



**UNIVERSITÀ DEGLI STUDI
DI SALERNO**

**Modelo para la clasificación del aguacate Hass en sus estados comerciales y de exportación,
mediante el uso de redes neuronales convolucionales**

Universidad Tecnológica de Pereira

Facultad de Ingenierías

Maestría en Ingeniería de Sistemas

Pereira

2020

**Modelo para la clasificación del aguacate Hass en sus estados comerciales y de exportación,
mediante el uso de redes neuronales convolucionales**

Proyecto para la obtención del título de Magíster en Ingeniería de Sistemas y Computación

Cristian Camilo Salazar Salazar

Director:

MSc César Augusto Jaramillo Acevedo

Universidad Tecnológica de Pereira

Facultad de Ingenierías

Maestría en Ingeniería de Sistemas

Pereira

2020

Dedicatoria

A mi familia, mi padre, mi madre, mis hermanos que sin lugar a duda siempre confiaron y creyeron en mis capacidades, a ellos por estar siempre presentes, y por ser la familia que aunque no escogí siempre le agradezco a la vida por tenerlos siempre conmigo, a ellos por hacerme lo que soy hoy.

Agradecimientos:

A Eduardo Pulido, por acompañarme y asesorarme en todo el proceso del desarrollo del documento, a los profesores César Augusto Jaramillo Acevedo y Jorge Iván Ríos por su compromiso en este proyecto.

A la Universidad Tecnológica de Pereira, que sin duda me dio todas las bases para culminar este proceso, a la Universidad de Salerno que nos acompañó y nos ayudó en el proceso, a los profesores Roberto Tagliaferri, Ornella Malandrino y Rosaria Cerrone, por todo el apoyo, la paciencia y la disposición de ayudarnos el tiempo que estuvimos en Italia.

A mis compañeros de maestría Francesco Masuccio, Michele Cipriano, Lucia Pascarella, Iliaria Avino, Alejandro Moreno, Santiago Marin, Fredy Sánchez y Fabio Andrés Jiménez, por apoyarme, ayudarme y regalarme tan lindos momentos.

A Natalia Montenegro, por la paciencia, el amor y sobre todo la fortaleza en este proceso de 2 años, muchas gracias.

Resumen.

El mercado del aguacate hass en Colombia ha tenido un crecimiento en los últimos años debido a su alta comercialización internacional, por este motivo se han gestado procesos de producción del aguacate hass en todo el territorio nacional con una pronunciada influencia de inversionistas extranjeros. En el caso particular del departamento de Risaralda se ha generado mucho interés por introducirse en el mercado del aguacate Hass en especial en pequeños y medianos productores de la agroindustria, ellos están optando por la producción masiva de este fruto, sin embargo, luego de un estudio de campo, se pudo determinar que para poder ser exportado, el aguacate debe cumplir con ciertas condiciones fisiológicas y físicas, este proceso de clasificación se realiza de manera manual en la mayoría de los casos.

La escala de clasificación según los requerimientos comerciales del fruto son:

1. Aguacates tipo exportación, este aguacate debe ajustarse a los estándares internacionales según lo indicado por el ICONTEC, END 094 FRUTAS FRESCAS, VARIEDAD AGUACATE HASS, ESPECIFICACIONES y además, las indicaciones que se encuentra en la página oficial de la FAO, NORMA DEL CODEX PARA EL AGUACATE (CODEXSTAN197-1995).
2. Aguacates para comercializar nacionalmente, estos se refieren directamente a los que no cumplen las normas de exportación, pero que se encuentra en condiciones óptimas de consumo.
3. Aguacates ya maduros que no cumplen ninguno de los requisitos anteriores.

Este proyecto propone realizar la identificación de cada categoría por medio del análisis de la apariencia externa del fruto mediante el procesamiento de imágenes y el uso de redes neuronales convolucionales (CNN's).

Para llevar a buen término la empresa iniciada se tomó como base:

- La herramienta desarrollada por facebook para la implementación de una Mask R-CNN.
- Para el entrenamiento de La red se usó un dataset de 311 imágenes tomadas después del proceso de clasificación manual realizado por una exportadora regional, las imágenes fueron divididas en 211(train), 63(val), 37(test) aproximadamente 70% 20% y 10% respectivamente.
- Las imágenes fueron tomadas a plena luz del día dentro de la bodega de la exportadora, sin el uso de ambientes controlados con una cámara Evil Samsung NX 300, con un lente fijo de 45 mm, apertura de diafragma de f 1.8.
- El etiquetado de las imágenes fue realizado con la aplicación web COCO Annotator, la cual retorna las anotaciones directamente en formato COCO para poder realizar el entrenamiento usando el Detectron2 y el cual retornó los siguientes resultados:
 - **Avocado_mature** (Aguacate maduro) = **91.238 %** de precisión¹.
 - **nationals** (Aguacate de comercialización nacional) = **97.585 %** de precisión.
 - **export** (Aguacate de comercialización internacional) = **97.227 %** de precisión.²

Palabras Clave: Deep Learning, Redes Neuronales, Redes Neuronales Convolucionales, R-CNN, CNN, Fast R-CNN, Avocado, Faster R-CNN, Mask R-CNN, Detectron, Detectron2, Classification.

¹ Cuando se refiere a precisión se hace referencia a AP (Average Precision).

² Si desean usar el dataset o necesitan ayuda para usar la herramienta detectron. escribir un correo a ccss@utp.edu.co

Abstract.

The hass avocado market in Colombia has grown in the last years due to its high international commercialization. For this reason, hass avocado production processes have been developed throughout the national territory with a pronounced influence of foreign investors. In the particular case of Risaralda department, a lot of interest has been generated to introduce in the market of Hass avocado, especially in small and medium producers of the agro industry, they are opting for the massive production of this fruit. However, after a field study, it was possible to determine that in order to be exported, the avocado must comply with certain physiological and physical conditions; this classification process is carried out manually in most cases.

The scale of classification according to the commercial requirements of the fruit are

1. Export type avocados, this avocado must comply with international standards as indicated by ICONTEC, END 094 FRESH FRUITS, AVOCADO HASS VARIETY, SPECIFICATIONS and also with the indications found in the official FAO website, CODEX STANDARD FOR AVOCADO (CODEXSTAN197-1995).
2. Avocados to be marketed nationally, these refer directly to those that do not comply with export standards, but are in optimum condition for consumption.
3. Ripe avocados that do not meet any of the above requirements.

This project proposes to carry out the identification of each category through the analysis of the external appearance of the fruit by means of image processing and the use of convolutional neural networks (CNN's).

To carry out this project successfully, the company was taken as a base:

- The tool developed by facebook for the implementation of a R-CNN Mask.
- For the training of The Network was used a dataset of 311 images taken after the manual classification process performed by a regional exporter, the images were divided into 211(train), 63(val), 37(test) approximately 70% 20% and 10% respectively.
- The images were taken in full daylight inside the exporter's warehouse, without the use of controlled environments with an Evil Samsung NX 300 camera, with a fixed 45 mm lens, f 1.8 aperture.
- The labeling of the images was done with the web application COCO Annotator, which returns the annotations directly in COCO format to be able to perform the training using the Detectron2 and which returned the following results:
 - **Avocado_mature** (ripe avocado) = **91.238 % AP**
 - **nationals** (National marketing avocado) = **97.585 % AP**.
 - **export** (international marketing avocado) = **97.227% AP**.

Keywords: Deep Learning, Neural Networks, Convolutional Neural Networks, R-CNN, CNN, Fast R-CNN, Avocado, Faster R-CNN, Mask R-CNN, Detectron, Detectron2, Classification

Tabla de contenido

- 1. Introducción**
- 2. Definición del problema**
 - 2.1. Planteamiento**
 - 2.1.1. Estudio de campo**
 - 2.1.2. Algunos estudios previos**
 - 2.2. Formulación de la pregunta problema**
 - 2.3. Sistematización**
- 3. Justificación**
- 4. Objetivos**
 - 4.1. Objetivo general**
 - 4.2. Objetivos específicos**
- 5. Marco de referencia**
- 6. Hipótesis**
- 7. Metodología**
 - 7.1. Tipo de investigación**
 - 7.2. Método de investigación**
 - 7.3. Técnicas para la recolección de información**
- 8. Marco teórico**
 - 8.1. Redes Neuronales**
 - 8.2. Redes Neuronales Convolucionales**
 - 8.3. Detección de Objetos**
 - 8.4. Avances en las Redes Neuronales Convolucionales**
 - 8.5. Detectron y Detectron2**
- 9. Desarrollo de la investigación**
 - 9.1. A Propósito del deep learning**
 - 9.1.1 Redes neuronales convolucionales**
 - 9.1.2 R-CNN**
 - 9.1.3 Fast/Faster R-CNN**

9.1.4	Mask R-CNN	
9.1.5	Detectron/Detectron2	
9.2.	Estudio de campo	
9.3.	Preparación del dataset	
9.3.1	COCO Format	
9.3.2	Labeling data	
9.3.2.1	COCO Annotator	
9.4.	Iniciando con Detectron2	
9.4.1	Entrenamiento usando Mask R-CNN	
9.5.	Resultados	
9.6.	Validación	
10.	Conclusiones	
11.	Trabajos futuros	
12.	Bibliografía	
12.1	Artículos académicos	
12.2	Webgrafía	
12.3	Softwares	
13.	Anexos.	

Índice de tablas

1.	Tabla 1	28
2.	Tabla 2	30

Índice de imágenes

1.	Imagen 1	14
2.	Imagen 2	16
3.	Imagen 3	17

4. Imagen 4	18
5. Imagen 5	20
6. Imagen 6	40
7. Imagen 7	46
8. Imagen 8	47
9. Imagen 9	48
10. Imagen 10	49
11. Imagen 11	52
12. Imagen 12	53
13. Imagen 13	53
14. Imagen 14	57
15. Imagen 15	58
16. Imagen 16	59
17. Imagen 17	60
18. Imagen 18	62
19. Imagen 19	63
20. Imagen 20	64
21. Imagen 21	65
22. Imagen 22	66
23. Imagen 23	67
24. Imagen 24	68
25. Imagen 25	69
26. Imagen 26	70

1. Introducción

La agroindustria en Colombia tiene un potencial productivo aún latente, de hecho para solucionar el problema de la coca y otros cultivos ilícitos el país debe pasar necesariamente por la exploración y puesta en marcha de políticas en el agro que permitan reemplazar estos cultivos por otros que se encuentren en el marco de la legalidad y, que además, generen ingresos suficientes para garantizar el buen vivir de los campesinos, tanto a los pequeños como a los medianos productores del campo.

En ese sentido el aguacate hass ha aparecido como una alternativa viable, esto con base a la factibilidad de exportación y comercialización que presenta actualmente el aguacate a nivel mundial, que ha atraído nuevos inversionistas internacionales y productores locales y ha reactivado la economía por medio de este sector al iniciar un proceso de cultivo de aguacate en Colombia, tan así que el ministerio de agricultura presenta cifras de un crecimiento del 89% en los últimos 5 años .

Se estima que aproximadamente 54 mil personas en todo el país son beneficiados directa o indirectamente gracias a este producto en todas sus etapas productivas y esta es solo la punta del Iceberg de un fruto que recién comienza a vincularse a la economía nacional.

Coherente con lo planteado anteriormente es necesario que la academia, es decir: la universidad, ponga todo su caudal científico y técnico al servicio de este sector que, a penas en su fase inicial, ya se muestra como un gran motor que dinamice el sector productivo agroindustrial del país.

En ese orden de ideas las ingenierías en todas sus áreas del saber y del hacer, y en especial la ingeniería de sistemas, debe acercarse para conocer cuales son las falencias que aún se encuentran en el proceso de producción del aguacate y construir herramientas, metodologías y sistemas que ayuden a optimizar su proceso, mejorar su calidad y así, contribuir en aumentar los beneficios para todos los que se encuentran vinculados a la cadena productiva de este fruto de manera directa y de manera indirecta

aportar un pequeño grano de arena en el mejoramiento de la calidad de vida de los miembros de la nación. Sobre lo dicho anteriormente versa esta investigación, la creación de un “Modelo para la clasificación del aguacate Hass en sus estados comerciales y de exportación, mediante el uso de redes neuronales convolucionales” que si bien es una obra de ingeniería, busca ser ante todo una apuesta ética.

2. Definición del problema

2.1. Planteamiento.

2.1.1. Estudio de campo

En la actualidad gran parte de la clasificación nacional del aguacate Hass se efectúa de manera artesanal, basada en criterios subjetivos de personal que lleva años de experiencia empírica en el campo, lo que implica que sea una medición sujeta al error humano. Esto es basado en un estudio de campo realizado en tres empresas de alto nivel en este producto: la exportadora más grande de Colombia, cuya operación principal se encuentra en el departamento de Risaralda (Cartama) y otras dos de las más grandes de la región (FLP) y (Hass Diamond). A continuación se presenta la siguiente tabla que contiene la información estadística de las empresas de aguacate hass más grandes del país, esta información es tomada del DANE organizada y facilitada por PROCOLOMBIA que es la entidad encargada de promover el Turismo, la Inversión Extranjera en Colombia, las Exportaciones no minero energéticas y la imagen del país, esta información aparece en la imagen 1.

Razón social	Exportaciones de aguacate USD 2018	Exportaciones de aguacate USD 2019 (ene-jun)
Avofruit Sas - Cartama	13.451.553	13.674.104
Westfalia Fruit Colombia S.A.S	10.820.855	10.682.582
Pacific Fruits S.A.S	7.085.851	3.200.762
Hass Colombia Sat	4.719.591	6.066.418
Jardin Exotics Sas	4.439.581	3.982.981
C.I Fruty Green S.A	3.840.282	3.248.803
Green West S.A.S	2.098.393	1.382.449
Hass Diamond	1.840.361	449.814
Mountain Avocado Sas - Antes Tropyfruits S.A.S	1.453.859	568.953
Hasspacol S.A.S	1.416.662	272.505
C.I. F&P Trading S.A.S	1.232.390	618.057
Frutales Las Lajas S.A	1.175.326	239.767
Flp Colombia S.A.S.	1.153.484	3.494.586
Wolf & Wolf Latin America Ltda C I	1.056.016	17.189
Avocate Sas	994.973	620.188
Sci Vivasvan Sas	688.933	135.467
Esmeralda Quality Fruit S.A.S	547.957	252.544
Agua Hass	443.603	-
Verd Fruits S.A.S	441.661	541.974
Zaboka S.A.S.	375.933	537.943
Ama Tropicals Fruits Sas	371.779	126.434
Tropic Organic Colombia Sas	365.393	329.462
Comercializadora Internacional Disant	282.655	68.790
Frugol Sas	250.982	241.512
Aguacates Florez S.A.S.	220.490	150.074

Imagen 1, tabla de clasificación de las principales empresas de aguacate Hass en Colombia, fuente:

Procolombia

En el estudio de campo se pudo hacer visible en todas las entidades consultadas el uso de tres categorías para la clasificación del aguacate:

4. Aguacates tipo exportación, este aguacate debe ajustarse a los estándares internacionales según lo indicado por el ICONTEC, END 094 FRUTAS FRESCAS, VARIEDAD AGUACATE HASS, ESPECIFICACIONES y además, las indicaciones que se encuentra en la página oficial de la FAO, NORMA DEL CODEX PARA EL AGUACATE (CODEXSTAN197-1995).
5. Aguacates para comercializar nacionalmente, estos se refieren directamente a los que no cumplen las normas de exportación, pero que se encuentra en condiciones óptimas de consumo.
6. Aguacates ya maduros que no cumplen ninguno de los requisitos anteriores.

Sin embargo se encontró que para realizar dicha clasificación FLP y Hass Diamond solo realizan un procedimiento mecánico que consiste en seleccionar el aguacate mediante un sistema de báscula automático que identifica, mediante el peso de la fruta en un primer momento, si este se encuentra en condiciones óptimas de exportación o si se deja para la comercialización nacional; pero los siguientes filtros hasta su evaluación final requieren de la clasificación en otros aspectos físicos del aguacate (color, textura, cicatrización, entre otros) en los cuales se tiene que llevar a cabo una verificación manual, que requiere de gran esfuerzo humano y representa un análisis subjetivo para la clasificación del mismo, como se muestra en la imagen 2, 3 y 4.



Imagen 2, báscula para la clasificación del aguacate por peso en la empresa Hass Diamond. Fuente:

imagen de autoría propia.



Imagen 3, fase de clasificación manual del aguacate en la empresa Hass Diamond, fuente: imagen de autoría propia.



Imagen 4, fase de empacado de los aguacates para exportación en la empresa hass Diamond, fuente: imagen de autoría propia.

