



Estimación de producción en cítricos usando técnicas de aprendizaje automático

O. E. Apolo-Apolo¹, J. Martínez-Guanter^{1,2}, G. Egea¹, M. Pérez-Ruiz¹

¹ Dpto. Ingeniería Aeroespacial y Mecánica de Fluidos, Universidad de Sevilla, Ctra. Sevilla-Utrera km1 Sevilla 41013.

² Agropanning Agricultura Inteligente S.L. Programa Torres-Quevedo 2017 (PTQ-17-09506); eapolo@us.es, martinezj@us.es, gegea@us.es, manuelopez@us.es

Resumen: Estimar con exactitud la cosecha de un cultivo representa una información muy relevante para agricultores y cooperativas/agentes encargados de gestionar y vender el producto. De esta estimación depende la organización y logística necesarias para la recolección, planificación del almacenaje, stock y abastecimiento de los mercados. Actualmente la estimación de la cosecha se realiza en campo con personal experimentado realizando una inspección visual y en base a datos históricos, proceso que tiene riesgo de presentar errores humanos. Las redes neuronales convolucionales (CNN) basadas en el Deep Learning (DL), constituyen una herramienta prometedora para hacer estimaciones de rendimiento basadas en el reconocimiento y conteo de frutos. El objetivo del este trabajo ha sido crear un modelo basado en CNNs a partir de una arquitectura de red existente y entrenada para contar el número de frutos y estimar la producción de la parcela. Una vez entrenados los modelos, se testaron sobre imágenes tomadas con un dron multirrotor sobre 20 árboles seleccionados al azar de una parcela de cítricos *Citrus sinensis* (L.) cv. Navelina. Durante las tres campañas anuales en las que se han realizado las estimaciones, el error medio absoluto obtenido con DL fue entre 4-6% y el del técnico especialista en el aforo entorno al 8-11% frente a la producción real de la parcela. Estos resultados vislumbran un gran potencial de la metodología para la predicción del rendimiento de árboles de cítricos.

Palabras clave: aforo, naranjas, machine learning, UAV, Deep learning

1. Introducción

En la actualidad, el cultivo de cítricos es uno de más relevantes en los países del arco mediterráneo, con una superficie cosechada de 513.602 ha, de las cuales el 57% (295.000 ha) se encuentran en España. En este contexto, España es hoy en día el mayor productor y exportador de naranja en fresco de Europa, a pesar de la fuerte competitividad en el mercado de países fuera de la UE [1]. Las operaciones de cosecha son una de las más costosas por su realización manual en la gran mayoría de explotaciones en España [2], debido al escaso nivel de mecanización de estas tareas en nuestro país [3]. La elección del momento óptimo de cosecha depende del índice de madurez, color de la fruta y tamaño entre otros factores [4]. Sin embargo, el reto más importante en este tipo de cultivo es estimar el número de frutos y su estado con el mayor grado de anticipación posible para organizar las operaciones de cosecha y predecir su precio y puesta en el mercado [5].

Tradicionalmente estas estimaciones del rendimiento en naranjos y otros cítricos se han basado en las condiciones agronómicas del cultivo, los datos históricos de rendimiento de los cultivos y las observaciones visuales del cultivo [6]. Estas observaciones visuales son altamente

subjetivas, y dependen de la experiencia y conocimiento de los técnicos, llegando a errores en la estimación de entre el 15% y 25% [7]. En las últimas décadas, gracias a los avances en campos como la informática, la robótica, la inteligencia artificial y los sensores o cámaras digitales, se han estudiado diferentes opciones dentro de la llamada Agricultura de Precisión para estimar el rendimiento de los cultivos. Estos métodos semi-automatizados utilizan sistemas de captación de imágenes de forma automática y posteriormente aplican técnicas de segmentación de imágenes [8]. Por ejemplo, se llevaron a cabo algunos estudios anteriores de reconocimiento de frutas para manzanas [9], cítricos [10][11] y tomates [12]. Sin embargo, muchas de estas técnicas utilizaban únicamente la respuesta espectral entre píxeles como una característica única para detectar y contar frutos. La gran mayoría de estos trabajos concluyen que el uso de imágenes RGB o multispectrales puede ser una más que notable aproximación a los resultados reales, si bien establecen unas barreras comunes como son la presencia de sombras, frutos ocluidos y orientaciones espaciales con respecto a la luz que no facilitan su identificación y posterior conteo. Por ello, en los últimos años se ha incorporado a estas metodologías para el análisis de imágenes, el uso de las técnicas de aprendizaje automático o Machine Learning (ML).

