AVOCADROID

APLICATIVO ANDROID PARA LA DETECCIÓN DE GRADO DE MADUREZ DEL AGUACATE HASS CON TENSORFLOW

Saturdays.Ai @Bosch Connectory

PRESENTAN

María Inés Calderón Zetter

Rosa Elena Rodríguez Robles

Carlos Saúl Hernández Nuño

Leonardo Covarrubias García

Este proyecto se basa en el procesamiento de imágenes para determinar el estado de madurez de frutas, en este caso se utilizará el aguacate, la determinación se hará mediante la captura de una imagen, de la cual se analizarán diferentes características que posee la fruta, de la cuales se tomará como principal su color promedio.

1. Eje rector del proyecto

Alimenticio, Financiero e Industrial; esta herramienta se puede adaptar para los tres usos, de forma:

1.1 Alimenticio

Dicha herramienta, fomenta el consumo oportuno del producto que con base en la existencia de variedades de aguacate y con una predilección del consumo por el Aguacate Hass, existe una alta demanda de este producto. México se posiciona como líder mundial en la exportación de esta variedad de aguacate.

1.2 Financiero

Con datos del 2016, las exportaciones mexicanas representaron un porcentaje muy significativo de las importaciones de aguacate en los siguiente países: Guatemala, 100%; Canada. 95.41%; Japón, 92.72%; Estados Unidos, 91.32%; y El Salvador, 90.23%

En el 2030, se estima un aumento de la demanda mundial de 2.84 a 4.24 MMt (un crecimiento acumulado de 48.98%), mientras que la producción nacional de aguacate tiene la capacidad de incrementarse a 1.89 a 3.16 MMt, lo cual representa un crecimiento acumulado de 67.28%. Ante dicho escenario es factible destinar 1.02 MMt al consumo nacional y 2.14 MMt a las exportaciones.

1.3 Industrial

En el contexto productivo, de las 205,250 hectáreas sembradas en el año 2016, el 88.31% de la superficie se encuentra mecanizada, 82.34% cuenta con tecnología aplicada a la sanidad vegetal, mientras que 76.65% del territorio sembrado con este cultivo contó con asistencia técnica. Por otro lado, 45.98% de la producción es de modalidad de riego y el resto de temporal; estos datos sustentan la innovación en la industria agroalimentaria de manera que nuestro aplicativo abordaría la perspectiva del consumidor.

2. Planteamiento del Problema

Se estima que un tercio de todos los alimentos producidos a nivel mundial se pierden o se desperdician. (1) En México, cada año cerca de 20 millones de toneladas de alimentos se pierden o se desperdician en nuestro país. El desperdicio ocurre durante la distribución y consumo de los alimentos, en tanto que las pérdidas se producen durante la producción, poscosecha, almacenamiento y transporte. (2)

De acuerdo a un estudio realizado por el Banco Mundial en el que se compararon 79 alimentos diferentes, el aguacate figura como el sexto alimento más desperdiciado en

México, con un total de 312 mil 812 toneladas desperdiciadas anualmente, lo que equivale al 39.8% del total que se produce de esta fruta. (3)

Debido a todas estas consideraciones se desarrolla este proyecto en pro de un manera más eficiente y confiable de determinar el estado de madurez de esta fruta, y poder llevar las tecnologías a estos campos y tener un mayor control en cuanto la calidad del producto lo permita.

3. Descripción de la solución a la problemática detectada.

Base, en el desarrollo del algoritmo de reconocimiento de imágenes se busca lograr determinar el grado de madurez (*inicialmente*) de un aguacate; esto con la diferenciación de tres clases: 01 Verde, 02 Maduro, 03 *Pasado*.

En general, contribuir en una reducción del desperdicio de alimentos "ref. refiere a la disminución en la cantidad o calidad de los alimentos como resultado de las decisiones y acciones de los minoristas, proveedores de servicios alimentarios y consumidores" FAO, 2019., facilitando el consumo oportuno desde la experiencia de compra.

