

**UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

**VISIÓN POR COMPUTADORA PARA EL MANEJO DE
PLAGAS Y ENFERMEDADES EN CULTIVOS DE PAPA**

Por:

Fabián Nicolás Fuentes Plaza

Profesor Guía:

Sebastián Niklitschek Soto

Concepción, junio 2021



Tesis presentada a la

**DIRECCIÓN DE POSTGRADO
UNIVERSIDAD DE CONCEPCIÓN**



Para optar al grado de

MAGÍSTER EN INGENIERÍA INDUSTRIAL

RESUMEN

VISIÓN POR COMPUTADORA PARA EL MANEJO DE PLAGAS Y ENFERMEDADES EN CULTIVOS DE PAPA

Fabián Nicolás Fuentes Plaza
Concepción, junio 2021

PROFESOR GUÍA: Sebastián Niklitschek Soto
PROGRAMA: Magíster en Ingeniería Industrial

La papa es el tercer cultivo más importante a nivel mundial con 300 millones de toneladas métricas producidas anualmente. En Chile corresponde al gremio más numeroso del país con casi 60 mil agricultores de este cultivo. La amenaza de una posible plaga o enfermedad que cause pérdidas o bajas en los rendimientos de los cultivos de papa, es una realidad que se vive día a día, siendo los principales afectados los agricultores a pequeña escala. Para poder combatir y tomar decisiones de manera correcta y oportuna, resulta esencial la disponibilidad de herramientas que permitan la detección temprana de enfermedades que afecten a los cultivos. En este estudio se formula el uso de técnicas de aprendizaje profundo como una alternativa viable para enfrentar esta problemática, sustentándose en la alta efectividad que poseen las redes neuronales convolucionales al momento de resolver la tarea de clasificar imágenes. Se utilizaron imágenes de hojas de papa obtenidas del conjunto de datos PlantVillage e imágenes de tubérculos obtenidas del conjunto de datos Fruit360 con adición de imágenes de recopilación propia, como entrada para modelos entrenados a partir de las arquitecturas VGG16 y DenseNet201. Como resultado se obtuvo un modelo capaz de diferenciar entre las clases sana, tizón temprano y tizón tardío con una precisión del 98.44% para el caso de las hojas, y un modelo capaz de diferenciar entre las clases sana, tizón tardío y sarna común con una precisión del 96.88% para el caso de los tubérculos, además, se generaron visualizaciones de los mapas de características, de las principales capas convolucionales de la red VGG16, permitiendo una interpretación y análisis visual.

Palabras Clave: Visión computacional, aprendizaje profundo, redes neuronales convolucionales, cultivo de papa, tizón tardío.

ABSTRACT

COMPUTER VISION FOR PEST AND DISEASE MANAGEMENT IN POTATO CROPS

Fabián Nicolás Fuentes Plaza
Concepción, June 2021

THESIS SUPERVISOR: Sebastián Niklitschek Soto
PROGRAM: Master in Industrial Engineering

Potatoes are the third most important crop worldwide with 300 million metric tons produced annually. In Chile it corresponds to the largest union in the country with almost 60 thousand farmers of this crop. The threat of a possible pest or disease that causes losses in potato crop yields is a reality that is experienced every day, with small-scale farmers being the main affected. To be able to combat and make decisions in a correct and timely manner, the availability of a tool that allows the early detection of diseases that affect crops is essential. In this study, the use of deep learning techniques is formulated as a viable alternative to face this problem, based on the high effectiveness of convolutional neural networks solving the task of classifying images. Potato leaf images obtained from the PlantVillage dataset and tuber images obtained from the Fruit360 dataset with the addition of self-collected images, were used as input for models trained from the VGG16 and DenseNet201 architectures. As a result, a model capable of differentiating between the healthy, early blight and late blight classes was obtained with a precision of 98.44% for the case of leaves, and a model capable of differentiating between the healthy, late blight and common scab classes with a 96.88% precision for the case of tubers, in addition, visualizations of the characteristic maps of the main convolutional layers of the VGG16 network were generated, allowing visual interpretation and analysis.

Key words: Computer vision, deep learning, convolutional neural network, potato crop, late blight.

Tabla de contenidos

1	Introducción.....	1
2	Antecedentes generales.....	5
2.1	Industria y Agricultura 4.0	5
2.2	Agricultura de la papa.....	7
2.3	Enfermedades que afectan los cultivos de papa	7
2.3.1	Tizón tardío	8
2.3.2	Tizón temprano	9
2.3.3	Sarna común.....	10
2.4	Visión computacional.....	10
3	Fundamentos teóricos.....	12
3.1	Redes neuronales convolucionales	12
3.1.1	Capa convolucional.....	13
3.1.2	Capa de pooling	15
3.1.3	Capa de clasificación (capa completamente conectada)	16
3.2	Red VGG16	17
3.3	Red DenseNet201	18
3.4	Transferencia de aprendizaje	19
3.5	Mapas de características.....	21
4	Revisión de la literatura.....	23
4.1	ImageNet Large Scale Visual Recognition Challenge	23
4.2	CNN y transferencia de aprendizaje para clasificar enfermedades en cultivos.....	24

5	Metodología.....	26
5.1	Configuración del sistema.....	26
5.2	PlantVillage.....	26
5.3	Fruit360.....	29
5.4	Transferencia de aprendizaje	32
5.5	Mapas de características	35
6	Resultados	37
6.1	Entrenamiento sobre conjunto de datos PlantVillage original	37
6.2	Entrenamiento sobre el conjunto PlantVillage segmentado.....	38
6.3	Entrenamiento de modelos clasificadores de tubérculos	39
6.4	Evaluación sobre conjuntos de prueba.....	40
6.4.1	Evaluación sobre conjunto PlantVillage original.....	40
6.4.2	Evaluación sobre el conjunto PlantVillage segmentado.....	41
6.4.3	Evaluación sobre el conjunto Fruit360 más web scraping.....	43
6.5	Mapas de características red VGG16.....	45
6.5.1	Mapa de características modelo clasificador de enfermedades en hojas de papa.....	45
6.5.2	Mapa de características modelo clasificador de enfermedades en tubérculos.....	47
7	Discusiones	50
8	Conclusiones y extensiones.....	54
9	Bibliografía	57
10	Anexos	61

Lista de figuras

Figura 1. Esquema de las primeras cuatro revoluciones industriales	6
Figura 2. Ejemplos de ataque de tizón tardío sobre hojas y tubérculos	9
Figura 3. Ejemplos de ataque de tizón temprano sobre hojas de papa	9
Figura 4. Ejemplos de ataque de sarna común en tubérculos	10
Figura 5. Tipos de problemas de percepción visual	11
Figura 6. Esquema del funcionamiento de los filtros en la capa convolucional .	14
Figura 7. Esquema del funcionamiento de la capa pooling.....	15
Figura 8. Esquema básico de la arquitectura de una CNN de una capa convolucional	16
Figura 9. Esquema de la arquitectura de la red VGG16	18
Figura 10. Esquema de la arquitectura de la red DenseNet201	19
Figura 11. Arquitectura básica de una CNN con adición de una nueva capa de clasificación adaptada para nuevas categorías	20
Figura 12. Muestra de imágenes de los 3 tipos de categorías y 2 tipos de formatos utilizados en las configuraciones experimentales de este estudio	27
Figura 13. Ejemplo de las transformaciones realizadas a las imágenes de la categoría sana durante los procesos de aumentación de datos	28
Figura 14. Muestra de imágenes de la categoría White potato del conjunto Fruit360.....	29
Figura 15. Ejemplo procesos segmentación de imagenes.....	30
Figura 16. Ejemplo de las transformaciones realizadas a las imágenes durante los procesos de aumentación de datos.....	31

Figura 17. Muestra de imágenes de las 3 categorías utilizadas en las configuraciones experimentales de este estudio	32
Figura 18. Esquema red VGG16 adaptada para clasificar sobre nuevas categorías	34
Figura 19. Esquema de la red DenseNet201 adaptada para clasificar sobre nuevas categorías.....	34
Figura 20. Precisión de validación de las redes VGG16 y DenseNet201 sobre el conjunto PlantVillage original durante las 20 épocas de entrenamiento	37
Figura 21. Precisión de validación de las redes VGG16 y DenseNet201 sobre el conjunto PlantVillage segmentado durante las 20 épocas de entrenamiento	38
Figura 22. Precisión de validación de las redes VGG16 y DenseNet201 sobre el conjunto Fruit360 más web scraping durante las 20 épocas de entrenamiento	39
Figura 23. Matriz de confusión creada a partir de la evaluación del modelo de la red VGG16 sobre el conjunto de prueba PlantVillage original	40
Figura 24. Matriz de confusión creada a partir de la evaluación del modelo de la red DenseNet201 sobre el conjunto de prueba PlantVillage original	41
Figura 25. Matriz de confusión creada a partir de la evaluación del modelo de la red VGG16 sobre el conjunto de prueba PlantVillage segmentado	42
Figura 26. Matriz de confusión creada a partir de la evaluación del modelo de la red DenseNet201 sobre el conjunto de prueba PlantVillage segmentado	43
Figura 27. Matriz de confusión creada a partir de la evaluación del modelo de la red VGG16 sobre el conjunto de prueba Fruit360 más web scraping.....	44
Figura 28. Matriz de confusión creada a partir de la evaluación del modelo de la red DenseNet201 sobre el conjunto de prueba Fruit360 más web scraping.....	44
Figura 29. Ejemplo de los mapas de características de la red VGG16 sobre imágenes de las 3 categorías del conjunto PlantVillage original.....	45

Figura 30. Superposición de los mapas de características de la red VGG16 con la imagen de entrada original de cada una de las 3 categorías del conjunto PlantVillage original	46
Figura 31. Ejemplo de los mapas de características de la red VGG16 sobre imágenes de las 3 categorías del conjunto Fruit360 más web scraping	47
Figura 32. Superposición de los mapas de características de la red VGG16 con la imagen de entrada original de cada una de las 3 categorías del conjunto Fruit360 más web scraping	48
Figura 33. Mapas de características categoría hoja sana bloque convolucional 1	61
Figura 34. Mapas de características categoría hoja sana bloque convolucional 2	61
Figura 35. Mapas de características categoría hoja sana bloque convolucional 3	62
Figura 36. Mapas de características categoría hoja sana bloque convolucional 4	62
Figura 37. Mapas de características categoría hoja sana bloque convolucional 5	63
Figura 38. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 1	63
Figura 39. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 2	64
Figura 40. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 3	64
Figura 41. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 4	65

Figura 42. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 5	65
Figura 43. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 1	66
Figura 44. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 2	66
Figura 45. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 3	67
Figura 46. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 4	67
Figura 47. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 5	68
Figura 48. Mapas de características imagen categoría tubérculo sano bloque convolucional 1	68
Figura 49. Mapas de características imagen categoría tubérculo sano bloque convolucional 2	69
Figura 50. Mapas de características imagen categoría tubérculo sano bloque convolucional 3	69
Figura 51. Mapas de características imagen categoría tubérculo sano bloque convolucional 4	70
Figura 52. Mapas de características imagen categoría tubérculo sano bloque convolucional 5	70
Figura 53. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 1	71
Figura 54. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 2	71



Figura 55. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 3	72
Figura 56. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 4	72
Figura 57. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 5	73
Figura 58. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 1	73
Figura 59. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 2	74
Figura 60. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 3	74
Figura 61. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 4	75
Figura 62. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 5	75





Capítulo 1

1 Introducción

Así como la máquina vapor marcó un punto de inflexión en la historia a finales de siglo XVIII, dando inicio a la primera revolución industrial, en las primeras décadas del siglo XXI, la convergencia de tecnologías digitales como el Internet de las Cosas, *Big Data*, Robótica, *Machine Learning* e Inteligencia Artificial, dan pie a la llegada de la cuarta revolución industrial, también denominada industria 4.0, impulsando una industria con procesos más óptimos, eficientes y de mayor productividad. Este concepto fue adoptado el 2011 en Alemania como “*Industrie 4.0*”, sin embargo, desde el año 2017, se considera un término global y rige muchas de las políticas de gobierno de países como China, Japón y Estados Unidos [1].

La agricultura es uno de los sectores que adoptó sistemas de recopilación y análisis de datos, implementando sensores, drones y software. La nueva agricultura 4.0, promueve una industria con mejor calidad de cultivos y menos consecuencias negativas propias de la industria en el medio ambiente [2].

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 [4].

El cultivo de la papa, es uno de los que se ve principalmente afectado por la presencia de agentes patógenos. Su reproducción asexual utilizando el tubérculo como semilla facilita la propagación de enfermedades, además, dependiendo de cuál sea el agente patógeno y si se trata de una bacteria, virus, hongo, insecto o nemátodo, su manifestación podría darse en el tubérculo, raíces u hojas, dificultando la detección temprana de alguna enfermedad o peste que afecte el cultivo. La propagación de enfermedades en un papal podría traer consigo desde baja de rendimiento en los cultivos hasta una pérdida total, o inclusive llevar a una región completa a cuarentena [3]. El caso más icónico es el ocurrido en Irlanda durante la gran hambruna o hambruna de la patata en la década de 1840, donde una plaga de *Phytophthora infestans* conocida como tizón tardío, acabo con un 75% de los cultivos, causando la muerte de un millón de personas y la emigración de otro millón más, disminuyendo la población de la isla en un 25% [19].

Históricamente la identificación de enfermedades en cultivos ha estado en manos de instituciones u organizaciones agrícolas, generando diagnósticos en base a conocimiento de expertos, que, a pesar de ser bastante efectivo en la tarea de identificar enfermedades, no resulta ser un método ágil, ni de fácil acceso para todos los agricultores [4], [5].

Los grandes avances en herramientas de *Machine Learning* capaces de clasificar imágenes entre miles de categorías [7], [12], [13] y el aumento exponencial al acceso de internet y *smartphones* [4], [6], permiten que los enfoques más recientes, centren sus esfuerzos en implementar modelos de reconocimiento en dispositivos móviles, capaces de realizar detección temprana de agentes patógenos en los cultivos.

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, con el objetivo principal de generar clasificadores de alta precisión entre distintos tipos de enfermedades en cultivos de papa, permitiendo democratizar el uso de estas nuevas tecnologías para el área de la agricultura, específicamente para los agricultores pequeños.

La hipótesis principal es que modelos entrenados a partir de nuevos conjuntos de datos con imágenes de cultivos sanos y enfermos de papa, alcanzaran precisiones a la par a los reportados por modelos de vanguardia del estado del arte actual, sustentándose en los numerosos casos de éxito reportados por la literatura al utilizar técnicas de transferencia de aprendizaje de redes de alta precisión, sobre nuevas categorías [12], [13], [14], [15].

Para construir los modelos clasificadores de cultivos de papa se utilizó la técnica de transferencia de aprendizaje, entrenando las arquitecturas de las redes VGG16 y DenseNet201, sobre imágenes de hojas de papa y tubérculos obtenidas de los conjuntos PlantVillage [15] y Fruit360 [16] con adición de imágenes de recolección propia mediante técnicas de *web scraping*.

Para evitar que el aprendizaje de las redes se convierta en una caja negra, se generaron los mapas de características de las principales capas convolucionales de la red VGG16, permitiendo visualizar y depurar los resultados obtenidos por los modelos implementados con esta arquitectura [17], [18].

Los antecedentes generales de la problemática abordada y la descripción del campo de la visión computacional son presentados en el Capítulo 2. La revisión del estado del arte actual de las CNN y su aplicación en distintos problemas del sector son presentados en el Capítulo 3 y 4. La metodología utilizada,

considerando la configuración del sistema, arquitecturas de redes y conjuntos de datos utilizados, son expuestos en el Capítulo 5. Los principales resultados son expuestos en el Capítulo 6 y las discusiones respectivas en el Capítulo 7. Las conclusiones finales del estudio y sus posibles extensiones, son presentadas en el Capítulo 8.



Capítulo 2

2 Antecedentes generales

2.1 Industria y Agricultura 4.0

La industria 4.0 es un término originalmente instaurado en Alemania el año 2011, que surge como una iniciativa de gobierno integrando el concepto “*industrie 4.0*” como una parte fundamental del plan estratégico “*High Technology Strategy for Germany 2020*”, que buscaba posicionar al país en la elite de la innovación tecnológica a nivel mundial [1]. A pesar de tratarse de un concepto relativamente nuevo, está basado en áreas de investigación como la automatización, inteligencia artificial, tecnologías de producción, *Big data* y computación en la nube. Desde el año 2017 ya se considera un pilar fundamental en las políticas de gobierno de países como China, Japón y Estados Unidos [20].

El proceso de migración al uso de nuevas tecnologías ocurre de manera de transversal para todos los sectores de producción. La escasez de recursos, el endurecimiento de restricciones medio ambientales y el constante cambio en las necesidades de los consumidores, afecta directa o indirectamente tanto el sector primario, secundario y terciario, esto trae consigo una necesidad generalizada de adaptarse al uso de estas nuevas tecnologías a fin de mantenerse competitivos y poder satisfacer las necesidades del mundo actual. La Figura 1 muestra un esquema de la evolución de la industria y detalla los principales aspectos de cada una de las revoluciones sucesivas que ha experimentado.

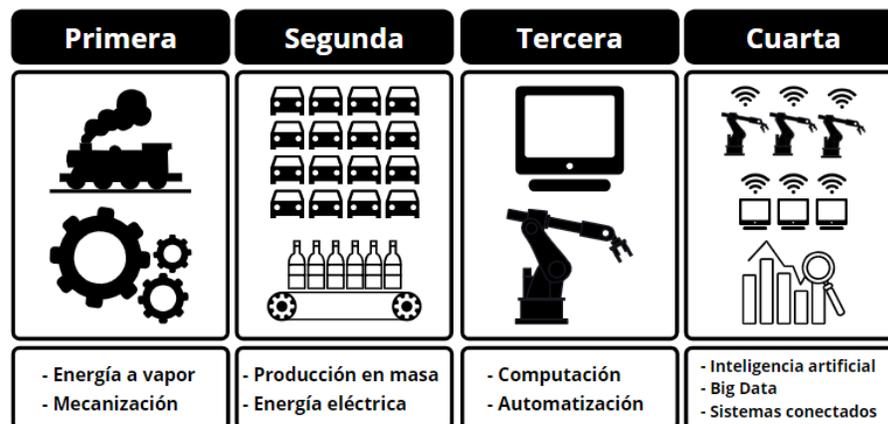


Figura 1. Esquema de las primeras cuatro revoluciones industriales. (Elaboración propia)

Uno de los sectores que ha adoptado sistemas de recopilación y análisis de datos es la agricultura, la implementación de sensores, drones y software, permiten el ahorro de suministros, recursos y tiempo, generando así una mejor calidad en los cultivos y reduciendo las consecuencias negativas propias de la industria en el medio ambiente. A corto plazo una de las principales ventajas corresponde a la disminución de horas de trabajo del suelo por parte de los agricultores, mientras que a mediano y largo plazo se observa el aumento de la productividad y efectividad de la explotación [2].

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, la magnitud del crecimiento de la industria y las tecnologías en general, es muy superior al de la agricultura, y por el momento sigue estando principalmente reservado para empresas o grandes firmas. Por esto factores como el cambio climático, declive de agentes polinizadores y enfermedades de cultivos, aun presentan una amenaza real para la estabilidad y sustentabilidad de la agricultura, especialmente para agricultores pequeños [2], [4].

2.2 Agricultura de la papa

La papa es el tercer cultivo más importante en términos de consumo humano después del arroz y el trigo, con 300 millones de toneladas métricas producidas y 1.4 mil millones de consumidores anualmente. La papa es esencial para la seguridad alimentaria frente al crecimiento de la población y el aumento de tasas globales de hambre. El principal consumidor de papas en el mundo es China y se espera que en 20 años el aumento de la demanda de alimentos de su población aumente en un 50%, en este escenario la agricultura de la papa jugará un rol fundamental [21].

En Chile la papa corresponde al cuarto cultivo más importante en cuanto a superficie con alrededor de 50 mil hectáreas cultivadas anualmente, además, corresponde al gremio agrícola más numeroso del país con 59.606 agricultores. La producción nacional es prácticamente completa para consumo interno. La papa no solo representa una pieza fundamental en cuanto a la dieta y economía, también, es importante por la gran variedad de especies nativas, siendo un patrimonio genético para el país [22].

2.3 Enfermedades que afectan los cultivos de papa

En el año 2011 los ingenieros agrónomos Ingrid Castro y Andrés Contreras, publicaron el manual “Manejo de plagas y enfermedades en el cultivo de papa” [3], describiendo con gran detalle la gran mayoría de las patologías por las que se pueden ver afectados los cultivos. En esta sección se recopila la información expuesta por los autores que resulta relevante para esta investigación.

La forma tradicional de reproducir una papa es de manera asexual, utilizando el tubérculo como semilla. Este método de reproducción vegetativa, requiere un proceso que mantenga la sanidad y preferiblemente posea certificación de semilla. El riesgo de convertir el tubérculo en un diseminador de enfermedades es grande, ya que a través del tiempo se va acumulando enfermedades de todo tipo. Existen muchas enfermedades y plagas que se alimentan de la planta de papa, provocando una baja del rendimiento, calidad y sanidad. El saber reconocer estas enfermedades que afectan el cultivo, resulta primordial para mantener una prevención y un control adecuado del producto.

Existe toda una gama de enfermedades que pueden afectar a la planta y al tubérculo. Los agentes patógenos, pueden estar presentes en el suelo o ser transmitidos por la papa semilla. Si el hospedero es abundante y permanente, se facilita la multiplicación del agente patógeno, por lo que, usar el suelo como monocultivo y se use la papa de mala calidad como semilla, se aumentará el inóculo y también las pérdidas debidas a un bajo rendimiento.

2.3.1 Tizón tardío

Esta enfermedad causada por el pseudohongo *Phytophthora infestans* es una de las más famosas e importantes a nivel mundial. Es responsable de la gran hambruna irlandesa, el caso más emblemático de alerta sanitaria sobre cultivos de papa del que se tenga registro. Es de distribución mundial, pudiéndose encontrar en cualquier lugar donde se cultive la papa. Un ataque fuerte de tizón puede destruir un papal en cuestión de días, ya que el hongo se multiplica rápidamente bajo las condiciones adecuadas. Los ataques se inician por las hojas desde que las plantas son pequeñas, si no existe control, podría afectar también los tallos de la planta, invadiéndola completamente hasta los tubérculos.

La Figura 2 muestra ejemplos de hojas y tubérculos afectados por esta enfermedad.

