

# Identificación Automática de Especies de Plantas de Costa Rica usando Visión por Computadora

Carranza-Rojas Jose\*, Mata-Montero Erick†

PARMA Research Group, Instituto Tecnológico de Costa Rica

Cartago, Costa Rica

Email: \*jcarranza@itcr.ac.cr, †emata@itcr.ac.cr

**Abstract**—En la última década, la identificación automática de plantas basada en técnicas de aprendizaje por computadora/visión artificial se ha convertido en un área de investigación muy activa. El enfoque típico usa imágenes de hojas de plantas y una perspectiva de visión por computadora tradicional basada en extractores de características morfológicas predeterminadas. Sin embargo, ha habido un cambio gradual hacia enfoques de aprendizaje profundo que incluyen el uso de imágenes tomadas *in-situ*. En este artículo, describimos el estado del arte y un resumen de nuestras investigaciones con respecto a algoritmos de visión por computadora para detectar automáticamente las especies de plantas a partir de sus imágenes.

Biodiversity Informatics, Computer Vision, Image Processing, Machine Learning, Automated Plant Identification

## I. INTRODUCCIÓN

La identificación automática de plantas basada en imágenes tiene como objetivo fundamental facilitar el desarrollo rápido de inventarios de plantas en una región, la enseñanza de la taxonomía, el ecoturismo y, consecuentemente, la conservación de la biodiversidad. Inicialmente, muchas investigaciones se enfocaron en el uso de imágenes de las hojas — para facilitar la segmentación — y en algoritmos de visión por computadora tradicionales basados en extractores de características morfológicas predeterminadas [1]–[3]. Sin embargo, los botánicos no sólo usan hojas sino otras partes de las plantas para identificarlas.

En el proceso de construir un conjunto de datos global con todas las especies posibles, se ha dado un fenómeno predecible: los conjuntos de datos están desbalanceados, tanto temáticamente, como geográfica y, por supuesto, taxonómicamente. Los herbarios son un obvio candidato para proporcionar imágenes de plantas de sus colecciones. Sin embargo, las imágenes de pliegos de herbarios son difíciles de trabajar por la gran cantidad de ruido visual y artefactos presentes [4].

Con el auge del aprendizaje profundo y la disponibilidad de múltiples GPU con enorme poder computacional, es posible trabajar con imágenes complejas e ir más allá de las imágenes de hojas para la identificación de plantas. Por lo tanto, el estado del arte se centra en modelos de aprendizaje profundo basados en convoluciones que permiten aprender qué patrones visuales

son importantes, en lugar de crear manualmente los extractores de características como se hizo anteriormente.

Por otro lado, normalmente las identificaciones se han hecho a nivel de especie. Sin embargo, en muchos casos la predicción de la especie da una exactitud muy baja debido a la variabilidad y similitud inter e intra-específicas. Por lo tanto, es imperativo explorar las jerarquías taxonómicas para la clasificación y regularización de los modelos de aprendizaje profundo, con el fin de hacer la identificación a varios niveles. Los enfoques jerárquicos y multi-etiqueta también podrían evitar clasificadores planos enormes, dado que se estima que la cantidad de especies de plantas en nuestro planeta es alrededor de 400.000.

El resto de este manuscrito se organiza de la siguiente manera: la sección II resume el trabajo relacionado más relevante. La sección III cubre el progreso realizado en nuestras investigaciones. Finalmente, la sección IV presenta las conclusiones y trabajo futuro.

## II. TRABAJO RELACIONADO

La investigación sobre identificación automatizada de plantas ha evolucionado desde el uso de técnicas tradicionales de visión por computadora a los llamados enfoques de aprendizaje profundo [5]. Este último proporciona, en general, mejor precisión [6] y evita la segmentación explícita mediante el uso de convoluciones en lugar de extractores de características predeterminados. En general, ofrece la oportunidad de clasificar imágenes muy complejas utilizando imágenes de hojas, flores y otros componentes de plantas, así como imágenes de todo el organismo. Sin embargo, normalmente requiere grandes conjuntos de imágenes y un poder computacional de mayor rendimiento que las técnicas tradicionales.