En Cartama, en cambio, poseen una máquina clasificadora, que según Ricardo Uribe, Gerente General de la empresa: “existen pocas máquinas de este tipo en el mundo”, información que es corroborada por Víctor Arboleda, gerente de Hass Diamond. Según sus fuentes: “esta máquina funciona por medio de la toma de fotografías y de análisis con rayos infrarrojos”, cabe recalcar lo dicho por Víctor Arboleda: “que el hecho de que hayan tan pocas en el mundo es debido a su alto costo”, por lo tanto la adquisición de dicho sistema de medición es inasequible para pequeños y medianos productores, que se ven forzados a contratar mano de obra extra para realizar procesos de clasificación

que generan sobrecostos y un alto margen de error en la clasificación, además no es posible conocer de manera detallada el funcionamiento de las máquinas de clasificación que se encuentran actualmente en el mercado, debido a que su acceso es restringido por cuestiones comerciales o industriales.

2.1.2. Antecedentes para la clasificación del aguacate Hass

Es necesario aclarar que debido a las especificidades para la clasificación del aguacate, el tipo de investigación desarrollada para este producto: clasificación mediante el uso de redes neuronales convolucionales y la novedad de la comercialización del fruto, especialmente en Colombia, es difícil encontrar fuentes que sirvan como referencia para la presente investigación de frutos distintos al aguacate. Sin embargo el ejercicio realizado por Choque & Salazar (2016) es una fuente de análisis vital para el proyecto iniciado, teniendo presente que en esta investigación ya se realizó un análisis de la clasificación no invasiva en distintos tipos de frutas. También se debe recalcar que la base para el proceso de desarrollo del sistema de clasificación del aguacate en el presente estudio, está atado a las normativas específicas para la exportación y comercialización nacional del aguacate, indicadas en ICONTEC, END 094 FRUTAS FRESCAS, VARIEDAD AGUACATE HASS, ESPECIFICACIONES y FAO, NORMA DEL CODEX PARA EL AGUACATE (CODEXSTAN197-1995) y que se diferencia en su totalidad a la clasificación basada en madurez planteada en anteriores investigaciones Choque & Salazar (2016) tal como lo muestra la siguiente imagen, tomada de INSTRUCTIVO CALIDAD DE EXPORTACIÓN Y CRITERIO DE COSECHA Hass Diamond.



Imagen 5, Escala de clasificación manual del aguacate de la empresa Hass Diamond, fuente: Hass Diamond.

En el momento se tienen diversos estudios para clasificación del aguacate Hass, entre ellos existe un prototipo móvil que logra clasificar los aguacates de acuerdo con su madurez, Choque & Salazar (2016), aunque su funcionamiento sigue siendo dispendioso y bajo entorno controlado. Hay otras investigaciones que también se basan en la clasificación del aguacate mediante su color y otros aspectos físicos con el uso de imágenes, así lo señalan tanto los trabajos de Arzate-Vásquez et al (2011) como los de Pinto, Rueda-Chacón y Arguello (2019) y los de Moreno, Forero, Gómez y Ramírez

(2019) trabajos que se abordarán con detalle más adelante. Algunos de los inconvenientes hallados en las investigaciones señaladas se centran en que todas realizan mediciones en ambientes controlados y tienen problemas en obtener la región de interés (ROI por sus siglas en inglés). Además algunos de estos sistemas desarrollados utilizan herramientas, materiales y softwares que no son libres, que requieren una inversión económica y por tales motivos, la industria continúa con una labor manual de selección del fruto, que no sólo incrementa los costos operativos, sino que además pone en riesgo la integridad de éste en el proceso de clasificación.

2.2. Formulación de la pregunta problema

¿Es posible desarrollar mediante el uso de redes neuronales convolucionales una clasificación del aguacate Hass?

2.3. Sistematización

- ¿Es posible la utilización de una red neuronal convolucional, en la clasificación del Aguacate Hass?
- ¿Se podrá implementar un sistema de clasificación del Aguacate Hass en ambientes no controlados, que cumpla con los estándares de comercialización?
- ¿Es posible clasificar el Aguacate Hass en una escala que obedezca a las necesidades actuales en la comercialización del fruto, nacional e internacional?

- ¿Es posible validar el modelo propuesto mediante un caso de prueba en una empresa exportadora?.

3. Justificación

Según recientes artículos, entre ellos Figueroa (2018), Colombia se posiciona como uno de los principales países en la exportación del aguacate hass, además de esto el departamento de Risaralda se está convirtiendo en referente a nivel nacional en términos de exportación de dicho producto Marín (2019), y sin embargo el conocimiento científico relacionado a la industria del aguacate se encuentra bastante atrasado Choque & Salazar (2016). Es necesario señalar que el proceso de clasificación de los aguacates, según el trabajo de campo realizado, requiere de un alto esfuerzo humano y unos costes de producción bastante elevados que dificultan el cultivo del mismo.

Debido a los altos costos de producción indicados anteriormente, quedan por fuera de la competencia pequeños y medianos productores, dado que no se encuentran en la capacidad económica de cultivar y comercializar este producto, por lo que tienden a consolidarse en este sector grandes emporios financieros y multinacionales, siendo estas las únicas capaces de llegar a producir en grandes cantidades ya que poseen el capital necesario para ello, tal como lo presenta Lucía, La (2019).

Si bien la inversión extranjera es considerada como fuente de crecimiento y empleo (Jimenez y Rendón 2012), es sabido de los inconvenientes que esta puede generar en el territorio nacional como explotación de mano de obra y principalmente pérdida de la soberanía, de esta forma lo plantean Ramírez y Florez (2006), permitiendo así que los pequeños y medianos productores se vean obligados a vender sus tierras a multinacionales extranjeras, debido a la falta de recursos y apoyos del gobierno nacional. No obstante en Colombia, todavía existen grandes extensiones de tierras con las cualidades

necesarias para sembrar aguacate hass, aproximadamente 200.000 hectáreas, sobre ello se puede apoyar en lo dicho por Morales (2016): “el 65,8 % de la tierra apta para sembrar en Colombia no se aprovecha”, tierras que serán disponibles para sembrar y que tienen condiciones óptimas en términos agropecuarios para la siembra de toda clase de productos agrícolas, pero la mayoría del territorio todavía se concentra en otros campos de la economía como la ganadería extensiva que ha sido implementada históricamente por las facilidades para su desarrollo y el bajo costo que implica su producción, así lo señala Posada (2002):

Quienes insisten en ver el predominio económico de la ganadería costeña como la manifestación de una actividad premoderna e irracional estarían pues ignorando las básicas limitaciones impuestas por los mismos factores de producción en el Caribe colombiano. Estarían desconociendo, más aún, los notables esfuerzos que se dieron también en la región por desarrollar otras actividades agrícolas -tabaco, azúcar, algodón, arroz y cacao-. Las frustraciones de todas estas experiencias debieron ser lamentablemente aleccionadoras, como lo fueron con frecuencia en otras partes del mundo en circunstancias en las que entonces “no pagaba arar, pero pagaba pastorear” (P. 76).

Teniendo presente lo consignado en las líneas anteriores, es necesario advertir que sobre la base de la creación de nuevas tecnologías que disminuyan los costos de producción de distintos procesos agropecuarios se puede volcar el campo hacia un sistema mucho más diversificado y productivo.

Por lo tanto se propone la implementación de un sistema que aproveche la tecnología de vanguardia en la detección de objetos, usando de esta forma los recientes estudios de los diferentes algoritmos basados en redes neuronales convolucionales como lo propone facebook con su herramienta Detectron2 que agrupa los últimos avances en algoritmos de detección de objetos:

Detectron2 is a ground-up rewrite of Detectron that started with maskrcnn-benchmark. The platform is now implemented in PyTorch. With a new, more modular design, Detectron2 is flexible and extensible, and able to provide fast training on single or multiple GPU servers. Detectron2 includes high-quality implementations of state-of-the-art object detection algorithms, including DensePose, panoptic feature pyramid networks, and numerous variants of the pioneering Mask R-CNN model family also developed by FAIR. Its extensible design makes it easy to implement cutting-edge research projects without having to fork the entire codebase. (Wu, Kirillov, Massa, Wan-Yen Y Girshick. 2019).

Aprovechando las ventajas que ofrece la tecnología referenciada en el párrafo anterior, se puede crear una herramienta para la clasificación del aguacate hass en una escala diferente a la propuesta por los estudios realizados anteriormente Choque y Salazar (2016), ya que la escala que proponen las investigaciones anteriores se han centrado principalmente en el estado de madurez del aguacate hass, el cual es perceptible principalmente según el color de la fruta, en cambio, de acuerdo al estudio de campo realizado, las necesidades industriales requieren de una clasificación mucho más profunda, que dé cuenta del estado óptimo de la fruta para su exportación o comercialización nacional, esta clasificación se basa en la aparición de “cicatrices”, manchas, deformidades, entre otras características físicas en el aguacate tal como lo plantea Alejandra Millán (Jefe de planta FLP) que requieren de distintos análisis además de la coloración, por lo que se hace necesario la utilización de algoritmos para la detección y clasificación de objetos, estos algoritmos ya se encuentran integrados en el framework Detectron2 en su sistema de entrenamiento, por lo que se convierte en un instrumento ideal para la solución de dicho problema.

Obedeciendo a las necesidades y posibilidades de la economía actual: la revitalización del sector agroindustrial y el nicho del mercado en el que se encuentra Colombia en términos del aguacate Hass y

basados en el trabajo de campo realizado en las principales empresas exportadoras de la región (FLP, Hass Diamond y Cartama). Se deduce que la agroindustria del aguacate requiere la creación de un nuevo sistema de clasificación automatizado y asequible para los pequeños y medianos productores que sea capaz de determinar la siguiente escala de clasificación del aguacate:

1. Aguacate Hass tipo exportación (aguacates que cumplen el estándar internacional y se puede exportar a diferentes países).
2. Aguacate Hass para comercialización nacional (aguacates que no cumplen el estándar internacional, pero que cumplen el estándar para comercializar en el territorio nacional).
3. Aguacates maduros que no cumplen con los estándares de comercialización nacional ni de exportación.

De esta forma se estaría acercando a los pequeños, medianos y grandes productores a tecnologías que están en la vanguardia internacional que sean de fácil acceso y que dejen instauradas las bases para siguientes y continuas investigaciones que puedan surgir a través del tema de exportación del aguacate hass y otros productos de la agroindustria.

4. Objetivos

4.1. Objetivo general

- Desarrollar un modelo mediante el uso de redes neuronales convolucionales para la clasificación del aguacate Hass en su estado de comercialización en ámbito internacional y nacional.

4.2 Objetivos específicos

- Utilizar una arquitectura de red neuronal convolucional que se adapte a las condiciones especiales del problema.
- Implementar un modelo de clasificación del aguacate Hass que funcione en ambientes no controlados.
- Proponer una escala de comercialización del fruto que obedezca a las necesidades del mercado vigente.
- Validar el modelo propuesto mediante un caso de prueba en una empresa exportadora.

5. Marco de referencia

Choque y Salazar 2016, en su trabajo: Desarrollo de algoritmo y prototipo móvil para medir el grado de madurez del aguacate hass mediante procesamiento digital de imágenes, con el cual obtuvieron el título de pregrado en Ingeniería en sistemas y computación crearon un prototipo móvil, que, mediante el procesamiento de imágenes, logra clasificar la madurez del aguacate Hass en 4 estados de madurez diferentes propuestos por ellos mismos:

1. Grado 1 o verde: es el estado primerizo del aguacate después de que es cortado del árbol, en este estado el aguacate aún no es apto para consumo.
2. Grado 2 : este estado representa los aguacates que han avanzado en su proceso de madurez y presentan un cierto grado de decoloración, en otras palabras se evidencia la pérdida del color verde, pero que aún no son aptos para el consumo.
3. Grado 3 o madurez de consumo: en esta sección se encuentran los aguacates que son óptimos para el consumo humano y la apariencia física ha tenido un cambio significativo en cuanto a la pérdida del color verde, tornándose de tonos marrones y rojizos.
4. Grado 4: este grado se representa por los aguacates que no son adecuados para el consumo humano pero que pueden ser utilizados para la fabricación de otros productos para el mercado, se representan por su coloración negra y su piel demasiado suave y sensible.

El prototipo para realizar la clasificación antes mencionada fue desarrollado bajo el seguimiento del proceso de madurez de una muestra de aguacates y la variación de la media de los componentes rojo, verde y azul (RGB por sus siglas en inglés) en dicho proceso. El procesamiento de imágenes encuentra el área de interés (ROI) usando dos algoritmos, filtro sobel y flood fill, los cuales se utilizaron de la siguiente forma:

- Filtro Sobel: el algoritmo de filtro de sobel fue utilizado para encontrar los bordes del aguacate Hass, una vez identificado, éste es el área de interés.
- Flood fill: una vez identificada el área de interés, lo que hacen es ubicar un punto dentro del área y aplicando el algoritmo flood fill empiezan a tomar todos los píxeles dentro del área de interés.

Con todos los píxeles del área de interés localizados, procedían a tomar la media total de sus componentes RGB, es decir suman el valor de cada componente y lo dividen por el número total de píxeles encontrados dentro del área de interés dando el resultado de la media de color del fruto en sus componentes $\overline{color} = (\overline{R}, \overline{G}, \overline{B})$.

Para la asignación de la escala previamente descrita se realizó, mediante una muestra de 50 aguacates, un seguimiento diario partiendo desde la recolección del árbol, hasta el último estado en condiciones de consumo. Luego por medio del análisis estadístico, se hallaron los rangos aceptados en cada uno de los componentes RGB, en síntesis: cada día durante este seguimiento se tomaron los valores retornados por el algoritmo previamente descrito, con los siguientes resultados:

Tabla 1. Datos obtenidos por medio del algoritmo con 50 aguacates en estado de madurez 1 o verde realizado por Choque y Salazar (2016)

Red	Green	Blue
59	66	43
67	69	47
66	74	47
71	76	48
52	60	40
65	72	43
61	66	45
57	59	41
67	75	46

60	68	44
60	76	40
66	73	49
58	60	45
67	76	48
52	60	39
54	56	41
60	73	45
60	60	42
60	60	42
59	59	48
47	62	36
64	69	36
57	49	41
60	64	48
63	66	48
64	76	45
71	74	48
69	76	50
68	84	48
59	68	45
64	79	46
64	79	47
65	81	47
63	79	47
67	83	49
64	81	48
66	66	41
60	74	45
70	76	44
72	72	46
64	66	46
65	59	36
67	60	41
58	59	42
60	61	45
66	72	48
64	71	50
57	78	44

67	73	47
64	68	49

Tabla 2. Análisis estadístico a 50 aguacates en estado de madurez 1 o verde

	Red	Green	Blue
Media	64	70	45
Desviación estándar	5.78	7.89	3.80
X	58-69	62-77	41-48

Donde X es el intervalo obtenido a partir de la media aritmética de la siguiente manera:

$$X = (Media - Desviación estándar, Media + Desviación estándar)$$

De esta forma Choque y Salazar (2016) lograron obtener los rangos aceptados para la clasificación del aguacate en cada uno de los estados propuesto por ellos.

Con los rangos de clasificación por cada estado de madurez obtenido, el procedimiento final fue:

1. Usando el algoritmo descrito, tomaban la media de color del aguacate hass.
2. Con los resultados del algoritmo, se procede a ubicar los componentes \bar{R} , \bar{G} , \bar{B} en alguno de los rangos aceptados para la clasificación del aguacate hass.

Las líneas anteriores explican de manera general el procedimiento realizado por Choque y Salazar (2016) para el desarrollo de su investigación y que son un primer insumo para el proyecto en curso. Finalmente se dejan expuestos algunos de los alcances y limitaciones identificados en el análisis de esta investigación:

- Lograron llevar este procedimiento a un prototipo funcional móvil, que haciendo uso de su cámara obtenían la foto del aguacate al cual deseaban clasificar y por medio de la conexión a un servidor que realizaba todo el procesamiento de imágenes y el algoritmo descrito anteriormente, retornaba la clasificación del fruto.

- Afrontan problemas cotidianos de la región de forma tecnológica y proponen soluciones al alcance de la mayoría.
- Uno de sus principales problemas fueron los factores ambientales como la iluminación y la sombra.
- El algoritmo sólo era preciso bajo un entorno controlado y con ciertas condiciones: el aguacate tenía que estar ubicado en el centro de la imagen obtenida.
- No realizaban un proceso de detección de objetos para identificar el área de interés.
- La escala propuesta, es una escala funcional, pero no está en sintonía con las necesidades del mercado regional y nacional.

Arzate-Vázquez, I et al (2011) en su artículo: Image Processing Applied to Classification of Avocado Variety Hass (Persea americana Mill.) During the Ripening Process, con el que lograron plantear o reconocer tres estados de madurez del aguacate Hass mediante el procesamiento de imágenes y el uso de algoritmos como k-nearest neighbor y principal component analysis (PCA).