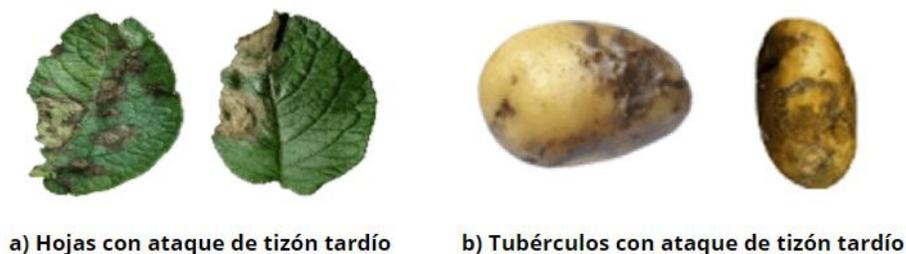


Figura 2. Ejemplos de ataque de tizón tardío sobre hojas y tubérculos.

2.3.2 Tizón temprano

Esta enfermedad es causada por el hongo *Alternaria solani* y se encuentra distribuida mundialmente. Generalmente se da en plantas desarrolladas con formación de tubérculos, siendo raro ver cultivos jóvenes atacados por el hongo. Es una enfermedad que se da comúnmente en las hojas, es muy raro ver ataques en tubérculos. Las papas atacadas no maduran bien y tienen problemas al momento de ser transportadas y almacenadas. Un ataque fuerte de esta enfermedad ocasiona graves mermas en la cosecha. La Figura 3 muestra ejemplos de hojas afectadas por esta patología.



Figura 3. Ejemplos de ataque de tizón temprano sobre hojas de papa.

2.3.3 Sarna común

La Sarna común es causada por dos tipos de bacterias *Streptomyces scabies* y *S. acidiscabies*, se ha introducido por la utilización de papa-semilla infectada en los suelos de todo el mundo. Es la cuarta enfermedad de más importancia en Estados Unidos y la principal en Chile con alrededor de un 80% de los tubérculos afectados. Esta bacteria no afecta el rendimiento de los cultivos, sin embargo, la calidad estética se ve afectada. El aspecto de la papa es uno de los factores más importantes al momento de comercializar la papa, por lo que el ataque de esta bacteria trae consigo daños económicos severos, disminuyendo el precio del producto. La Figura 4 muestra ejemplos de tubérculos afectados por esta enfermedad.



Figura 4. Ejemplos de ataque de sarna común en tubérculos.

2.4 Visión computacional

La visión computacional es de los campos más populares del *machine learning* y uno de los que tiene aplicación más concreta y utilizada. Las áreas de investigación que utilizan esta herramienta son variadas y van desde de la medicina hasta la ingeniería. Su aplicación abarca desde complejos sistemas de conducción autónoma de automóviles hasta la rutinaria acción de desbloquear un Smartphone utilizando reconocimiento facial.

El objetivo principal de las técnicas de visión computacional es transferir a un computador la capacidad de mirar, identificar y comprender el mundo visual, de la forma que lo hacen los humanos. Los investigadores del campo de la visión computacional aspiran a desarrollar algoritmos que permitan resolver tres tareas principales: clasificación de imágenes, detección de objetos, segmentación de imágenes [23]. La Figura 5 muestra las diferencias de los tipos de problemas de percepción visual, utilizando el ejemplo de un conjunto de tubérculos.

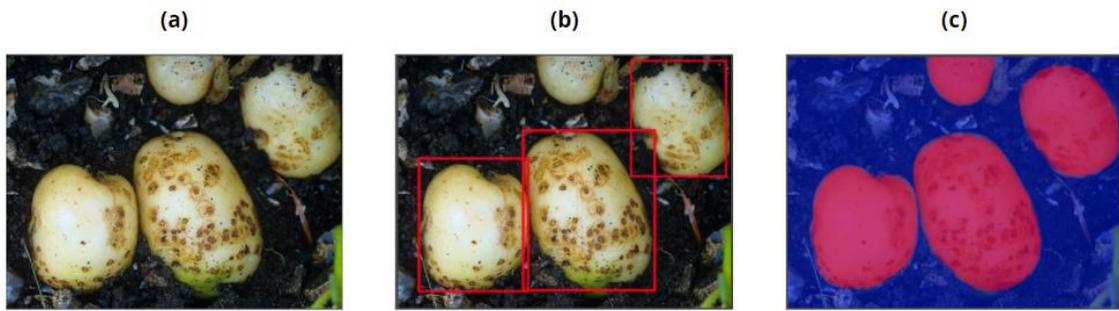


Figura 5. a) Imagen original. b) Clasificación y detección. c) Segmentación.

La Clasificación busca identificar los objetos o clases se encuentran presentes en una determinada imagen. La detección va un paso más allá con el objetivo de identificar e indicar la zona de la imagen en la que encontró cada categoría, como se muestra en la Figura 5 (b). Finalmente, la segmentación busca diferenciar entre objetos, fondos y perspectivas, un ejemplo de esto se observa en la Figura 5 (c), donde se diferencia entre cada uno de los tubérculos y se segmentan del fondo de la imagen. La detección y segmentación resultan ser de suma importancia en casos donde se tiene una gran cantidad objetos agrupados o sobrepuestos que se buscan diferenciar y reconocer, como lo serian árboles en un bosque o productos en un carro de supermercado.

Capítulo 3

3 Fundamentos teóricos

3.1 Redes neuronales convolucionales

Las redes neuronales convolucionales (CNN) son una técnica de aprendizaje profundo idénticas a las redes neuronales artificiales (ANN), están compuestas por neuronas, pesos y sesgos de aprendizaje. Sin embargo, las ANN presentan una dificultad al momento de clasificar imágenes de gran tamaño [25]. Si una ANN recibe una imagen en blanco y negro de calidad 8K (7680 × 4320 píxeles) necesitaría un total de 33.177.600 neuronas en la capa de entrada, y si se considerara una imagen con colores esto se multiplicaría, lo que implicaría además un gran número de parámetros. Por esto, enfrentar problemas de clasificación de imágenes de mayor complejidad con ANN, es computacionalmente muy costoso, además, existe una gran posibilidad de un sobre ajuste del modelo.

Las CNN permiten extraer las principales características de una imagen y bajar su dimensión sin perder sus principales cualidades, en palabras simples se tienen filtros que permiten agrupar los píxeles con los de su vecindario en un cierto rango. Esto es una analogía al modo de como el ser humano logra aprender a diferenciar y reconocer cosas, analizando no cada punto en específico de lo que rodea, sino, líneas, formas y patrones, que le permiten diferenciar, entender e interpretar lo que ve.

Una CNN típica está compuesta por las siguientes capas principales: Capa convolucional (Kernel), capa de pooling y capa de clasificación (completamente conectada) [26].

3.1.1 Capa convolucional

El objetivo principal de las capas convolucionales es realizar extracciones de las características de la imagen de entrada mediante la aplicación de filtros. Una CNN puede tener más de una capa convolucional, mientras que las primeras se encargan de extraer las características de bajo nivel como formas, líneas, color, las capas más profundas son encargadas de extraer las características de alto nivel, mucho más abstractas pero que son las que finalmente permiten a la red tener un entendimiento de las imágenes avanzado similar al de los humanos.

Una imagen es un conjunto de píxeles que se puede representar como una matriz de Largo x Ancho, por lo que una imagen de 5 x 5 píxeles en blanco y negro (1 canal de color), se puede representar como una matriz de 5 x 5. Si se utiliza un filtro con valor de kernel igual a 3 ($k=3$), el proceso de convolución se puede ejemplificar como lo muestra el esquema de la Figura 6.

importante recalcar que si la imagen tuviera más de un canal (imagen en colores), el proceso descrito se realizaría 3 veces, una por cada canal de color, teniendo así 3 kernel distintos.

3.1.2 Capa de pooling

La capa de pooling es la encargada de bajar la dimensión espacial de las características extraídas por la convolución, esto disminuye el poder computacional requerido para procesar los datos. Existen dos tipos de capa de pooling: MaxPooling y Average pooling. La capa de MaxPooling devuelve el valor más alto de la porción de la imagen cubierta por el filtro. La capa de average pooling devuelve el promedio de todos los valores cubiertos por el filtro. La combinación de las capas convolucional y pooling, corresponden a los bloques convolucionales de la red y permiten extraer características y disminuir el costo computacional de entrenar los modelos. La Figura 7 muestra un ejemplo del uso de las capas MaxPooling y Average pooling.

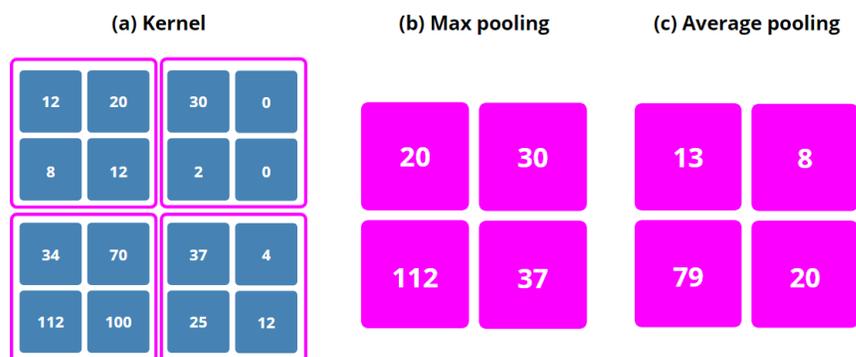


Figura 7. Esquema del funcionamiento de la capa pooling. (Elaboración propia)

En la Figura 7 (b) se observa como la capa de MaxPooling se queda directamente con el valor más alto de cada zona cubierta con el kernel descrito

en (a). En (c) se representa como la capa de Average pooling toma el valor del promedio de las zonas cubiertas por el kernel.

3.1.3 Capa de clasificación (capa completamente conectada)

Tras pasar por los bloques convolucionales y reducir la dimensión de la imagen de entrada, ya se encuentran las condiciones para utilizar el perceptrón multicapa convencional de las ANN. Se transforma la imagen en un vector de columna con una capa flatten y luego se ingresa a una capa completamente conectada que permite aprender las combinaciones no lineales de las características de salida de las capas convolucionales. Finalmente se tiene una capa softmax que permite clasificar en el número de clases que tenga el conjunto de datos de entrenamiento. La Figura 8 muestra un esquema de una CNN básica de una sola capa convolucional, que posee las capas principales que se han descrito en esta sección.

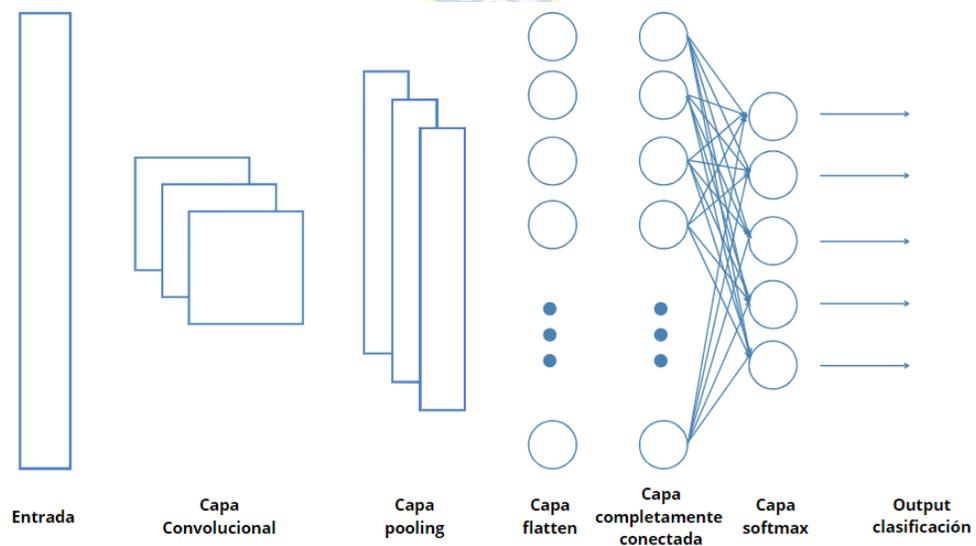


Figura 8. Esquema básico de la arquitectura de una CNN de una capa convolucional.

(Elaboración propia)

No existe un número establecidos de capas convolucionales que debe poseer una red, las distintas arquitecturas de CNN en la literatura realizan un balance entre la profundidad de la red (número de capa convolucionales) y el tamaño de los filtros. La línea principal en el último tiempo se basa modelos muchos más profundos, con arquitecturas de bloques sucesivos de capas convolucionales. El enfoque de modelos mucho más profundos permite generar modelos capaces de aprender características de alto nivel con mucho mayor detalle, sin embargo, genera un número de parámetros mayor y un costo computacional más elevado.

3.2 Red VGG16

La red VGG16 fue introducida en el año 2014 por Simonyan et al. [12] y junto a la red VGG19 pertenecen a variantes de la familia de arquitecturas VGG. Esta arquitectura alcanza una precisión del 92.7% clasificando entre las 1.000 categorías del conjunto de datos ImageNet, siendo ganadora del ILSVRC [24] el año 2014.

Esta red presentó un precedente al incorporar una arquitectura más profunda y capas convolucionales con filtros más pequeños. Este nuevo enfoque incorporó 16 capas y filtros de 3 x 3, duplicando el número de capas y disminuyendo a más de la mitad el tamaño de los filtros, de su antecesora la red AlexNet propuesta por Krizhevsky et al. [14] con 8 capas y filtros de 7x7.

La red VGG 16, está compuesta de 5 bloques de 2 ó 3 capas convolucionales, con cada capa convolucional seguida de una capa ReLU, y al final de cada bloque una capa de MaxPooling. Al final de la red se tiene 3 capas completamente conectadas y por último una capa softmax que permite clasificar

entre las 1.000 clases del conjunto ImageNet. La Figura 9 muestra un esquema de la arquitectura de la red VGG16.

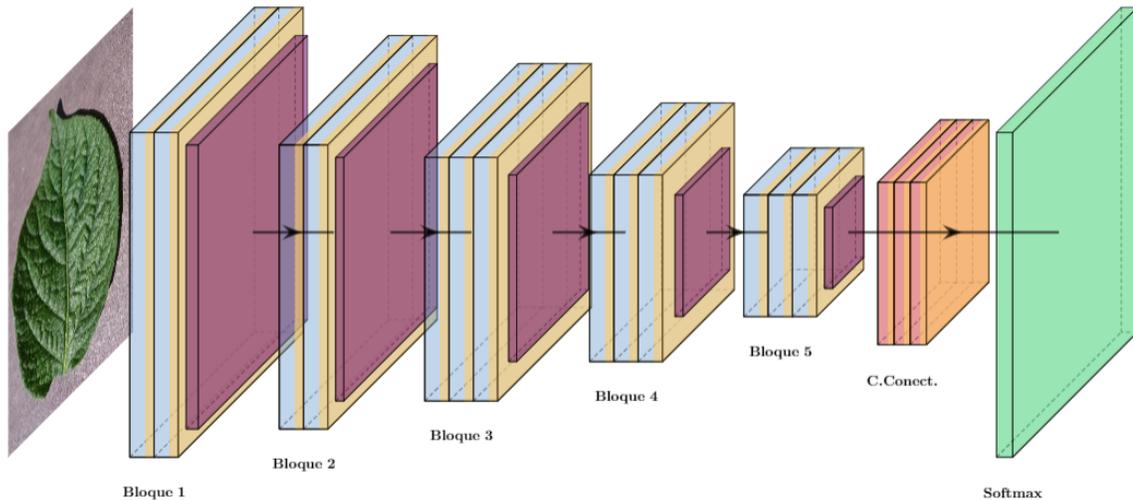


Figura 9. Esquema de la arquitectura de la red VGG16. (Elaboración propia)

3.3 Red DenseNet201

La red DenseNet201 fue introducida en el año 2017 por Huang et al. [13] y corresponde a una de las variantes de la familia de arquitecturas de redes neuronales convolucionales densamente conectadas. Los autores reportan resultados a la par con los modelos de la familia de redes ResNet [7] evaluando el desempeño de los modelos sobre distintos benchmarks en igualdad de condiciones. Lo anterior incluye al conjunto de datos ImageNet, donde presenta una precisión del 97.43% clasificando entre las 1.000 categorías disponibles, resultados que la llevaron a declararse triunfadora del ILSVRC [24] el año 2015.

La red DenseNet201 está compuesta por 2 tipos de conjuntos de capas: bloques densos y capas de transición. Los bloques densos corresponden a

secuencias de capas Batch normalization, ReLU y Convolutacional con filtros de 1x1 y 3x3, una tras otra, con la característica de que cada capa recibe todos los mapas de características de las capas convolucionales anteriores dentro del bloque. Las capas de transición están formadas por una secuencia de capas Batch normalization, ReLU, Convolutacional con filtros de 1x1 y una capa de Average pooling. La Figura 10 muestra un esquema de la arquitectura de la red DenseNet201.

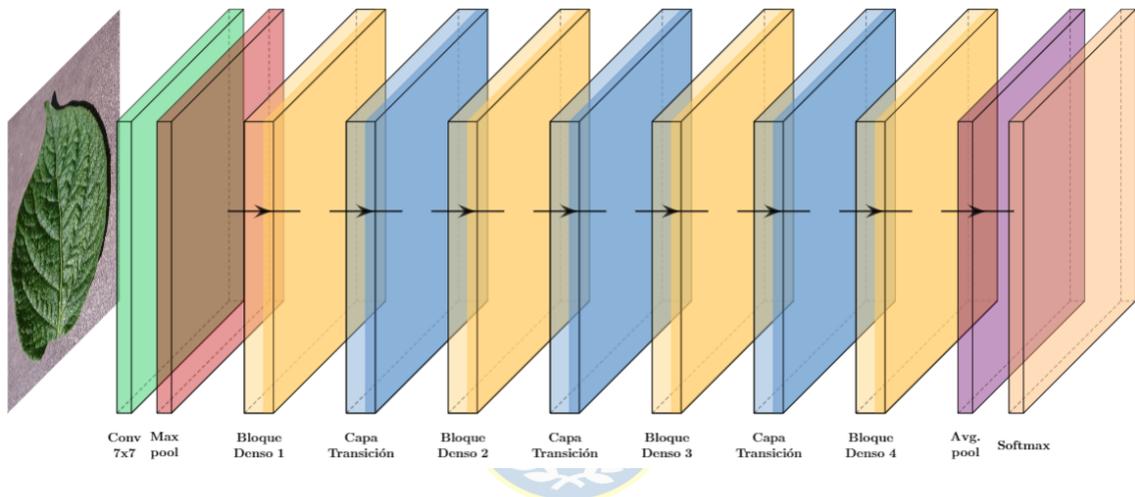


Figura 10. Esquema de la arquitectura de la red DenseNet201. (Elaboración propia)

3.4 Transferencia de aprendizaje

El auge en el rendimiento de los modelos de aprendizaje profundo expandió el alcance de los estudios basados en estas técnicas, extendiendo su aplicación a diversos ámbitos y campos de estudio. Bajo este escenario la técnica de transferencia de aprendizaje toma un rol importante, permitiendo utilizar arquitecturas pre-entrenadas como base para crear nuevos modelos capaces de clasificar nuevas categorías.

La idea principal de la transferencia de aprendizaje es sacar provecho de modelos que ya reporten buenos resultados y poseen buena adaptación para aprender características de bajo nivel que son comunes para todo tipo de imágenes (líneas, formas, colores, etc.). De esta forma solo es necesario entrenar la red sobre un nuevo conjunto de datos, para que aprenda las características específicas de las nuevas categorías [27]. Principalmente se incorpora una nueva capa de clasificación sobre la salida original de la arquitectura, incorporando una nueva capa completamente conectada y una capa softmax de clasificación. La Figura 11 muestra un esquema básico de la arquitectura resultante de añadir una nueva capa de clasificación a una arquitectura previamente entrenada.

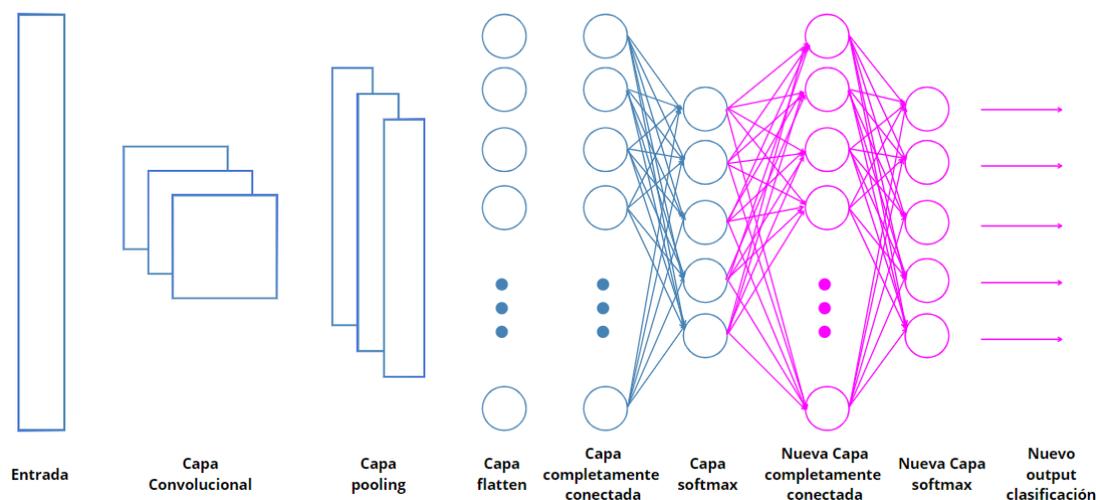


Figura 11. Arquitectura básica de una CNN con adición de una nueva capa de clasificación adaptada para nuevas categorías. (Elaboración propia)

En la Figura 11 se representa en color morado la nueva capa de clasificación incorporada a la red, esta nueva arquitectura conserva los pesos de la red original entrenada previamente sobre otro conjunto de datos representada en color azul, permitiendo así mantener capacidad de aprendizaje de la red original. Solo se entrena con un nuevo conjunto de datos, durante un cierto

número de épocas los parámetros de la capa de clasificación añadida, permitiendo así clasificar nuevas categorías.

3.5 Mapas de características

Los modelos de redes neuronales en general funcionan como una caja negra, por lo que no se tiene una explicación clara y específica del porqué el modelo tomo una decisión o realizó cierta predicción. Esto genera cierta incredulidad de parte de las áreas de la ciencia que históricamente se han basado en modelos empíricos. Para el caso de las CNN que se centran en el trabajo con imágenes, existen métodos propuestos por diversos autores que permitirían obtener una retroalimentación visual del proceso discriminativo de los modelos.

Las capas convolucionales de una CNN presentan el comportamiento de un localizador de objetos sin supervisión, permitiendo identificar visualmente la discriminación respecto del porqué el modelo asocio la entrada a una categoría específica. Sin embargo, esta capacidad de localizar se pierde una vez que las salidas de las capas convolucionales ingresan a la capa completamente conectada al momento de clasificar [29].

Zhou et al. [17] en su publicación “Learning Deep Features for Discriminative Localization”, propusieron la técnica de mapeo de activación de clases (CAM), para en la cual los puntajes obtenidos por la categoría predicha se mapean en la última capa convolucional de la red, esto permite resaltar las regiones discriminativas de la imagen.