### A. Técnicas Tradicionales de Identificación Automática de Plantas

Los enfoques tradicionales de visión por computador se basan en extractores de características y clasificadores predeterminados [1], [3], [7]. La mayoría de los autores dividen el proceso de reconocimiento de plantas en cinco fases, a saber,

adquisición de imágenes de plantas, mejora de imágenes, segmentación de imágenes, extracción de características y clasificación. La mayor parte del trabajo se ha centrado en la identificación de especies de plantas a partir de imágenes foliares.

a) *Adquisición de la imagen de la planta:* Muchos conjuntos de datos fueron desarrollados usando fondos uniformes para una segmentación más fácil, particularmente para el reconocimiento de la hoja. Por ejemplo, el conjunto de datos LeafSnap [3] consta de imágenes de 184 especies de árboles del noreste de USA. Incluye 23.916 imágenes de hojas frescas con fondos blancos. El set de datos Flavia [1] comprende 32 especies con 3.621 imágenes de hojas con fondo blanco. El conjunto de datos de especies de Costa Rica usado en [2] incluye 255 especies y aproximadamente 7.5k imágenes .

Hasta donde sabemos, pocos estudios han creado su conjunto de datos directamente a partir de imágenes de pliegos de herbario, *e.g.*, [8]. Sin embargo, miles de imágenes de pliegos de herbarios están disponibles ahora gracias a sistemas como iDigBio <sup>1</sup>.

b) *Mejora de la imagen:* Una vez que las imágenes son adquiridas, la siguiente fase consiste en pre-procesar la imagen para destacar las características importantes [9]. Este paso incluye la conversión de escala de grises, la reducción de ruido y otras conversiones de dominio de color [3].

c) *Segmentación de la planta:* El tercer paso consiste en extraer el elemento de interés de la imagen, por ejemplo, una hoja [3]. La mayoría de los estudios tratan con imágenes de hojas limpias con fondos uniformes y utilizan técnicas de agrupación de colores para extraer los píxeles de las hojas. Por ejemplo, en algunos estudios se utiliza un umbral muy básico [10]–[12], mientras que Expectation-Maximization (EM) se usa en [3] y [2].

d) *Extracción de características:* Se ha utilizado varias características morfológicas predeterminadas para describir las imágenes de plantas; por ejemplo, área, perímetro de la hoja y rectangularidad, entre otras [10], [11], [13], [14]. Enfoques posteriores usan formas o contornos con diferentes descriptores como Histogram of Curvature over Scale (HCoS) [1], [3]. Además, los datos de textura se han utilizado como información alternativa o complementaria utilizando, por ejemplo, Local Binary Patterns (LBP) y Local Binary Pattern Variance (LBPV) [1], [2].

e) *Clasificación:* Support Vector Machines (SVM) se han utilizado frecuentemente como clasificadores binarios para determinar si una imagen tiene una hoja presente [3], [13], [15]. Esto es en realidad una pre-clasificación para evitar perder el tiempo clasificando imágenes que no son de plantas. Artificial Neural Network (ANN) [1], [13], [16] y k-Nearest Neighbors (kNN) [3], [14], que se alimentan con las características extraídas en pasos previos, se han usado en esta etapa final para predecir la especie.