Estas técnicas de aprendizaje profundo permiten el entrenamiento de un modelo en base a sets de datos previamente obtenidos (y etiquetados, en nuestro caso), y generar a modo de salidas de dicho modelo la clasificación o detección de elementos en dichas imágenes de forma automática. Una revisión de la literatura al respecto de uso de estas técnicas para estimar los rendimientos de los diferentes tipos de cultivos, la mayoría de estos métodos utilizan imágenes tomadas mediante plataformas terrestres o sensores a nivel del suelo [13]. Estas plataformas producen resultados precisos, pero consumen más tiempo y tienen mayores costes, lo que las hace inviables en las explotaciones de naranjos.

Por ello, los vehículos aéreos no tripulados (UAVs), capaces de portar sensores y cámaras de muy alta resolución, proporcionan una alternativa que puede ser implementada rápidamente y de forma barata [14]. Estos tipos de vehículos generan una gran cantidad de datos, generalmente en forma de imágenes o secuencias de vídeo, y permiten patrones de vuelo flexibles y adaptados a las tareas que se les demanda [15]. En este contexto, la combinación del potencial tecnológico de los drones y las técnicas de machine learning proporciona perspectivas e información únicas que, de otro modo, sería muy costoso obtener utilizando técnicas tradicionales [16].

En vista de lo anterior, el objetivo principal de este estudio fue desarrollar y validar durante 3 campañas, un sistema de detección de naranjas usando imágenes RGB obtenidas por vuelo de un UAV en combinación con una red neuronal artificial para su aforo de forma automatizada. El desarrollo de esta nueva metodología para estimar el rendimiento de los naranjos con alta precisión o al menos con un error menor que el de los métodos visuales puede contribuir a aumentar la rentabilidad económica de los agricultores y reducir los costos de logística y operacionales.

2. Materiales y métodos

2.1. Localización de ensayos y obtención de imágenes aéreas

Los ensayos se llevaron a cabo durante tres campañas de producción (2017, 2018 y 2019) en un área de 4 hectáreas de una explotación comercial de cítricos (*Citrus sinensis* L. cv. Navelina) cerca de Sevilla, suroeste de España (Lat: 37.512574; Long: -5.956659). El muestreo se llevó a cabo en 20 árboles individuales seleccionados al azar de un total de 1.654 árboles. Las imágenes fueron adquiridas bajo condiciones de luz natural empleando la cámara RGB de un dron comercial Phantom 3 (SZ DJI Technology Co.). La resolución de las imágenes de la cámara era de 4.000 × 3.000 píxeles. La altura de vuelo se estableció en seis metros sobre el suelo (tres sobre la copa) y se utilizó la opción de vuelo manual. Se obtuvieron imágenes por los lados izquierdo y derecho de cada árbol

(Figura 1). Además, se empleó una regla corta de madera calibrada en centímetros durante la toma de fotografías para tratar de escalar las imágenes y determinar el tamaño medio del fruto.

Se consideró como verdad-terreno (la producción efectiva de naranjas por árbol) los kilogramos de naranjas, de cada uno de los 20 árboles seleccionados, que se recogieron de forma manual y que fueron pesados de forma individualizada, obteniendo los datos de producción real para luego ser evaluados frente a las estimaciones del modelo.

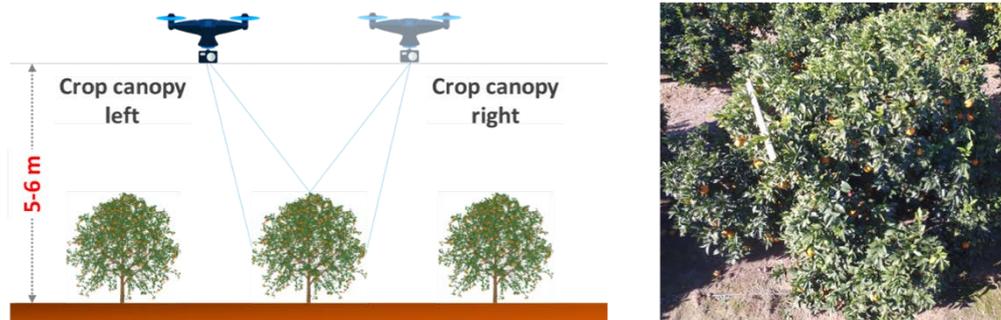


Figura 1. Ilustración del proceso de captura de imágenes mediante un sensor RGB a bordo de un dron por cada uno de los lados del naranjo (izquierda). Ejemplo de imagen capturada del naranjo para ser analizada (derecha).