La determinación del estado de maduración de las frutas por medio de dispositivos computacionales o también llamada visión artificial se está volviendo fundamental debido a las ventajas que presenta para identificar diferentes características a objetos de diferentes formas, colores y tamaños. Estas ventajas ofrecidas por estos dispositivos se emplean debido a su rapidez y la exactitud con la cual arrojará los resultados mediante la detección y la clasificación de aspectos de interés de la fruta obtenidos mediante la visión artificial.

Debido al avance tecnológico de los últimos años se ha permitido llevar todas estas tecnologías de visión artificial a diferentes áreas de la industria, una de estas es la de la exportación de frutas y en la actualidad la mayoría de clasificación de estas frutas se hacen manualmente y utilizando un buen observador y conocedor de estas. No es fácil a simple vista ver las cualidades que presenta el aguacate y es preciso tener un contacto de forma tangible para determinar el estado de madurez en el que se encuentra, y esto conlleva a que el aguacate se vea sometido a alteración de su integridad

4. Hipótesis

"El uso de una red neural convolucional entrenada en un set imágenes lo suficientemente grande de aguacates hass en diferentes estados de maduración permitirá reconocer mediante una aplicación móvil y el uso de la cámara de un smartphone el grado de madurez de un aguacate hass"

"detectar el grado de madurez de un aguacate con reconocimiento de imágenes soportado con un algoritmo de machine learning para dispositivos móviles"

Abordamos la problemática buscando que cualquier persona que cuente con un smartphone pueda conocer el grado de madurez de una fruta, proponiendo inicialmente el aguacate. Si contamos con una aplicación móvil de fácil uso para catalogar los frutos dependiendo su grado de madurez, podremos otorgarle a los usuarios una herramienta de selección

responsable para su consumo, para que de esta forma una persona que adquiera este tipo de fruto pueda conocer su grado de madurez y así definir qué tipo de consumo desea, reduciendo el desperdicio de este fruto.

5. Metodología

5.1 Descripción de los Datos

Los datos que se utilizaron como base de entrenamiento incluyen:

- 1. Fotografías de aguacate hass adquiridos en Guadalajara, Jal.
- 2. Fotografías de aguacate hass encontrados en internet: WSID-100 (4) y Fruit-Images-Dataset (5).
- 3. Fotografías de diversos objetos (*no aguacate hass*)

Las fotografías no cuentan con alguna característica especial de luz, acomodo, pixelaje, o ensamble de la misma, únicamente especificamos si existe o no existe un aguacate en la misma y en qué grado de madurez suponemos que se encuentra a la vista.

Adicionalmente, previo a la integración de las fotografías realizamos un proceso de tratamiento de imágenes para reducir el peso en bytes de las mismas dejando un promedio de ~800 Kb utilizando la herramienta RIOT (Rapid image optimizer tool) (6)

5.2 Descripción del modelo

El modelo desarrollado consiste en una Red Neuronal Convolucional implementada en el lenguaje de programación Python mediante el uso de la biblioteca y plataforma TensorFlow para la creación de algoritmos, modelos y arquitecturas de Machine Learning.

Para el desarrollo del modelo, se empleó la técnica conocida como *Transferencia de Aprendizaje* o "*Transfer Learning*", la cual permite desarrollar una red neuronal basándose en el conocimiento aprendido por otra red neural la cual está implementada en un modelo pre-entrenado, y que sirve de base para sobre ella construir la arquitectura de red que se desea crear y que hará uso de los pesos ya aprendidos y el entrenamiento ya realizado sobre el modelo base.

Al utilizar transfer learning para la clasificación de imágenes, el modelo utilizado como base generalmente es un modelo pre-entrenado sobre un set de datos muy grande, típicamente para en una tarea de clasificación de imágenes a gran escala. La intuición detrás de la técnica de transfer learning es que si un modelo es entrenado en un set de datos lo suficientemente grande y general, este modelo servirá efectivamente como un modelo genérico del mundo visual. En base a esto, la intención es aprovechar los mapas de características aprendidos por este modelo para evitar iniciar desde cero entrenando un modelo modelo grande en un set de datos grande. En la práctica el correcto uso de la técnica de Transfer Learning puede ayudar a reducir el tiempo de implementación y entrenamiento de un modelo de clasificación de imágenes.