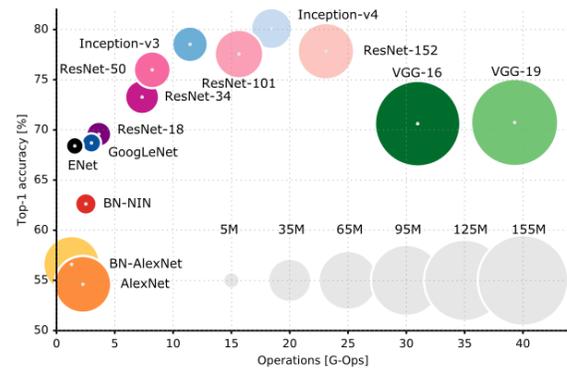


Figure 1: Comparación de modelos de aprendizaje profundo [21]

### B. Deep Learning para la Identificación Automática de Plantas

El desafío PlantCLEF es parte de uno más amplio denominado LifeCLEF [17]. LifeCLEF inició en 2011 con el fin de hacer que los científicos compitan para lograr una mejor detección de organismos, en particular, plantas, aves y peces. El desafío en PlantCLEF no sólo incluye imágenes de hojas sino también otros componentes tales como frutas, tallos, flores y otros. Desde 2014, los mejores resultados se han obtenido mediante el uso de modelos de aprendizaje profundo, en particular Convolutional Neural Networks (CNN) [6]. Desde entonces, esto se ha convertido en la norma. Por otro lado, en 2012, el desafío ImageNet [18], que es un desafío más generalizado, demostró por primera vez el poder del aprendizaje profundo con el trabajo de [19], el cual bajó la tasa de error del 26.2% al 15,3% en clases de naturaleza general (no en plantas).

Varios modelos se han desarrollado para identificar plantas, particularmente con imágenes complejas tomadas *in-situ*. GoogleNet es un modelo muy popular, así como VGGNet. Estos modelos se originaron en el desafío de ImageNet y fueron migrados para convertirse en sistemas de identificación de plantas con algunas adaptaciones. La Figura 1 muestra diferentes modelos populares actuales. Se puede observar que existe una relación directamente proporcional entre la precisión top-1 usando ImageNet y la cantidad de operaciones realizadas por los modelos. Inception-v4 [20] permanece como uno de los mejores modelos probados con ImageNet. Sin embargo, hasta la fecha, este modelo no se ha utilizado para la identificación de plantas.

En [22], se estudia cómo funcionan los sistemas de visión por computadora de vanguardia utilizando un subconjunto del conjunto de datos PlantCLEF 2014. Se utiliza 500 especies de árboles, hierbas y helechos, con un total de 19.504 observaciones. Las imágenes se obtuvieron utilizando la aplicación Pl@ntNet [23] y crowdsourcing.

En el desafío PlantCLEF 2016, los resultados fueron muy ilustrativos [24]. El conjunto de entrenamiento utilizado en el

<sup>1</sup><https://www.idigbio.org/>

tenía 113.205 imágenes pertenecientes a 41.794 observaciones de 1.000 especies, proveídas por 8.960 contribuyentes. Una observación consiste en varias imágenes del mismo espécimen. El conjunto de prueba fue construido a partir de datos de usuarios de Pl@ntNet [23] y se agregaron imágenes no vegetales. Además, se agregaron imágenes de plantas que no corresponden a las 1.000 especies en el conjunto de entrenamiento. El conjunto de pruebas está compuesto por 8.000 imágenes, 4.633 etiquetados como las 1.000 especies conocidas, 3.367 como nuevas clases desconocidas. Los 26 competidores con el mejor rendimiento se basaron en CNN. La mejor precisión media alcanzada fue de 71.8 % por el equipo japonés [25] que usó un modelo profundo VGGNet [26] con Spatial Pyramid Pooling (SPP). En general, al añadir imágenes no vegetales y especies desconocidas al conjunto de datos de desafío, el rendimiento del modelo se degrada [24], como se espera en un escenario de la vida real.

### III. METODOLOGÍA Y TRABAJO REALIZADO EN COSTA RICA

En una primera fase, el grupo de investigación PARMA<sup>2</sup> trabajó usando enfoques tradicionales de identificación de plantas. En la segunda fase se ha basado en el aprendizaje profundo.