En 2019, Brownlee [18] implementó una metodología para extraer los filtros de cada una de las 13 capas convolucionales de la arquitectura original de la red VGG16. Aplicando estos filtros a una imagen de entrada se obtienen los mapas

de características de la red para dicha imagen. En palabras simples se crean sub-modelos a partir de la red original, que toman una imagen como entrada y la salida corresponde al mapa de características. Independiente de la dimensión de la imagen del mapa de características, la relación espacial se mantiene respecto de la imagen de entrada original, permitiendo superponerlas con la imagen original.

Mientras más profunda sea la capa convolucional, más abstracta se tornará la imagen de la salida y las características capturadas por los filtros, a tal punto que para capas muy profundas ya sea imposible interpretar el resultado de los sub modelos bajo la lógica racional de un ser humano.

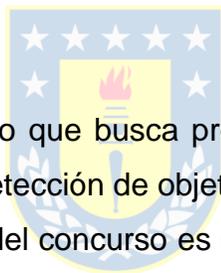


Capítulo 4

4 Revisión de la literatura

4.1 ImageNet Large Scale Visual Recognition Challenge

El auge del uso de técnicas de aprendizaje profundo como la alternativa número uno para abordar problemas de clasificación de imágenes, se debe principalmente a la creación del conjunto de datos ImageNet [24], al concurso “*ImageNet Large Scale Visual Recognition Challenge*” y a la aparición del modelo AlexNet [14].



El ILSVRC es un concurso que busca promover en los investigadores el desarrollo de algoritmos para detección de objetos y clasificación de imágenes a gran escala. La tarea principal del concurso es generar algoritmos que permitan clasificar entre 1.000 categorías distintas con 1.4 millones de imágenes disponibles para entrenamiento, pertenecientes al conjunto de datos ImageNet.

El año 2012 el modelo AlexNet basado en arquitecturas de CNN, alcanzó una precisión del 83.6% dejando muy por detrás al segundo lugar y disminuyendo en gran porcentaje respecto del 74.2% obtenido por el ganador del año anterior. Desde ese momento las CNN se volvieron el enfoque principal para resolver este tipo de problemas, alcanzando precisiones inesperadas y que solo aumentaban año tras año, alcanzando el 97.43% de precisión el año 2015. Otro factor importante que incide, no solo, en el auge de las CNN, sino, de las redes neuronales en general es el significativo aumento en el poder de cómputo de los procesadores actuales.

4.2 CNN y transferencia de aprendizaje para clasificar enfermedades en cultivos

En la agricultura, dos tercios de las pérdidas que se tienen anualmente el sector se deben a enfermedades que afectan los cultivos. En el marco de la búsqueda de alguna solución para esta problemática, las CNN se han tornado como la alternativa número uno en los últimos años, jugando un rol importante en la creación de modelos capaces de reconocer enfermedades en distintos tipos de cultivos.

El año 2015 Mohanty et al. [4], entrenaron las arquitecturas AlexNet [14] y GoogleNet [28] con el conjunto de datos PlantVillage [15], creando nuevos modelos capaces de clasificar entre 26 enfermedades que afectan a 14 tipos de cultivos distintos, obteniendo para el mejor modelo implementado una de 99.35%.

En 2018 Rangarajan et al. [8], enfocaron su trabajo en la clasificación de enfermedades en cultivos de tomate. Utilizando un subconjunto del conjunto de datos PlantVillage [15], entrenaron las redes AlexNet [14] y VGG16 [12] para clasificar entre 6 enfermedades que afectan cultivos de tomate, obteniendo precisiones de 97.49%.

En 2019 Elsharif et al. [9], enfocaron su trabajo en diferenciar distintos tipos y variedades de tubérculos. Implementaron una CNN de arquitectura propia para clasificar entre 4 tipos de especies papa, reportando precisiones de un 99.5%.

Too et al. [11], evaluaron el desempeño de las redes VGG16 [12], Inception V4 [28], ResNet [7] y DenseNet [13], para identificar enfermedades en cultivos, sobre el mismo conjunto que Mohanty et al. [14], alcanzando una precisión de 99.75% para la red DenseNet.

En 2020 Da costa et al. [10], utilizaron la red pre entrenada ResNet [7] para identificar defectos en una industria procesadora de tomates en Brasil, permitiendo clasificar el estado estético de los tomates entre aptos para venderse como tal o ser destinado a industria de salsa de tomate o derivados, obteniendo una precisión de 94.6%.



Capítulo 5

5 Metodología

5.1 Configuración del sistema

El estudio se llevó a cabo con un sistema operativo Windows 10, CPU Intel(R) Core(TM) i5-7300HQ @ 2.50GHz, GPU NVIDIA GeForce GTX 1050 2GB y 8GB de RAM. Los modelos de CNN fueron implementados en Python versión 3.8.5, utilizando las librerías Tensorflow versión 2.3 y Keras versión 2.4.3.



5.2 PlantVillage

Para entrenar los modelos clasificadores de hojas de papa se utilizó parte del conjunto de datos PlantVillage [15], 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. Este conjunto de datos fue creado en el marco de un proyecto que buscaba poner a disposición de la comunidad un repositorio de acceso libre a imágenes, con el fin de facilitar el desarrollo de herramientas móviles capaces de detectar enfermedades. Todas las imágenes del conjunto de datos fueron tomadas en estaciones de investigación asociadas a universidades Land Grant en Estados Unidos por técnicos especializados utilizando una cámara Sony DSC - Rx100/13

20.2 megapíxeles. 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, etiquetadas con las categorías sana, *Alternaria solani* (Tizón temprano) y *Phytophthora Infestans* (Tizón tardío), en sus versiones original y segmentada. La Figura 12 muestra un ejemplo de cada categoría de los conjuntos de datos seleccionados.

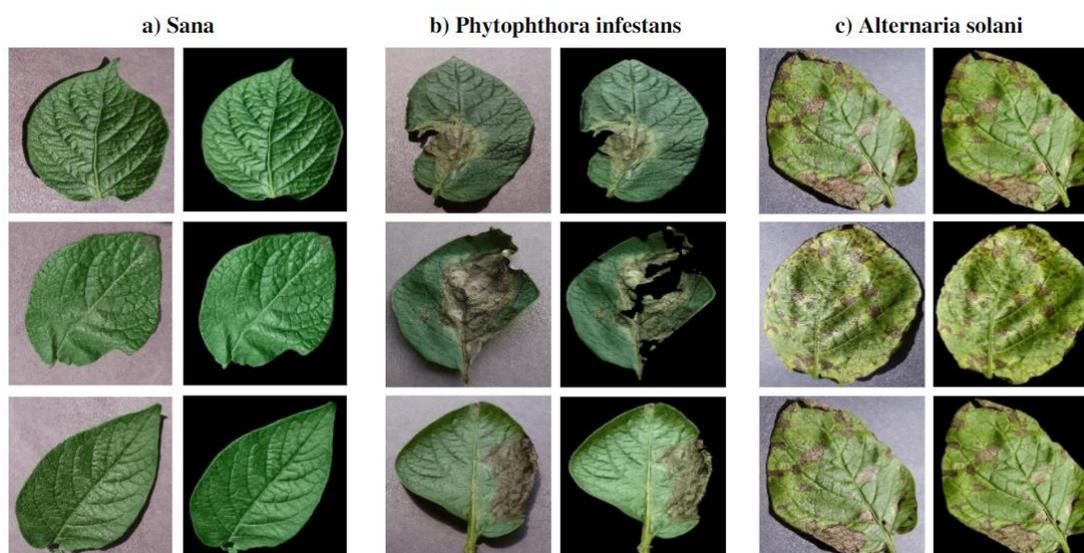


Figura 12. Muestra de imágenes de los 3 tipos de categorías y 2 tipos de formatos utilizados en las configuraciones experimentales de este estudio.

El conjunto de datos original se encontraba desbalanceado respecto de la relación de la categoría sana con las que poseían alguna patología. La categoría de tizón temprano y tizón tardío estaban compuestas por 1.000 imágenes cada una, mientras que la categoría sana solo poseía 100 imágenes. Para solucionar este problema se realizó una aumentación de los datos correspondientes a la categoría sana, aplicando una rotación en sentido anti horario de 90, 180 y 270 grados, seguido de una reflexión horizontal al resultado de cada una de las

rotaciones. La Figura 13 muestra un ejemplo de las transformaciones realizadas durante la aumentación de datos.

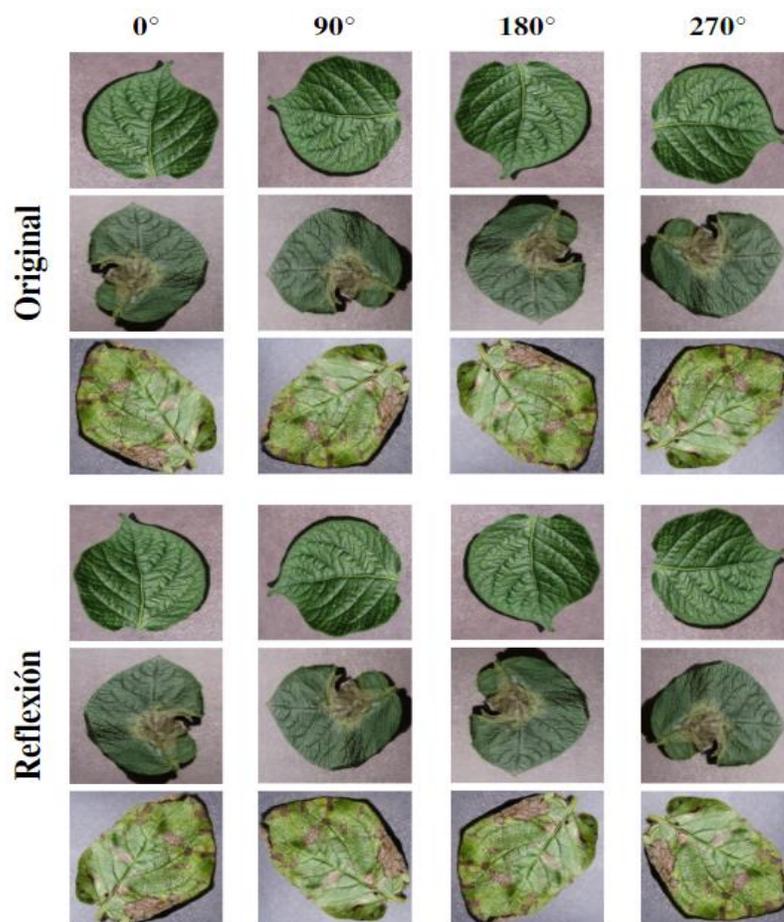


Figura 13. Ejemplo de las transformaciones realizadas a las imágenes de la categoría sana durante los procesos de aumentación de datos.

Finalmente se obtuvo un conjunto de datos con 1.216 imágenes para la categoría sana, 1.000 imágenes para la categoría tizón temprano y 1.000 imágenes para la categoría tizón tardío.

5.3 Fruit360

Para entrenar los modelos clasificadores de tubérculos se utilizó parte del conjunto de datos Fruit360 [16], el cual está compuesto por frutas y vegetales, correspondientes a 90.483 imágenes. El conjunto de datos tiene en total 131 categorías, de las cuales 4 corresponden a tipos de papa: white potato, red potato, washed red potato y sweet potato. No posee categorías que incluyan patologías, ya que este no fue el objetivo principal con el que fue creado.

Todas las imágenes fueron tomadas con una cámara Logitech C920, a estas se les realizó una substracción del para crear un fondo uniforme de color blanco. Para este estudio se seleccionaron las imágenes correspondientes a la categoría White potato. La Figura 14 muestra un ejemplo de las imágenes de esta categoría.

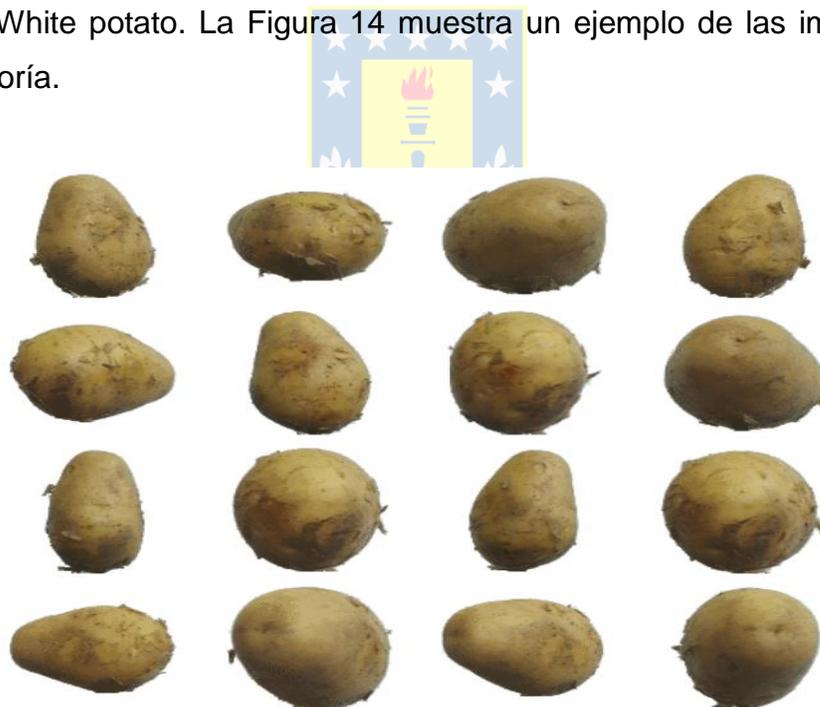


Figura 14. Muestra de imágenes de la categoría White potato del conjunto Fruit360.

Para la adquisición de imágenes de papa con alguna patología, se utilizó técnicas de web scraping. Se buscó imágenes que contuvieran tubérculos afectados con *Streptomyces scabies* (sarna común) y *Phytophthora infestans* (tizón tardío), para esto se realizaron dos procesos de búsqueda. La primera búsqueda se realizó utilizando las palabras clave: White potato, papa, common scab, *Streptomyces scabies* y sarna común. La segunda búsqueda se utilizó las palabras clave: White potato, papa, late blight, *Streptomyces scabies* y tizón tardío. Para aumentar la variedad de imágenes de tubérculos sanos, también, se realizó una búsqueda de imágenes para esta categoría. Las imágenes obtenidas, fueron filtradas para evitar duplicidad y revisadas manualmente a modo de eliminar registros mal etiquetados. Todas las imágenes encontradas fueron segmentadas, re escaladas y sustraídas del fondo, para hacerlas coincidir con el formato del conjunto de datos Fruit 360. La Figura 15 muestra un ejemplo del proceso de segmentación que se realizó sobre las imágenes encontradas.



Figura 15. a) Imagen original. b) Segmentación de imagen. c) Tubérculo segmentado y sustraído del fondo.

En la Figura 15 (a) se puede observar ejemplos de imágenes recolectadas con las técnicas de web scraping. En (b) se representa la selección y segmentación de los tubérculos útiles presentes en la imagen. En (c) se muestran ejemplos de la imagen del tubérculo final, segmentado y sustraído del fondo.

El número de imágenes recolectadas tras el web scraping no fue el suficiente como para entrenar el modelo clasificador. Para solucionar este problema se aplicó el mismo procedimiento de aumentación de datos que al conjunto PlantVillage descrito en la Sección 5.2. La Figura 16 muestra un ejemplo de las transformaciones realizadas al conjunto de datos resultante.

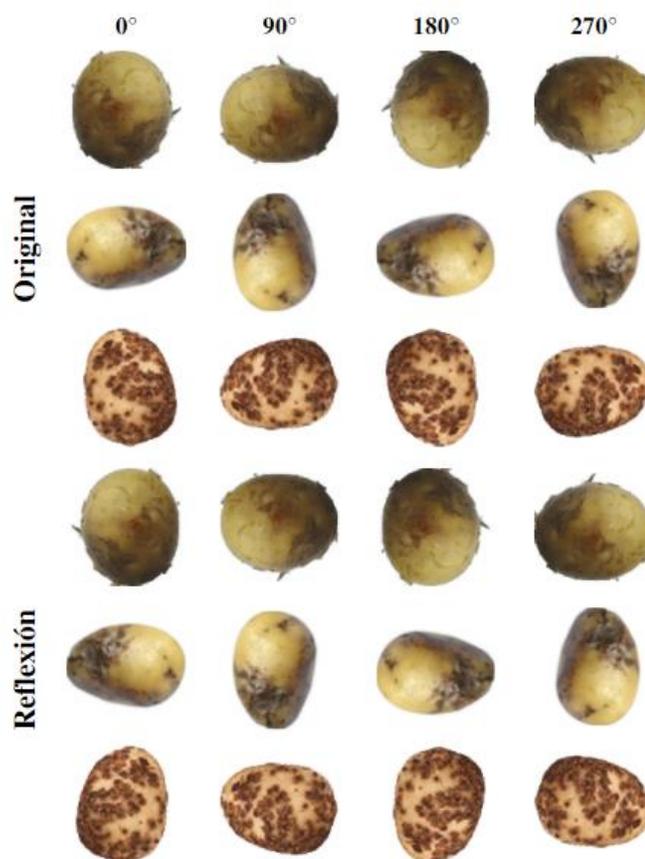


Figura 16. Ejemplo de las transformaciones realizadas a las imágenes durante los procesos de aumentación de datos.

Finalmente se obtuvo un conjunto de datos con 1.000 imágenes para la categoría sana, 1.000 para la categoría sarna común y 800 para la categoría tizón tardío. La Figura 17, muestra un ejemplo de cada tipo de categoría del conjunto de datos.

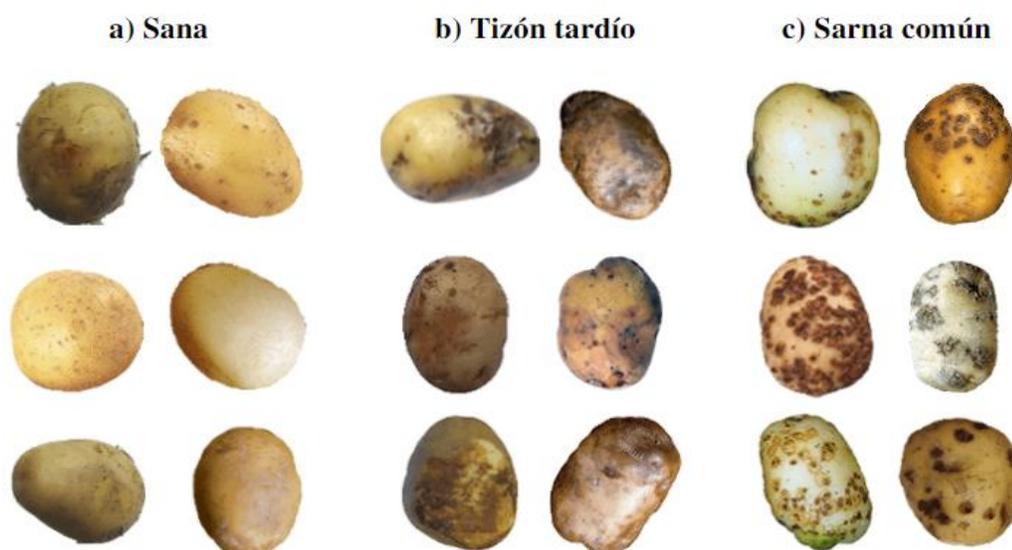


Figura 17. Muestra de imágenes de las 3 categorías utilizadas en las configuraciones experimentales de este estudio.

5.4 Transferencia de aprendizaje

Para implementar los modelos clasificadores de hojas de papa y tubérculos, se utilizó la técnica de transferencia de aprendizaje sobre arquitecturas de CNN que han demostrado versatilidad al momento de ser re-entrenadas sobre nuevos conjuntos de datos, para clasificar nuevas categorías. Para esto se seleccionó las arquitecturas de los modelos VGG16 y DenseNet201, ambos con buenos resultados en los principales benchmarks utilizados para comparar precisión en la tarea de clasificar objetos y que reportan buena adaptación para clasificar específicamente enfermedades en cultivos.

Para los modelos implementados a partir de la red VGG16, se utilizó la arquitectura original de la red con 14.714.688 parámetros y con los pesos de pre entrenamiento con el conjunto de imágenes de ImageNet. A esta arquitectura se le añadió una nueva capa de clasificación, la salida de la red original entra a una capa completamente conectada y finalmente a una nueva capa de clasificación softmax. La arquitectura resultante conto con 75.267 nuevos parámetros que permitieron rentrenar la red sobre nuevos conjuntos de datos y clasificar nuevas categorías.

Para los modelos implementados a partir de la red DenseNet201, se utilizó la arquitectura original de la red con 18.321.984 parámetros y con los pesos de pre entrenamiento con el conjunto de imágenes de ImageNet. A esta arquitectura se le añadió una nueva capa de clasificación, la salida de la red original entra a una capa completamente conectada y finalmente a una nueva capa de clasificación softmax. La arquitectura resultante conto con 282.243 nuevos parámetros que permitieron rentrenar la red sobre nuevos conjuntos de datos y clasificar nuevas categorías.

Se entrenó ambas arquitecturas sobre los conjuntos de datos PlantVillage original, PlantVillage segmentado y Fruit360 más adición de imágenes con web scraping, detallados en las secciones anteriores. De esta forma se obtuvieron seis modelos en total, donde cuatro corresponden a clasificadores de hojas de papa capaces de identificar las clases sana, tizón temprano y tizón tardío. Y dos modelos correspondientes a clasificadores de tubérculos capaces de identificar las clases sana, sarna común y tizón tardío.

Las Figura 18 y Figura 19 muestran un esquema de las redes VGG16 y DenseNet201 respectivamente, donde se representa en color azul las

arquitecturas originales de las redes y en color morado las nuevas capas de clasificación que se le añadió a cada una.

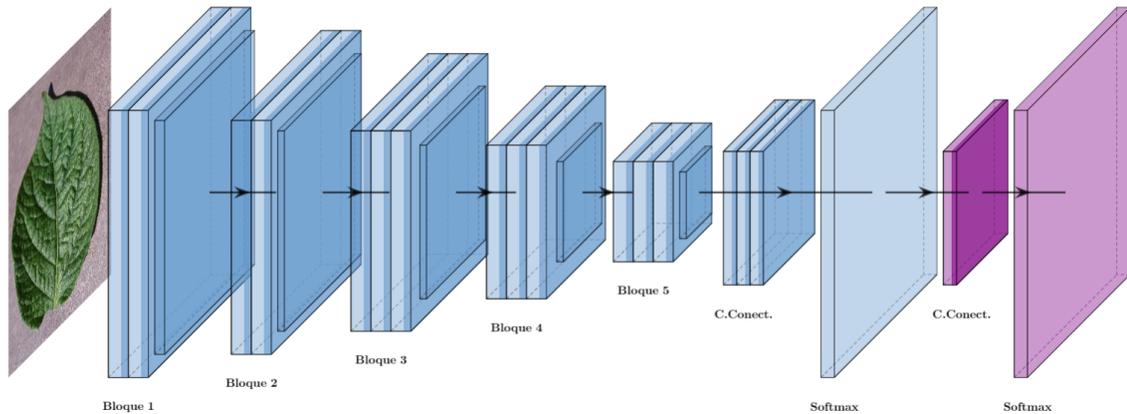


Figura 18. Esquema de la red VGG16 adaptada para clasificar sobre nuevas categorías.