El desarrollo del proyecto fue estructurado de la siguiente forma:

- Seguimiento de la madurez de una muestra de aguacates tomadas de la misma finca por 18 días consecutivos.
- Toma de fotografías usando un entorno controlado con lámparas leds y posicionando la cámara de forma vertical con respecto al fruto.
- Una análisis de pérdida de grasa.

Encontraron que los primeros doce días no hubo ningún cambio significativo en los componentes de color del aguacate y en los seis días posteriores, se vieron resultados más optimistas en cuanto al cambio de color y textura para la clasificación en los tres estados que los mismos plantearon.

Las líneas anteriores explican de manera general el procedimiento realizado por Arzate-Vázquez, I et al (2011) para el desarrollo de su investigación.

Finalmente se dejan expuestos algunos de los alcances y limitaciones identificados en el análisis de esta investigación:

- Comparan los resultados obtenidos en cuanto al color con la pérdida de grasa del aguacate.
- Usan entorno controlados para la toma de resultados.
- No presentan propuestas que acerquen este tipo de investigaciones al uso cotidiano de la sociedad.
- No se especifican claramente los estados de madurez propuestos.
- tienen un porcentaje de error aproximadamente del 20%.
- Encontraron cambios en la rugosidad del aguacate como factor de madurez.

6. Hipótesis

¿Es posible implementar un modelo mediante el uso de redes neuronales convolucionales neuronales para una clasificación del aguacate Hass en mercados nacionales e internacionales?.

7. Metodología

7.1. Tipo de investigación

El tipo de investigación es cuantitativa – experimental, realizando inferencias a partir de los resultados obtenidos sobre los datos muestrales que se utilizarán como fundamento a esta investigación.

7.2. Método de investigación

- Realizar la escala de clasificación según necesidades de mercado.
- Seleccionar una o varias muestras de imágenes de aguacate Hass.
- Realizar una clasificación en rangos escogidos, sobre el estado del aguacate Hass.
- Segmentar el conjunto de imágenes, en tres conjuntos para el entrenamiento y evaluación de las redes neuronales, es decir, un conjunto de entrenamiento, otro de validación y uno final para evaluar los resultados.
- Obtener resultados clasificatorios, utilizando un mismo conjunto de imágenes de evaluación.
- Realizar pruebas sobre ambientes con alto ruido.

7.3. Técnicas para la recolección de información

- Se realizó un trabajo de campo para la obtención del conjunto de imágenes de aguacate Hass.
- Se realizaron consultas a personas especializadas en el comercio y exportación del aguacate Hass, para la supervisión de la segmentación del conjunto de imágenes, según su estado (Exportación, Nacional, Maduro).
- El conjunto de datos para el entrenamiento son obtenidos bajo los estados naturales de procesamiento del fruto, bien sea dentro de la bodega bajo iluminación natural diurna o con iluminación artificial al interior de esta, por ende la recolección de información no será en un espacio controlado, por lo que se espera poder realizar clasificaciones en espacios no controlados por completo más cercanas a las necesidades de la industria.

8. Marco teórico

8.1. Redes neuronales

Las redes neuronales son un sistema complejo con base en la interconexión de recursos en el aprendizaje automático para la creación de modelos que buscan como finalidad la predicción, este sistema trata de asemejar el comportamiento de una neurona animal en la cual su principal característica es la interconectividad.

El proceso en el que se basa una red neuronal consiste en la entrada de un peso o un conocimiento previo con el cual se puedan generar unas predicciones significativamente precisas, entonces se tiene que a una red neuronal le llegan unos estímulos de entradas que se puedan considerar como inputs, y los cuales mediante el procesamiento de las neuronas (cada neurona puede funcionar como un modelo aparte) debe retornar resultados de predicción; el conocimiento previamente adquirido llamado peso logra descomponer estas entradas en unas capas clasificatorias y al final entregar una predicción.

Ahora el aprendizaje de una red neuronal está basado en asignarle o definir unos costos de función es decir: darle un (cost function) y utilizar un gradiente descendiente (gradient descent) para minimizar estos costos, de esta manera a las predicciones menos acertadas se les asignan los costos más altos y las predicciones más acertadas van a tener el menor costo. Lo que busca el aprendizaje es que en cada iteración se puedan reducir dichos costos.

Resumiendo lo anterior, entregando a la red neuronal unos conjuntos de datos de entrenamiento en los que se le puedan asignar los costos de cada transacción, la red neuronal va ir construyendo automáticamente estos costos como aristas entre las entradas, las neuronas y las salidas esperadas.

8.2. Redes neuronales convolucionales

Las redes neuronales convolucionales (CNN's) están basadas en el uso de redes neuronales pero que enfatizan en imágenes, o sea que tienen unas entradas y salidas similares a las de las redes neuronales pero su funcionamiento se basa en la predicción de características explícitas en las imágenes.

Debido a que las imágenes contienen tantos datos por ejemplo una simple imagen de 64×64 contiene un total de 12.288 datos ($64 \times 64 \times 3$), al momento de tratar de esta cantidad de datos como una red neuronal normal el desempeño de dicha red se ve altamente reducido debido a la cantidad de datos que debe procesar en una simple imagen, es así como las CNN's se convierten en uso esencial para el procesamiento de imágenes.

Las CNN's se caracterizan por una serie de pasos para poder funcionar con la cantidad de datos que contiene una imagen, los cuales se enumeran así:

1. La entrada para la red neuronal va a mantenerse con los valores de cada píxel.
2. La capa de agrupamiento se caracteriza por la reducción de la imagen en pequeños segmentos de la imagen, reduciendo el número de parámetros entregados a la red es decir, si una imagen viene de ($64 \times 64 \times 3$) la imagen se va a dividir en pequeños segmentos, los cuales contendrán menos datos que la imagen original.
3. Capas totalmente conectadas: Se centra en reproducir el mismo funcionamiento de las redes neuronales para generar clasificaciones y predicciones.

Por lo tanto el entrenamiento de una CNN necesita un poco más de tiempo en comparación a las redes neuronales tradicionales, debido al número de pasos y procedimientos que se tienen que realizar, pero que logran solucionar el problema inicial de la cantidad de datos en una imagen.

8.3. Avances en las redes neuronales convolucionales

Las redes neuronales convolucionales (CNN por sus siglas en inglés) han tenido unos avances significativos en los últimos años debido a su alto costo computacional, si bien las CNN's tradicionales usan una capa de agrupamiento, que en otras palabras es seccionar la imagen en pequeños fragmentos, el tiempo de procesamiento se incrementa demasiado, por lo cual llega Region-CNN (R-CNN) que en un primer momento divide la imagen en 2000 áreas de interés (ROI) como cuadros delimitadores (bounding box) y por cada una de estas áreas de interés se hace una CNN convencional y finalmente cada área de interés se redefine usando una regresión.

De esta forma se convierten todos las regiones de interés (bounding box) en una búsqueda selectiva que se basa en agrupaciones de color, el contenido de las regiones y el tamaño.

Se puede afirmar que la búsqueda selectiva de regiones de interés es un buen método para la clasificación y la detección de objetos dentro de una imagen, sin embargo los costos computacionales en vez de disminuir tienden a aumentar debido a la cantidad de procesamiento que tienen que realizar, para superar esto llega Faster R-CNN, que lo que hace, principalmente, es cambiar la forma de obtener las regiones de interés en comparación a como lo hace la búsqueda selectiva, agregando una nueva red neuronal para obtener la propuestas de regiones (Region Proposals) llamada red de propuestas de región (Region Proposal Network RPN) que realiza la detección de áreas de interés mucho más rápido que R-CNN.

Mask R-CNN a diferencia de la Faster R-CNN usa una red en pirámide de características (Feature Pyramid Networks). Mask R-CNN es un algoritmo que se basa en dos estados, el primer estado toma la imagen y retorna las regiones de interés y el segundo estado clasifica cada región de interés y retorna todos los píxeles a los que esta pertenece.

8.4. Detección de objetos

La detección de objetos en una imagen consiste en identificar las características relevantes de la imagen o los objetos que se encuentran en la misma, es decir, las personas pueden identificar cada objeto o característica presente en una imagen basados en el conocimiento de cómo es la apariencia del objeto, pero ¿cómo se podrían expresar estos conocimientos en datos válidos para una máquina?

La detección de objetos se puede lograr gracias a los últimos avances de las redes neuronales convolucionales y su optimización en la obtención de las regiones de interés, que permiten una extracción de características más profundas en la imagen y enfocadas en las regiones de interés.

9. Desarrollo de la investigación

En este apartado se explicará de manera detallada el desarrollo de la investigación en la implementación del modelo para la clasificación del aguacate Hass en sus estados comerciales y de exportación, mediante el uso de redes neuronales convolucionales.

Se explica todo el proceso desde el porqué se eligió el tipo de red neuronal, el estudio de campo, la preparación del dataset, como también las herramientas utilizadas, todo el proceso del entrenamiento de la red neuronal y la obtención de resultados.

9.1. A propósito del deep learning

En la actualidad se ha vuelto indispensable el uso de redes neuronales para la clasificación, esto debido a su alta precisión y el mejoramiento continuo que implica el uso del deep learning en los problemas de clasificación. Los mundos económicos y científicos se ven hoy beneficiados con el desarrollo de nuevas tecnologías y algoritmos centrados en la clasificación y estimación de diferentes problemas entre los cuales se pueden destacar: estimaciones de mercado, el cambio de una moneda con respecto a otra, cirugías basadas en redes neuronales totalmente automáticas Minar et all (2018) y Délano R., P. (2019).

9.1.1. Redes neuronales convolucionales

Si bien las redes neuronales se basan en la predicción con base en una información inicial para producir un resultado clasificadorio, cuando a imágenes se refiere, esta cantidad de datos aumenta a tal punto que sería imposible el procesamiento de una imagen por medio de redes neuronales, por ejemplo: en una simple imagen de 64 x 64 píxeles el peso que recae sobre una sola neurona sería de 12.288 datos

para analizar³, lo que implicaría una sobrecarga de información y por lo tanto un aumento considerable en el tiempo de ejecución O'Shea, K., & Nash, R. (2015). Para el manejo de imágenes en deep learning es común escuchar hoy en día el uso de redes neuronales convolucionales, las cuales lo que hacen, a groso modo, es dividir la imagen en pequeñas secciones de imagen para que el procesamiento de la cantidad de datos provenientes sean adecuados para el análisis en una red neuronal O'Shea, K., & Nash, R. (2015).

9.1.2. R-CNN

Con el uso de redes neuronales convolucionales sin embargo el procesamiento de imágenes continúa siendo dispendioso:

“Features matter. The last decade of progress on various visual recognition tasks has been based considerably on the use of SIFT and HOG . But if we look at performance on the canonical visual recognition task, PASCAL VOC object detection, it is generally acknowledged that progress has been slow during 2010-2012, with small gains obtained by building ensemble systems and employing minor variants of successful methods” Girshick, et al. (2014),

Lo que proponen los autores de la cita anterior es que para la obtención de las regiones de interés (ROI's) se puede usar el método R-CNN (Regions with CNN feature). Por medio de este método se divide la imagen en, aproximadamente, 2000 regiones de interés.

“Our object detection system consists of three modules. The first generates category-independent region proposals. These proposals define the set of candidate detections available to our detector.

³ O'Shea, K., & Nash, R. (2015) detallan este proceso de la siguiente manera: With this dataset a single neuron in the first hidden layer will contain 784 weights ($28 \times 28 \times 1$ where 1 bare in mind that MNIST is normalised to just black and white values), which is manageable for most forms of ANN. If you consider a more substantial coloured image input of 64×64 , the number of weights on just a single neuron of the first layer increases substantially to 12, 288.

The second module is a large convolutional neural network that extracts a fixed-length feature vector from each region. The third module is a set of class specific linear SVMs. In this section, we present our design decisions for each module, describe their test-time usage, detail how their parameters are learned, and show detection results on PASCAL VOC 2010-12 and on ILSVRC2013.”

VOC 2010 test	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv	mAP
DPM v5 [20] [†]	49.2	53.8	13.1	15.3	35.5	53.4	49.7	27.0	17.2	28.8	14.7	17.8	46.4	51.2	47.7	10.8	34.2	20.7	43.8	38.3	33.4
UVA [39]	56.2	42.4	15.3	12.6	21.8	49.3	36.8	46.1	12.9	32.1	30.0	36.5	43.5	52.9	32.9	15.3	41.1	31.8	47.0	44.8	35.1
Regionlets [41]	65.0	48.9	25.9	24.6	24.5	56.1	54.5	51.2	17.0	28.9	30.2	35.8	40.2	55.7	43.5	14.3	43.9	32.6	54.0	45.9	39.7
SegDPM [18] [†]	61.4	53.4	25.6	25.2	35.5	51.7	50.6	50.8	19.3	33.8	26.8	40.4	48.3	54.4	47.1	14.8	38.7	35.0	52.8	43.1	40.4
R-CNN	67.1	64.1	46.7	32.0	30.5	56.4	57.2	65.9	27.0	47.3	40.9	66.6	57.8	65.9	53.6	26.7	56.5	38.1	52.8	50.2	50.2
R-CNN BB	71.8	65.8	53.0	36.8	35.9	59.7	60.0	69.9	27.9	50.6	41.4	70.0	62.0	69.0	58.1	29.5	59.4	39.3	61.2	52.4	53.7

Table 1: Detection average precision (%) on VOC 2010 test. R-CNN is most directly comparable to UVA and Regionlets since all methods use selective search region proposals. Bounding-box regression (BB) is described in Section C. At publication time, SegDPM was the top-performer on the PASCAL VOC leaderboard. [†]DPM and SegDPM use context rescoring not used by the other methods.

Imagen 6, Resultados de la implementación de R-CNN en VOC 2010 test, fuente: Girshick, et all.

(2014)

Es notorio el mejoramiento que las R-CNN aportan al manejo de imágenes en redes neuronales aumentando su capacidad y velocidad, desde entonces el avance de las redes neuronales convolucionales (CNN’s) ha estado en aumento y ha conllevado también al desarrollo de diferentes investigaciones de las cuales se destacan los métodos de obtener las regiones de interés (ROI’s)

9.1.3. Fast/Faster R-CNN

También nuevos estudios en CNN’s en los cuales llegan Fast R-CNN y Faster R-CNN que proponen mejoras al estudio previo de R-CNN

“The network first processes the whole image with several convolutional (conv) and max pooling layers to produce a conv feature map. Then, for each object proposal a region of interest (RoI) pooling layer extracts a fixed-length feature vector from the feature map. Each feature vector is fed into a sequence of fully connected (fc) layers that finally branch into two sibling output layers: one that produces softmax probability estimates over K object classes plus a catch-all “background” class and another layer that outputs four real-valued numbers for each of the K object classes. Each set of 4 values encodes refined bounding-box positions for one of the K classes.” R. Girshick, (2015)

Sin Embargo en Faster R-CNN Ren, S., et all. (2017) entendieron que la base de las redes neuronales convolucionales se centra en la obtención de las regiones de interés (ROI's) e implementaron una novedosa forma de obtener el ROI usando otra red neuronal llamada Región Proposal Network (RPN).

“One may note that fast region-based CNNs take advantage of GPUs, while the region proposal methods used in research are implemented on the CPU, making such runtime comparisons inequitable. An obvious way to accelerate proposal computation is to reimplement it for the GPU. This may be an effective engineering solution, but re-implementation ignores the down-stream detection network and therefore misses important opportunities for sharing computation. In this paper, we show that an algorithmic change—computing proposals with a deep convolutional neural network—leads to an elegant and effective solution where proposal computation is nearly cost-free given the detection network’s computation. To this end, we introduce novel Region Proposal Networks (RPNs) that share convolutional layers with state-of-the-art object detection networks. By sharing convolutions at test-time, the marginal cost for computing proposals is small (e.g., 10ms per image). Our observation is that the convolutional feature maps used by region-based detectors, like

Fast RCNN, can also be used for generating proposals. On top of these convolutional features, we construct an RPN by adding a few additional convolutional layers that simultaneously regress region bounds and objectness scores at each location on a regular grid. The RPN is thus a kind of fully convolutional network (FCN) and can be end-to-end specifically for the task for generating detection proposals” Ren, S., et al. (2017)

Es así como Faster R-CNN se posiciona como una de las herramientas más óptimas para el entrenamiento de redes neuronales mediante el procesamiento de imágenes.

9.1.4. Mask R-CNN

En síntesis y como medio final para esta investigación se aborda Mask R-CNN que a diferencia de Faster/Fast R-CNN se basa en usar el Instance segmentation Romera-Paredes, B. and Torr, P. (2016), con lo cual el fin es clasificar y detectar cada objeto dentro de una imagen.