1) *Enfoque Tradicional de Identificación de Plantas:* Nuestra primera publicación usa un conjunto de datos con 66 especies de Costa Rica y estudia cómo la textura es una mejor característica comparada con la curvatura de la hoja [2]. Este estudio usa imágenes solo de hojas con fondos uniformes y emplea tanto imágenes "ruidosas" como "limpias". Se extrae la textura con LBPV y la curvatura con HCoS. La mejora en la exactitud cuando se utiliza la textura es de 25.6% para las imágenes limpias y 42.5% para las imágenes ruidosas, lo que revela que la textura puede lidiar mejor con imágenes ruidosas. La Figura 2 muestra los resultados de la ejecución HCoS contra la textura y varias combinaciones de ambos en imágenes ruidosas de hojas. Relacionado con este estudio, también se publicó una versión extendida [27] que estudia más profundamente la exactitud para cada una de las 66 especies.

Otro estudio realizado fue un survey que resume los desafíos y oportunidades actuales en el dominio de identificación de plantas [4]. En este estudio, se propone el uso de pliegos de herbarios para construir un conjunto de datos global, lo cual estamos explorando actualmente.

2) *Aprendizaje profundo para la identificación de plantas:* Nuestro trabajo de aprendizaje profundo comenzó durante una pasantía doctoral de Carranza-Rojas en CIRAD e INRIA en Montpellier, Francia. Desarrollamos un modelo de aprendizaje profundo para medir la exactitud obtenida al usar imágenes de herbarios. Mediante el uso de datos de iDigBio, creamos un conjunto de datos con más de 250k imágenes y 1.225 especies, adecuadas para enfoques de aprendizaje profundo.

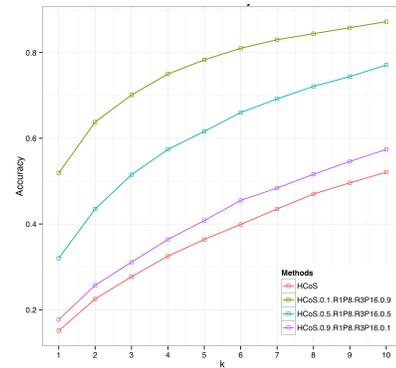


Figure 2: Exactitud de HCoS versus métodos combinados con extracción de textura, tomado de [2].

Ejecutamos una versión modificada del modelo de GoogleNet usando batch-normalization y Parametric Rectified Linear Unit (PRELU). Se realizó un total de 16 experimentos, algunos utilizando sólo el conjunto de datos de herbarios, otros tratando de hacer *transfer learning* de los herbarios a las imágenes silvestres y otros tratando de utilizar las características aprendidas de herbarios de una región para identificar especies de herbarios de otra región. Este trabajo se publicará este año.

3) *Regularización jerárquica:* Hasta donde sabemos, ningún autor ha intentado formalmente utilizar las taxonomías de las plantas para clasificar automáticamente los organismos no sólo a nivel de especie sino de otros taxones de nivel superior. Tampoco se ha utilizado conocimiento de otros taxones para realizar la clasificación de especies.

Muy pocos estudios han abordado el problema de la clasificación jerárquica. Silla et al [28] presentan un survey muy detallado sobre las diferentes técnicas utilizadas para la clasificación jerárquica y también diseñan un marco unificador para clasificar los enfoques existentes. Sin embargo, no se centra en el aprendizaje profundo, sino en el aprendizaje automático tradicional.

Una posible solución para la clasificación jerárquica es usar un enfoque multi-etiqueta, donde cada ítem tiene varias etiquetas asociadas con diferentes clases o nivel de clases [28]. En realidad, no es necesario ni siquiera tener niveles de una jerarquía. En cierto sentido, un enfoque multi-etiqueta no es un enfoque jerárquico, pero a menudo se utiliza como un primer intento de abordar problemas jerárquicos [28]. En nuestra investigación actual, proponemos utilizar el conocimiento previo de la jerarquía taxonómica vegetal para agregar términos adicionales a la función de pérdida/error basada en los niveles de taxón. Además, añadimos niveles a la definición de jerarquía propuesta en [29]. Por limitación de espacio, la formalización de este enfoque no la podemos presentar en esta publicación, pero sí en la presentación en JOCICI 2017.