2.2. Desarrollo de la arquitectura de la red neuronal

La potencia y capacidad de cómputo de los equipos informáticos ha sido desde el inicio una de las barreras para el desarrollo del aprendizaje profundo, ya que se necesita trabajar con grandes cantidades de datos complejos. Para el desarrollo del modelo capaz de procesar las imágenes, se empleó en primer lugar la librería de visión por computador OpenCV [17] del lenguaje de programación Python.

El desarrollo de la red neuronal ha sido hasta ahora algo realmente complejo, debido a las complicadas arquitecturas computacionales que se debían desarrollar prácticamente de cero. La irrupción de las grandes compañías informáticas y la puesta a disposición de sus herramientas lo ha convertido en algo relativamente asequible. En nuestro caso, la red neuronal se generó empleando las herramientas de código abierto Keras [18] y Google TensorFlow [19], que facilitan la construcción y entrenamiento de los modelos. Estas herramientas crean capas de abstracción e interfaces de diseño y evaluación rápida de las arquitecturas Deep Learning [20].

2.3. Generación del dataset y entrenamiento del modelo

Para entrenar un modelo de red neuronal artificial, especialmente una red convolucional, se requiere de un set de datos lo más extenso posible, de manera que el sistema tenga cuantos más ejemplos de aprendizaje mejor. También se necesita que los ejemplos sean lo suficientemente dispares para evitar el sobreaprendizaje (*overfitting*). En este trabajo, esto se aborda con un dataset que incluye frutos de color naranja, frutos que no están maduros (verdes), con diferentes tamaños y formas y frutos ocluidos por ramas y hojas (Figura 2). En este trabajo se generó un primer conjunto de datos con 400 imágenes utilizando la API de TensorFlow, como se sugiere en [21].

Los modelos proporcionados por TensorFlow (<https://github.com/tensorflow/models>) fueron pre-entrenados y seleccionados en base a su precisión y la velocidad del entrenamiento.

De forma adicional a este primer dataset, se añadieron las imágenes aéreas de alta resolución obtenidas por el UAV. Un total de 50 de estas imágenes fueron etiquetadas de forma manual empleando el software libre de etiquetado gráfico LabelImg [22] (Figura 2).

Para evitar procesar manualmente todo el conjunto de imágenes de entrenamiento, se generó el modelo basado en la red neuronal, y se inició el proceso de entrenamiento automático. Una vez

entrenados los modelos, se utilizó el mejor modelo para detectar frutos de naranja en las imágenes tomadas con el UAV (Figura 3). Cada fruta anaranjada se etiquetó con una caja verde que mostraba la precisión o bondad del algoritmo en el proceso de reconocimiento de cada naranja.

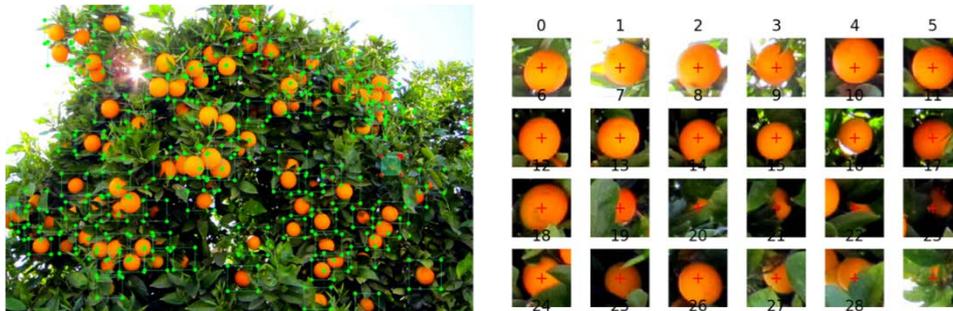


Figura 2. Etiquetado de imágenes sobre imagen completa (izquierda) y diferentes casos de naranjas etiquetadas (derecha).