(Elaboración propia)

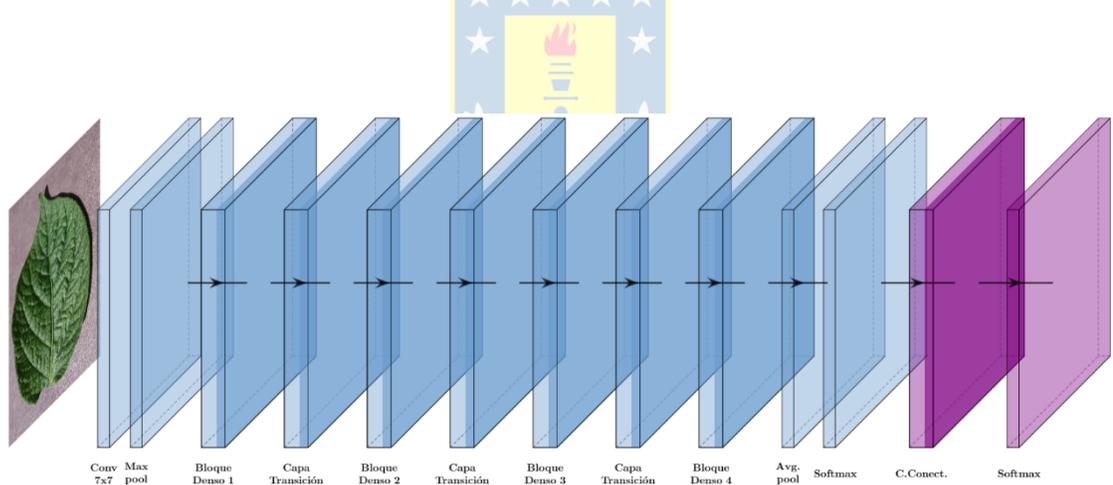


Figura 19. Esquema de la red DenseNet201 adaptada para clasificar sobre nuevas categorías.

(Elaboración propia)

Los conjuntos de datos PlantVillage original, PlantVillage segmentado y Fruit360 con adición de web scraping, fueron divididos en 80% entrenamiento, 10% validación y 10% prueba. Los conjuntos se armaron de forma que las imágenes obtenidas a partir del proceso de aumentación se encontraran en el

mismo conjunto que la imagen original. Es decir, si una imagen fue seleccionada para uno de los conjuntos (entrenamiento, validación o prueba) cualquier otra imagen que corresponda a una rotación o reflexión de esta, fue seleccionada para el mismo conjunto. Esto se realizó a fin de evitar un modelo mecanizado para reconocer los mismos patrones ya extraídos por los filtros de convolución durante el entrenamiento.

Las imágenes fueron re-escaladas a 224x224 píxeles, para que coincidieran con las dimensiones de entrada estándar de las redes pre-entrenadas. Para todas las distintas configuraciones de experimentos, los entrenamientos de los modelos se realizaron durante 20 épocas. Los modelos implementados a partir de la red VGG16 tomaron un tiempo de entrenamiento aproximado de 360 segundos por época con el conjunto de datos PlantVillage y 323 segundos por época con el conjunto Fruit360 más web scraping. Los modelos implementados a partir de la red DenseNet201 tomaron un tiempo de entrenamiento aproximado de 280 segundos por época con el conjunto de datos PlantVillage y 238 segundos por época con el conjunto Fruit360 más web scraping.

5.5 Mapas de características

Se replicó la metodología descrita en la Sección 3.5 por Brownlee [18] creando una serie de sub modelos a partir de la red VGG16. Se seleccionó los filtros de las principales capas de la red, correspondientes a la última capa convolucional de cada uno bloque de convolución de la descritos en la Sección 3.2.

Los filtros se aplicaron a imágenes al azar de cada una de las categorías de los conjuntos de prueba PlantVillage original y Fruit360 más adición de imágenes

por web scraping. Finalmente se superpuso los principales mapas de características de las primeras capas convolucionales con la imagen de entrada al modelo, generando nuevas visualizaciones.



Capítulo 6

6 Resultados

6.1 Entrenamiento sobre conjunto de datos PlantVillage original

Los modelos entrenados con el conjunto de datos PlantVillage original alcanzaron una precisión de validación 95.7% en la época número 20 para el caso de la red VGG16 y un 98.4% en la época número 17 para el caso de la red DenseNet201. La Figura 20 muestra la precisión de validación alcanzada por los modelos VGG16 y DenseNet201 sobre el conjunto PlantVillage original en todo el horizonte de épocas de entrenamiento.

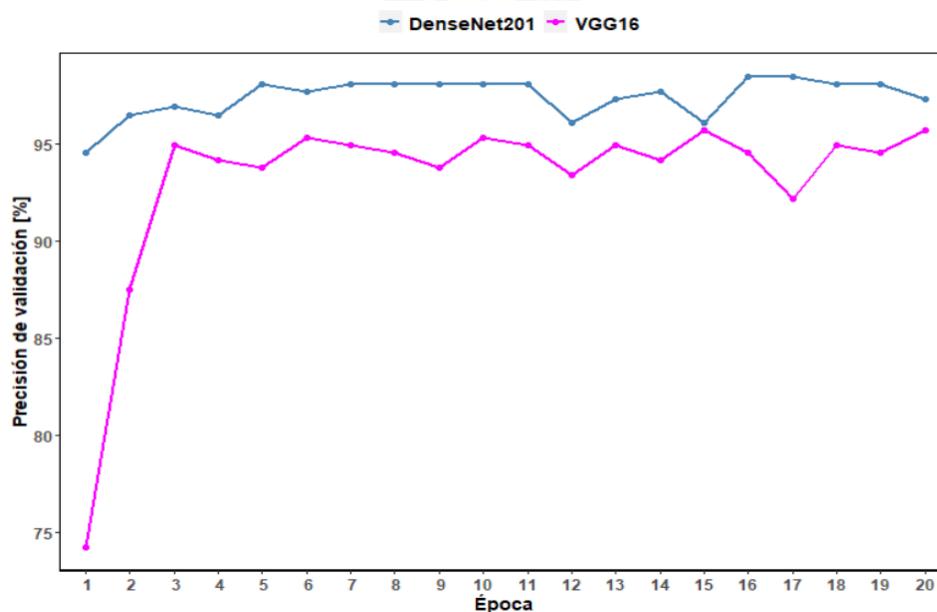


Figura 20. Precisión de validación de las redes VGG16 y DenseNet201 sobre el conjunto PlantVillage original durante las 20 épocas de entrenamiento.

6.2 Entrenamiento sobre el conjunto PlantVillage segmentado

Los modelos entrenados con el conjunto de datos PlantVillage segmentado alcanzaron una precisión de validación 95.7% en la época número 18 para el caso de la red VGG16 y un 99.2% en la época número 14 para el caso de la red DenseNet201. La Figura 21 muestra la precisión de validación alcanzada por los modelos VGG16 y DenseNet201 sobre el conjunto PlantVillage segmentado en todo el horizonte de épocas de entrenamiento.

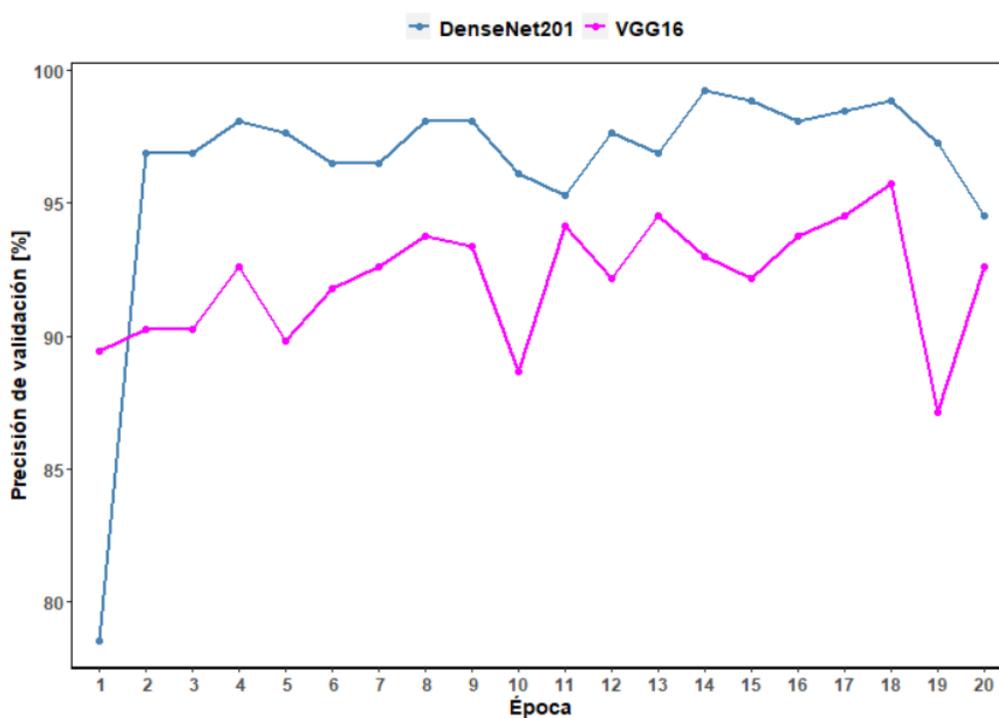


Figura 21. Precisión de validación de las redes VGG16 y DenseNet201 sobre el conjunto PlantVillage segmentado durante las 20 épocas de entrenamiento.

6.3 Entrenamiento de modelos clasificadores de tubérculos

Los modelos entrenados con el conjunto de datos Fruit360 más web scraping alcanzaron una precisión de validación 96.8% en la época número 11 para el caso de la red VGG16 y un 89.8% en la época número 5 para el caso de la red DenseNet201. La Figura 22 muestra la precisión de validación alcanzada por los modelos VGG16 y DenseNet201 sobre el conjunto Fruit360 más web scraping en todo el horizonte de épocas de entrenamiento.

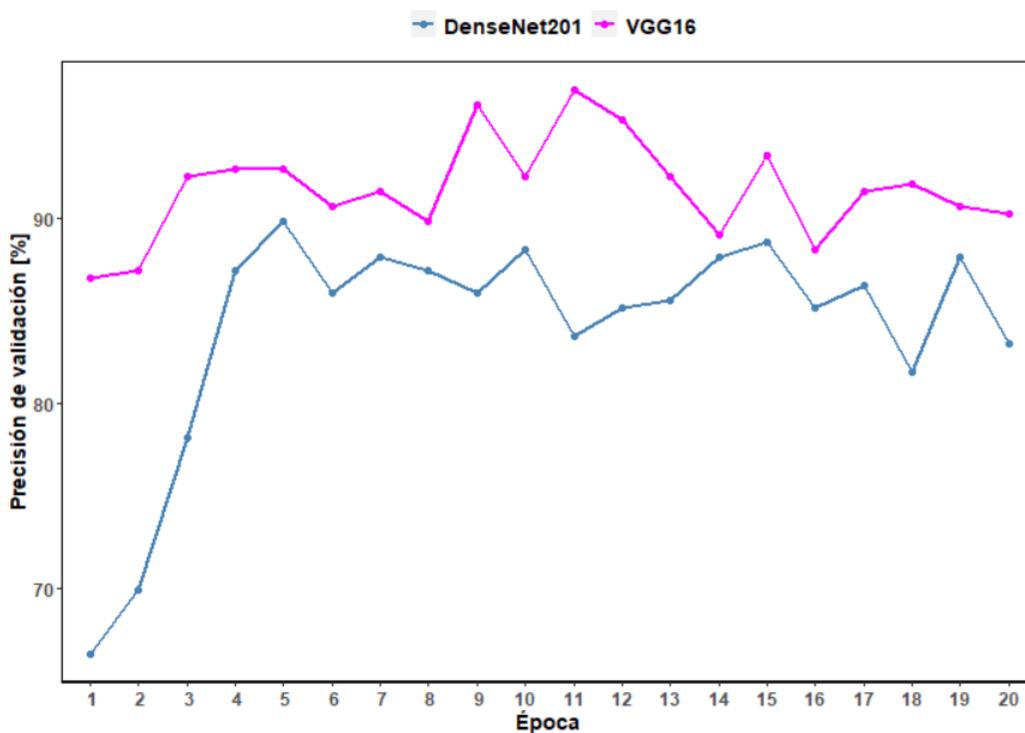


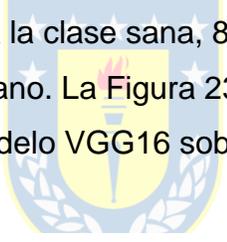
Figura 22. Precisión de validación de las redes VGG16 y DenseNet201 sobre el conjunto Fruit360 más web scraping durante las 20 épocas de entrenamiento.

6.4 Evaluación sobre conjuntos de prueba

6.4.1 Evaluación sobre conjunto PlantVillage original

Los modelos con mejores resultados de precisión de validación entrenados con el conjunto PlantVillage original, se evaluaron sobre el conjunto de prueba. El modelo entrenado con la red VGG16 obtuvo un 94.7% de precisión sobre el conjunto de prueba y el modelo entrenado con la red DenseNet201 obtuvo un 99.1% de precisión sobre el conjunto de prueba.

Para el modelo VGG16, se obtuvo una predictibilidad del 92.5% para la clase sana, 96% para la clase tizón tardío y 96% para la clase tizón temprano. Una sensibilidad del 97.3% para la clase sana, 88.8% para la clase tizón tardío y 97.9% para la clase tizón temprano. La Figura 23 muestra la matriz de confusión creada a partir de evaluar el modelo VGG16 sobre el conjunto de prueba.



Etiqueta Predicha

		Etiqueta Predicha		
		Tizón Temprano	Sana	Tizón Tardío
Etiqueta Real	Tizón Temprano	96	1	3
	Sana	0	111	9
	Tizón Tardío	2	2	96

Figura 23. Matriz de confusión creada a partir de la evaluación del modelo de la red VGG16 sobre el conjunto de prueba PlantVillage original.

Para el modelo DenseNet201, se obtuvo una predictibilidad del 100% para la clase sana, 98% para la clase tizón tardío y 99% para la clase tizón temprano. Se obtuvo una sensibilidad del 97.5% para la clase sana, 100% para la clase tizón tardío y 100% para la clase tizón temprano. La Figura 24 muestra la matriz de confusión creada a partir de evaluar el modelo DenseNet201 sobre el conjunto de prueba.

		Etiqueta Predicha		
		Tizón Temprano	Sana	Tizón Tardío
Etiqueta Real	Tizón Temprano	99	1	0
	Sana	0	120	0
	Tizón Tardío	0	2	98

Figura 24. Matriz de confusión creada a partir de la evaluación del modelo de la red DenseNet201 sobre el conjunto de prueba PlantVillage original.

6.4.2 Evaluación sobre el conjunto PlantVillage segmentado

Los modelos con mejores resultados de precisión de validación entrenados con el conjunto PlantVillage segmentado, se evaluaron sobre el conjunto de prueba. El modelo entrenado con la red VGG16 obtuvo un 92.2% de precisión sobre el conjunto de prueba y el modelo entrenado con la red DenseNet201 obtuvo un 87.2% de precisión sobre el conjunto de prueba.

Para el modelo VGG16, se obtuvo una predictibilidad del 97.5% para la clase sana, 84% para la clase tizón tardío y 94% para la clase tizón temprano. Se obtuvo una sensibilidad del 88.6% para la clase sana, 92.3% para la clase tizón tardío y 96.9% para la clase tizón temprano. La Figura 25 muestra la matriz de confusión creada a partir de evaluar el modelo VGG16 sobre el conjunto de prueba.

		Etiqueta Predicha		
		Tizón Temprano	Sana	Tizón Tardío
Etiqueta Real	Tizón Temprano	94	2	4
	Sana	0	117	3
	Tizón Tardío	3	13	84

Figura 25. Matriz de confusión creada a partir de la evaluación del modelo de la red VGG16 sobre el conjunto de prueba PlantVillage segmentado.

Para el modelo DenseNet201, se obtuvo una predictibilidad del 93.3% para la clase sana, 67% para la clase tizón tardío y 100% para la clase tizón temprano. Se obtuvo una sensibilidad del 94.9% para la clase sana, 74% para la clase tizón tardío y 100% para la clase tizón temprano. La Figura 26 muestra la matriz de confusión creada a partir de evaluar el modelo DenseNet201 sobre el conjunto de prueba.

		Etiqueta Predicha		
		Tizón Temprano	Sana	Tizón Tardío
Etiqueta Real	Tizón Temprano	100	0	0
	Sana	8	112	0
	Tizón Tardío	27	6	67

Figura 26. Matriz de confusión creada a partir de la evaluación del modelo de la red DenseNet201 sobre el conjunto de prueba PlantVillage segmentado.

6.4.3 Evaluación sobre el conjunto Fruit360 más web scraping

Los modelos con mejores resultados de precisión de validación entrenados con el conjunto Fruit360 más web scraping, se evaluaron sobre el conjunto de prueba. El modelo entrenado con la red VGG16 obtuvo un 91.1% de precisión sobre el conjunto de prueba y el modelo entrenado con la red DenseNet201 obtuvo un 93.9% de precisión sobre el conjunto de prueba.

Para el modelo VGG16, se obtuvo una predictibilidad del 100% para la clase sana, 64% para la clase tizón tardío y 91% para la clase sarna común. Se obtuvo una sensibilidad del 98% para la clase sana, 90.1% para la clase tizón tardío y 85% para la clase sarna común. La Figura 27 muestra la matriz de confusión creada a partir de evaluar el modelo VGG16 sobre el conjunto de prueba.

		Etiqueta Predicha		
		Sarna Común	Sana	Tizón Tardío
Etiqueta Real	Sarna Común	91	2	7
	Sana	0	100	0
	Tizón Tardío	16	0	64

Figura 27. Matriz de confusión creada a partir de la evaluación del modelo de la red VGG16 sobre el conjunto de prueba Fruit360 más web scraping.

Para el modelo DenseNet201, se obtuvo una predictibilidad del 100% en la clase sana, 67% en la clase tizón tardío y 96% en la clase sarna común. se obtuvo una sensibilidad del 100% en la clase sana, 94.3% en la clase tizón tardío y 89.7% en la clase sarna común. La Figura 28 muestra la matriz de confusión creada a partir de evaluar el modelo DenseNet201 sobre el conjunto de prueba.

		Etiqueta Predicha		
		Sarna Común	Sana	Tizón Tardío
Etiqueta Real	Sarna Común	96	0	4
	Sana	0	100	0
	Tizón Tardío	13	0	67

Figura 28. Matriz de confusión creada a partir de la evaluación del modelo de la red DenseNet201 sobre el conjunto de prueba Fruit360 más web scraping.

6.5 Mapas de características red VGG16

6.5.1 Mapa de características modelo clasificador de enfermedades en hojas de papa

Las capas convolucionales de la red VGG16 poseen 64 filtros cada una, aplicando los filtros correspondientes a las salidas de cada uno de los 5 bloques convolucionales de la red a una imagen de entrada, se obtuvieron 320 mapas de características por una imagen de cada categoría. La Figura 29 muestra un ejemplo de los mapas de características de las salidas de los bloques convolucionales 1, 3 y 5 de la red aplicados sobre imágenes de las 3 categorías disponibles en el conjunto de datos. En los anexos se encuentran disponibles los 320 mapas de características generados para esta combinación de modelo e imágenes del conjunto.

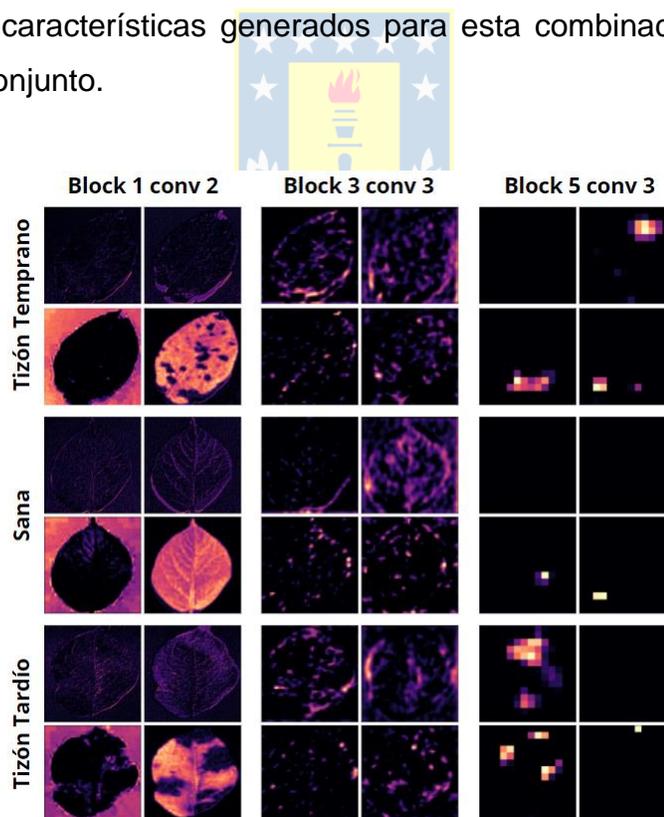


Figura 29. Ejemplo de los mapas de características de la red VGG16 sobre imágenes de las 3 categorías del conjunto PlantVillage original.

Se puede observar en los mapas de características las cualidades que el modelo detecta para imagen, permitiendo ver la hoja y las zonas con signos distintivos dentro de ella. Los filtros de la red mantienen la relación espacial, esto permite superponer los mapas de características con la imagen de entrada original. La Figura 30 muestra un ejemplo de aplicar los filtros de los primeros bloques convolucionales de la red VGG16 y superponerla con la imagen de entrada original.