<sup>2</sup><http://parma.ic-itcr.ac.cr/>

#### IV. CONCLUSIONES Y TRABAJO FUTURO

Ha habido un gran progreso en la identificación automática de plantas basada en imágenes. Los nuevos enfoques que utilizan el aprendizaje profundo superan a las técnicas tradicionales basadas en características predeterminadas ya que proporcionan mejores resultados con imágenes más complejas, en particular, las imágenes de plantas tomadas en ambientes naturales. Hemos logrado progresar también en el uso de imágenes de herbarios para la identificación de plantas, cuyos resultados están cerca de ser publicados. Además, hasta donde sabemos, nuestra investigación da los primeros pasos para formalizar y tomar en cuenta información en varios niveles de las jerarquías para la clasificación de especies.

Con la formalización actual y las capas profundas de regularización jerárquica, tenemos actualmente un primer intento de implementación de una capa de pérdida jerárquica basada en una taxonomía de plantas predefinida. Se necesita más trabajo para generalizar esto a cualquier nivel de taxón y también para agregar pesos limitados a cada término de nivel de taxón. Esto constituye la línea base para nuestros experimentos con funciones de pérdida jerárquica y será la línea base para futuros experimentos.

#### REFERENCES

- [1] S. Wu, F. Bao, E. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in *Signal Processing and Information Technology, 2007 IEEE International Symposium on*, Dec 2007, pp. 11–16.
- [2] E. Mata-Montero and J. Carranza-Rojas, "A texture and curvature bimodal leaf recognition model for identification of costa rican plant species," in *Computing Conference (CLEI), 2015 Latin American*, Oct 2015, pp. 1–12.
- [3] N. Kumar, P. Belhumeur, A. Biswas, D. Jacobs, W. Kress, I. Lopez, and J. Soares, "Leafsnap: A computer vision system for automatic plant species identification," in *Computer Vision – ECCV 2012*, ser. Lecture Notes in Computer Science, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Springer Berlin Heidelberg, 2012, pp. 502–516.
- [4] E. Mata-Montero and J. Carranza-Rojas, *Automated Plant Species Identification: Challenges and Opportunities*. Springer International Publishing, 2016, pp. 26–36.
- [5] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning," 2016, book in preparation for MIT Press. [Online]. Available: <http://www.deeplearningbook.org>
- [6] A. Joly, H. Goëau, H. Glotin, C. Spampinato, P. Bonnet, W.-P. Vellinga, R. Planque, A. Rauber, R. Fisher, and H. Müller, *LifeCLEF 2014: Multimedia Life Species Identification Challenges*. Cham: Springer International Publishing, 2014, pp. 229–249.
- [7] M. S. K. M. Z. Rashad, B.S.el-Desouky, "Plants images classification based on textural features using combined classifier," *International Journal of Computer Science and Information Technology*, vol. 3, no. 4, 2011.
- [8] D. Wijesingha and F. Marikar, "Automatic detection system for the identification of plants using herbarium specimen images," *Tropical Agricultural Research*, vol. 23, no. 1, 2012.
- [9] V. Metre and J. Ghorpade, "An overview of the research on texture based plant leaf classification," *CoRR*, vol. abs/1306.4345, 2013.
- [10] K.-B. Lee, K.-W. Chung, and K.-S. Hong, "An implementation of leaf recognition system based on leaf contour and centroid for plant classification," in *Ubiquitous Information Technologies and Applications*, ser. Lecture Notes in Electrical Engineering, Y.-H. Han, D.-S. Park, W. Jia, and S.-S. Yeo, Eds. Springer Netherlands, 2013, vol. 214, pp. 109–116.
- [11] K.-B. Lee and K.-S. Hong, "An implementation of leaf recognition system using leaf vein and shape," *International Journal of Bio-Science and Bio-Technology*, pp. 57–66, Apr 2013.