Con base a lo anterior, para el presente proyecto se decidió utilizar esta técnica para reducir el tiempo de desarrollo del modelo. Como modelo base para la transferencia de

aprendizaje, se utilizó el modelo *MobileNetV2*, el cual es un modelo de Red Neural para visión por computadora creado por Google y entrenado en el set de datos de clasificación de imágenes IMAGENET (ILSVRC-2012-CLS), y que tiene como objetivo servir como un modelo de visión por computadora de propósito general lo suficientemente eficiente para ser ejecutado en aplicaciones móviles y en dispositivos de bajo poder de cómputo y restringidos recursos de energía.

Tomando como base el modelo MobileNetV2, se decidió construir sobre éste una arquitectura de red neuronal convolucional compuesta por 4 capas:

- Capa convolucional 2D
- Capa de dropout
- Capa de reducción (pooling) global
- Capa clasificadora totalmente conectada

5.3 **Delimitantes**

- El manejo del aguacate durante y después de la cosecha debe ser cuidadoso para garantizar al consumidor la calidad del fruto que ellos requieren.
- Los operarios que laboran en el campo y la planta empacadora deben conocer bien el producto, su calidad y principales defectos para que no sean considerados como factor de rechazo.
- Identificar fuentes de contaminación del fruto y tomar medidas correctivas para llevar al consumidor productos seguros para consumir.
- Conocer cuales son las mejores condiciones para su manejo, que permita extender su vida comercial.
- Los procesos de cosecha y acondicionamiento del aguacate deben tomar en cuenta los requerimientos de los clientes y consumidores finales en el mercado meta, así como el tiempo de cosecha hasta la exhibición en los puntos de venta y los cambios esperados durante el transporte.

5.4 De datos

En función, del modelo utilizado nuestros datos para la alimentación del database se traduce en fotografías que por términos del mismo entrenamiento requiere de cantidades superiores de 700 fotografías; aquí se presenta una de las limitantes para un flujo de trabajo constante debido a que la recolección inicialmente se determinó como "una a una".

Con avance de trabajo, se encontraron distintos datasets con fotografías que cubren parcialmente los requerimientos, dicho aporte facilitó el proceso reduciendo únicamente en evaluar las fotografías encontradas para posteriormente someterlas en tratamiento.

5.5 Económicas

Los precios del aguacate fluctúan generalmente con la temporada y se dividen entre temporada baja y alta. (7)

5.6 **Temporales**

La temporada alta de cosecha se produce entre los meses de octubre y febrero, mientras que la temporada media es entre marzo y mayo, y la baja entre junio y septiembre. (7)

6. Marco teórico

Madurez Fisiológica: la fruta se encuentra fisiológicamente madura cuando ha logrado un desarrollo normal y posee un tamaño en el cual puede continuar madurando después de estar cosechado.

Madurez Hortícola: en este estado la fruta se encuentra en un estado apto para el consumo final. Madurez Organoléptica o Consumo: es este estado la fruta tiene características (color, sabor, aroma, textura) para su consumo.

Procesamiento de Imágenes: Es el conjunto de las técnicas que se aplican a las imágenes digitales con la finalidad de mejorar la calidad o la extracción de información de interés.

El área de la ciencia y tecnología de alimentos es uno de los campos donde más se han empleado las técnicas de procesamiento de imágenes. A continuación se mencionan algunos trabajos que han sido publicados en este campo:

Mendoza y Aguilera (2004) implementaron un sistema de visión por computadora para identificar las etapas de maduración de plátanos basándose en el color, el desarrollo de puntos cafés y en la textura de la imagen del fruto. Los resultados obtenidos por ellos, mostraron que existen pequeñas variaciones en los datos de color y de apariencia, y que esta es una técnica simple de clasificación muy confiable para identificar las etapas de maduración del plátano en comparación de la percepción visual. Usando las coordenadas del espacio de color CIELab (L*, a* y b*), el porcentaje de área color café y el contraste, fue posible clasificar una muestra de 49 plátanos en 7 etapas de maduración con una exactitud del 98%.