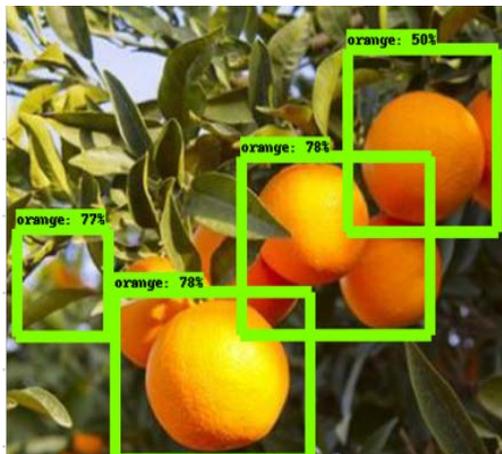


Figura 3. Detección automática de naranjas y precisión del algoritmo en cada detección. Nótese que algunas de éstas no son detectadas por presentar un alto grado de oclusión por otros frutos.

2.4. Análisis de la precisión del modelo.

Durante las tres campañas en las que tuvo lugar el estudio, cada árbol se consideró como una unidad experimental y el número de frutos detectado fue determinado de forma individual. Se realizó un análisis del error estadístico y la varianza con el paquete estadístico Statgraphics (Statgraphics Centurion XV). Las imágenes de cada uno de los lados del árbol se agruparon, y se comparó el número de frutos de cada árbol con el recuento visual de frutos mediante análisis de regresión lineal. Para comparar el rendimiento real con el rendimiento estimado, se desarrolló un modelo de regresión lineal considerando el número de frutos detectados y el peso del fruto para cada árbol. Finalmente, se compararon los rendimientos totales obtenidos a partir de la estimación del rendimiento mediante visión artificial y la estimación visual realizada por un técnico profesional durante la temporada de cosecha para obtener los datos reales de rendimiento.

3. Resultados y discusión

3.1. Conteo automático de frutos

Tras la generación del modelo, se empleó la red neuronal para la detección automática de frutos en los 20 árboles del estudio. Se comparó el número de frutos obtenidos con el mejor modelo de conjunto de datos con el número de frutos contados por métodos visuales en cada conjunto de imágenes. Evaluamos las métricas de píxeles naranjas en un modelo con cerca de 10.000 iteraciones. La red de detección de estos píxeles produce un mapa de probabilidad, obteniendo las detecciones mostradas en la figura 3. El valor estadístico de R^2 obtenido de 0.96 indica que el modelo entrenado tiene una buena precisión en términos de los frutos detectados en las imágenes de ambos lados de los árboles (Figura 4).

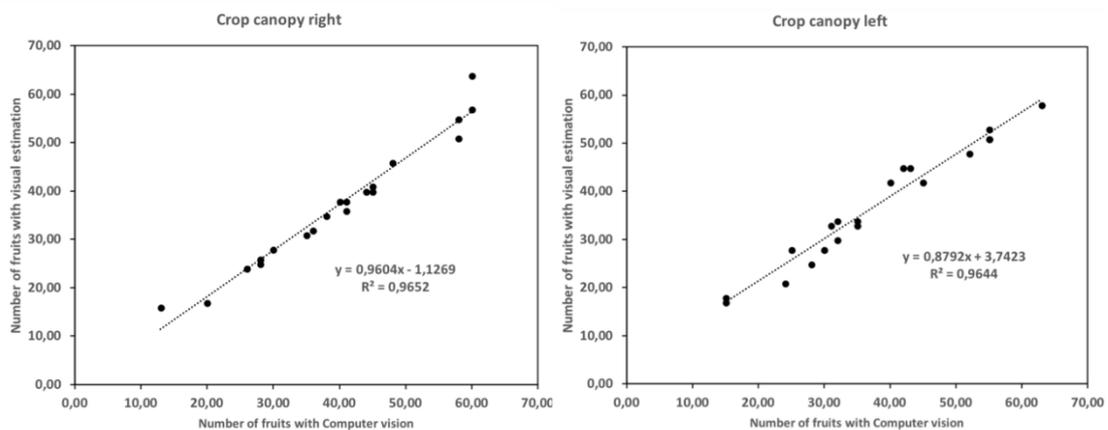


Figura 4. Número de frutas detectadas de forma automática en las imágenes del lado derecho e izquierdo del árbol, comparadas con el número de naranjas detectadas visualmente en las imágenes.

El valor del ajuste de los resultados obtenidos ($R^2=0.96$) mejora ligeramente al obtenido en [13], en el que además se evaluó la bondad del modelo con una curva ROC. Nuestro flujo de trabajo con el modelo de Deep Learning ya entrenado, generó resultados de identificación de las naranjas con un coste de tiempo de computación por imagen de alrededor de 30 segundos. Este bajo coste computacional hará atractiva una posible aplicación en el plano práctico de este tipo de modelos.

