

INVESTIGACIÓN SOBRE EL RECONOCIMIENTO AUTOMÁTICO FOLIAR DE LA LAURISILVA CANARIA.

La investigación que se resume en el presente artículo, da evidencia de la posibilidad de lograr exitosamente una taxonomía de laurásea canaria de forma automática, usando muestras de hojas. Imágenes digitalizadas por medio de un escáner, de un muestrario de hojas, son usadas como elementos de clasificación. Éstas son procesadas, extrayendo parámetros discriminantes que, por medio de diferentes modelos de aprendizaje automático, los identifican como descriptivos de pertenencia a una especie. El resultado se presenta con un valor de evidencia, medición que permite evaluar el peso de parámetros o conjuntos de parámetros en la clasificación. El sistema podría ser usado para determinar parámetros taxonómicos que permitan establecer un sistema de patentes de plantas desarrolladas o descubiertas, de interés económico.

This paper summarizes a research that establishes the possibility of an automatic successful taxonomic task, for laurisilvae canariensis trees, using simple leaves. Images of scanned leaves are used as means of classification. These images are processed obtaining discriminating parameters that, by means of automatic learning models, are identified as descriptors of class belonging. The obtained result is weighted with an evidence value, a measure that allows for valuation of parameters or parameters set impact on the classifying task. The system may be used as mean of taxonomic parameter identification, allowing for a registration board or a plant patent chart setting, of developed or discovered valuable species.

Juan Carlos Briceño Lobo

Carlos Manuel Travieso González

Jesús Bernardino Alonso Hernández

Miguel Ángel Ferrer Ballester

PRESENTACIÓN

Este artículo resume una labor de investigación sobre el reconocimiento automático de hojas de laurisilva canaria con una tasa de éxito del 88 %.

INTRODUCCIÓN

El reconocimiento y taxonomía de especies biológicas, tiene importancia no sólo desde el punto de vista académico, sino también económico. Distinguir inequívocamente la pertenencia de un espécimen o la determinación de una nueva variedad, para las compañías productoras de especies vegetales, puede significar ingresos considerables (Silvey, 1984). A título de ejemplo, (Warren, 1997), cita el estableci-

miento de un PBR (Plant Breeders Rights) en el Reino Unido. Esta normativa de estandarización, protege los derechos de los criadores que desarrollan nuevas variedades de plantas, de la misma forma que los inventores patentan máquinas y procesos. Con el fin de aligerar el procedimiento taxonómico de plantas efectuado por un especialista, o con fines de establecer un PBR en el registro de especies de rédito ayudando a establecer los parámetros discriminantes que conducen a su clasificación; una herramienta alternativa y adecuada es el establecimiento de un “reconocedor automático”.

Tal sistema es un procedimiento informático que, partiendo de evidencia sobre la descripción de un

Esta normativa de estandarización, protege los derechos de los criadores que desarrollan nuevas variedades de plantas, de la misma forma que los inventores patentan máquinas y procesos.

especimen, puede efectuar una clasificación. Es además deseable que dicha clasificación sea calificada con algún grado de certidumbre que nos permita evaluar la necesidad de “una segunda opinión” o de la búsqueda de nueva información que, eventualmente, sería conducente a una clasificación más probable o de certidumbre mayor. Una clasificación calificada tiene ventajas, ya que el sistema puede usarse como herramienta de evaluación en el comportamiento de los parámetros. Esto se logra en la medida en que se pueda “seguir” el peso específico de cada parámetro o conjunto de parámetros, de manera a contrastar en qué punto una especie se distingue de otra; o cómo una especie, por medio de muestras recolectadas en el tiempo, muta en otra particular y es identificada por el clasificador como distinta o como nueva.

PROCEDIMIENTO

El procedimiento utilizado para la taxonomía automática foliar comienza con la creación de una base de datos con las diferentes especies recogidas seguido de la parametrización de una forma óptima de las diferentes características de las hojas. El siguiente paso consistirá en insertarlas en un sistema basado en redes neuronales capaz de reconocer dichas especies. Posteriormente se comprobará que la tasa de reconocimiento es válida, dando por buena la aplicación de este método de estudio.

