca: carga una imagen que ya existe en el almacenamiento del dispositivo.

Para obtener mejores resultados tomar la fotografía de la hoja sobre un fondo blanco y evitar las condiciones de luminosas extremas (mucho o muy poca iluminación).



Instalación

Descargar la aplicación desde la tienda oficial de Android en Google Play:

<https://play.google.com/store/apps/details?id=com.iica.hueanalysis>

Presionar el botón “Instalar”.

Requerimientos

Para ejecutar la aplicación se requiere como mínimo tener Android 4.0 (Ice Cream Sandwich), equivalente al API nivel 14.

Atribución

Esta aplicación se basa en el ejemplo publicado por 'TheAppGuruz'.

Anexo 4: taller de divulgación de resultados por medio de videoconferencia

The screenshot shows a Google Meet interface during a video conference. The main content is a presentation slide with the following text:

CRIA
Programa Consorcios Regionales de Investigación Agropecuaria

La investigación en consorcio implica salirse del marco o la caja tradicional de generar, validar y transferir conocimientos y tecnologías, porque La naturaleza y la sociedad demandan investigación aplicada para solucionar los problemas del presente e investigación básica para comprenderlos y replantear la visión hacia el futuro. H. Villafuerte

Gracias por participar

3 regiones • 7 departamentos • 13 productos • 18 cadenas

Logos for GOBIERNO DE COLOMBIA, USDA, and IICA are visible in the top right of the slide.

Participant thumbnails on the right include:

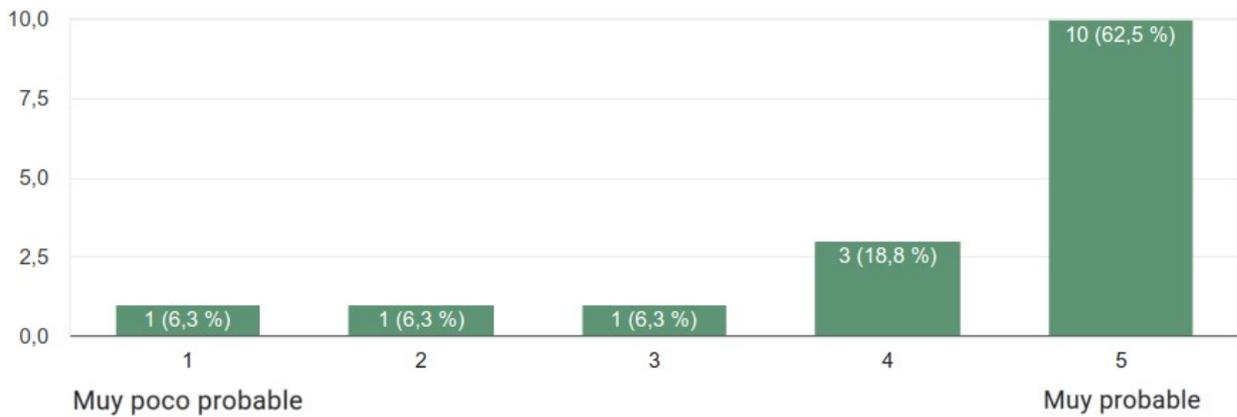
- Hugo Ronaldo Villafuerte V. (Speaker icon)
- Suleyna Sandoval
- nery galdamez

At the bottom of the slide, a sharing notification reads: "meet.google.com está compartiendo tu pantalla. Dejar de compartir Ocultar".

Anexo 5: resultados de la encuesta de retroalimentación sobre las aplicaciones por parte de los participantes del taller

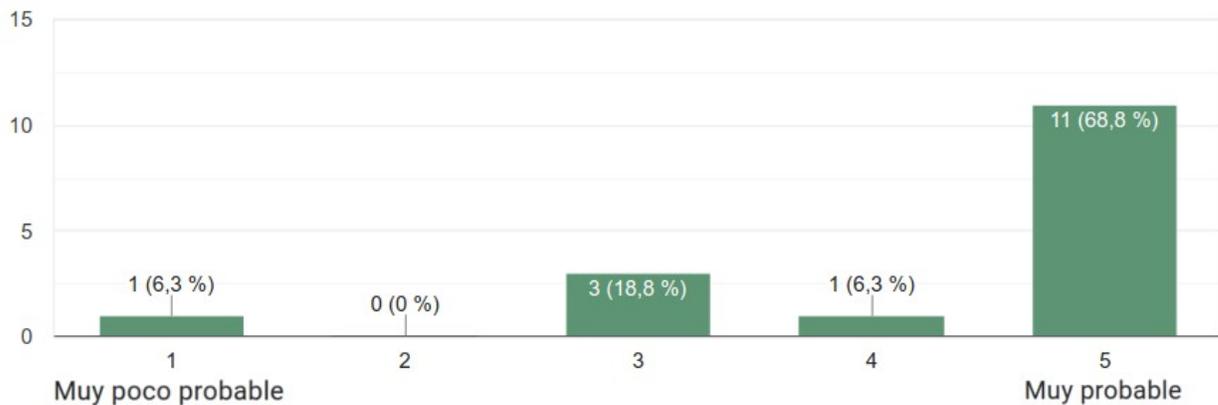
Si tuviera que trabajar en un proyecto relacionado con frijol, ¿Qué tan probable es que utilice alguna de las aplicaciones como apoyo?

16 respuestas



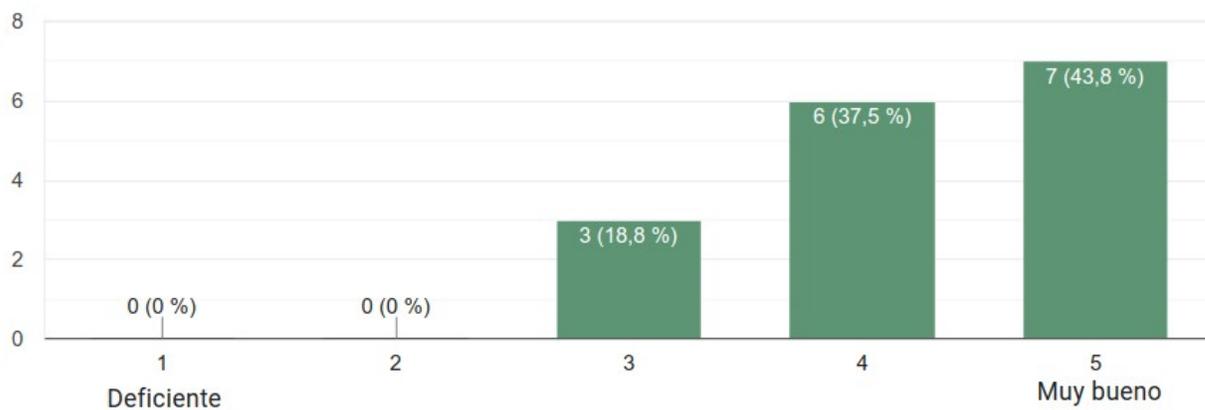
¿Qué tan probable es que recomiende alguna de las aplicaciones a alguien que trabaje con cultivos de frijol?

16 respuestas



¿Cómo califica el aspecto visual (estético) de las aplicaciones?

16 respuestas



¿Cómo califica la facilidad de uso de las aplicaciones?

16 respuestas