“The vision community has rapidly improved object detection and semantic segmentation results over a short period of time. In large part, these advances have been driven by powerful baseline systems, such as the Fast/Faster RCNN and Fully Convolutional Network (FCN) frameworks for object detection and semantic segmentation, respectively. These methods are conceptually intuitive and offer flexibility and robustness, together with fast training and inference time. Our goal in this work is to develop a comparably enabling framework for instance segmentation. Instance segmentation is challenging because it requires the correct detection of all objects in an image while also precisely segmenting each instance. It therefore combines elements from the classical computer vision tasks of object detection, where the goal is to classify individual objects and localize each using a bounding box, and semantic RoIAlign segmentation, where the goal is to classify each pixel into a fixed set of categories without differentiating object instances.¹ Given this, one might expect

a complex method is required to achieve good results. However, we show that a surprisingly simple, flexible, and fast system can surpass prior state-of-the-art instance segmentation results. Our method, called Mask R-CNN, extends Faster R-CNN by adding a branch for predicting segmentation masks on each Region of Interest (RoI), in parallel with the existing branch for classification and bounding box regression (Figure 1). The mask branch is a small FCN applied to each RoI, predicting a segmentation mask in a pixel-to-pixel manner. Mask R-CNN is simple to implement and train given the Faster R-CNN framework, which facilitates a wide range of flexible architecture designs. Additionally, the mask branch only adds a small computational overhead, enabling a fast system and rapid experimentation. In principle Mask R-CNN is an intuitive extension of Faster R-CNN, yet constructing the mask branch properly is critical for good results. Most importantly, Faster RCNN was not designed for pixel-to-pixel alignment between network inputs and outputs. This is most evident in how RoIPool, the de facto core operation for attending to instances, performs coarse spatial quantization for feature extraction. To fix the misalignment, we propose a simple, quantization-free layer, called RoIAlign, that faithfully preserves exact spatial locations. Despite being” He, Kaiming et al. (2017).

Con lo dicho hasta el momento se puede concluir que Mask R-CNN es una herramienta fundamental en la clasificación y detección de objetos, proporcionando así una forma de usar las redes neuronales convolucionales para clasificar objetos dentro de una imagen, y además de esto, retornando los segmentos en los que se encuentra ubicado dicho objeto.

9.1.5. Detectron/Detectron2.

Detectron es un framework desarrollado por Facebook que compila diversos algoritmos de CNN's para detección de objetos, que permite de manera fácil e integrada utilizar diferentes algoritmos como Faster R-CNN, Mask R-CNN usando instance segmentation.

Detectron2 es una mejora a Detectron que implementa el estado del arte en detección de objetos y adhiere nuevas herramientas como Panoptic Segmentation Kirillov, Alexander et all. (2018).

“Progress in AI is a community effort that includes individuals, large and small labs, academia, and industry. The problems we aim to solve go far beyond what any individual or group can achieve in isolation. For this reason, we believe strongly in sharing code that enables reproducible research, rapid experimentation, and development of new ideas. By releasing Detectron2, we hope to further accelerate research in the areas of object detection, segmentation, and human pose understanding.

New research starts with understanding, reproducing, and verifying previous results in the literature.

With Detectron2, we aim to provide high-quality reference implementations for many state-of-the-art algorithms in order to democratize this phase of the research process.

The library's modular design also enables researchers to implement new projects with clean separation from standard detection library functionality. As an example, Mesh R-CNN, FAIR's recent work on predicting per-object instance 3D meshes from 2D images, was developed in Detectron2. Detectron2's modular design enabled the researchers to easily extend Mask R-CNN to work with complex data structures representing 3D meshes, integrate new data sets, and design novel evaluation metrics.

Detectron2 can be easily shared between research-first use cases and production-oriented use cases.

Because the library is built in PyTorch, new models can be implemented rapidly and then transferred to production.” Wu, Y, et all. (2019)

Con Detectron2 se pueden desarrollar investigaciones usando el “state-of-the-art” para la detección de objetos con el cual se pueden alcanzar resultados óptimos en los que no sería tan sencillo llegar actuando de forma individual.

9.2. Estudio de campo.

El estudio de campo fue realizado en tres empresas de alto nivel en la exportación de aguacate hass: la primera de ellas es Cartama, la exportadora más grande de Colombia, cuya operación principal se encuentra en el departamento de Risaralda y otras dos de las más grandes de la región y el país: FLP y Hass Diamond. Para constatar la importancia de las empresas sobre las cuales se desarrolló el estudio de campo a continuación se presenta la siguiente tabla que contiene la información estadística de las empresas de aguacate hass más grandes del país, esta información es tomada del DANE organizada y facilitada por PROCOLOMBIA que es la entidad encargada de promover el Turismo, la Inversión Extranjera en Colombia, las Exportaciones no minero energéticas y la imagen del país, esta información aparece en la imagen 1.

Se deduce que la agroindustria del aguacate requiere la creación de un nuevo sistema de clasificación automatizado y asequible para los pequeños y medianos productores que sea capaz de determinar la siguiente escala de clasificación del aguacate:

1. Aguacates tipo exportación, este aguacate debe ajustarse a los estándares internacionales según lo indicado por el ICONTEC, END 094 FRUTAS FRESCAS, VARIEDAD AGUACATE HASS, ESPECIFICACIONES y además, las indicaciones que se encuentra en la página oficial de la FAO, NORMA DEL CODEX PARA EL AGUACATE (CODEXSTAN197-1995).
2. Aguacates para comercializar nacionalmente, estos se refieren directamente a los que no cumplen las normas de exportación, pero que se encuentra en condiciones óptimas de consumo.

3. Aguacates ya maduros que no cumplen ninguno de los requisitos anteriores.

ver imagen 5



Imagen 7, Proceso de preselección y pesado por la máquina en Hass Diamond, fuente: autoría propia

Además de lo expuesto anteriormente se encontró que para realizar dicha clasificación FLP y Hass Diamond solo realizan un procedimiento mecánico que consiste en seleccionar el aguacate mediante un

sistema de báscula automático que identifica, mediante el peso de la fruta en un primer momento, si este se encuentra en condiciones óptimas de exportación o si se deja para la comercialización nacional; pero los siguientes filtros hasta su evaluación final requieren de la clasificación en otros aspectos físicos del aguacate (color, textura, cicatrización, entre otros) en los cuales se tiene que llevar a cabo una verificación manual, que requiere de gran esfuerzo humano y representa un análisis subjetivo para la clasificación del mismo.



Imagen 8, proceso después del pesado del aguacate, es transportado y depositado en su respectivo calibre, fuente: autoría propia



Imagen 9, Proceso de verificación manual del fruto (ver imagen 5), fuente: autoría propia.

Los siguientes filtros después del pesado del aguacate, son totalmente manuales, en este punto el personal tiene que seleccionar el fruto según sus condiciones físicas externas y clasificar si cumple con las condiciones para ser exportado, o por el contrario se deja para ser comercializado nacionalmente (ver imagen 9).



Imagen 10, proceso de empaquetado del aguacate para exportación, fuente: autoría propia

Durante el estudio de campo se realizó también la adquisición del dataset por medio de fotografías tomadas a los aguacates que ya habían sido previamente clasificados por la compañía exportadora en sus diferentes estados (tipo exportación, para comercialización nacional o rechazados), dichas fotografías fueron tomadas en entornos normales dentro de la misma planta de exportación a plena luz del día, no hubo en ningún momento algún ambiente controlado.

La cámara usada para la toma de las fotografías fue realizada con una cámara tipo Evil Samsung NX 300, con un lente fijo de 45 mm, apertura de diafragma de $f. 1.8$.

9.3. Preparación del dataset

Para preparar el dataset lo primero que se necesita es tener las etiquetas de las imágenes en un formato que sea reconocido por Detectron como COCO format o pascal VOCC.

9.3.1 COCO FORMAT

Una de las principales metas de la detección de objetos es el manejo de imágenes y el entendimiento de las imágenes, es decir ¿qué objetos están presentes en dicha imagen?, ¿cómo es el contexto de la imagen?, ¿qué clase de objetos están presentes? y ¿cuáles son las posiciones de dichos objetos?

Microsoft COCO: Common Objects in Context es un estudio reciente que aporta, además de un dataset con imágenes cotidianas del día a día, un formato que permite que las imágenes puedan ser etiquetadas o en otras palabras: las anotaciones del dataset para poder realizar entrenamientos a las redes neuronales.

“We introduce a new large-scale dataset that addresses three core research problems in scene understanding: detecting non-iconic views (or non-canonical perspectives [12]) of objects, contextual reasoning between objects and the precise 2D localization of objects. For many categories of objects, there exists an iconic view. For example, when performing a web-based image search for the object category “bike,” the top-ranked retrieved examples appear in profile, unobstructed near the center of a neatly composed photo. We posit that current recognition systems perform fairly well on iconic views, but struggle to recognize objects otherwise – in the background, partially occluded, amid clutter [13] – reflecting the composition of actual everyday scenes. We verify this experimentally; when evaluated on everyday scenes, models trained on our data perform better than those trained with prior datasets. A challenge is finding natural images that contain multiple objects. The identity of many objects can only be resolved using context, due to small size or ambiguous appearance in the image. To push research in contextual reasoning, images depicting

scenes [3] rather than objects in isolation are necessary. Finally, we argue that detailed spatial understanding of object layout will be a core component of scene analysis. An object's spatial location can be defined coarsely using a bounding box [2] or with a precise pixel-level segmentation [14], [15], [16]. As we demonstrate, to measure either kind of localization performance it is essential for the dataset to have every instance of every object category labeled and fully segmented. Our dataset is unique in its annotation of instance-level segmentation masks, Fig. 1.

To create a large-scale dataset that accomplishes these three goals we employed a novel pipeline for gathering data with extensive use of Amazon Mechanical Turk. First and most importantly, we harvested a large set of images containing contextual relationships and noniconic object views. We accomplished this using a surprisingly simple yet effective technique that queries for pairs of objects in conjunction with images retrieved via scene-based queries [17], [3]. Next, each image was labeled as containing particular object categories using a hierarchical labeling approach [18]. For each category found, the individual instances were labeled, verified, and finally segmented. Given the inherent ambiguity of labeling, each of these stag” Lin, Tsung-Yi et al. (2014)

Es así como en Lin, Tsung-Yi et al. (2014), plantean un nuevo formato para etiquetar las imágenes dentro de un dataset el cual es, hoy en día, usado para el entrenamiento de las redes neuronales como Mask/Faster/Fast R-CNN.

9.3.2 Labeling data.

Como se expuso en el apartado 9.3, uno de los principales factores para poder usar detectron/detectron2 para el entrenamiento de la red neuronal es que las anotaciones del dataset deben tener un formato específico, en particular el formato elegido es COCO FORMAT 9.3.1, después del estudio de campo realizado 9.2, y de las imágenes obtenidas en él, lo siguiente es proceder con la

preparación del dataset, del cual se obtuvieron 311 imágenes que fueron divididas en 211(train), 63(val), 37(test) aproximadamente 70% 20% y 10% respectivamente.

Después de tener el dataset dividido en train/validation/test, lo siguiente es proceder a etiquetar las imágenes para las anotaciones en formato coco.

9.3.2.1 COCO Annotator.

COCO Annotator, es una aplicación web desarrollada por Brooks, J. (2019), la cual permite hacer el etiquetado de las imágenes y exportar las anotaciones en COCO FORMAT.

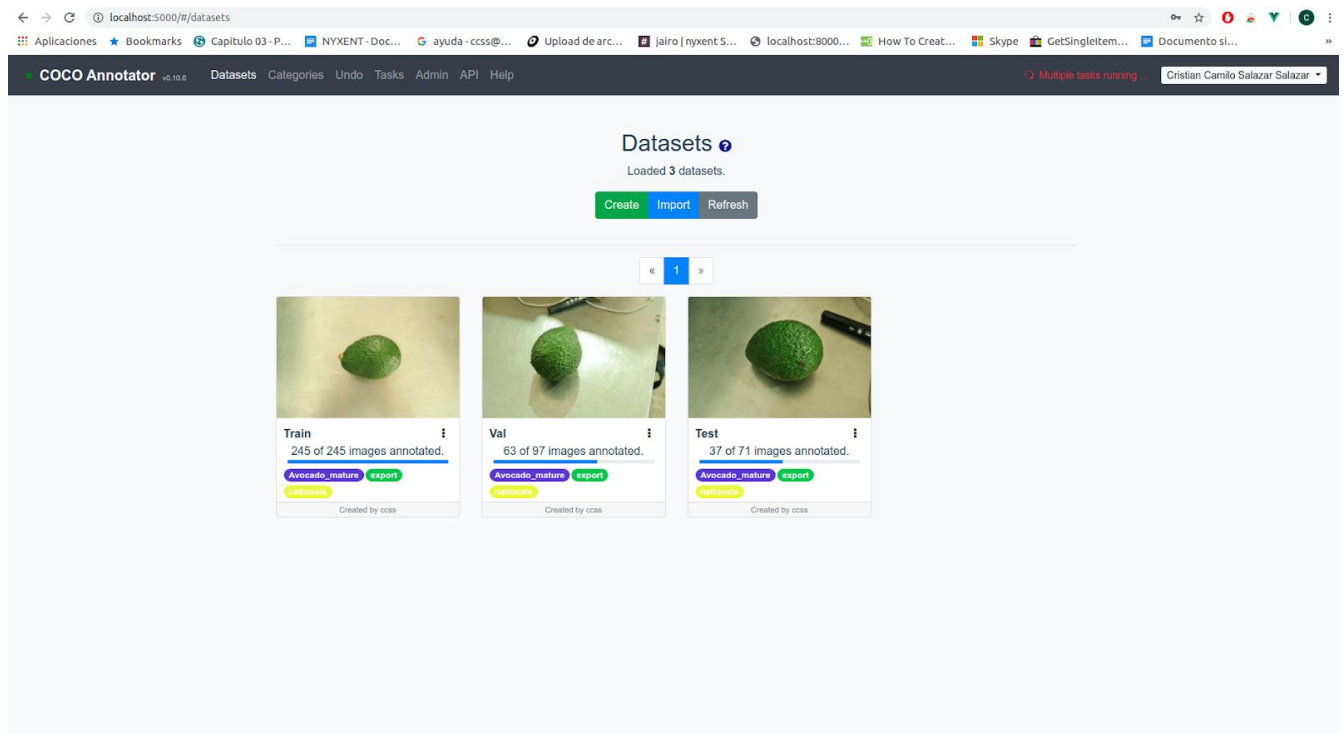


Imagen 11, inicio de la aplicación COCO Annotator, fuente: Autoría propia

En la imagen 11 se puede notar la división del dataset entre Train/Val/Test respectivamente.