Recientemente, Pérez-Nieto y col. (2010) aplicaron técnicas de procesamiento de imágenes para estudiar los cambios de la estructura de la masa y su relación con la temperatura de la masa, la pérdida de masa y la altura de pan durante la etapa de cocción. Los resultados mostraron que durante la cocción, se presentó un fenómeno de coalescencia entre las burbujas, y se observó un máximo de dicho fenómeno en las primeras etapas de cocción (<250 s). Este fenómeno fue asociado con un incremento en la altura de la masa, sin un cambio significativo en la temperatura del centro de la masa. Además, mencionaron que la dimensión fractal de la textura de la miga, la dimensión fractal del contorno de las células y las características de la miga de grano fueron indicadores útiles para estudiar la coalescencia, así como también para distinguir el tiempo de cocción necesario para que las características de la miga se mantuvieran constantes.

También las técnicas de procesamiento de imágenes han sido utilizadas para estudiar los procesos de oscurecimiento enzimático de frutos, como por ejemplo Quevedo y col. (2011) emplearon el análisis de textura de imágenes para cuantificar el oscurecimiento de tres variedades de aguacates. En base a los resultados obtenidos, concluyen que en

general, el parámetro de dimensión fractal puede ser usado para evaluar las cinéticas de oscurecimiento y discriminar entre variedades de aguacates.

De más está decir que el sentido de la visión es uno de los grandes prodigios de la naturaleza. En fracciones de segundos, podemos identificar objetos dentro de nuestro campo de visión, sin siquiera detenernos a pensar en ello. Pero no sólo podemos nombrar estos objetos que observamos, sino que también podemos percibir su profundidad, distinguir perfectamente sus contornos, y separarlos de sus fondos. De alguna manera los ojos captan datos de píxeles, pero el cerebro transforma esa información en características más significativas - líneas, curvas y formas - que podrían indicar, por ejemplo, que estamos mirando a una persona.

En función que el área del cerebro responsable de la visión es una de las zonas más estudiadas y que más conocemos; sabemos que la corteza visual contiene una disposición jerárquica compleja de neuronas. Por ejemplo, la información visual es introducida en la corteza a través del área visual primaria, llamada V1. Las neuronas de V1 se ocupan de características visuales de bajo nivel, tales como pequeños segmentos de contorno, componentes de pequeña escala del movimiento, disparidad binocular, e información básica de contraste y color. V1 luego alimenta de información a otras áreas, como V2, V4 y V5. Cada una de estas áreas se ocupa de los aspectos más específicos o abstractas de la información. Por ejemplo, las neuronas en V4 se ocupan de objetos de mediana complejidad, tales como formas de estrellas en diferentes colores. La corteza visual de los animales es el más potente sistema de procesamiento visual que conocemos, por lo que suena lógico inspirarse en ella para crear una variante de redes neuronales artificiales que ayude a identificar imágenes; es así como surgen las redes neuronales convolucionales.

Red Neuronal Convolucional: las redes neuronales convolucionales son muy similares a las redes neuronales ordinarias como el perceptrón multicapa; se componen de neuronas que tienen pesos y sesgos que pueden aprender. Cada neurona recibe algunas entradas. realiza un producto escalar y luego aplica una función de activación. Al igual que en el perceptrón multicapa también vamos a tener una función de pérdida o costo (por ejemplo SVM / Softmax) sobre la última capa, la cual estará totalmente conectada. Lo que diferencia a las redes neuronales convolucionales es que suponen explícitamente que las entradas son imágenes, lo que nos permite codificar ciertas propiedades en la arquitectura; permitiendo ganar en eficiencia y reducir la cantidad de parámetros en la red. Las redes neuronales convolucionales vienen a solucionar el problema de que las redes neuronales ordinarias no escalan bien para imágenes de mucha definición; por ejemplo en el problema de MNIST, las imágenes son de 28x28; por lo que una sola neurona plenamente conectado en una primera capa oculta de una red neuronal ordinaria tendría 28 x 28 = 784 pesos. Esta cantidad todavía parece manejable, pero es evidente que esta estructura totalmente conectado no funciona bien con imágenes más grandes. Si tomamos el caso de una imagen de mayor tamaño, por ejemplo de 200x200 con colores RGB, daría lugar a neuronas que tienen 200 x 200 x 3 = 120.000 pesos. Por otra parte, el contar con tantos parámetros, también sería un desperdicio de recursos y conduciría rápidamente a sobreajuste. (8)