- [12] M. G. Larese, R. Namías, R. M. Craviotto, M. R. Arango, C. Gallo, and P. M. Granitto, "Automatic classification of legumes using leaf vein image features," *Pattern Recogn.*, vol. 47, no. 1, pp. 158–168, Jan. 2014.
- [13] Y. Herdiyeni and M. Santoni, "Combination of morphological, local binary pattern variance and color moments features for indonesian medicinal plants identification," in *Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference*, Dec 2012, pp. 255–259.
- [14] A. Bhardwaj, M. Kaur, and A. Kumar, "Recognition of plants by leaf image using moment invariant and texture analysis," *International Journal of Innovation and Applied Studies*, vol. 3, no. 1, pp. 237–248, 2013.
- [15] Q. Nguyen, T. Le, and N. Pham, "Leaf based plant identification system for android using surf features in combination with bag of words model and supervised learning," in *International Conference on Advanced Technologies for Communications (ATC)*, October 2013.
- [16] A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Leaf classification using shape, color, and texture features," *International Journal of Computer Trends and Technology*, 2011.
- [17] A. Joly, H. Goëau, H. Glotin, C. Spampinato, P. Bonnet, W. Vellinga, R. Planquè, A. Rauber, S. Palazzo, B. Fisher, and H. Müller, "Lifeclef 2015: Multimedia life species identification challenges," in *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 6th International Conference of the CLEF Association, CLEF 2015, Toulouse, France, September 8-11, 2015, Proceedings*, 2015, pp. 462–483.
- [18] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.
- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, P. Bartlett, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., 2012, pp. 1106–1114.
- [20] C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, inception-resnet and the impact of residual connections on learning," *CoRR*, vol. abs/1602.07261, 2016. [Online]. Available: <http://arxiv.org/abs/1602.07261>
- [21] A. Canziani, A. Paszke, and E. Culurciello, "An analysis of deep neural network models for practical applications," *CoRR*, vol. abs/1605.07678, 2016. [Online]. Available: <http://arxiv.org/abs/1605.07678>
- [22] P. Bonnet, A. Joly, H. Goëau, J. Champ, C. Vignau, J.-F. Molino, D. Barthélémy, and N. Boujemaa, "Plant identification: man vs. machine," *Multimedia Tools and Applications*, pp. 1–19, 2015.
- [23] A. Joly, H. Goëau, P. Bonnet, V. Bakić, J. Barbe, S. Selmi, I. Yahiaoui, J. Carré, E. Mouysset, J.-F. Molino, N. Boujemaa, and D. Barthélémy, "Interactive plant identification based on social image data," *Ecological Informatics*, vol. 23, pp. 22 – 34, 2014, special Issue on Multimedia in Ecology and Environment. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S157495411300071X>
- [24] H. Goëau, P. Bonnet, and A. Joly, "Plant identification in an open-world (lifeclef 2016)," in *Working Notes of CLEF 2016 - Conference and Labs of the Evaluation forum, Évora, Portugal, 5-8 September, 2016.*, 2016, pp. 428–439. [Online]. Available: <http://ceur-ws.org/Vol-1609/16090428.pdf>
- [25] S. Choi, "Plant identification with deep convolutional neural network: Snumedinfo at lifeclef plant identification task 2015," in *CLEF*, 2015.
- [26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [27] J. Carranza-Rojas and E. Mata-Montero, "Combining Leaf Shape and Texture for Costa Rican Plant Species Identification," *CLEI Electronic Journal*, vol. 19, pp. 7 – 7, 04 2016.
- [28] C. N. Silla and A. A. Freitas, "A survey of hierarchical classification across different application domains," *Data Min Knowl Disc*, vol. 22, pp. 31–72, 2011.
- [29] F. Wu, J. Zhang, and V. Honavar, *Learning Classifiers Using Hierarchically Structured Class Taxonomies*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 313–320. [Online]. Available: [http://dx.doi.org/10.1007/11527862\\_4](http://dx.doi.org/10.1007/11527862_4)