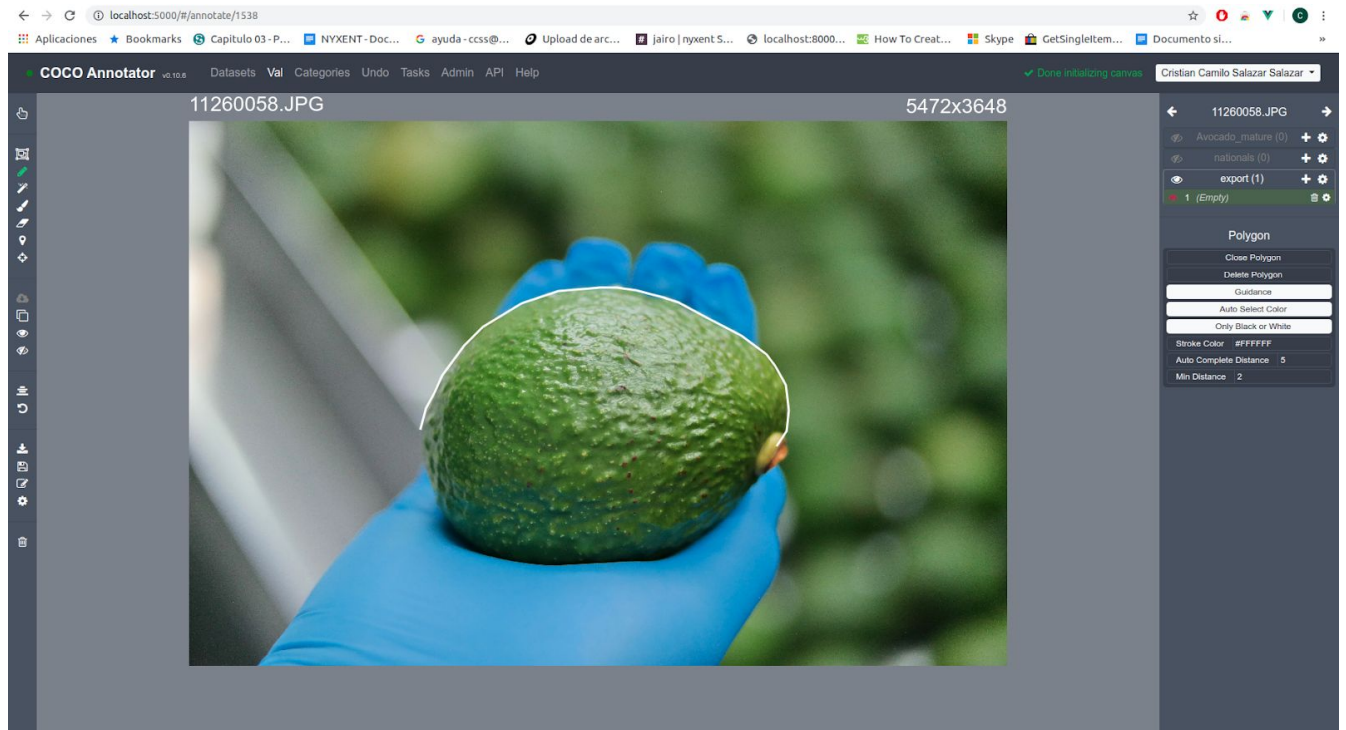


Imagen 12, Proceso de etiquetado de una imagen usando COCO Annotator, fuente: autoría propia

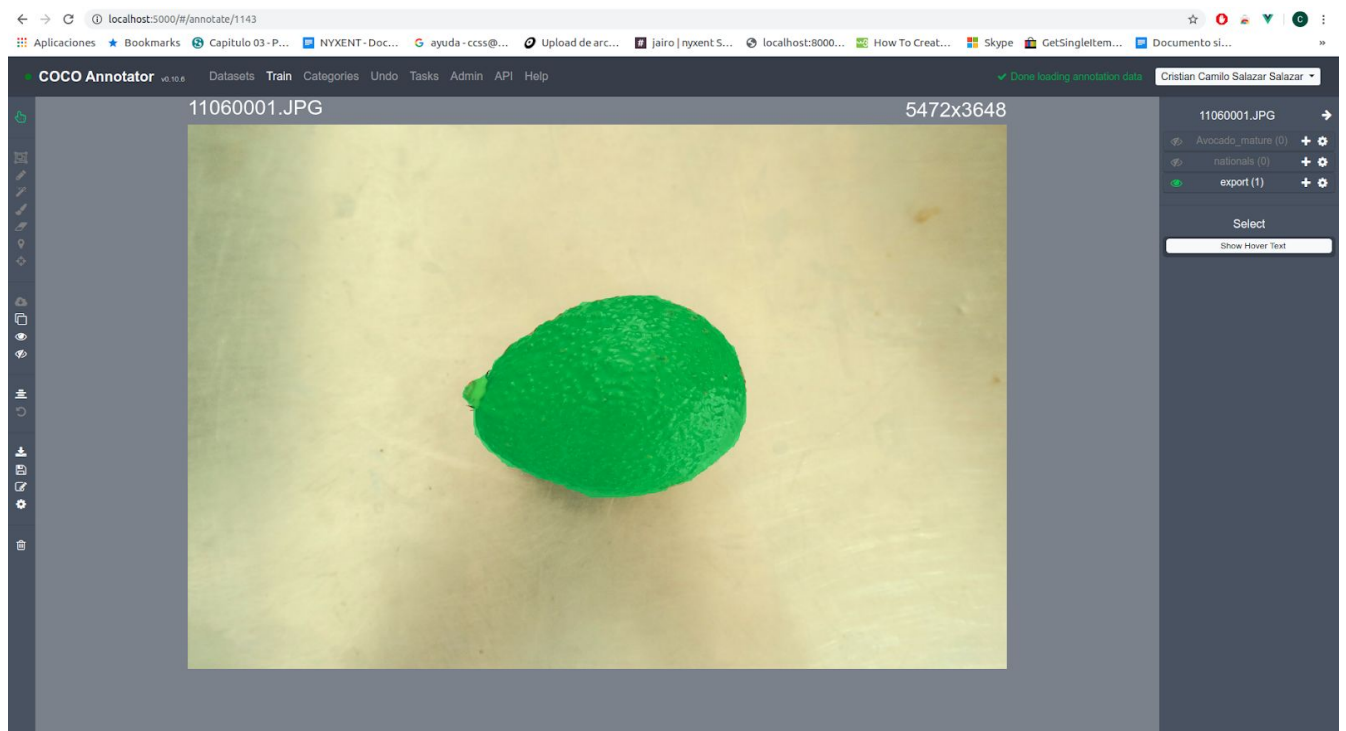


Imagen 13, Imagen etiquetada usando el método de polígono, fuente: Autoría propia.

9.4. Iniciando con Detectron2

Para iniciar con Detectron2, se debe seguir el tutorial [Welcome to detectron2's documentation! — detectron2 0.1 documentation. \(2019\)](#), en este se encuentra el paso a paso de la instalación del framework hasta el entrenamiento de la red neuronal y la obtención de resultados.

9.4.1 Entrenamiento usando Mask R-CNN

Para el entrenamiento de una red neuronal usando Detectron2 con un dataset diferente al de COCO, o personalizado, se deben registrar los datasets en el catálogo de detectron2 usando el siguiente proceso:

```
# Some basic setup
# Setup detectron2 logger
import detectron2
from detectron2.utils.logger import setup_logger
setup_logger()

# import some common libraries
import numpy as np
import cv2
from google.colab.patches import cv2_imshow

# import some common detectron2 utilities
from detectron2.engine import DefaultPredictor
from detectron2.config import get_cfg
from detectron2.utils.visualizer import Visualizer
from detectron2.data import MetadataCatalog

import os
import numpy as np
import json
from detectron2.structures import BoxMode
```

```

import itertools

from detectron2.data.datasets import register_coco_instances
from detectron2.data.datasets import load_coco_json

register_coco_instances("Avocados_train", {}, "avocados/Train/Train-2.json",
"avocados/Train")
register_coco_instances("Avocados_test", {}, "avocados/Test/Test-3.json",
"avocados/Test")
register_coco_instances("Avocados_val", {}, "avocados/Val/Val-4.json",
"avocados/Val")

```

Quando se hayan registrado los datasets en detectron se puede proceder al entrenamiento de la red neuronal, de la siguiente forma:

```

from detectron2.engine import DefaultTrainer
from detectron2.config import get_cfg

cfg = get_cfg()

cfg.merge_from_file("./detectron2_repo/configs/COCO-InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml")

cfg.DATASETS.TRAIN = ("Avocados_train", "Avocados_val")

cfg.DATASETS.TEST = () # no metrics implemented for this dataset

cfg.OUTPUT_DIR = 'drive/My Drive/WEIGHTS/output_without_w_train_1000_iter'

cfg.MODEL.WEIGHTS = "drive/My Drive/WEIGHTS/output_without_w_train_1000_iter/model_final.pth"

cfg.DATALOADER.NUM_WORKERS = 2

cfg.SOLVER.IMS_PER_BATCH = 2

cfg.SOLVER.BASE_LR = 0.00025

```

```

cfg.SOLVER.MAX_ITER = 3000    # 3000 iterations seems good enough, but you
can certainly train longer

cfg.MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE = 256    # faster, and good enough
for this dataset

cfg.MODEL.ROI_HEADS.NUM_CLASSES = 3

os.makedirs(cfg.OUTPUT_DIR, exist_ok=True)

trainer = DefaultTrainer(cfg)

trainer.resume_or_load(resume=False)

trainer.train()

```

Para probar el modelo final se corre el “inference” que proporciona detectron2:

```

cfg.merge_from_file("../detectron2_repo/configs/COCO-InstanceSegmentation/mas
k_rcnn_R_50_FPN_3x.yaml")

cfg.OUTPUT_DIR = 'drive/My Drive/WEIGHTS/output_without_w_train_1000_iter'

cfg.MODEL.WEIGHTS = "drive/My
Drive/WEIGHTS/output_without_w_train_1000_iter/model_final.pth"

cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST = 0.8    # set the testing threshold
for this model

cfg.DATASETS.TEST = ("Avocados_val", "Avocados_test")

predictor = DefaultPredictor(cfg)

import random

from detectron2.utils.visualizer import ColorMode

dataset_dicts = load_coco_json('avocados/Val/Val-4.json', 'avocados/Val',
'Avocados_val')

for d in random.sample(dataset_dicts, 3):

    im = cv2.imread(d["file_name"])

    outputs = predictor(im)

```



```
v = Visualizer(im[:, :, ::-1],  
              metadata=MetadataCatalog.get("Avocados_val"),  
              scale=0.8  
            )  
  
v = v.draw_instance_predictions(outputs["instances"].to("cpu"))  
cv2_imshow(v.get_image()[:, :, ::-1])
```

Lo que arroja los siguientes resultados:

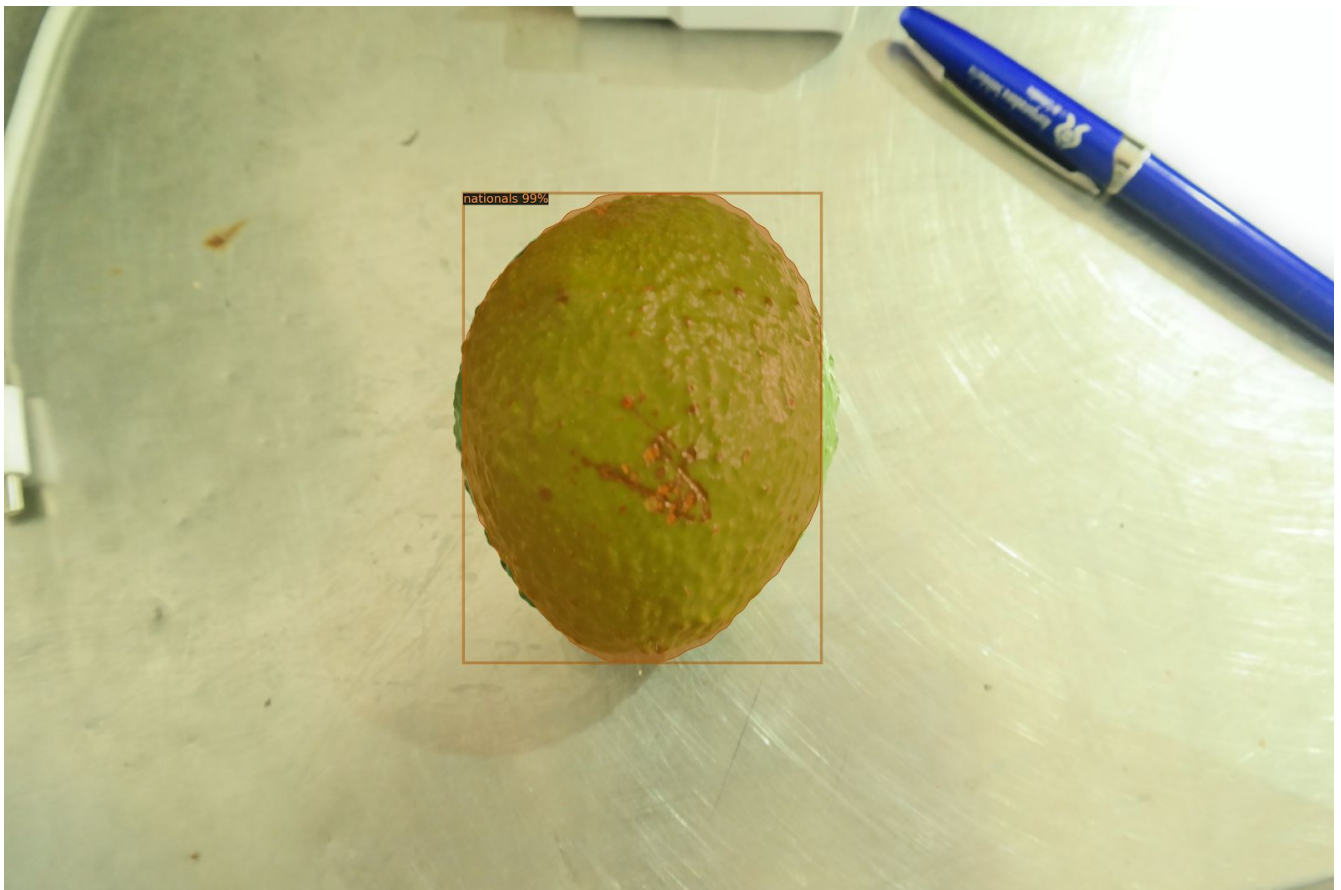


Imagen 14, Resultado del inference usando el modelo final después del entrenamiento detección de aguacate para comercialización nacional, fuente: autoría propia

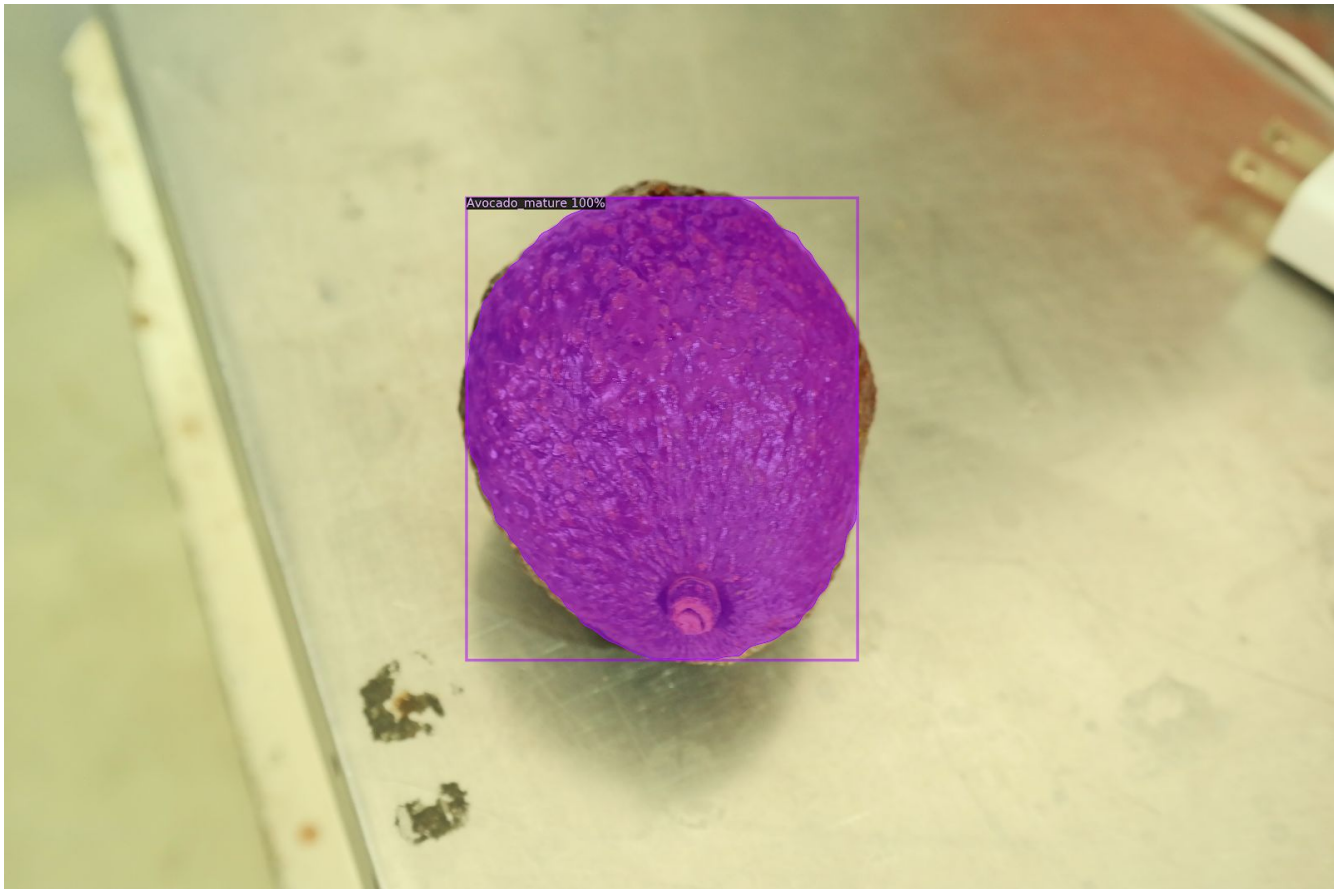


Imagen 15, Resultado del inference usando el modelo final después del entrenamiento detección de aguacate maduro, fuente: autoría propia