Las redes neuronales convolucionales trabajan modelando de forma consecutiva pequeñas piezas de información, y luego combinando esta información en las capas más profundas de la red. Una manera de entenderlas es que la primera capa intentará detectar los bordes y establecer patrones de detección de bordes. Luego, las capas posteriores trataran de combinarlos en formas más simples y, finalmente, en patrones de las diferentes posiciones de los objetos, iluminación, escalas, etc. Las capas finales intentarán hacer coincidir una imagen de entrada con todas los patrones y arribar a una predicción final como una suma ponderada de todos ellos. De esta forma las redes neuronales convolucionales son capaces de modelar complejas variaciones y comportamientos dando predicciones bastantes precisas.

En general, las redes neuronales convolucionales van a estar construidas con una estructura que contendrá 3 tipos distintos de capas:

Capa Convolucional.

Como dijimos anteriormente, lo que distingue a las redes neuronales convolucionales de cualquier otra red neuronal es utilizan un operación llamada convolución en alguna de sus capas; en lugar de utilizar la multiplicación de matrices que se aplica generalmente. La operación de convolución recibe como entrada o input la imagen y luego aplica sobre ella un filtro o kernel que nos devuelve un mapa de las características de la imagen original, de esta forma logramos reducir el tamaño de los parámetros. La convolución aprovecha tres ideas importantes que pueden ayudar a mejorar cualquier sistema de machine learning, ellas son:

- Interacciones dispersas, ya que al aplicar un filtro de menor tamaño sobre la entrada original podemos reducir drásticamente la cantidad de parámetros y cálculos;
- Parámetros compartidos, que hace referencia a compartir los parámetros entre los distintos tipos de filtros, ayudando también a mejorar la eficiencia del sistema; y
- Representaciones equivariante, que indican que si las entradas cambian, las salidas van a cambiar también en forma similar.

Capa de Reducción o Pooling

La capa de reducción o pooling se coloca generalmente después de la capa convolucional. Su utilidad principal radica en la reducción de las dimensiones espaciales (ancho x alto) del volumen de entrada para la siguiente capa convolucional. No afecta a la dimensión de profundidad del volumen. La operación realizada por esta capa también se llama reducción de muestreo, ya que la reducción de tamaño conduce también a la pérdida de información. Sin embargo, una pérdida de este tipo puede ser beneficioso para la red por la siguiente razón:

I. La disminución en el tamaño conduce a una menor sobrecarga de cálculo para las próximas capas de la red; también trabaja para reducir el sobreajuste.

La operación que se suele utilizar en esta capa es max-pooling, que divide a la imagen de entrada en un conjunto de rectángulos y, respecto de cada subregión, se va quedando con el máximo valor.

Capa Clasificadora Totalmente Conectada

Al final de las capas convolucional y de pooling, las redes utilizan generalmente capas completamente conectados en la que cada píxel se considera como una neurona separada al igual que en una red neuronal regular. Esta última capa clasificadora tendrá tantas neuronas como el número de clases que se debe predecir. (9)

TensorFlow: biblioteca de código abierto para aprendizaje automático a través de un rango de tareas, y desarrollado por Google para satisfacer sus necesidades de sistemas capaces de construir y entrenar redes neuronales para detectar y descifrar patrones y correlaciones, análogos al aprendizaje y razonamiento usados por los humanos.

TensorFlow es una plataforma informática para la compilación de modelos de aprendizaje automático. Ofrece diversos kits de herramientas que te permiten crear modelos con el nivel de abstracción que prefieras. Puedes usar API de nivel inferior para compilar modelos definiendo una serie de operaciones matemáticas. O bien, puedes usar API de nivel superior (como *tf.estimator*) para especificar arquitecturas predefinidas, como regresores lineales o redes neuronales.