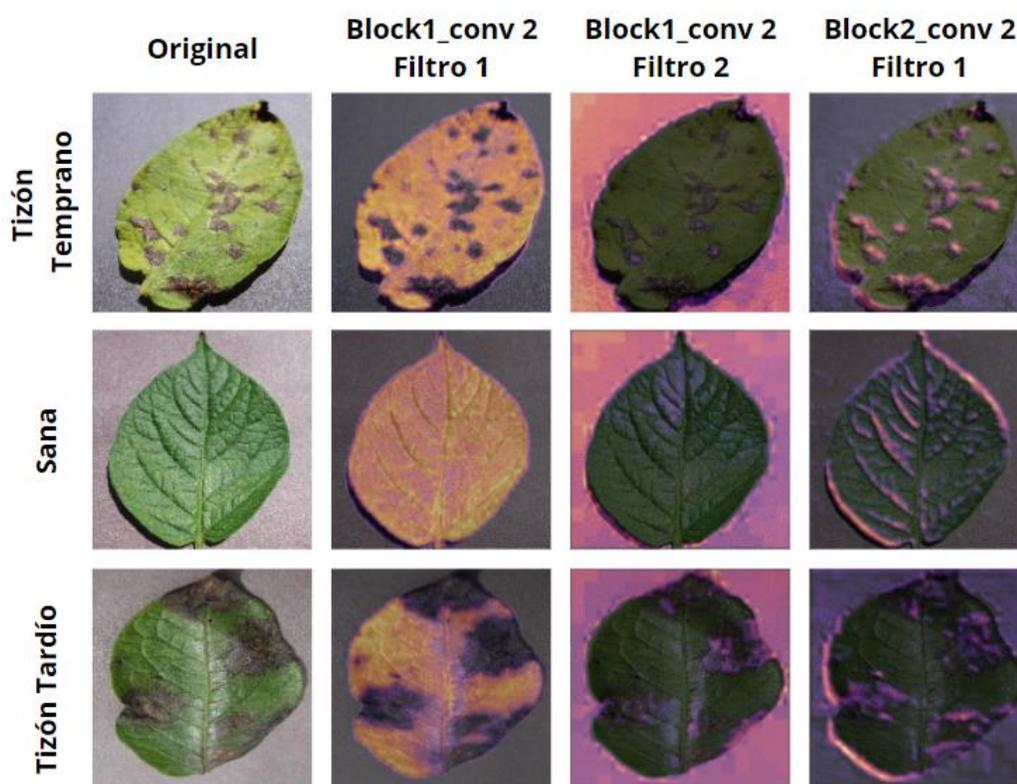


Figura 30. Superposición de los mapas de características de la red VGG16 con la imagen de entrada original de cada una de las 3 categorías del conjunto PlantVillage original.

El superponer la imagen de entrada al modelo con los mapas de características respectivos, permite visualizar que parte de la imagen guió la decisión tomada por el modelo. En la Figura 30 se puede observar como en las

clases que presentan alguna patología la red logra diferenciar estas zonas del resto de la hoja.

6.5.2 Mapa de características modelo clasificador de enfermedades en tubérculos

Las capas convolucionales de la red VGG16 poseen 64 filtros cada una, aplicando los filtros correspondientes a las salidas de cada uno de los 5 bloques convolucionales de la red a una imagen de entrada, se obtuvieron 320 mapas de características por una imagen de cada categoría. La Figura 31 muestra un ejemplo de los mapas de características de las salidas de los bloques convolucionales 1, 3 y 5 de la red. Aplicados sobre imágenes de las 3 categorías disponibles en el conjunto de datos. En los anexos se encuentran disponibles los 320 mapas de características generados para esta combinación de modelo e imágenes del conjunto.

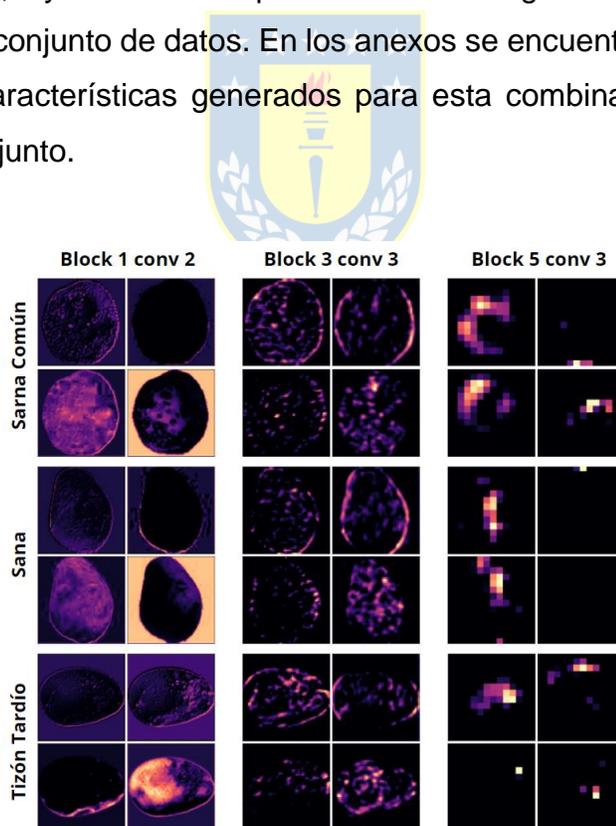


Figura 31. Ejemplo de los mapas de características de la red VGG16 sobre imágenes de las 3 categorías del conjunto Fruit360 más web scraping.

Se puede observar en los mapas de características las cualidades que el modelo detecta para imagen, permitiendo ver el tubérculo y las zonas con signos distintivos dentro de el. Los filtros de la red mantienen la relación espacial, esto permite superponer los mapas de características con la imagen de entrada original. La Figura 32. muestra un ejemplo de aplicar los filtros de los primeros bloques convolucionales de la red VGG16 y superponerla con la imagen de entrada original.

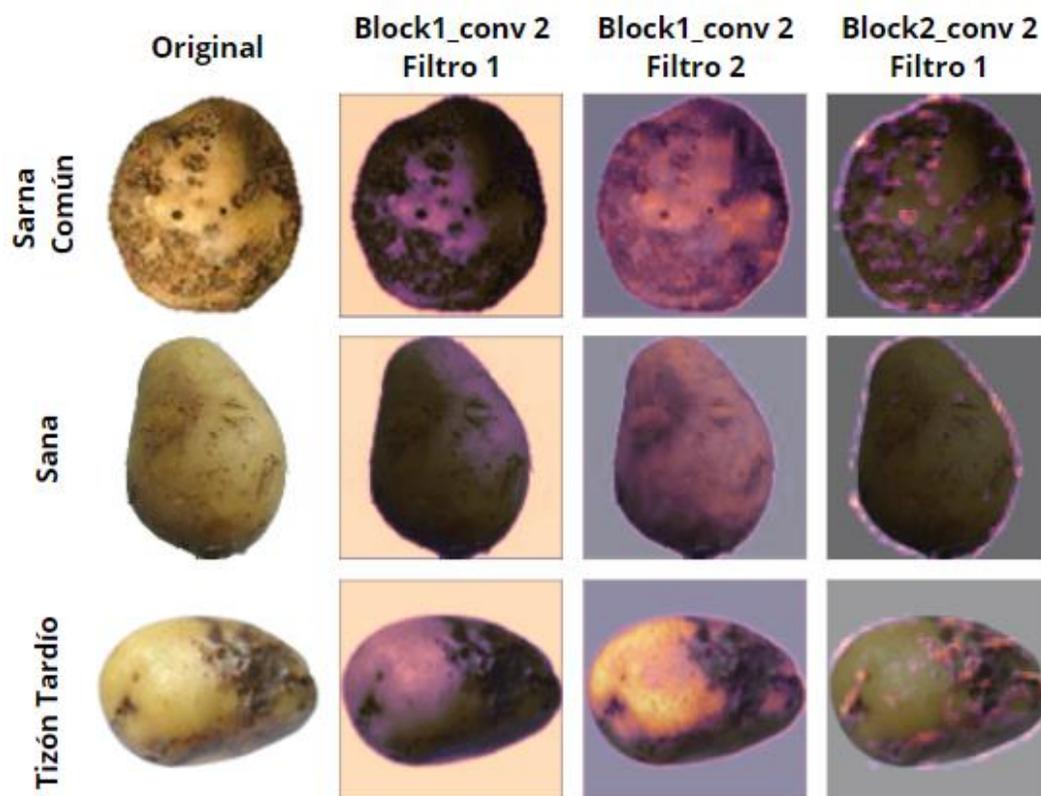


Figura 32. Superposición de los mapas de características de la red VGG16 con la imagen de entrada original de cada una de las 3 categorías del conjunto Fruit360 más web scraping.

El superponer la imagen de entrada al modelo con los mapas de características respectivos, permite visualizar que parte de la imagen guió la decisión tomada por el modelo. En la Figura 32 se puede observar como en las

clases que presentan alguna patología la red logra diferenciar estas zonas del resto del tubérculo.



Capítulo 7

7 Discusiones

Las arquitecturas utilizadas en este estudio presentan una buena adaptación de aprendizaje para clasificar las nuevas categorías de los conjuntos PlantVillage original, PlantVillage segmentado y Fruit360 más web scraping. Todas las configuraciones de redes y conjuntos de datos alcanzaron una precisión de validación del 90% a más tardar en la época 5 de entrenamiento.

El mejor modelo clasificador de enfermedades en hojas de papa fue el entrenado con la red DenseNet201 y el conjunto de datos PlantVillage original, con un 97.5% de precisión de validación y un 99.1% al evaluarse sobre el conjunto de prueba. Con precisiones individuales del 100% para la clase sana, 98% para la clase tizón tardío y 99% para la clase tizón temprano, además, de una sensibilidad del 97.5% para la clase sana, 100% para la clase tizón tardío y 100% para la clase tizón temprano. De las métricas obtenidas al evaluar el modelo sobre el conjunto de prueba se tiene que para todas las clases se alcanzan valores sobre el 98% tanto en precisión como sensibilidad, mostrando una gran adaptación para diagnosticar de buena manera todas las clases por igual, entregando así resultados confiables y permitiendo atribuir la gran precisión general del modelo a cada una de las clases.

A pesar de que el modelo entrenado sobre la misma red con el conjunto PlantVillage segmentado alcanzó una precisión de validación mayor del 99.2%, su desempeño baja un poco al evaluarse sobre el conjunto de prueba con un 87.2% de precisión. Se obtuvieron precisiones individuales de clase del 93.3%

para la clase sana, 67% para la clase tizón tardío y 100% para la clase tizón temprano. La principal hipótesis es que esto se podría deber a que la segmentación del fondo en el conjunto de datos PlantVillage se realizó en color negro, el cual coincide con las principales zonas de la hoja donde se presenta esta enfermedad.

El mejor modelo clasificador de enfermedades en tubérculos fue el entrenado con la red VGG16 y el conjunto de datos Fruit360 más web scraping, con un 96.8% de precisión de validación y un 91.1% al evaluarse sobre el conjunto de prueba. Con precisiones individuales del 100% para la clase sana, 64% para la clase tizón tardío y 91% para la clase sarna común, además, se obtuvo una sensibilidad del 98% para la clase sana, 90.1% para la clase tizón tardío y 85% para la clase sarna común. De las métricas obtenidas a partir de evaluar el modelo con el conjunto de prueba, se tiene que la sensibilidad de la clase sarna común es la única que baja del 90%, lo cual se podría atribuir principalmente a la relación 5:4 existente respecto del número de datos de la clase tizón tardío. Esto lleva al modelo a privilegiar el etiquetar registros con esta clase, lo cual aumenta el desempeño general del modelo, en desmedro de la clase menos representada de tizón tardío. Esto podría explicar la menor precisión individual de la clase tizón tardío respecto de las otras clases. Sin embargo, esto no genera una gran problemática para la tarea de clasificación de enfermedades de papa, ya que la clase sana presenta una sensibilidad y precisión superiores al 98%, por lo que la tasa de “falsos negativos” (tubérculos enfermos etiquetados como sanos) es casi nula. Además, al presentar la clase sarna común la sensibilidad más baja no se generaría un problema serio, ya que el etiquetar un tubérculo con sarna como tizón, no tendría la gravedad de etiquetar un tubérculo con tizón como sarna, lo cual en este específico problema se podría considerar como un falso positivo debido a la criticidad que tiene la enfermedad de tizón.

Las visualizaciones de los mapas de características de sobre las imágenes muestran las características que extraen los filtros de capas convolucionales a lo largo de la red. En la Figura 29 y Figura 31 se puede observar como los filtros de las primeras capas capturan las líneas, formas y fondos, y se puede percibir a partir de los filtros que la imagen corresponde a una hoja o un tubérculo. A medida que los filtros pertenecen a capa más profundas, las visualizaciones se tornan abstractas y difícil de interpretar.

Los mapas de características no entregan una medición cuantitativa, pero sin resultan una representación visual que permite sacar conclusiones y entender mejor el funcionamiento de la red. Al superponer los filtros con la imagen de entrada original, podemos darnos una idea general de las características que el modelo va considerando relevantes para cada categoría.

En la Figura 30 se puede observar como los filtros de la capa convolucional de salida del primer y segundo bloque de la red VGG16, claramente activan y resaltan las características de la clase tizón temprano para la que tiene un 100% de precisión en el conjunto de prueba.

En la Figura 32 se puede observar como los filtros de la capa convolucional de salida del segundo bloque de la red VGG16, claramente activa y resalta características de las clases enfermas respecto de la clase sana.

A pesar de que solo se trata de una interpretación y que finalmente las capas de pooling de la red, toman la decisión de cuál de todos los filtros son los predominantes y la capa completamente conectada realizará la clasificación final, los mapas de activación constituyen una herramienta útil para comprender aquellas características que resultan relevantes para la clasificación, en

particular, nos permiten entender el por qué los modelos confunden clases o tomar decisiones respecto a si es necesario agregar variedad al conjunto de datos.



Capítulo 8

8 Conclusiones y extensiones

En este estudio, se abordó el problema de la detección de enfermedades en cultivos de papa. Para resolver este problema de clasificación se propuso un enfoque de aprendizaje profundo, mediante la utilización de modelos de redes neuronales convolucionales, utilizando la técnica de transferencia de aprendizaje.

Los resultados muestran una gran adaptación por parte de las arquitecturas para clasificar las nuevas categorías de los conjuntos de datos utilizados. Los mejores modelos obtenidos con la red VGG16, alcanzaron un 95.7%, 95.7% y 96.9% de precisión de validación, entrenándose sobre los conjuntos de datos PlantVillage original, PlantVillage segmentado y Fruit360 más web scraping, respectivamente. Y los mejores modelos obtenidos con la red DenseNet201, alcanzaron un 98.4%, 99.2% y 89.4% de precisión de validación, entrenándose sobre los conjuntos de datos PlantVillage original, PlantVillage segmentado y Fruit360 más web scraping, respectivamente.

Los modelos entrenados a partir de la red VGG16 con mejores resultados evaluados sobre los conjuntos de prueba, obtuvieron un 94.7%, 92.2% y un 91.1% de precisión de prueba, sobre los conjuntos PlantVillage original, PlantVillage segmentado y Fruit360 más web scraping, respectivamente. Los mejores modelos entrenados con la red DenseNet201, obtuvieron un 99.1%, 87.2% y un 93.9% de precisión de prueba, sobre los conjuntos PlantVillage original, PlantVillage segmentado y Fruit360 más web scraping, respectivamente.

En conclusión, los modelos implementados a partir de redes neuronales convolucionales, adaptados para clasificar enfermedades en cultivos de papa, presentan grandes resultados con precisiones superiores a 90% en todas las configuraciones experimentales. Se igualan los resultados reportados por la literatura en la aplicación de estas técnicas a otras categorías. Y los mapas de características muestran que el modelo logra detectar correctamente en la imagen las zonas donde existe una patología. Por lo que se refuerza y demuestra, la factibilidad de la utilización de modelos de aprendizaje profundo como una alternativa viable para la solución de este problema.

Una alternativa de mejora para obtener un modelo más calibrado y fino, al momento de diagnosticar tubérculos, sería calibrar el modelo mediante un ajuste de la función de pérdida mediante la inclusión de ponderadores. El asignar pesos a cada clase en base a la relación del número de datos que poseen los conjuntos de entrenamiento y validación, permitiría eliminar la predilección del modelo por clasificar correctamente las clases más numerosas en desmedro de las clases con menor cantidad de datos.

Para estudios futuros se busca incorporar nuevas categorías, para disponer de un mayor número de patologías a identificar. Además, de crear conjuntos de datos utilizando imágenes tomadas a partir de tubérculos reales con supervisión de un ente experto. Finalmente, incorporar este modelo en una aplicación móvil que permita democratizar el uso de las herramientas de inteligencia artificial a todos los niveles y ayude a los agricultores pequeños a realizar una detección temprana de posibles enfermedades que afecten sus cultivos.

Se espera la aplicación móvil que permita utilizar los modelos resultantes de este estudio de manera práctica, se encuentre disponible lo más pronto posible una vez publicado el trabajo. Para poder generar un aporte y ayudar a promover

esta línea de investigación, en el tiempo de espera a que la aplicación sea liberada de forma web o móvil, se pondrá a disposición de la comunidad un repositorio de github de libre acceso, donde se podrá acceder a los modelos, imágenes de prueba e información respecto al estudio y sus futuros proyectos. (Enlace: https://github.com/Fabfuentes16/CNN_Potato_Crop_Diseases)



9 Bibliografía

[1] Kosacka-Olejnik, M., & Pitakaso, R. (2019). Industry 4.0: state of the art and research implications. *LogForum*, 15(4).

[2] Zambon, I., Cecchini, M., Egidi, G., Saporito, M. G., & Colantoni, A. (2019). Revolution 4.0: Industry vs. agriculture in a future development for SMEs. *Processes*, 7(1), 36.

[3] Castro Urrutia, I., & Contreras Méndez, A. (2016). Manejo de plagas y enfermedades en el cultivo de la papa.

[4] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.

[5] Ehler, L. E. (2006). Integrated pest management (IPM): definition, historical development and implementation, and the other IPM. *Pest management science*, 62(9), 787-789.

[6] Número de usuarios de smartphones a nivel mundial desde 2016 hasta 2021. (2020, 16 septiembre). Rosa Fernandez. <https://es.statista.com/estadisticas/636569/usuarios-de-telefonos-inteligentes-a-nivel-mundial>.

[7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

- [8] Rangarajan, A. K., Purushothaman, R., & Ramesh, A. (2018). Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia computer science*, 133, 1040-1047.
- [9] Elsharif, A. A., Dheir, I. M., Mettleq, A. S. A., & Abu-Naser, S. S. (2019). Potato Classification Using Deep Learning. *International Journal of Academic Pedagogical Research*, 1-8.
- [10] da Costa, A. Z., Figueroa, H. E., & Fracarolli, J. A. (2020). Computer vision based detection of external defects on tomatoes using deep learning. *Biosystems Engineering*, 190, 131-144.
- [11] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, 272-279.
- [12] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [13] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [14] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.

- [15] Hughes, D., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv preprint arXiv:1511.08060.
- [16] Mureşan, H., & Oltean, M. (2017). Fruit recognition from images using deep learning. arXiv preprint arXiv:1712.00580.
- [17] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2921-2929).
- [18] How to Visualize Filters and Feature Maps in Convolutional Neural Networks. (2019, 5 julio). Jason Brownlee. <https://machinelearningmastery.com/how-to-visualize-filters-and-feature-maps-in-convolutional-neural-networks>.
- [19] Irish Potato Famine. (2019, 7 junio). HISTORY. <https://www.history.com/topics/immigration/irish-potato-famine>.
- [20] da Silva, V. L., Kovaleski, J. L., & Pagani, R. N. (2019). Technology transfer in the supply chain oriented to industry 4.0: a literature review. *Technology Analysis & Strategic Management*, 31(5), 546-562.
- [21] Datos y cifras de la papa. CIP. <https://cipotato.org/es/potato>.
- [22] Papas. Odepa. <https://www.odepa.gob.cl/rubros/papas-y-tuberculos>.
- [23] Feng, X., Jiang, Y., Yang, X., Du, M., & Li, X. (2019). Computer vision algorithms and hardware implementations: A survey. *Integration*, 69, 309-320.

[24] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255).

[25] A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. (2018, 15 diciembre). Sumit Saha. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.

[26] Convolutional Neural Network. (2019, 24 febrero). dshahid380. <https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529>.

[27] A Gentle Introduction to Transfer Learning for Deep Learning. (2019, 16 octubre). Jason Brownlee. <https://machinelearningmastery.com/transfer-learning-for-deep-learning>.

[28] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[29] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2014). Object detectors emerge in deep scene cnns. arXiv preprint arXiv:1412.6856.

10 Anexos

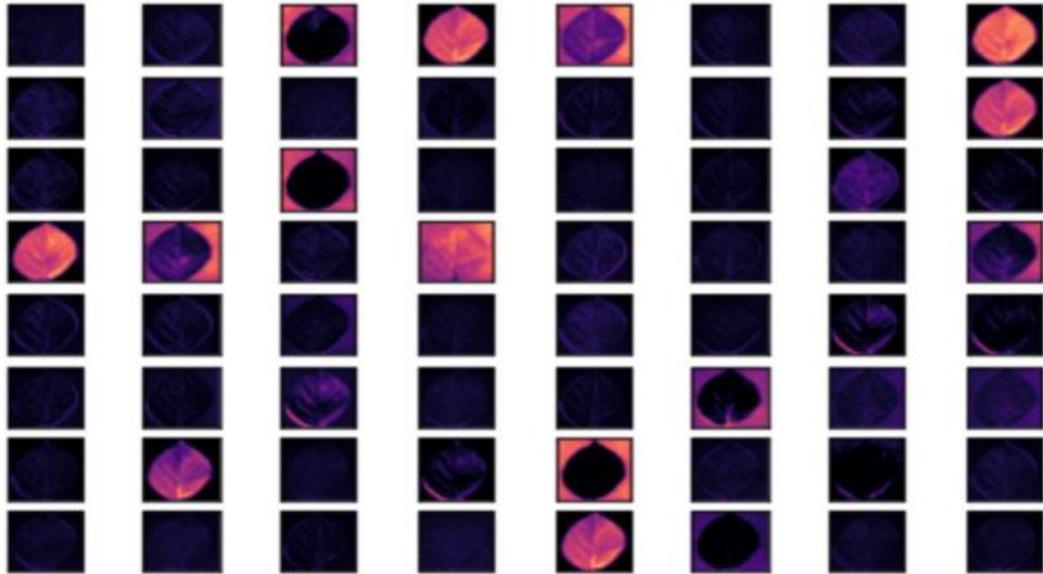


Figura 33. Mapas de características imagen categoría hoja sana bloque convolucional 1.