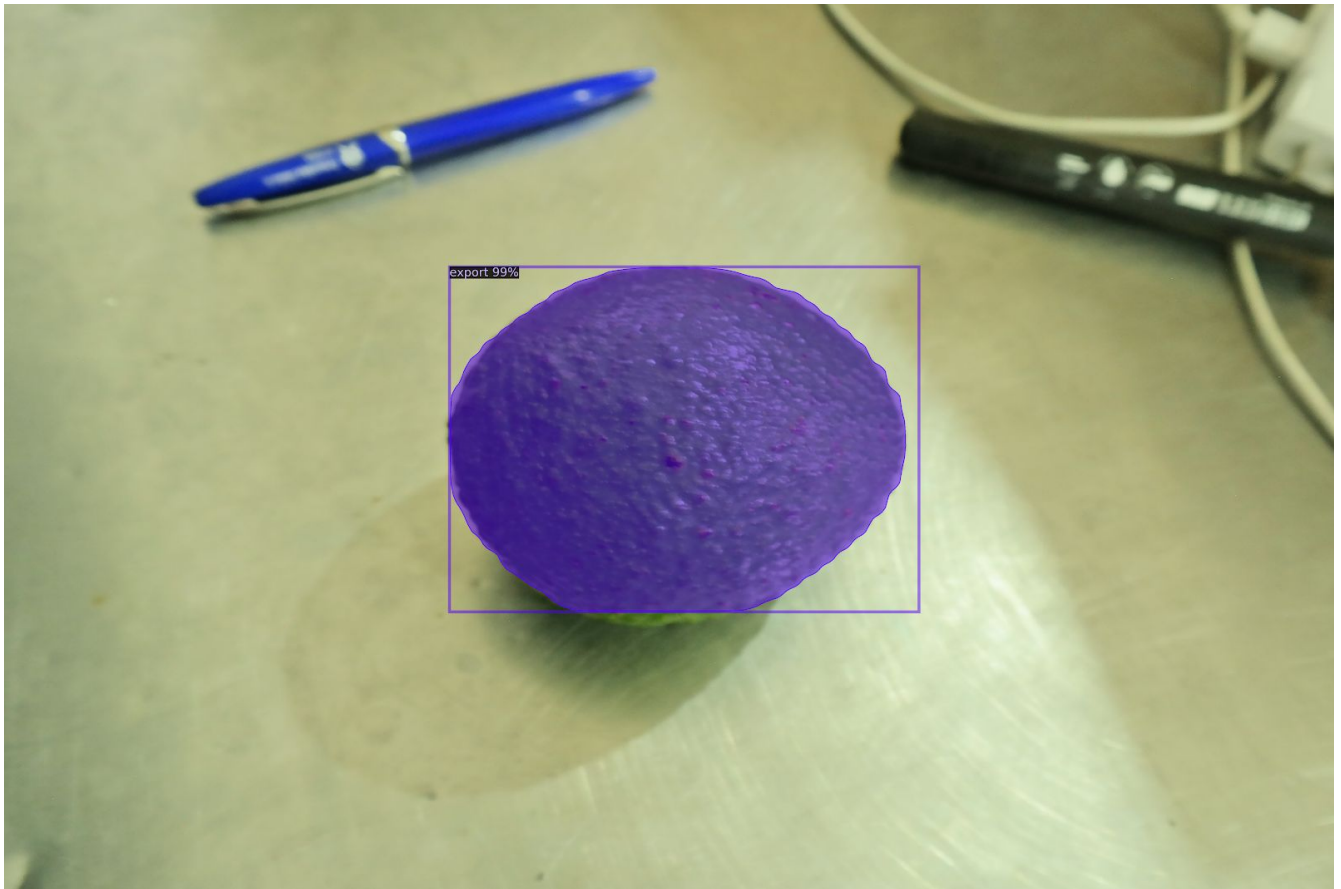


Imagen 16, Resultado del inference usando el modelo final después del entrenamiento detección de aguacate para exportación, fuente: autoría propia.

9.5. Resultados.

Los resultados en la detección de objetos se miden según su mAP (mean Average Precision).

“Average precision (AP), the area under the recall-precision (RP) curve, is the standard performance measure for object detection. Despite its wide acceptance, it has a number of shortcomings, the most important of which are (i) the inability to distinguish very different RP curves, and (ii) the lack of directly measuring bounding box localization accuracy.” Oksuz et all. (2018)

```

Evaluate annotation type *segm*
DONE (t=0.10s).
Accumulating evaluation results...
DONE (t=0.02s).
Average Precision (AP) @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.953
Average Precision (AP) @[ IoU=0.50      | area= all | maxDets=100 ] = 0.991
Average Precision (AP) @[ IoU=0.75      | area= all | maxDets=100 ] = 0.991
Average Precision (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.953
Average Recall (AR) @[ IoU=0.50:0.95 | area= all | maxDets= 1 ] = 0.862
Average Recall (AR) @[ IoU=0.50:0.95 | area= all | maxDets= 10 ] = 0.971
Average Recall (AR) @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.971
Average Recall (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.971
[01/19 20:51:19 d2.evaluation.coco_evaluation]: Evaluation results for segm:
| AP | AP50 | AP75 | APs | APm | APl |
|:---:|:---:|:---:|:---:|:---:|:---:|
| 95.350 | 99.097 | 99.097 | -100.000 | -100.000 | 95.350 |
[01/19 20:51:19 d2.evaluation.coco_evaluation]: Per-category segm AP:
| category | AP | category | AP | category | AP |
|:---:|:---:|:---:|:---:|:---:|:---:|
| Avocado_mature | 91.238 | nationals | 97.585 | export | 97.227 |
OrderedDict([('bbox',
  {'AP': 79.04568051645326,
   'AP-Avocado_mature': 76.81631956299077,
   'AP-export': 84.15165118996372,
   'AP-nationals': 76.16907079640522,
   'AP50': 95.96598107624644,
   'AP75': 94.66860995982445,
   'APl': 79.04568051645326,
   'APm': -100.0,
   'APs': -100.0}),
 ('segm',
  {'AP': 95.34982420184558,
   'AP-Avocado_mature': 91.23762376237624,
   'AP-export': 97.22705603893722,
   'AP-nationals': 97.58479280422321,
   'AP50': 99.09714337820763,
   'AP75': 99.09714337820763,
   'APl': 95.34982420184558,
   'APm': -100.0,
   'APs': -100.0})])

```

Imagen 17, Resultado arrojado después del entrenamiento, fuente autoría propia.

En la imagen 17 podemos apreciar los resultados por cada categoría en la escala propuesta.

Avocado_mature (Aguacate maduro) = **91.238 %** de precisión.

nationals (Aguacate de comercialización nacional) = **97.585 %** de precisión.

export (Aguacate de comercialización internacional) = **97.227 %** de precisión.

9.5.2 Precision

Presición, esta da cuenta de qué tanta precisión hay en las predicciones, en otras palabras el porcentaje de las predicciones que son verdaderas:

$$Precision = TP / (TP + FP)$$

TP (true positive) FP (false positives)

9.5.3 Recall.

Recall es la medida de que tan bien se encuentran los positivos.

$$Recall = TP / (TP + FN)$$

FN (False Negatives)

9.5.4 IoU (Intersection over union).

La intersección sobre la unión mide que tan desfasada está la predicción del objeto real, por lo regular a esta predicción se le asigna un threshold en algunos datasets:

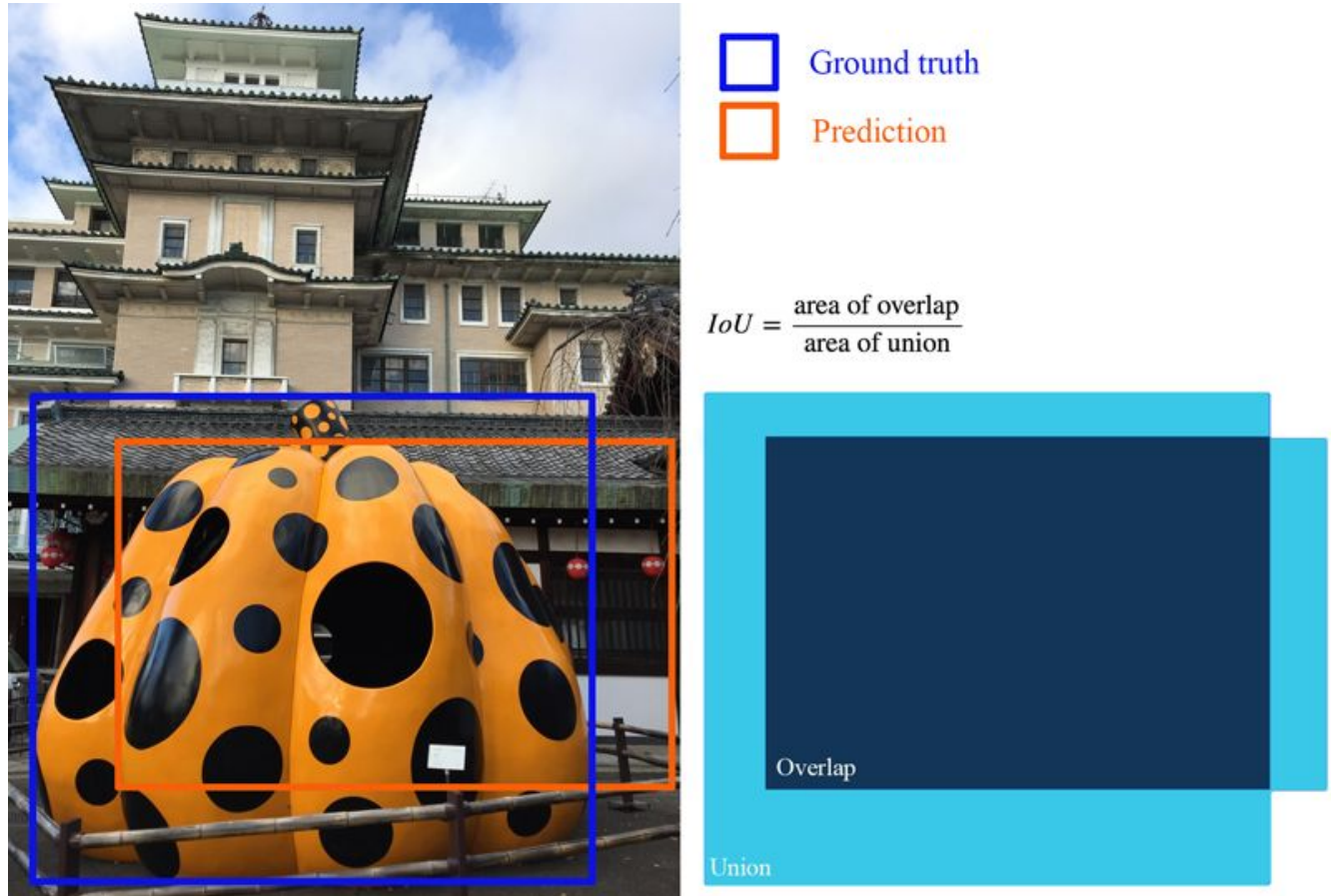


Imagen 18, como se calcula la intersección sobre la unión, fuente :

https://medium.com/@jonathan_hui/map-mean-average-precision-for-object-detection-45c121a31173

9.6. Validación.

La validación de modelo fue realizada el día 26 de noviembre del 2019 en una visita técnica a la exportadora Hass Diamond, en la que se realizó la toma de fotografías a 34 aguacates ya previamente

clasificados por Hass Diamond (ver imagen 5), las 34 imágenes fueron tomadas con las 3 categorías diferentes (Aguacates maduros, Aguacates para comercializar a nivel nacional, Aguacates para exportar). A continuación se presentan los resultados del modelo funcionando en dichas imágenes:



Imagen 19: validación del modelo final aguacate tipo exportación, resultado verdadero, fuente: autoría propia.



Imagen 20: validación del modelo final aguacate tipo exportación, resultado verdadero, fuente: autoría propia.

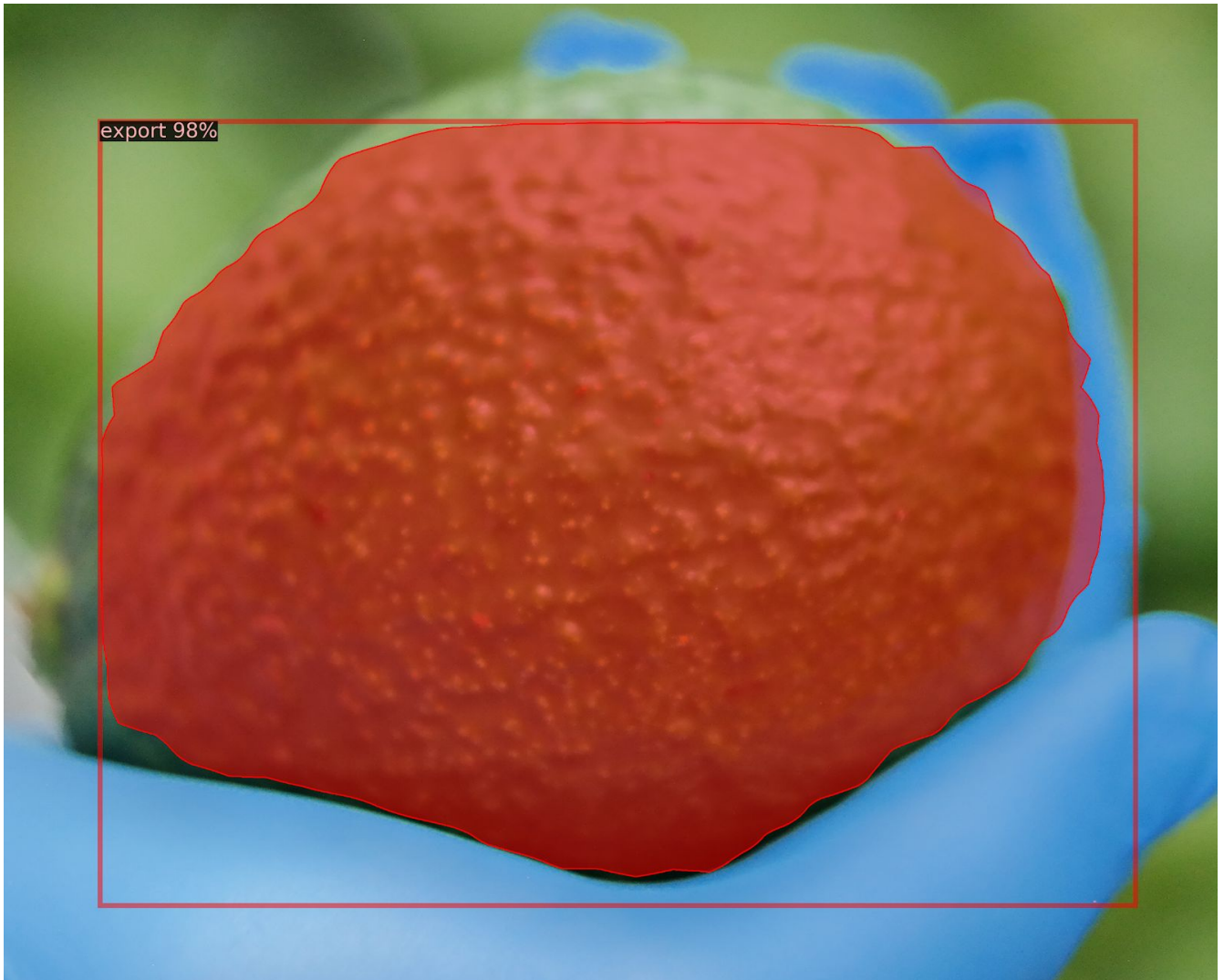


Imagen 21: validación del modelo final aguacate tipo exportación, resultado verdadero, fuente: autoría propia.



Imagen 22: validación del modelo final aguacate tipo exportación, resultado verdadero, fuente: autoría propia.