3.2. Estimación del rendimiento

Al término de cada una de las campañas se cuantificó la producción total de la explotación, obteniendo un rendimiento medio de 217.000 kg. Durante estas campañas, las estimaciones medias del técnico profesional encargado del aforo fueron de 200.000 kg. La tabla 1 a continuación muestra el número de frutos detectados por el modelo en las imágenes de cada uno de los lados del árbol (detecciones en el lado derecho: N_D y lado izquierdo: N_I) y el rendimiento estimado por cada árbol en kg (Y_{est}).

El rendimiento total de la explotación estimado se puede calcular considerando el número de árboles de la parcela (1654) y el rendimiento medio obtenido por cada árbol (132.96 kg), resultando en un rendimiento total estimado de 219.909 kg. La comparación del rendimiento total estimado mediante el modelo generado con el rendimiento total obtenido como verdad-terreno, arroja un error del 1.54%. De la misma forma, la comparación de la estimación realizada por el técnico aforador con el rendimiento final obtenido genera un error del 7.83%. Sin duda esto conlleva una mejora en la estimación, reduciendo la incertidumbre del aforo y permitiendo automatizar el conteo de frutos de forma rápida, repetible y relativamente asequible.

Tabla 1. Comparación de resultados obtenidos tras la cosecha de forma manual y por el modelo generado. YO = Rendimiento obtenido en la cosecha de forma manual en kg/árbol, WO= Peso medio de los frutos obtenidos por árbol, NL= Número de frutos detectados por el modelo en las imágenes del lado izquierdo, NR= Número de frutos detectados por el modelo en las imágenes del lado derecho de cada árbol, Yest = Rendimiento estimado basado en el número de frutos detectados por el modelo en cada árbol calculado en kg/árbol.

Árbol	Yo (kg/árbol)	WO (kg/fruto)	NL	NR	Yest (kg/árbol)
1	187.00	0.55	45	42	154.80
2	145.80	0.45	40	42	144.48
3	199.10	0.45	52	48	172.00
4	120.00	0.43	32	34	123.84
5	139.10	0.49	32	30	120.40
6	162.80	0.27	63	58	208.12
7	122.10	0.24	55	53	197.80
8	123.20	0.25	55	51	197.80
9	133.50	0.45	35	34	130.72
10	139.30	0.50	31	33	123.84
11	116.30	0.64	25	28	77.40
12	137.40	0.49	43	45	122.12
13	126.30	0.48	35	33	108.36
14	132.40	0.30	55	51	194.36
15	80.60	0.29	30	28	111.80
16	112.80	0.58	24	21	86.00
17	146.40	0.58	28	25	110.08
18	85.70	0.72	15	17	48.16
19	145.20	0.28	42	45	149.64
20	110.10	0.27	15	18	77.40
Media	133.26	0.43	37.60	39.70	132.96

4. Conclusiones

Los resultados de este trabajo sobre la estimación del aforo mediante un modelo basado en redes neuronales e imágenes a partir de drones fueron prometedores. Sin embargo, se pone de manifiesto que las principales limitaciones en estas detecciones automáticas de frutos son la capacidad de cómputo, la dificultad para discernir entre frutos con una gran casuística (oclusiones, sombras, frutos no maduros) y la dificultad en el desarrollo de modelos complejos que permitan afinar en mayor medida los resultados obtenidos. El modelo generado en las sucesivas campañas tratará de integrar estas frutas ocluidas o no maduras como parámetros adicionales para mejorar la detección. La estimación del rendimiento del modelo fue más cercana al rendimiento real que las estimaciones visuales realizadas por un técnico profesional. Estos resultados son prometedores y nos hacen ser optimistas sobre la posibilidad de desarrollar un modelo más preciso para estimar el rendimiento en explotaciones comerciales de cítricos de manera robusta y fiable.

5. Agradecimientos

Este trabajo ha sido posible gracias al proyecto *DronFruit* (Referencia: GOP21-SE-16-0048) financiado por la Junta de Andalucía. Los autores agradecen al técnico de DronSAP, Pablo Agüera Requena la gran ayuda prestada en el desarrollo del trabajo. Los autores también quieren agradecer el contrato Torres Quevedo (PTQ-17-09506) concedido a JMG y a la Beca de investigación predoctoral para el desarrollo del programa de I+D+i de la Universidad de Sevilla (IV.3 2017).