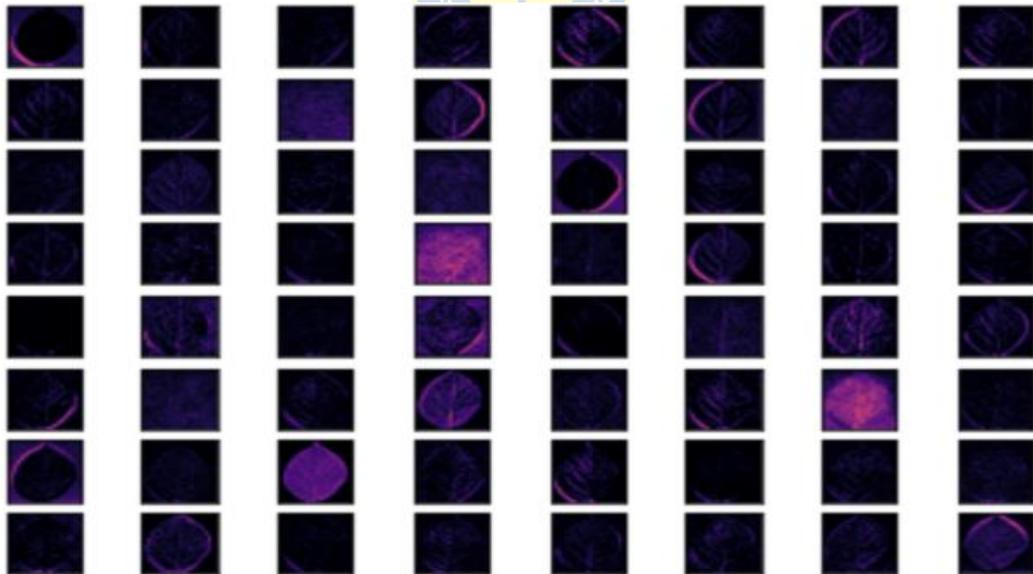


Figura 34. Mapas de características imagen categoría hoja sana bloque convolucional 2.

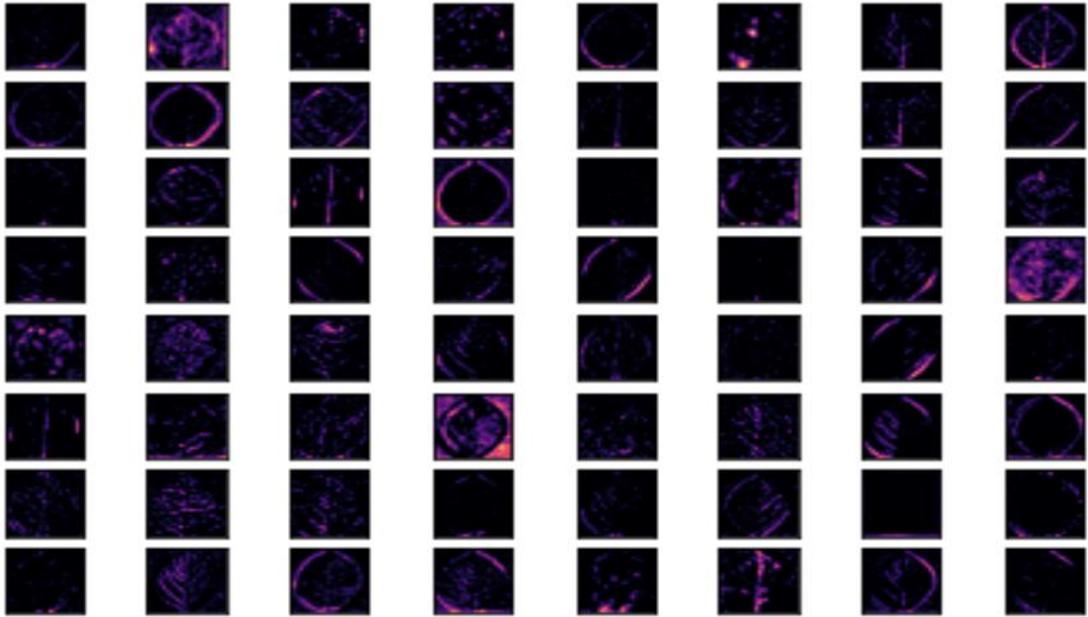


Figura 35. Mapas de características imagen categoría hoja sana bloque convolucional 3.

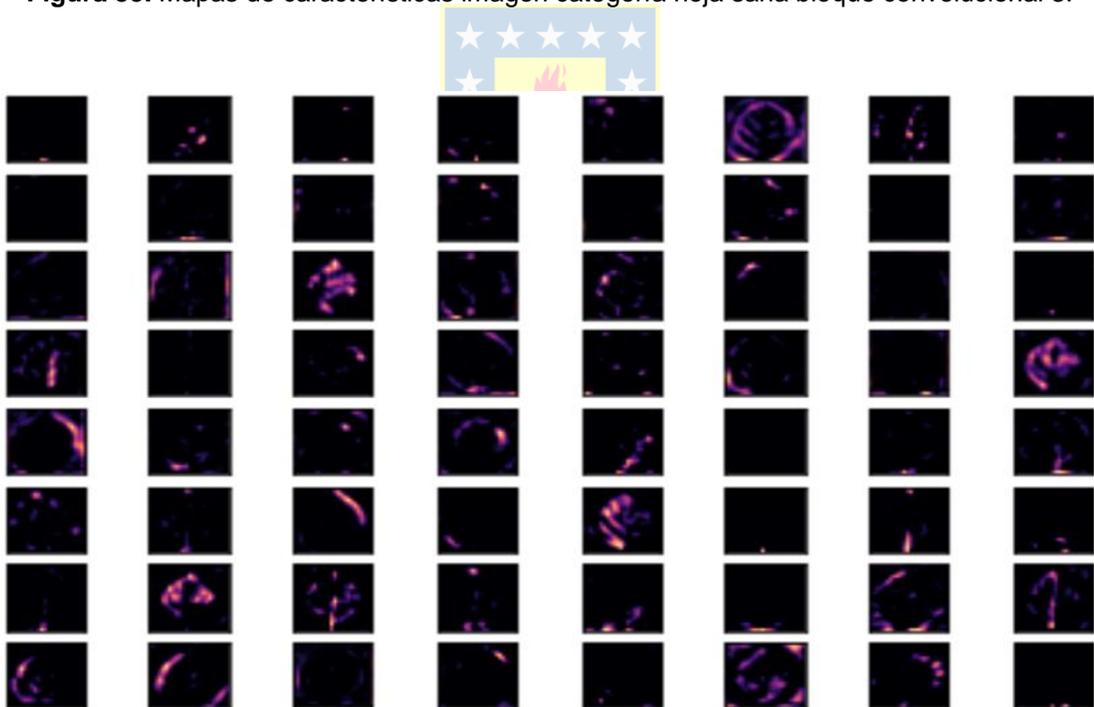


Figura 36. Mapas de características imagen categoría hoja sana bloque convolucional 4.

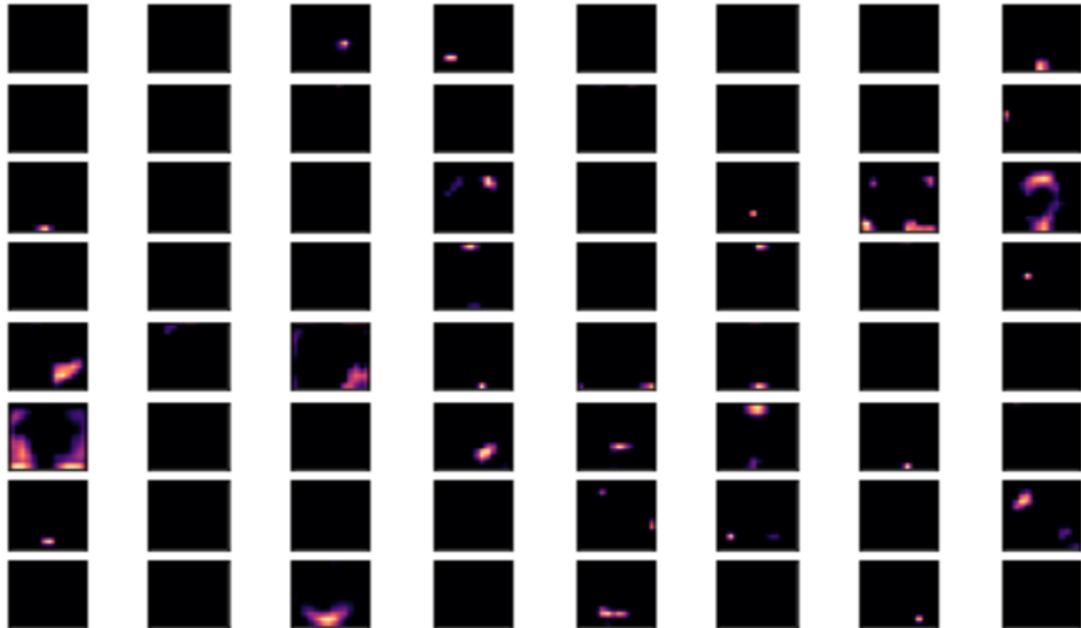


Figura 37. Mapas de características imagen categoría hoja sana bloque convolucional 5.

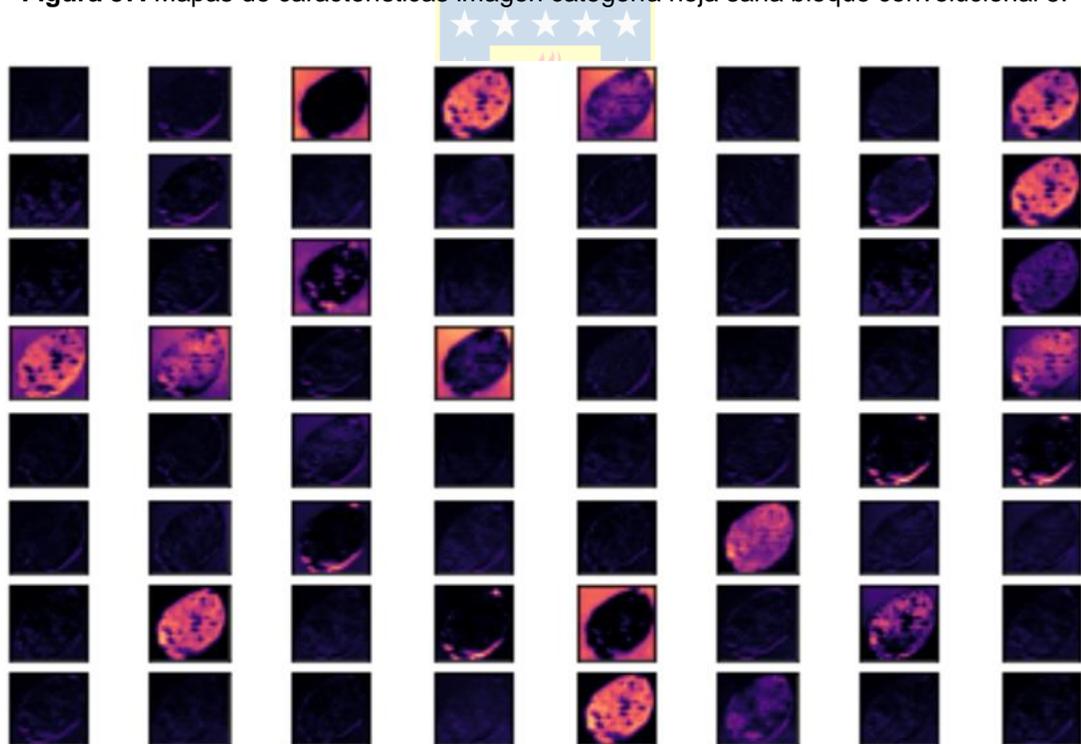


Figura 38. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 1.

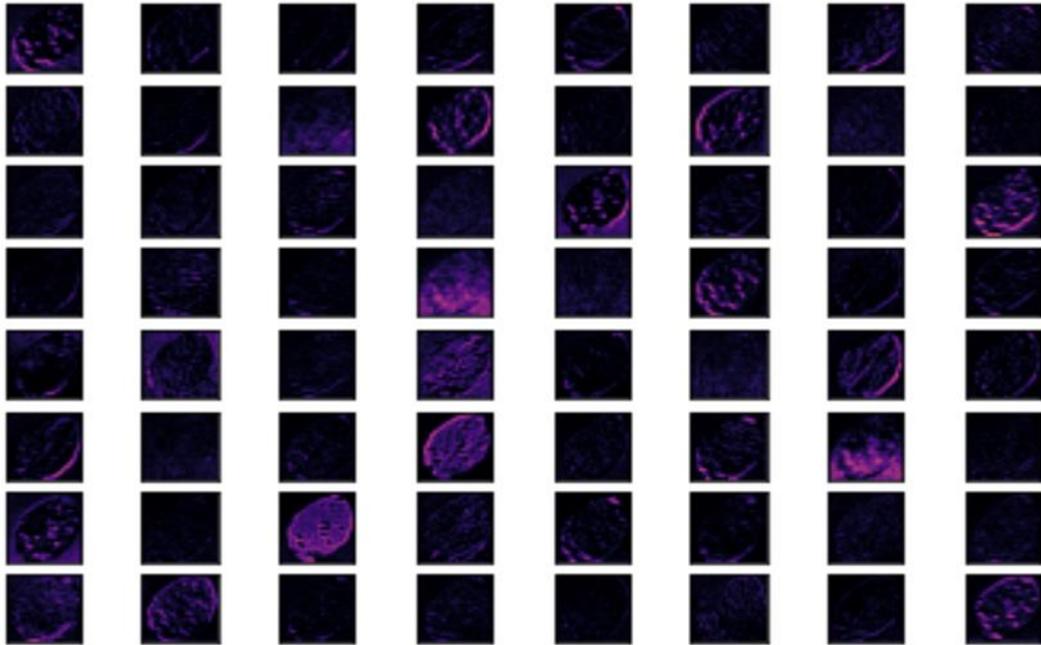


Figura 39. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 2.

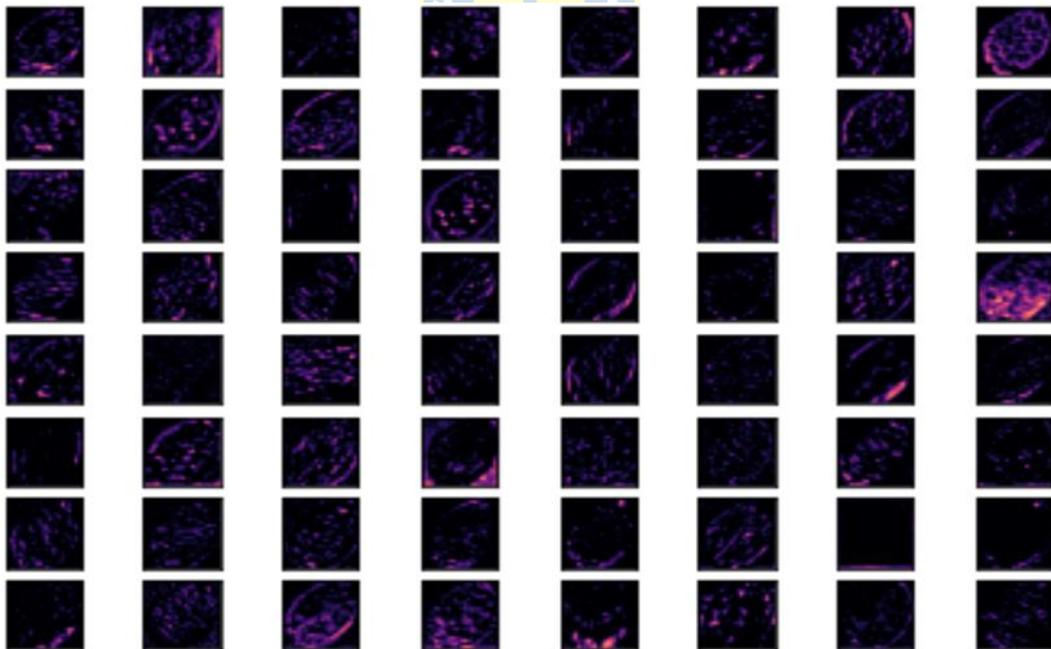


Figura 40. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 3.

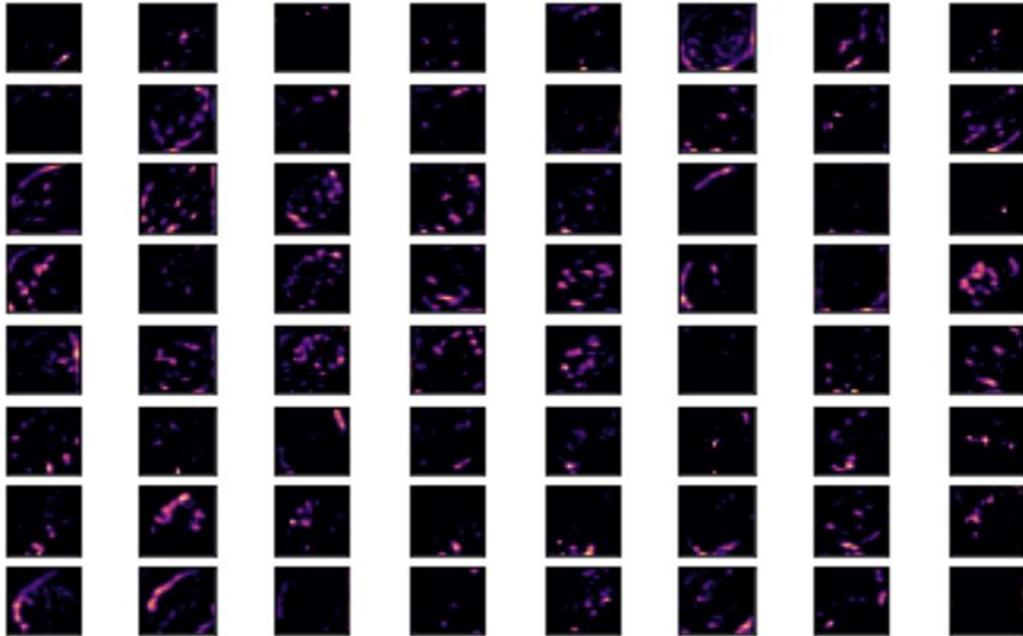


Figura 41. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 4.

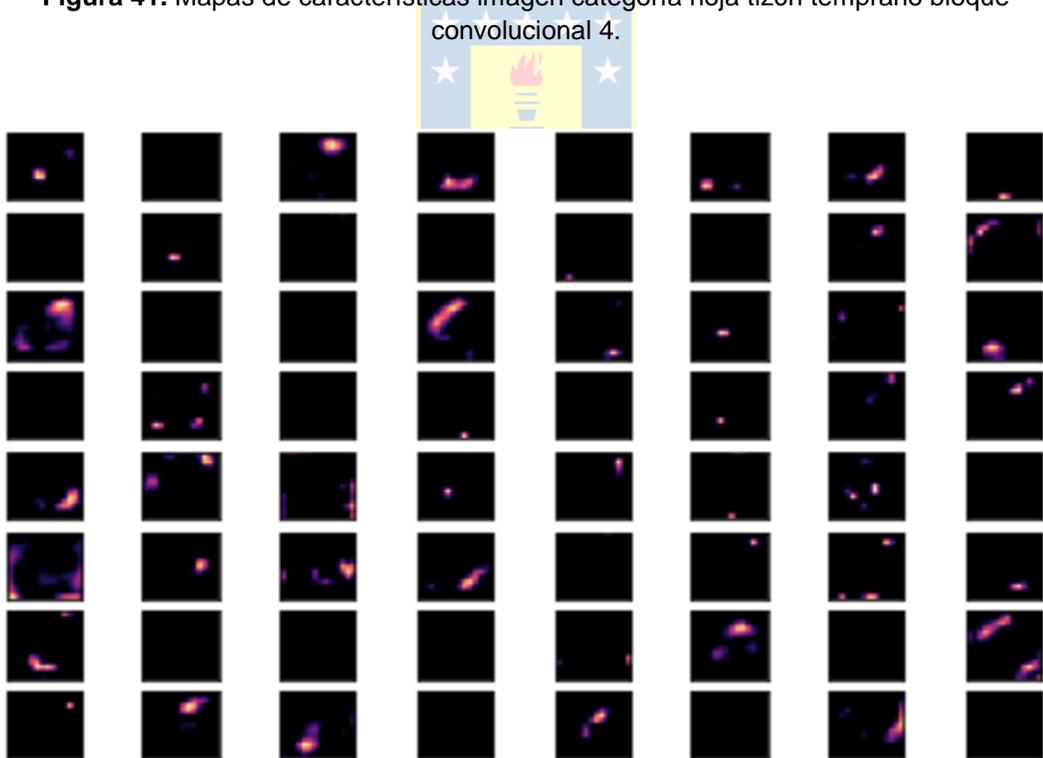


Figura 42. Mapas de características imagen categoría hoja tizón temprano bloque convolucional 5.

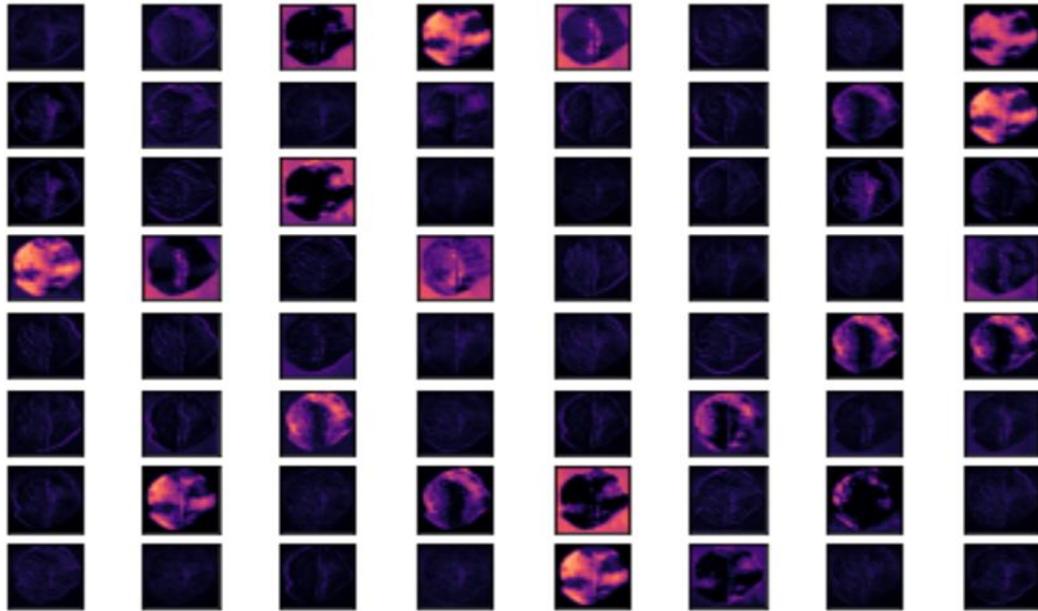


Figura 43. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 1.