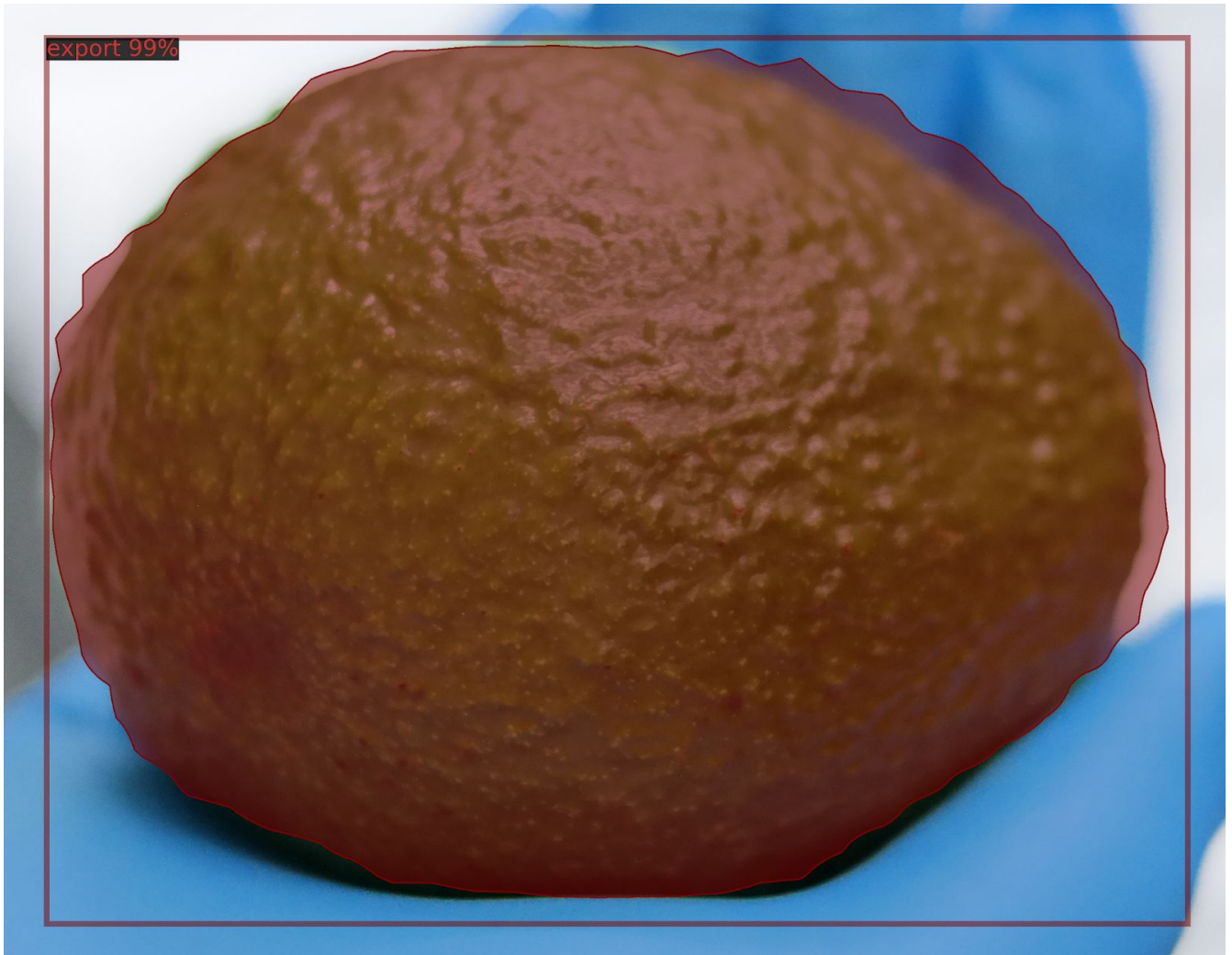


Imagen 23: validación del modelo final aguacate tipo exportación, resultado verdadero, fuente: autoría propia.

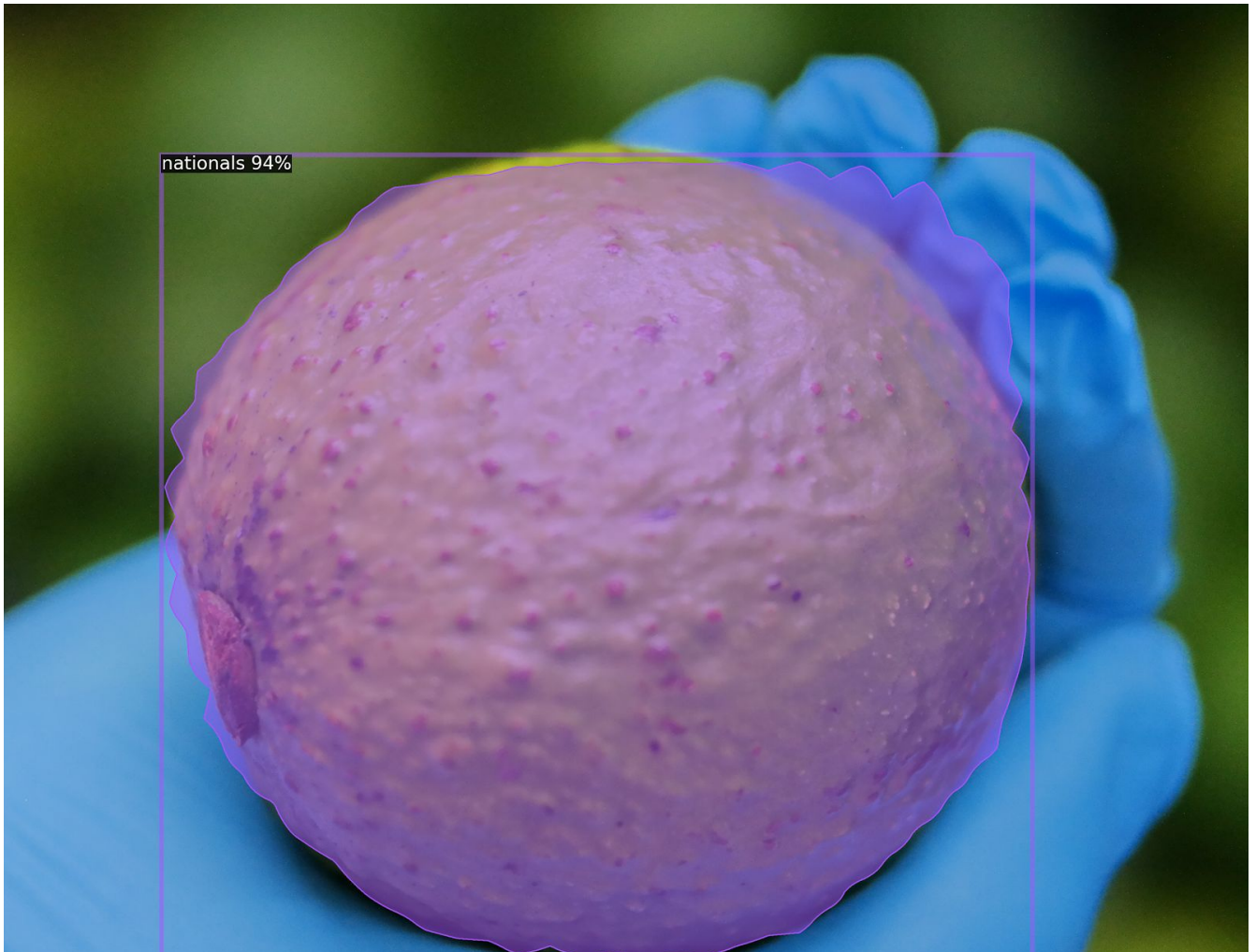


Imagen 24: validación del modelo final aguacate tipo nacionales, resultado verdadero, fuente: autoría propia.

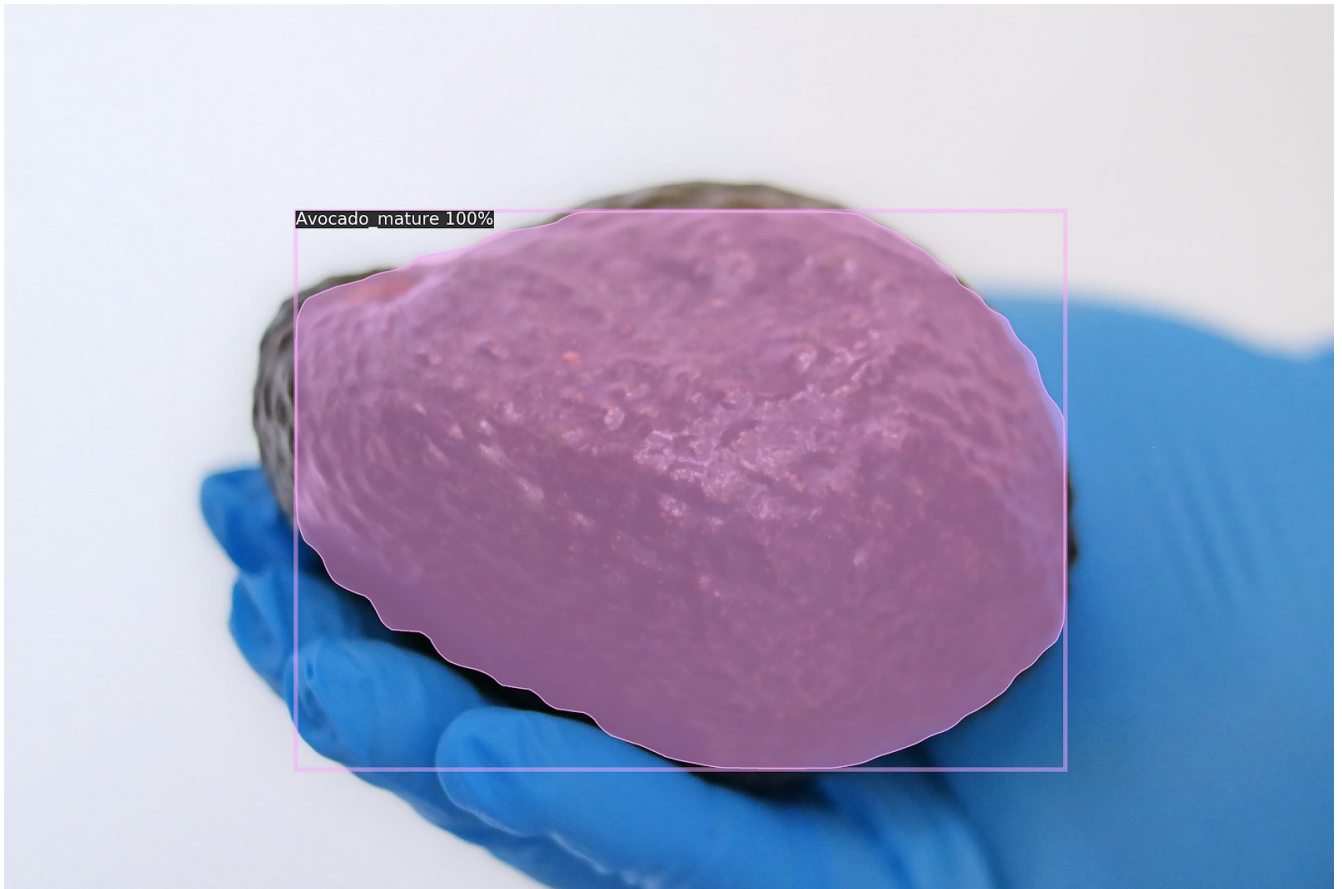


Imagen 25: validación del modelo final aguacate tipo nacionales, resultado verdadero, fuente: autoría propia.

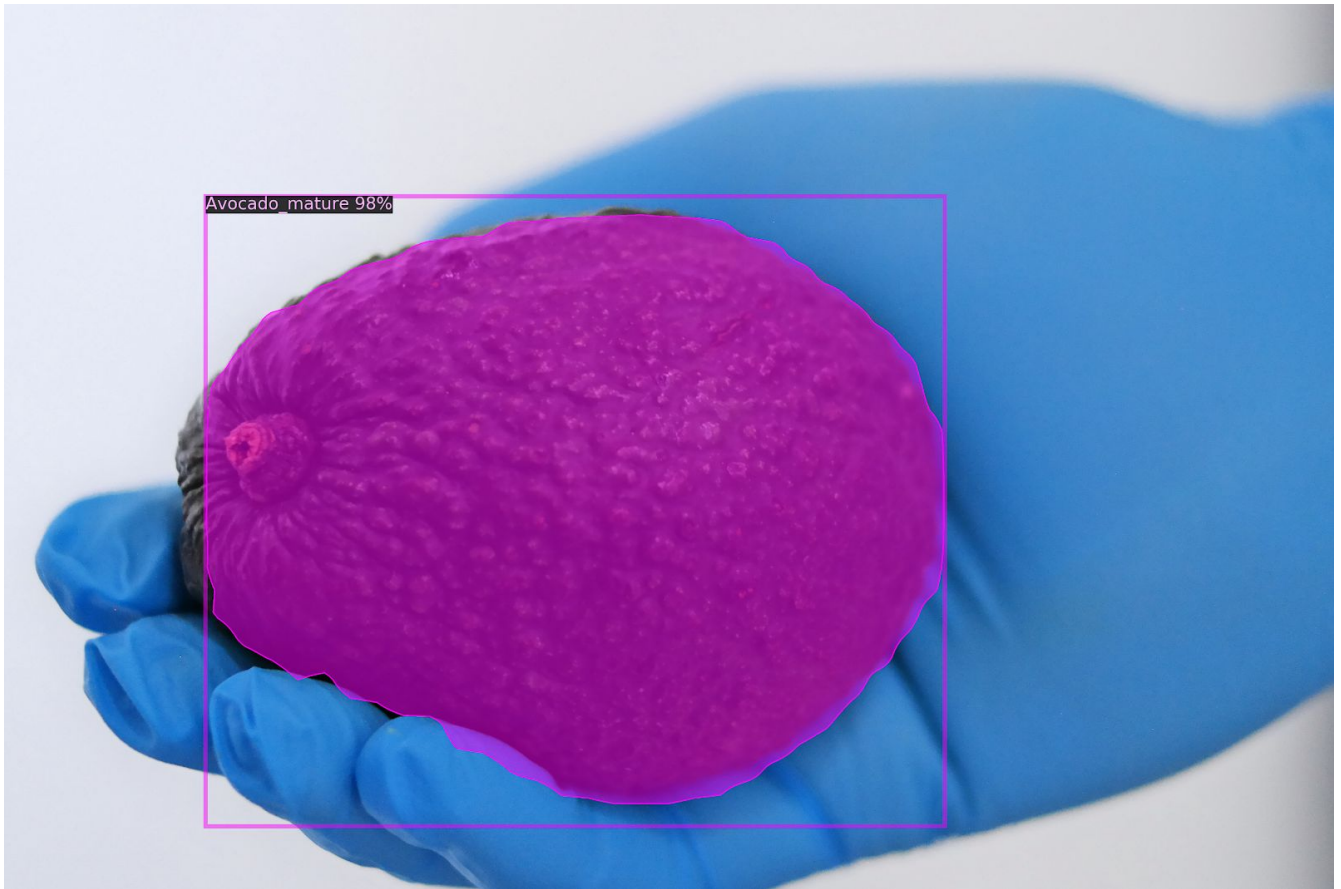


Imagen 26: validación del modelo final aguacate tipo maduro, resultado verdadero, fuente: autoría propia.

10. Conclusiones

1. Es posible usar el método propuesto para clasificar el aguacate hass en sus diferentes estados para comercialización usando las redes neuronales convolucionales.
2. Es posible realizar la clasificación del aguacate hass en sus estados comerciales sin la necesidad de tener entornos controlados con el método propuesto en este proyecto.
3. Se propuso una escala de clasificación que obedece a las necesidades de comercialización en la agroindustria del aguacate actual.

11. Trabajos futuros

1. Es necesario continuar enriqueciendo el dataset para poder mejorar la red neuronal.
2. La presente investigación deja las bases para la construcción de un sistema robusto en el que puedan intervenir diferentes áreas interdisciplinarias y poder así continuar en la creación de soluciones a las necesidades agroindustriales en Colombia.

12. Bibliografía

12.1 Artículos académicos (papers)

Arzate-Vázquez, I., Chanona-Pérez, J. J., de Perea-Flores, M. J., Calderón-Domínguez, G., Moreno-Armendáriz, M. A., Calvo, H., ... Gutiérrez-López, G. (2011). Image Processing Applied to Classification of Avocado Variety Hass (Persea americana Mill.) During the Ripening Process. Food and Bioprocess Technology, 1307-1313. <https://doi.org/10.1007/s11947-011-0595-6>

Baek, S., Jang, J. and Yoon, S. (2019). End-to-End Blood Pressure Prediction via Fully Convolutional Networks. IEEE Access, 7, pp.185458-185468.

Basha, S., Dubey, S., Pulabaigari, V. and Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. Neurocomputing, 378, pp.112-119.

Bi, L., Feng, D. and Kim, J. (2018). Dual-Path Adversarial Learning for Fully Convolutional Network (FCN)-Based Medical Image Segmentation. The Visual Computer, 34(6-8), pp.1043-1052.

Choque, W. & Salazar, C. (2016). Desarrollo de algoritmo y prototipo móvil para medir el grado de madurez del aguacate hass mediante procesamiento digital de imágenes (tesis de pregrado) Universidad Tecnológica de Pereira, Pereira, Colombia. Recuperado de:

<http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/8174/00642C549.pdf?sequence=1&isAllowed=y>

Délano R., P. (2019). Inteligencia artificial en otorrinolaringología. Revista de otorrinolaringología y cirugía de cabeza y cuello, 79(1), pp.7-7.

Fish, K., Barnes, J. and AikenAssistant, M. (1995). Artificial neural networks: A new methodology for industrial market segmentation. *Industrial Marketing Management*, 24(5), pp.431-438.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P. and Garcia-Rodriguez, J. (2018). A survey on deep learning techniques for image and video semantic segmentation. *Applied Soft Computing*, 70, pp.41-65.