Referencias

1. Cardeñosa, V., Medrano, E., Lorenzo, P., Sánchez-Guerrero, M. C., Cuevas, F., Pradas, I., & Moreno-Rojas, J. M.; Effects of salinity and nitrogen supply on the quality and health-related compounds of strawberry fruits (*Fragaria × ananassa* cv. Primoris); *Journal of the Science of Food and Agriculture*; 2015; 95(14), 2924–2930.
2. Torregrosa Mira, A.; Albert Gil, FE.; Aleixos Borrás, MN.; Ortiz Sánchez, MC.; Blasco Ivars, J. Analysis of the detachment of citrus fruits by vibration using artificial vision. *Biosystems Engineering*, 2014, 119:1-12
3. Fonfría, M.; *Citricultura*; Mundi Prensa Libros S.A, Ed; 2012; Madrid.
4. Lado, J., Rodrigo, M. J., & Zacarías, L.; Maturity indicators and citrus fruit quality. *Stewart Postharvest Review*; 2014; 10(2).
5. Chinchuluun, R., Lee, W. S., Burks, T. F., & Florida State Horticultural, S.; Machine vision-based citrus yield mapping system; *Proceedings of the 119th Annual Meeting of the Florida State Horticultural Society*; 2006; 119, 142–147.
6. Srivastava, A.K., Huchche, A.D., Ram,L., Singh, S.; Yield prediction in intercropped versus monocropped citrus orchards; *Scientia Horticulturae*; 2007; Volume 114, Issue 1, 67-70
7. Apolo-Apolo, O. E., Martínez-Guanter, J., Egea, G., Pérez-Ruiz, M.; UAV and Machine Learning Techniques Applied to Yield Estimation of Orange Trees, *Proceedings of the 2018 European Agricultural Engineering Conference Wageningen, The Netherlands*, 2018.
8. Teixidó, M., Font, D., Pallejà, T., Tresanchez, M., Nogués, M., Palacín, J.; Definition of linear color models in the RGB vector color space to detect red peaches in orchard images taken under natural illumination; *Sensors*; 2012; 12, 7701–7718.
9. Linker, R.; Cohen, O.; Naor, A.; Determination of the number of green apples in RGB images recorded in orchards; *Computers and Electronics in Agriculture*; 2012; 81, 45-57
10. Hannan, M.W.; Burks, T.F.; Bulanon, D.M.; A Machine Vision Algorithm Combining Adaptive Segmentation and Shape Analysis for Orange Fruit Detection. *Agric. Eng Int. CIGR J.*, XI; 2009; 1281.
11. Kurtulmus, F.; Lee, W.S.; Vardar, A. Green citrus detection using “eigenfruit”, color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture* 2011; 78 (2), 140–149
12. Yamamoto, K., Guo, W., Yoshioka, Y., & Ninomiya, S.; On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors*, 2014; 14(7), 12191-12206.
13. Chen, S.W., Shivakumar, S.S., Dcunha, S., Das, J., Okon, E., Qu,C., Taylor, C.J., Kumar,V.; *IEEE Robotics and Automation Letters*; 2017; Vol 2, No. 2. 781-788
14. Martínez, J., Egea, G., Agüera, J., & Pérez-Ruiz, M. A cost-effective canopy temperature measurement system for precision agriculture: a case study on sugar beet. *Precision Agriculture*, 2017; 18(1), 95–110
15. Martínez-Guanter, J., Agüera, P., Agüera, J., Pérez-Ruiz, M.; Spray and economics assessment of a UAV-based ultra-low-volume application in olive and citrus orchards, *Precision Agriculture*, 2019, pp 1-18.
16. Yin, X., Lan, Y., Wen, S., Zhang, J., & Wu, S. Natural UAV tele-operation for agricultural application by using Kinect sensor. *International Journal of Agricultural and Biological Engineering*, 2018, 11(4), 173–178.
17. Bradski, G. *The OpenCV Library*. Dr. Dobb’s *Journal of Software Tools*, 2000.
18. Chollet, F. *Keras*. 2015
19. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. *Tensorflow: A system for large-scale machine learning*. 12th *Symposium on Operating Systems Design* 2016 (pp. 265-283).
20. Murata, T., Onuki, Y., Nukui, S., Inagi, S., Qiu, X., Watanabe, M., Okamoto, H., *Predicting Relations Between RDF Entities by Deep Neural Network*. *The Semantic Web 2017*
21. Rosebrock, A., *Deep Learning for Computer Vision with Python, Practitioner Bundle*, 2017. Pyimagesearch.com.
22. Tzutalin. *LabelImg*. 2015 <https://github.com/tzutalin/labelImg>