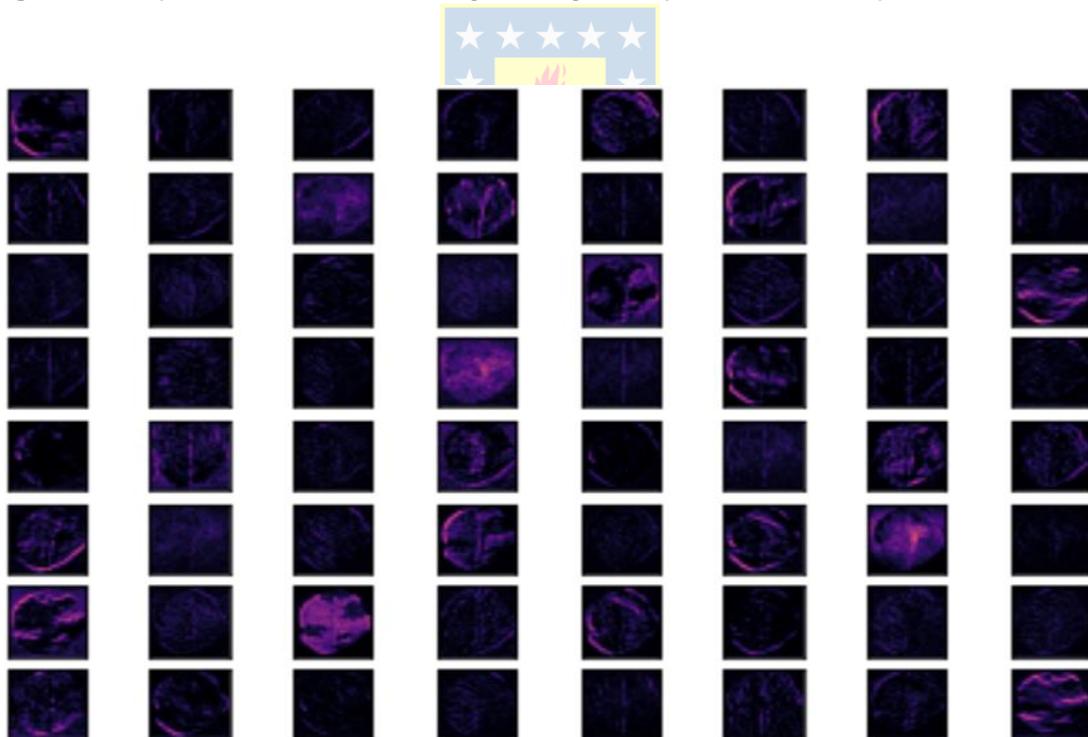


Figura 44. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 2.

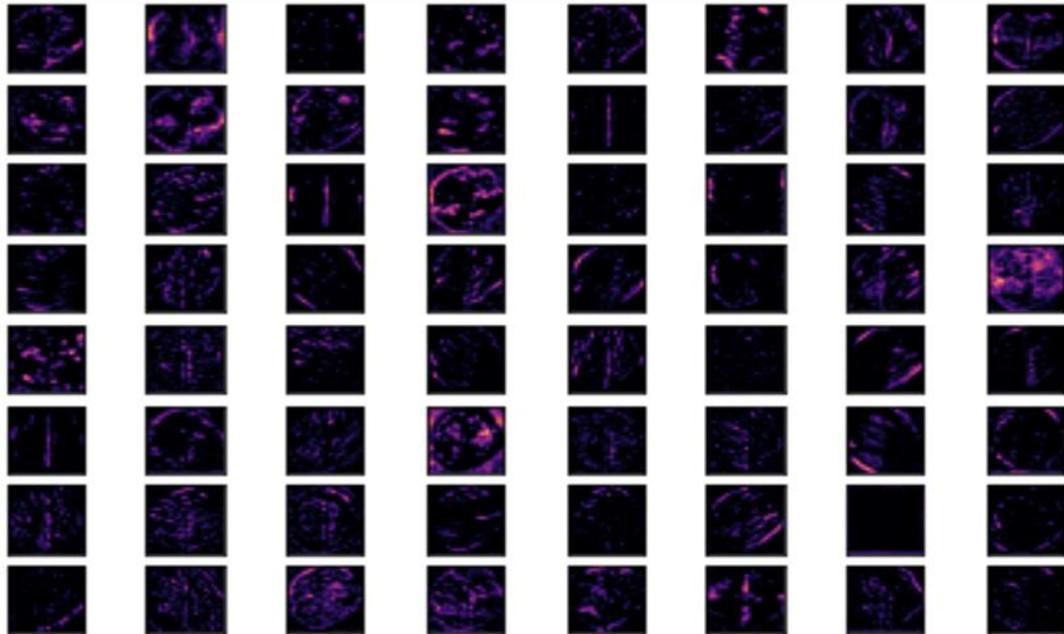


Figura 45. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 3.

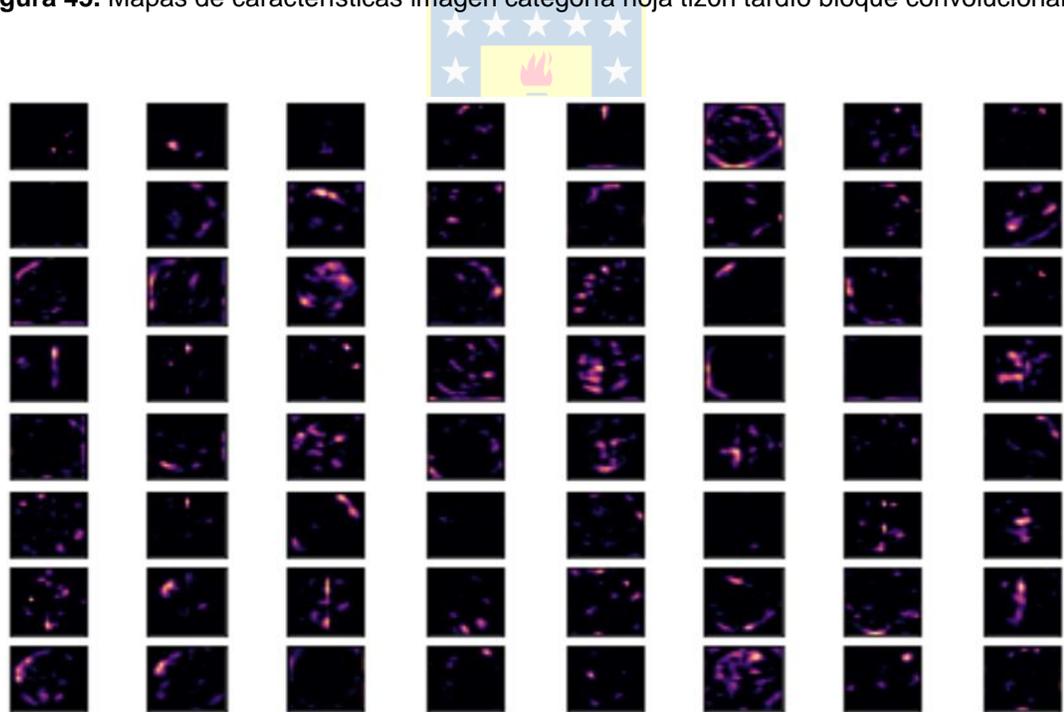


Figura 46. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 4.

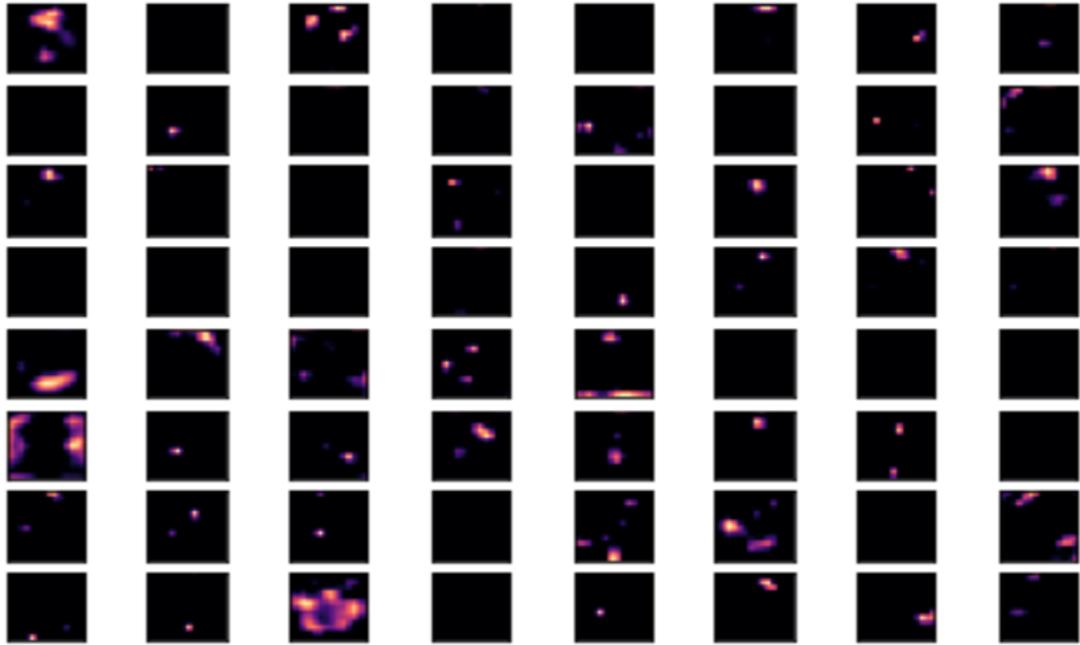


Figura 47. Mapas de características imagen categoría hoja tizón tardío bloque convolucional 5.

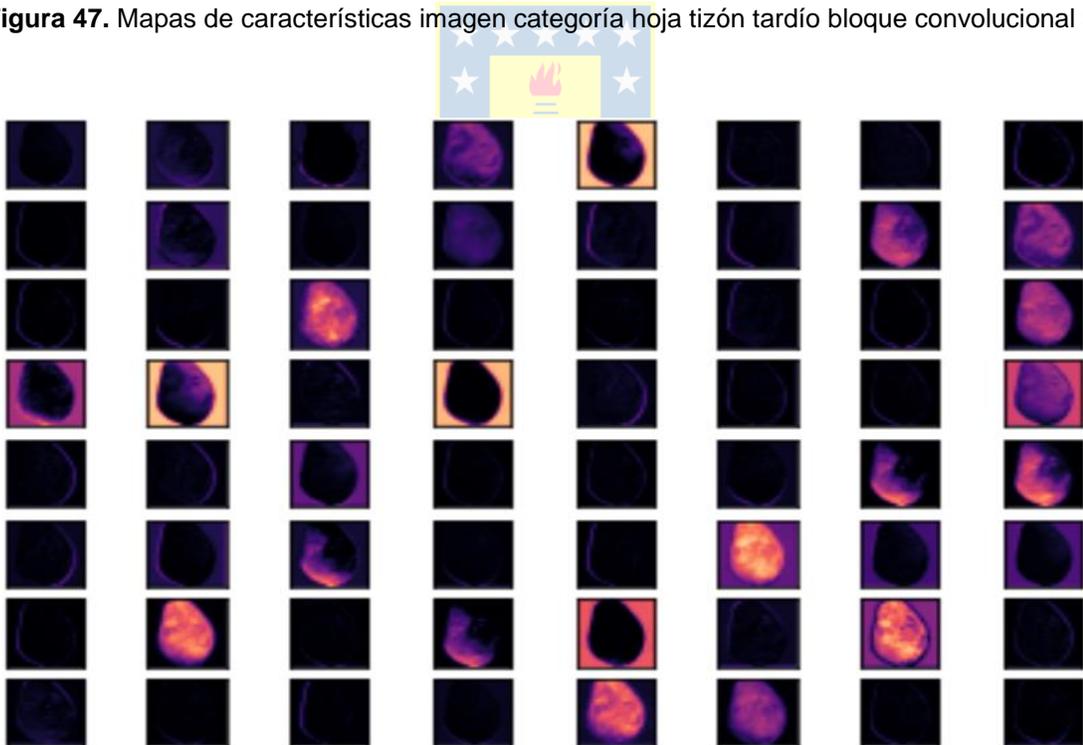


Figura 48. Mapas de características imagen categoría tubérculo sano bloque convolucional 1.

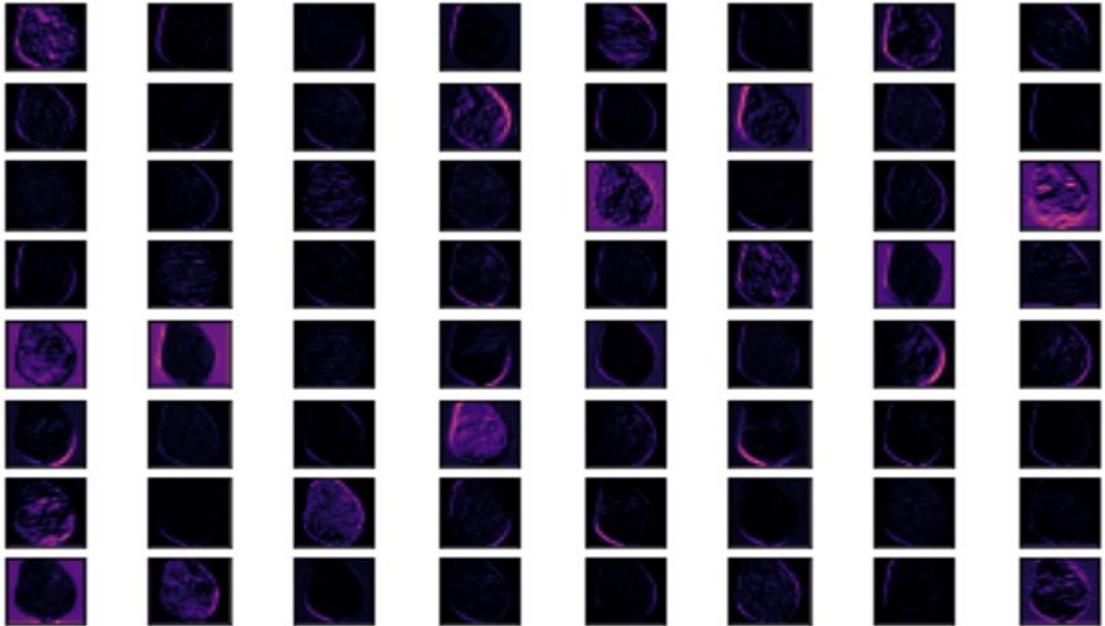


Figura 49. Mapas de características imagen categoría tubérculo sano bloque convolucional 2.

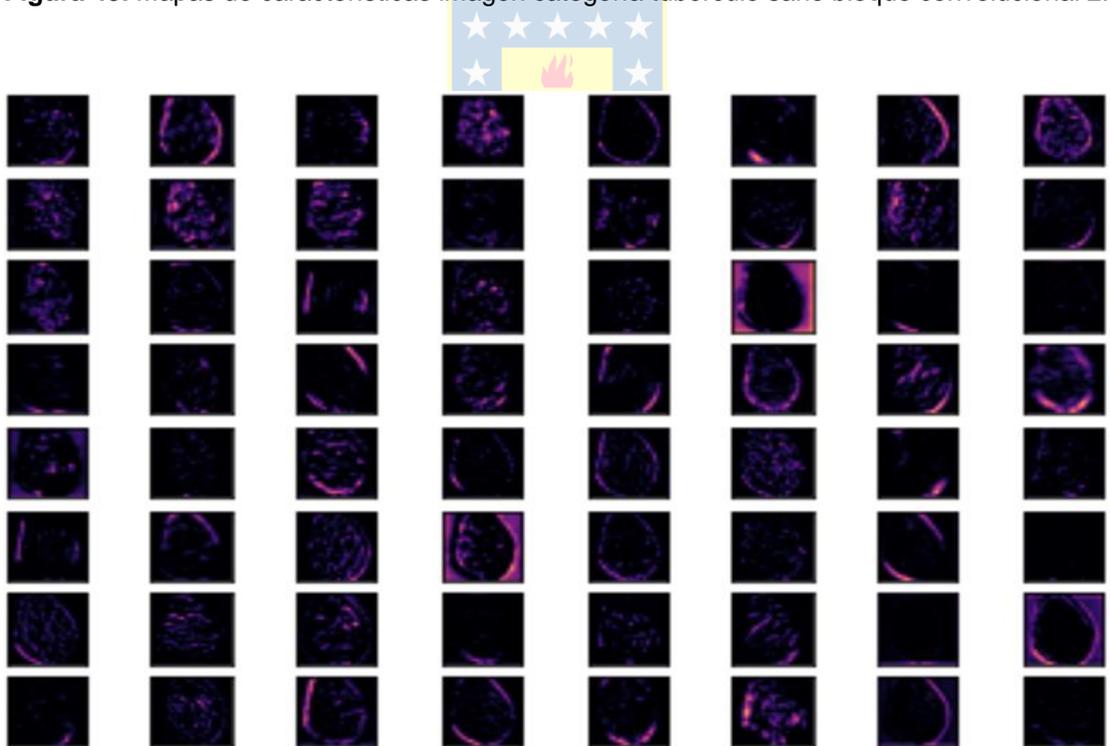


Figura 50. Mapas de características imagen categoría tubérculo sano bloque convolucional 3.

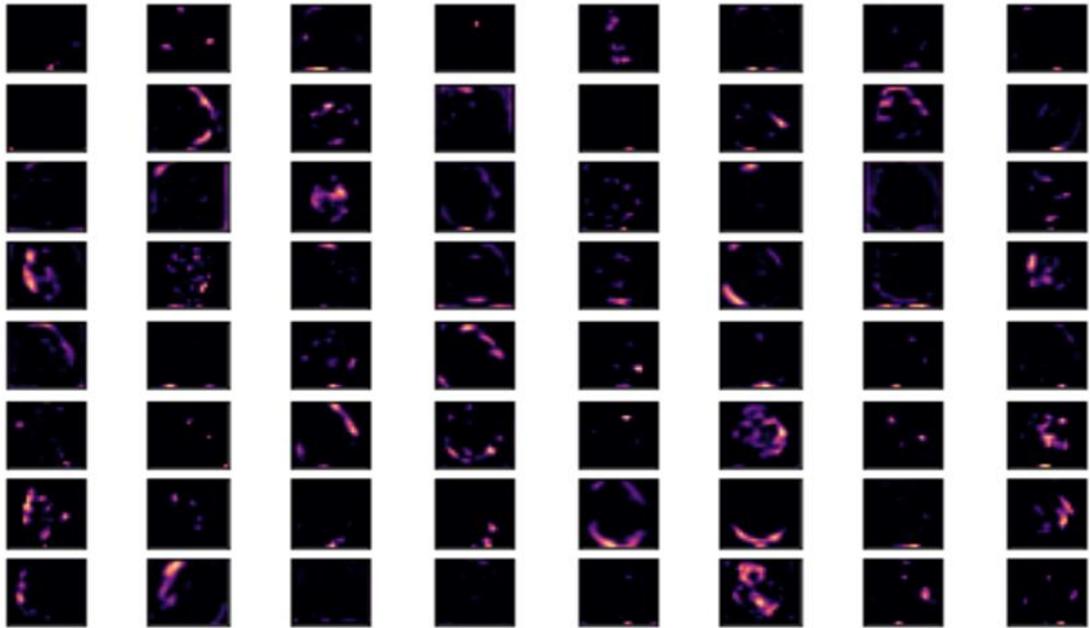


Figura 51. Mapas de características imagen categoría tubérculo sano bloque convolucional 4.

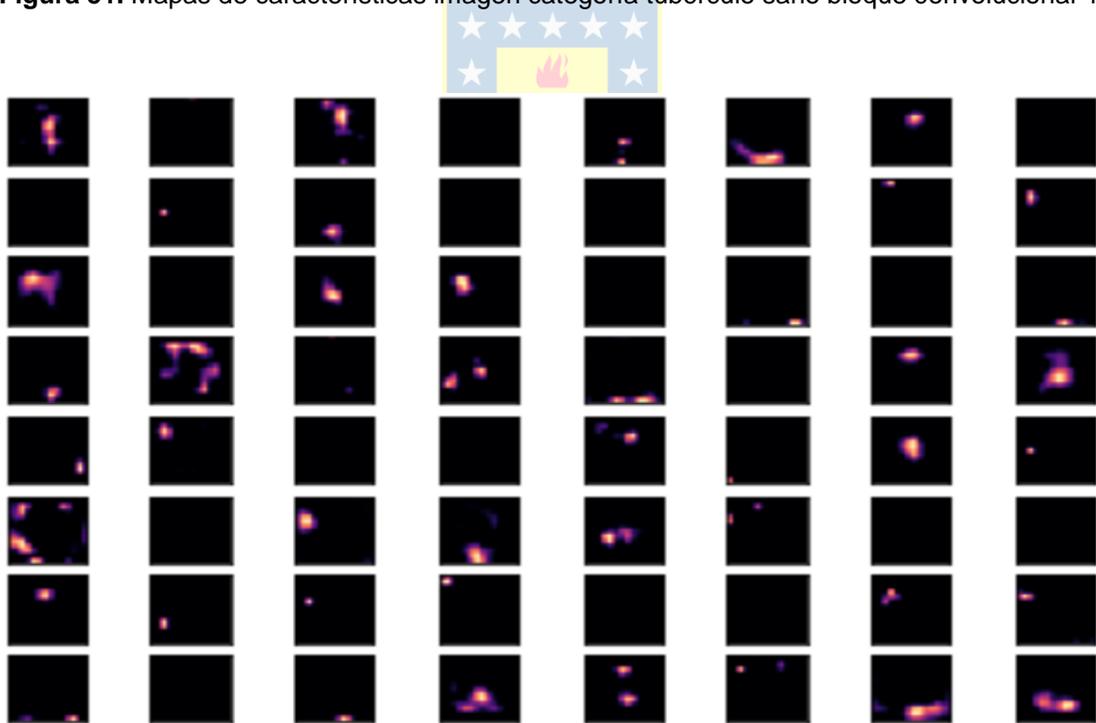


Figura 52. Mapas de características imagen categoría tubérculo sano bloque convolucional 5.

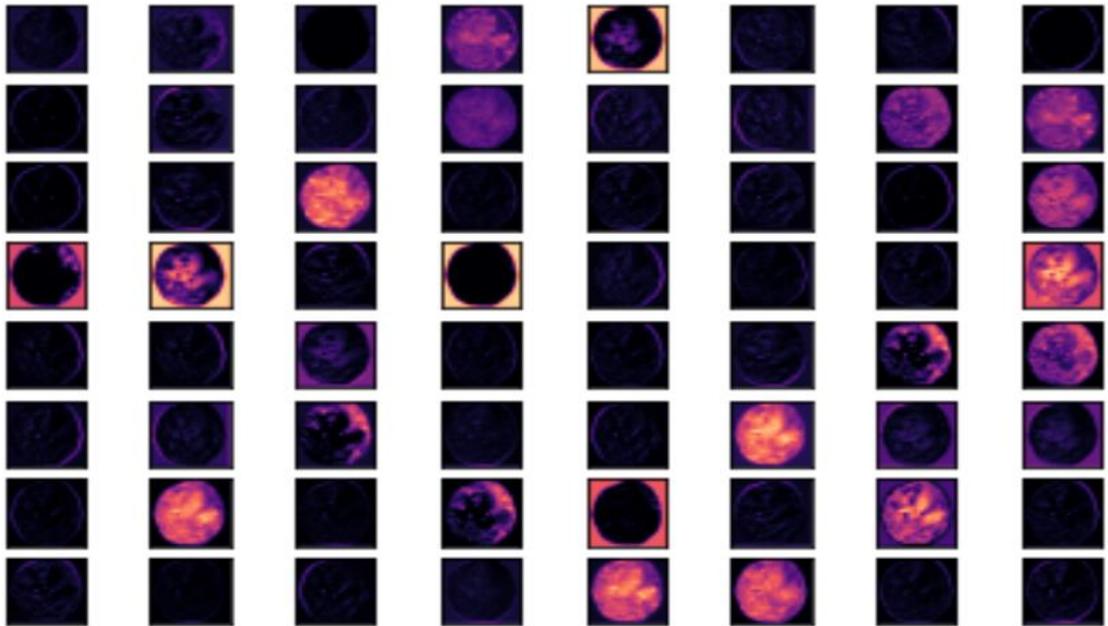


Figura 53. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 1.