Girshick, R., Donahue, J., Darrell, T., Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. Recuperado de: <https://arxiv.org/pdf/1311.2524.pdf>

Girshick, Ross. (2015). Fast r-cnn. 10.1109/ICCV.2015.169.

He, Kaiming & Gkioxari, Georgia & Dollar, Piotr & Girshick, Ross. (2017). Mask R-CNN. 2980-2988. 10.1109/ICCV.2017.322.

Inoue, K. (2019). Semantic segmentation of breast lesion using deep learning. *Ultrasound in Medicine & Biology*, 45, p.S52.

Jiménez, D., & Rendón, H. (2012). Determinantes y efectos de la Inversión Extranjera Directa: revisión de literatura. Recuperado de: <https://revistas.unal.edu.co/index.php/ede/article/view/35873/43643>

Kaijun, G., Shaoyuan, S., Guangshun, Y. and Haitao, Z. (2017). Semantic segmentation of night vision images for unmanned vehicles based on deep learning. *Journal of Applied Optics*, 38(3), pp.567-572.

Kang, W., Xiang, Y., Wang, F., Wan, L. and You, H. (2018). Flood Detection in Gaofen-3 SAR Images via Fully Convolutional Networks. *Sensors*, 18(9), p.2915.

Khalifa, N., Taha, M., Hassanien, A. and Mohamed, H. (2019). Deep Iris: Deep Learning for Gender Classification Through Iris Patterns. *Acta Informatica Medica*, 27(2), p.96.

Kirillov, Alexander & He, Kaiming & Girshick, Ross & Rother, Carsten & Dollár, Piotr. (2018). Panoptic Segmentation.

Lapko, A. and Lapko, V. (2019). Face DetectNet: face detection via fully-convolutional network. *Computer Optics*, 43(2), pp.238-244.

Lateef, F. and Ruichek, Y. (2019). Survey on semantic segmentation using deep learning techniques. *Neurocomputing*, 338, pp.321-348.

Li, C., Zhang, X., Liao, L., Jin, L. and Yang, W. (2019). Skeleton-Based Gesture Recognition Using Several Fully Connected Layers with Path Signature Features and Temporal Transformer Module. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, pp.8585-8593.

Lin, C., Li, Y. and Lee, C. (2020). Using Fully Convolutional Networks for Floor Area Detection. *Sensors and Materials*, 32(1), p.159.

Lin, Tsung-Yi & Maire, Michael & Belongie, Serge & Hays, James & Perona, Pietro & Ramanan, Deva & Dollár, Piotr & Zitnick, C.. (2014). Microsoft COCO: Common Objects in Context.

Liu, T., Abd-Elrahman, A., Morton, J. and Wilhelm, V. (2018). Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GIScience & Remote Sensing*, 55(2), pp.243-264.

Méndez, R., Auat, F. y Rosell-Polo, J. (2017). Flexible System of Multiple RGB-D Sensors for

Measuring and Classifying Fruits in Agri-food Industry. Department of Electronic Engineering.

Universidad Técnica Federico Santa María, Valparaíso, Chile. doi: 10.1016/j.compag.2017.05.014.

Minar, Matiur Rahman & Naher, Jibon. (2018). Recent Advances in Deep Learning: An Overview.

10.13140/RG.2.2.24831.10403.

Moreno, G., Forero, M., Gómez, F., Ramírez, M. (2019) Hass avocado classification by color and

volume using a Kinect sensor, Proc. SPIE 11137, Applications of Digital Image Processing XLII,

111371P ; doi: 10.1117/12.2529843.

Negri, G., Nazário, F., Oliveira, J. and Nied, A. (2018). Back-emf Based Rotor Position Estimation

For Low Cost Pmsm Drive Using Fully Connected Cascade Artificial Neural Networks. *Eletrônica de*

Potência, 23(1), pp.69-77.

Oksuz, Kemal & Cam, Baris & Akbas, Emre & Kalkan, Sinan. (2018). Localization Recall

Precision (LRP): A New Performance Metric for Object Detection.

O'Shea, K., & Nash, R. (2015). An Introduction to Convolutional Neural Networks. ArXiv,

abs/1511.08458.

Pinto, J., Rueda-Chacón, H., Arguello, H. (2019). Classification of Hass avocado (*persea americana*

mill) in terms of its ripening via hyperspectral images. vol. 22, no. 45, pp. 109-128, doi:

10.22430/22565337.1232

Posada, E. (2002). Empresarios y ganaderos en la Costa Atlántica (1850-1950). En: *Empresas y*

empresarios en la historia de Colombia. Siglos XIX y XX. Una colección de estudios recientes. Dávila,

C. (comp.). Bogotá: Colombia, Ediciones Norma. P. 61-82.

Ren, Y., Zhu, C. and Xiao, S. (2018). Object Detection Based on Fast/Faster RCNN Employing Fully Convolutional Architectures. *Mathematical Problems in Engineering*, 2018, pp.1-7.

Shekastehband, B., Azaraxsh, A. and Showkati, H. (2018). Experimental seismic study on shear walls with fully-connected and beam-only-connected web plates. *Journal of Constructional Steel Research*, 141, pp.204-215.

R. Girshick, (2015), "Fast R-CNN," *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 2015, pp. 1440-1448. doi: 10.1109/ICCV.2015.169

Ramirez, C., & Florez, L. (2006). apuntes de inversión extranjera directa: definiciones, tipología y casos de aplicación colombianos. *apuntes de economía*. No 8. Recuperado de:

http://www.icesi.edu.co/departamentos/images/departamentos/FCAE/economia/apuntesEconomia/apuntes_de_inversion_extranjera_directa.pdf

Ren, S., He, K., Girshick, R. and Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), pp.1137-1149.

Romera-Paredes, B. and Torr, P. (2016). Recurrent Instance Segmentation. *Computer Vision – ECCV 2016*, pp.312-329.

Uijlings, J., Sande, K., Gevers, T. & Smeulders, A. (2012). Selective search for object recognition. Recuperado de: <http://www.huppelen.nl/publications/selectiveSearchDraft.pdf>

Vieux, R., Benois-Pineau, J., Domenger, J. and Braquelaire, A. (2010). Segmentation-based multi-class semantic object detection. *Multimedia Tools and Applications*, 60(2), pp.305-326.

- Wang, J., Luo, J., Liu, B., Feng, R., Lu, L. and Zou, H. (2019). Automated diabetic retinopathy grading and lesion detection based on the modified R-FCN object-detection algorithm. *IET Computer Vision*.
- Xia, X., Persello, C. and Koeva, M. (2019). Deep Fully Convolutional Networks for Cadastral Boundary Detection from UAV Images. *Remote Sensing*, 11(14), p.1725.
- Xie, X., Du, D., Li, Q., Liang, Y., Tang, W., Ong, Z., Lu, M., Huynh, H. and Goh, R. (2017). Exploiting Sparsity to Accelerate Fully Connected Layers of CNN-Based Applications on Mobile SoCs. *ACM Transactions on Embedded Computing Systems*, 17(2), pp.1-25.
- Xu, Q. and Pan, G. (2017). SparseConnect: regularising CNNs on fully connected layers. *Electronics Letters*, 53(18), pp.1246-1248.
- Xu, Q., Zhang, M., Gu, Z. and Pan, G. (2019). Overfitting remedy by sparsifying regularization on fully-connected layers of CNNs. *Neurocomputing*, 328, pp.69-74.
- Xue, Y. and Li, Y. (2018). A Fast Detection Method via Region-Based Fully Convolutional Neural Networks for Shield Tunnel Lining Defects. *Computer-Aided Civil and Infrastructure Engineering*, 33(8), pp.638-654.
- Zhang, G. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 30(4), pp.451-462.
- Zhang, J., He, J., Chen, T., Liu, Z. and Chen, D. (2019). Abnormal region detection in cervical smear images based on fully convolutional network. *IET Image Processing*, 13(4), pp.583-590.
- Zhang, Y. and Mu, Z. (2017). Ear Detection under Uncontrolled Conditions with Multiple Scale Faster Region-Based Convolutional Neural Networks. *Symmetry*, 9(4), p.53.

Zhao, Z., Zheng, P., Xu, S. & Wu, X. (2019). Object Detection with Deep Learning: A Review.

Recuperado de: <https://arxiv.org/pdf/1807.05511.pdf>

12.2 Webgrafía

Aguacate: el oro verde de la economía colombiana. (30 de marzo de 2017). Dinero. Recuperado de:

<https://www.dinero.com/edicion-impresia/informe-especial/articulo/aguacate-exportacion-y-mercado-en-colombia/243434>

“Colombia será uno de los líderes mundiales en aguacate hass”: Iván Duque. (29 de octubre de 2019). Dinero. Recuperado de:

<https://www.dinero.com/pais/articulo/colombia-quiere-ser-uno-de-los-mayores-exportadores-de-aguacate/278520>

FAO. (1995). NORMA DEL CODEX PARA EL AGUACATE. Recuperado de:

http://www.fao.org/input/download/standards/321/CXS_197s.pdf

Figuroa, H (03 de junio de 2018). Colombia podría ser el segundo productor mundial de aguacate Hass. El Universal. Recuperado de:

<https://www.eluniversal.com.co/economica/colombia-podria-ser-el-segundo-productor-mundial-de-aguacate-hass-279989-HBeu395882>

Hass Diamond. (2019). INSTRUCTIVO CALIDAD DE EXPORTACIÓN Y CRITERIO DE COSECHA.

ICONTEC. (2018). END 094 FRUTAS FRESCAS, VARIEDAD AGUACATE HASS. Recuperado de: <http://www.analdex.org/wp-content/uploads/2016/02/END-094.pdf>

Lucia, La. (13 de marzo de 2019). Camposol se expande a Colombia: más hectáreas para la producción de paltas. SemanaEconomica. Recuperado de:

<https://semanaeconomica.com/sectores-empresas/agropecuario/271230-camposol-se-expande-a-colombia-mas-hectareas-para-la-produccion-de-paltas>

Marin, L. (08 de noviembre de 2019). Primera exportación de aguacate hass a Japon sale desde Risaralda. Recuperado de:

<https://www.rcnradio.com/colombia/eje-cafetero/primera-exportacion-de-aguacate-hass-japon-sale-desde-risaralda>

Medium. (2020). What's New in Deep Learning Research: How Google Uses Reinforcement Learning to Ask All the Right.... [online] Available at:

<https://towardsdatascience.com/whats-new-in-deep-learning-research-how-google-uses-reinforcement-learning-to-ask-all-the-right-69c172f113c4>

Morales, M., (23 de mayo del 2016). El 65,8 % de la tierra apta para sembrar en Colombia no se aprovecha. Recuperado de:

<https://www.eltiempo.com/archivo/documento/CMS-16601436>

Procolombia. ¿Qué es PROCOLOMBIA?. recuperado de:

<http://www.procolombia.co/nosotros/que-es-procolombia>

Siguen creciendo las exportaciones de aguacate. (13 de septiembre de 2019). Dinero. Recuperado de: <https://www.dinero.com/economia/articulo/siguen-creciendo-las-exportaciones-de-aguacate/276829>

12.3 Softwares.

Brooks, J. (2019). COCO Annotator. <https://github.com/jsbroks/coco-annotator/>.

Welcome to detectron2's documentation! — detectron2 0.1 documentation. (2019). Retrieved 19 October 2019, from <https://detectron2.readthedocs.io/>

Wu, Y., Kirillov, A., Massa, F., Wan-Yen, L., & Girshick, R. (10 de octubre de 2019). Detectron2: A PyTorch-based modular object detection library. Recuperado de:

<https://ai.facebook.com/blog/-detectron2-a-pytorch-based-modular-object-detection-library/>

13. Anexos.

En este apartado se anexara el código completo desde la instalación hasta el entrenamiento de la red neuronal.

Anexo 1: Instalación dtectron2

```
# install dependencies
pip install -U torch torchvision cython

pip install -U 'git+https://github.com/facebookresearch/fvcore.git'
'git+https://github.com/cocodataset/cocoapi.git#subdirectory=PythonAPI'

import torch, torchvision

torch.__version__
```



```
git clone https://github.com/facebookresearch/detectron2 detectron2_repo
pip install -e detectron2_repo
```

Anexo 2: Código

```
# Code

# Some basic setup

# Setup detectron2 logger

import detectron2

from detectron2.utils.logger import setup_logger

setup_logger()

# import some common libraries

import numpy as np

import cv2

from google.colab.patches import cv2_imshow

# import some common detectron2 utilities

from detectron2.engine import DefaultPredictor
from detectron2.config import get_cfg
from detectron2.utils.visualizer import Visualizer
from detectron2.data import MetadataCatalog

import os

import numpy as np

import json

from detectron2.structures import BoxMode

import itertools
```

```

from detectron2.data.datasets import register_coco_instances

from detectron2.data.datasets import load_coco_json

register_coco_instances("Avocados_train", {}, "avocados/Train/Train-2.json",
"avocados/Train")

register_coco_instances("Avocados_test", {}, "avocados/Test/Test-3.json",
"avocados/Test")

register_coco_instances("Avocados_val", {}, "avocados/Val/Val-4.json",
"avocados/Val")

MetadataCatalog.get("Avocados_val").thing_classes = ["Avocado_mature",
"nationals", "export"]

MetadataCatalog.get("Avocados_train").thing_classes = ["Avocado_mature",
"nationals", "export"]

MetadataCatalog.get("Avocados_test").thing_classes = ["Avocado_mature",
"nationals", "export"]

import random

avocados_dict = load_coco_json('avocados/Val/Val-4.json', 'avocados/Val',
'Avocados_val')

for d in random.sample(avocados_dict, 3):

    img = cv2.imread(d["file_name"])

    visualizer = Visualizer(img[:, :, ::-1],
metadata=MetadataCatalog.get("Avocados_val"), scale=0.5)

    vis = visualizer.draw_dataset_dict(d)

    cv2_imshow(vis.get_image()[:, :, ::-1])

```

```
# Training

from detectron2.engine import DefaultTrainer

from detectron2.config import get_cfg

cfg = get_cfg()

cfg.merge_from_file("./detectron2_repo/configs/COCO-InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml")

cfg.DATASETS.TRAIN = ("Avocados_train", "Avocados_val")

cfg.DATASETS.TEST = () # no metrics implemented for this dataset

cfg.OUTPUT_DIR = 'drive/My Drive/WEIGHTS/output_without_w_train_1000_iter'

cfg.MODEL.WEIGHTS = "drive/My Drive/WEIGHTS/output_without_w_train_1000_iter/model_final.pth"

cfg.DATALOADER.NUM_WORKERS = 2

cfg.SOLVER.IMS_PER_BATCH = 2

cfg.SOLVER.BASE_LR = 0.00025

cfg.SOLVER.MAX_ITER = 3000 # 3000 iterations seems good enough, but you
can certainly train longer

cfg.MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE = 256 # faster, and good enough
for this dataset

cfg.MODEL.ROI_HEADS.NUM_CLASSES = 3

os.makedirs(cfg.OUTPUT_DIR, exist_ok=True)

trainer = DefaultTrainer(cfg)

trainer.resume_or_load(resume=False)

trainer.train()

# Validation

cfg.merge_from_file("./detectron2_repo/configs/COCO-InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml")
```

```

cfg.OUTPUT_DIR = 'drive/My Drive/WEIGHTS/output_without_w_train_1000_iter'

cfg.MODEL.WEIGHTS = "drive/My
Drive/WEIGHTS/output_without_w_train_1000_iter/model_final.pth"

cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST = 0.8 # set the testing threshold
for this model

cfg.DATASETS.TEST = ("Avocados_val", "Avocados_test")

predictor = DefaultPredictor(cfg)

import random

from detectron2.utils.visualizer import ColorMode

dataset_dicts = load_coco_json('avocados/Val/Val-4.json', 'avocados/Val',
'Avocados_val')

for d in random.sample(dataset_dicts, 3):

    im = cv2.imread(d["file_name"])

    outputs = predictor(im)

    v = Visualizer(im[:, :, ::-1],

                    metadata=MetadataCatalog.get("Avocados_val"),

                    scale=0.8

    )

    v = v.draw_instance_predictions(outputs["instances"].to("cpu"))

    cv2_imshow(v.get_image()[:, :, ::-1])

# Results

from detectron2.evaluation import COCOEvaluator, inference_on_dataset

from detectron2.data import build_detection_test_loader

evaluator = COCOEvaluator("Avocados_val", cfg, False, output_dir="./drive/My
Drive/Wdrive/My
Drive/WEIGHTS/output_without_w_train_1000_iter/model_final.pthEIGHTS/output_
without_w_train_1000_iter/")

```

```
val_loader = build_detection_test_loader(cfg, "Avocados_val")  
inference_on_dataset(trainer.model, val_loader, evaluator)
```