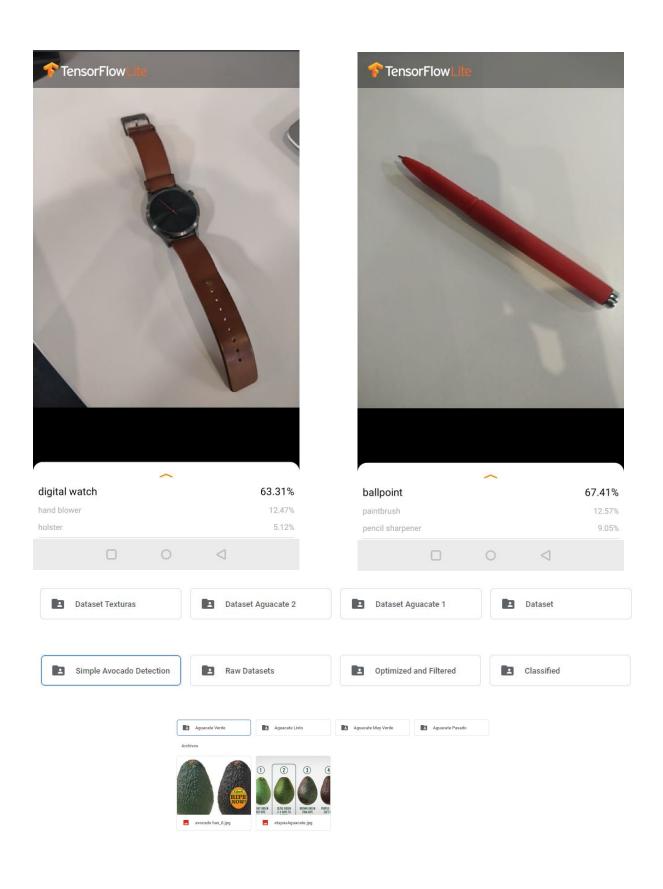


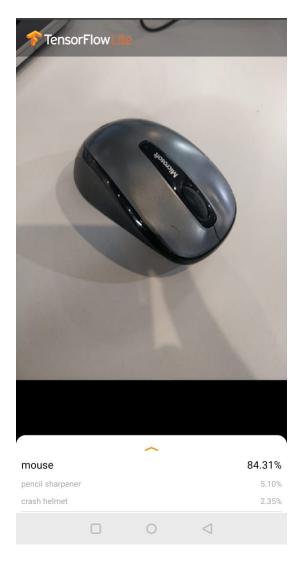
Imagen: jerarquía de TensorFlow

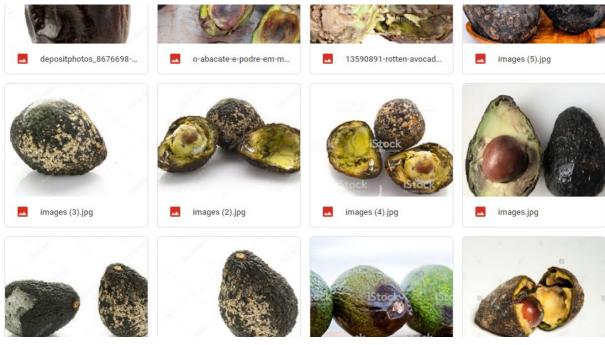
7. Resultados

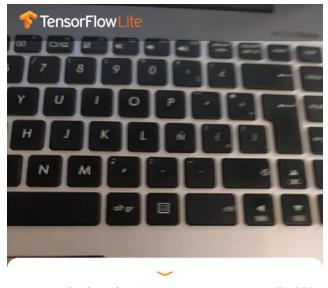
Se puede trabajar con diferentes tipos de aguacates ya que entre ellos existen demasiadas diferencias, es decir unos varían en color más que otros, maduran en muchos más días, tienen diferentes formas y diferentes tamaños, también se podrá hacer una comparación entre los resultados obtenidos mediante diferentes métodos de clasificación y determinar cuál es el que tiene más exactitud y menos porcentaje de error, y poder de esta manera emplear un método más confiable para este tipo de pruebas. De manera que, en esta primera fase de trabajo se delimitó únicamente en el tipo: hass.