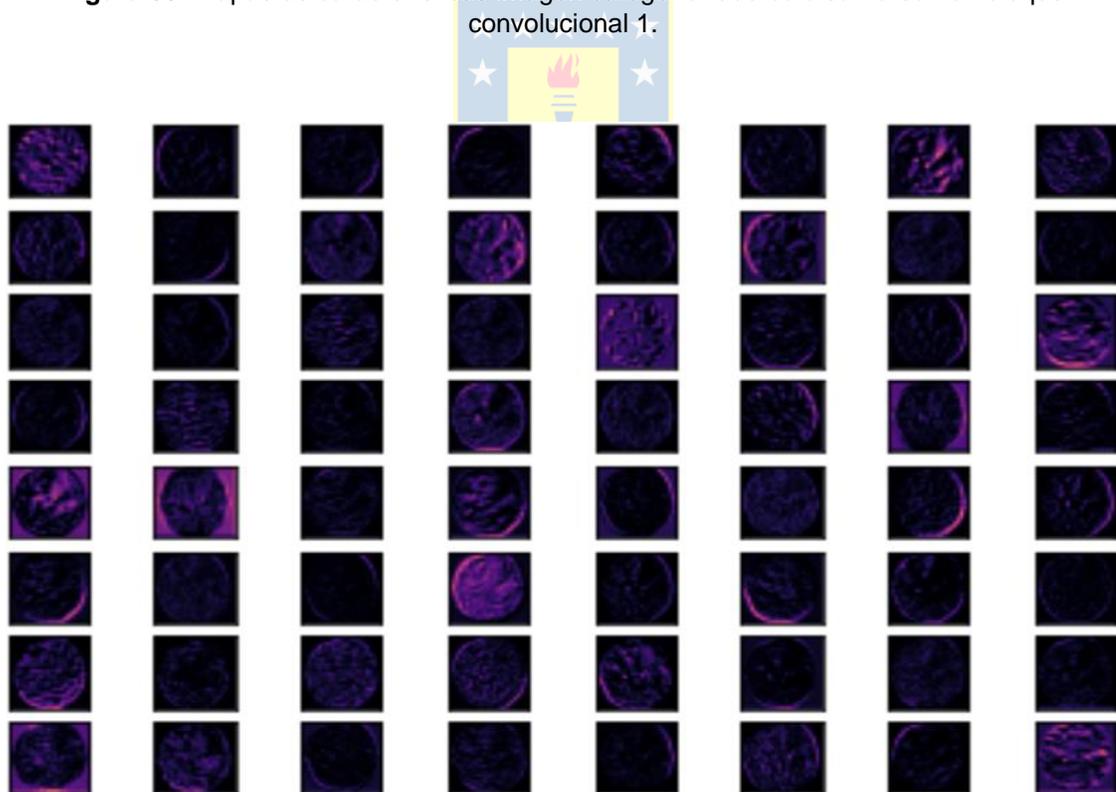


Figura 54. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 2.

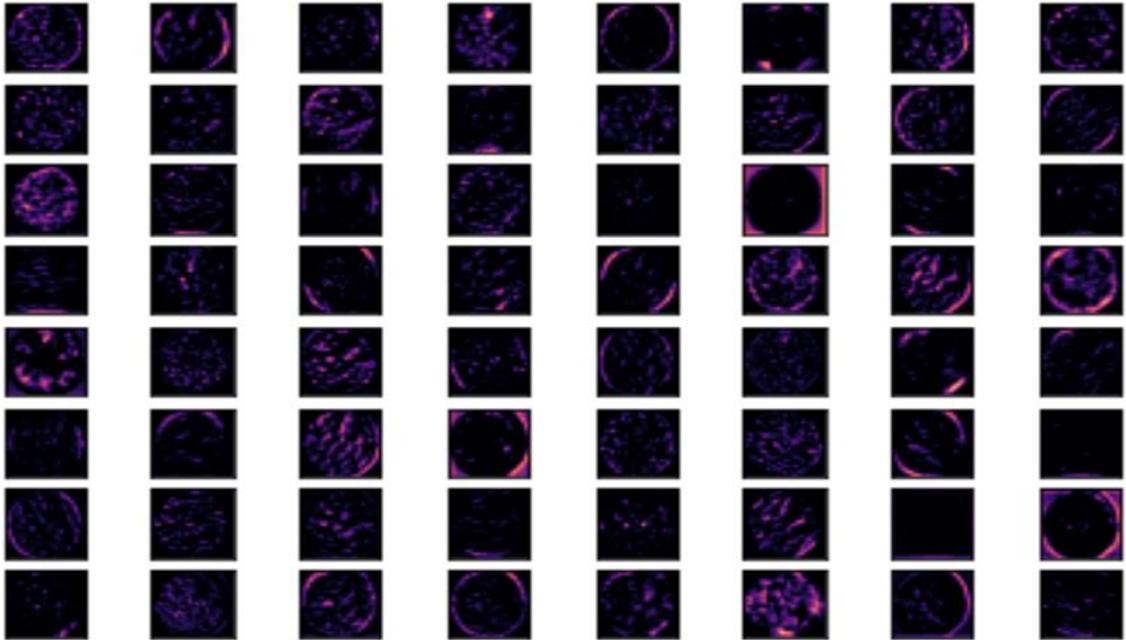


Figura 55. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 3.

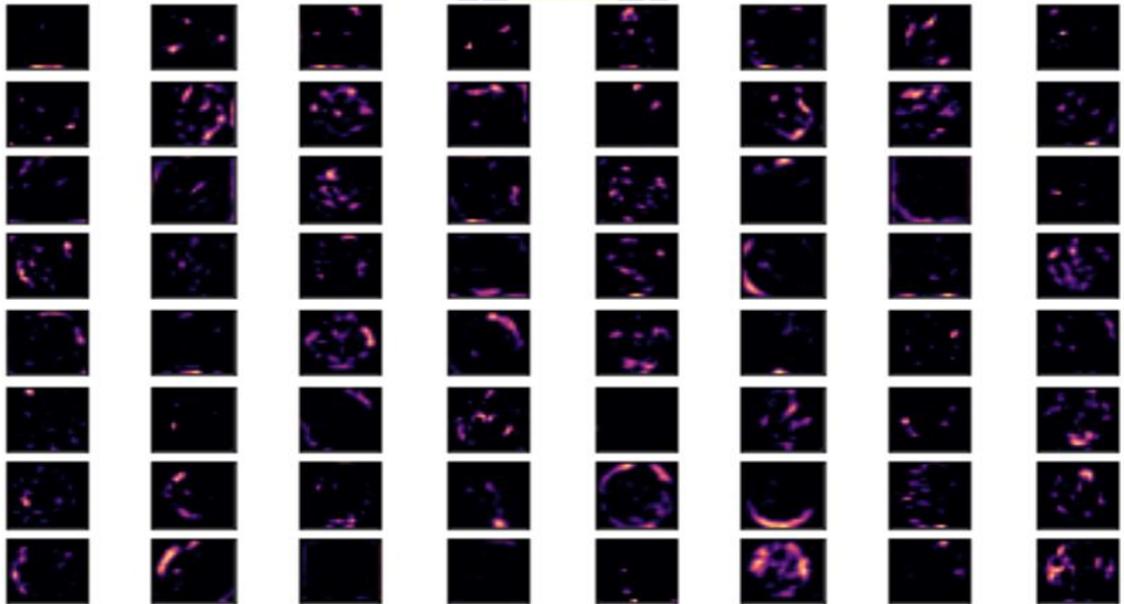


Figura 56. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 4.

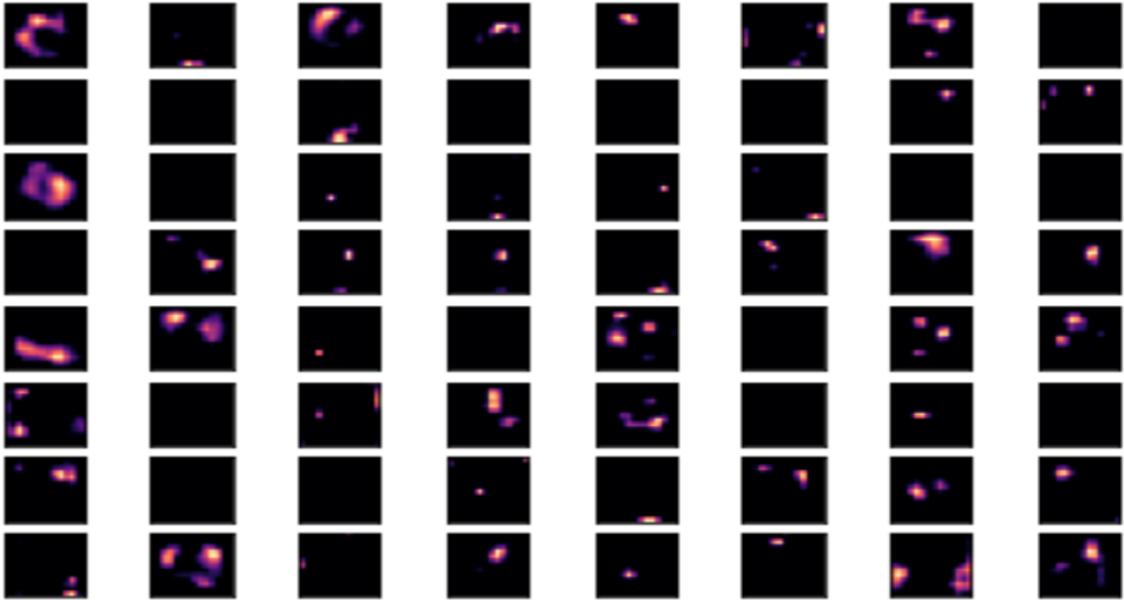


Figura 57. Mapas de características imagen categoría tubérculo sarna común bloque convolucional 5.

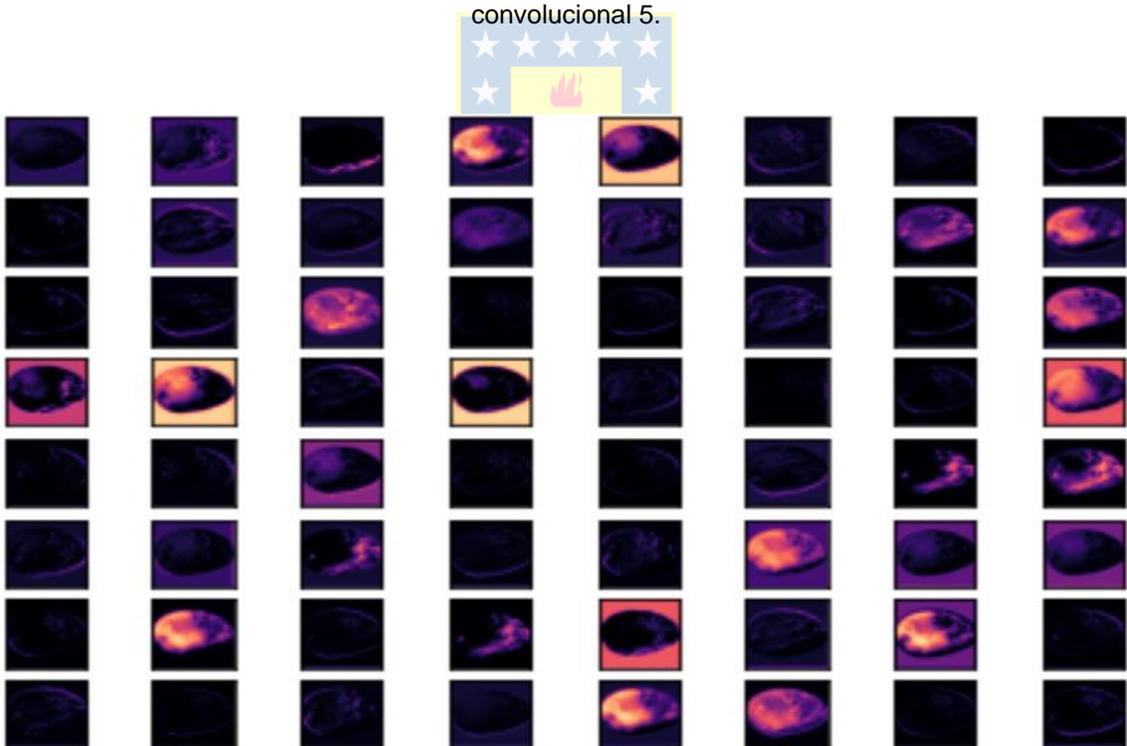


Figura 58. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 1.

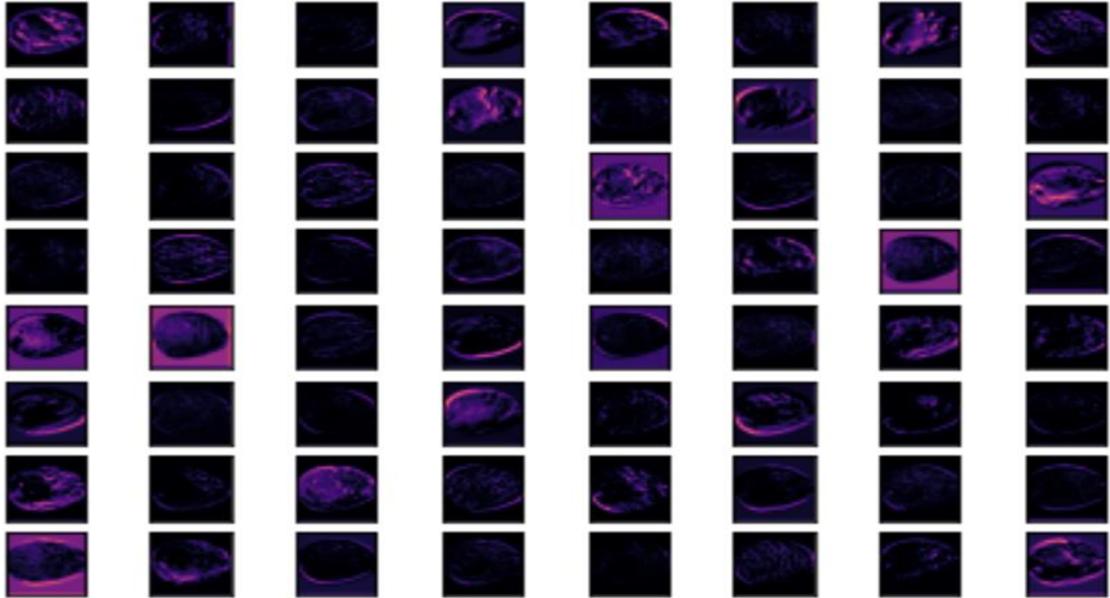


Figura 59. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 2.

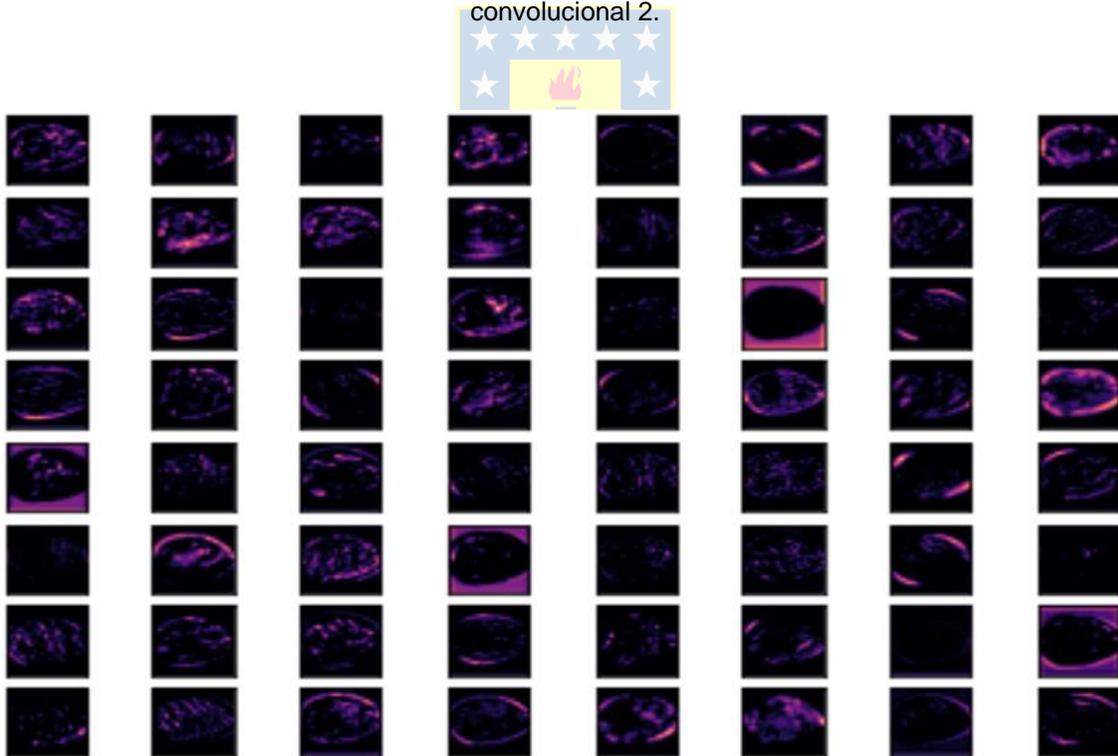


Figura 60. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 3.

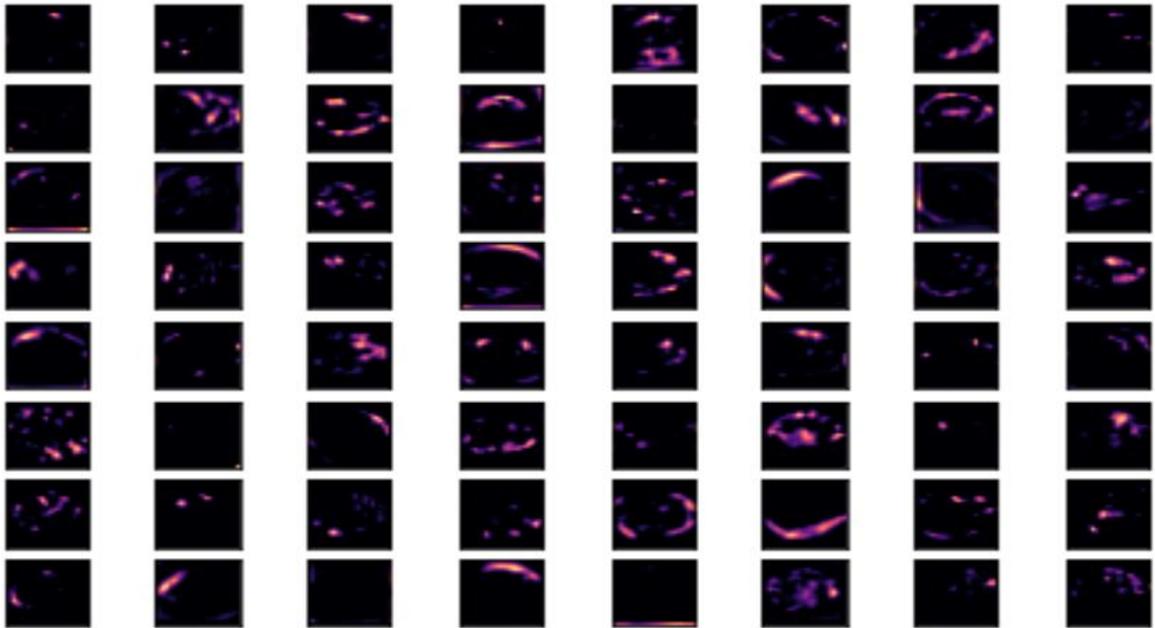


Figura 61. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 4.

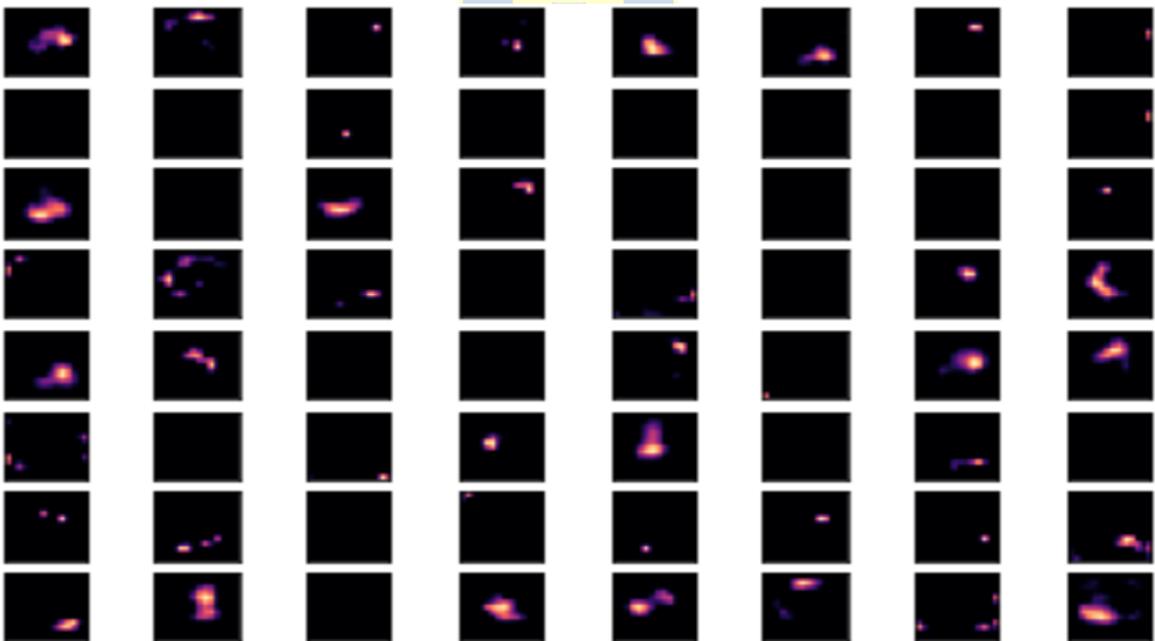


Figura 62. Mapas de características imagen categoría tubérculo tizón tardío bloque convolucional 5.