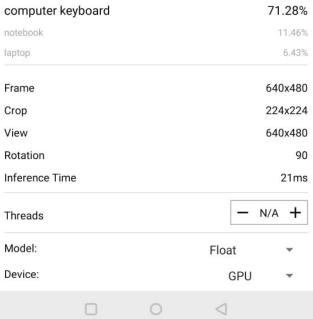
Anexamos fotografías del entrenamiento del modelo para la distinción de distintos objetos, agregando el accuracy de efectividad del mismo así como referencia del dataset generado para el entrenamiento.













8. Conclusiones

Se logra determinar el estado de madurez con cierto grado de error, ya que no fue posible garantizar que los aguacates que fueron empleados como base de datos perteneciera a una misma cosecha y de un mismo palo, ya que levemente varían en su estado de madurez, determinando esto por su dureza, y variación de color. Se desarrolló una interfaz gráfica por medio de la cual muestra al usuario el estado de maduración en el que se encuentra el aguacate sometido al análisis en tiempo real.

Con base a este proyecto realizado se podrá implementar mejoras que permita brindar mayor precisión, una de estas mejoras es incorporar un sensor de presión dentro del ambiente controlado para determinar la dureza del aguacate y así obtener otra característica determinística. Otra consideración para implementar sobre este proyecto sería la de la detección de defectos en el aguacate, para permitir su selectividad partiendo de los resultados obtenidos y así garantizar una mayor calidad de los aguacates al momento de comercializarlos.

9. Bibliografía

- 1. Organización de las Naciones Unidas para la Alimentación y la Agricultura, F. A. O. (s.f.). Pérdida y desperdicio de alimentos. Recuperado 24 septiembre, 2019, de http://www.fao.org/food-loss-and-foodwaste/es/
- 2. Milenio, M. (2017, 26 julio). México busca reducir pérdida y desperdicio de alimentos. Recuperado 24 septiembre, 2019, de https://www.milenio.com/politica/mexico-busca-reducir-perdida-desperdicioalimentos
- 3. El Financiero. (2018, 16 octubre). Aguacate y tortillas, entre los alimentos que más desperdician los mexicanos. Recuperado 23 septiembre, 2019, de https://www.elfinanciero.com.mx/nacional/aguacate-ytortillas-entre-los-10-alimentos-que-mas-s-se-desperdician-en-mexico
- 4. <u>Towards Automatic Construction of Diverse</u>, <u>High-quality Image Dataset</u>, Yazhou Yao, Jian Zhang, Fumin Shen, Dongxiang Zhang, Zhenmin Tang and Heng Tao Shen, arXiv:1708.06495 [cs.CV], 2017. Recuperado 16 noviembre, 2019, de http://www.multimediauts.org/dataset/WSID-100.html
- 5. Horea Muresan, Mihai Oltean, Fruit recognition from images using deep learning, Acta Univ. Sapientiae, Informática Vol. 10, Issue 1, pp. 26-42, 2018.Recuperado 16 noviembre, 2019, de https://github.com/Horea94/Fruit-Images-Dataset
- 6. RIOT Rapid Image Optimizer Tool https://riot-optimizer.com/
- 7. Cristóbal Delgadillo Pérez, C. (2013, 2 enero). Industria de aguacate mexicano fuerte y floreciente Hortalizas. Recuperado 16 noviembre, 2019, de https://www.hortalizas.com/cultivos/industria-de-aguacate-mexicano-fuerte-y-floreciente/

8. James Cowley. (2018). *Redes Neuronales Convolucionales*. Recuperado 18 noviembre 2019.

https://www.ibm.com/developerworks/ssa/library/cc-convolutional-neural-network-vision-recognition/index.html

9. Raul E. Lopez Briega. (2016). *Redes Neuronales Convolucionales con TensorFlow.* Recuperado 18 de noviembre 2019.

https://relopezbriega.github.io/2016/08/02/redes-neuronales-convolucionales-con-tensorflow/