Para ello será necesario disponer un ordenador personal equipado con un escáner que sirva de interfase entre las hojas recogidas para la base de datos y los ficheros guardados en algún formato digital para la imagen. Este conjunto de ficheros formará la base de datos de partida a la cual habrá que realizarle una extracción de parámetros que relacione unívocamente cada muestra a cada especie. Una vez extraídos estos parámetros, se pasará al siguiente paso que consistirá en entrenar una red neuronal dimensionada para tal cometido. La elección de la red neuronal más adecuada para lograr nuestro propósito ha sido estudiada en este trabajo. El entrenamiento consistirá en introducirle un número adecuado de muestras de la base de datos que sirvan de patrón de conocimiento. El último paso es testear la red, es decir, el resto de muestras de nuestra base de datos se introducirá en el sistema diseñado y se comprobará la efectividad de nuestra red neuronal (ver figura 1).

CREACIÓN DE LA BASE DE DATOS

Para poder realizar un sistema de reconocimiento de diversas especies vegetales, es absolutamente imprescindible construir una base de datos. Esa base de datos deberá contener las muestras de las distintas especies que se pretenden estudiar. El número de muestras

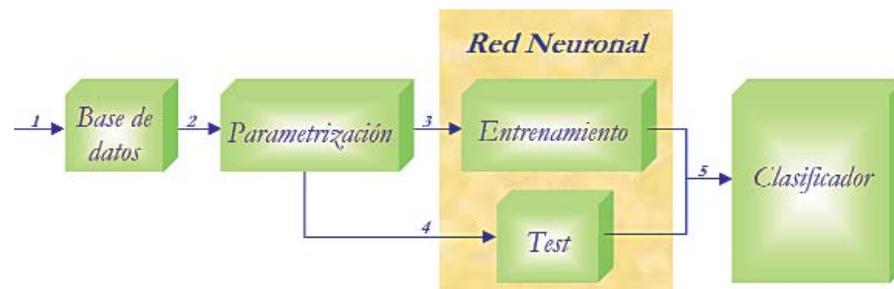


Figura 1. Sistema de taxonomía foliar automática.

será lo suficientemente grande como para, primero, poder entrenar la red neuronal con garantías y, segundo, testar dicha red neuronal con el fin de examinar los resultados obtenidos. Además, la cantidad de muestras escogidas por especie vegetal debe abarcar la mayor cantidad de formas y estructuras que puede tomar una única especie. De esta manera se asegura un estudio más robusto de las distintas especies vegetales.

Atendiendo a este razonamiento, la recogida de las muestras se realizó en diferentes épocas del año, pretendiendo con ello abarcar todos los colores y formas que las hojas van tomando a lo largo de las cuatro estaciones. Además, se prestó especial atención en desear aquellas muestras que se encontraron degradadas, intentando siempre que las muestras seleccionadas presentaran un buen estado de conservación.

Las especies que se eligieron tienen la particularidad de ser autóctonas de las Islas Canarias, endémi-

cas o no, que están encuadradas en las zonas bajas (Mocán) y las zonas montañosas húmedas de la Laurisilva y del Fayal-Brezal. Se suelen situar entre los 500 metros sobre el nivel del mar y un poco más de los 1500 metros (ver figura 2) cubriendo variadas características altitudinales. Se puede encontrar a menor altitud como en el caso de la isla de La Palma, debido a que posee grandes zonas húmedas que se extienden prácticamente hasta el nivel del mar (Santos, 1979).

La base de datos se compone de 16 especies distintas, con 75 muestras de cada una de ellas. A continuación se listan todas las especies que forman parte del estudio: *Especie 1*: Hija (*Prunus lusitanica* L. spp. hixa (Will.) Franco), *Especie 2*: Acebiño (*Ilex canariensis* Poir.), *Especie 3*: Follao (*Viburnum rigidum* Vent.), *Especie 4*: Naranja salvaje (*Ilex perado* Ait. spp. *platyphylla* Webb & Berth.), *Especie 5*: Til o Tilo (*Ocotea foetens* (Ait.) Benth. & Hook. f.), *Especie 6*: Laurel o Loro (*Laurus azorica* (Seub.) Franco), *Especie 7*: Palo blanco (*Picconia*

Las especies que se eligieron tienen la particularidad de ser autóctonas de las Islas Canarias, endémicas o no, que están encuadradas en las zonas bajas y las zonas montañosas húmedas de la Laurisilva y del Fayal-Brezal

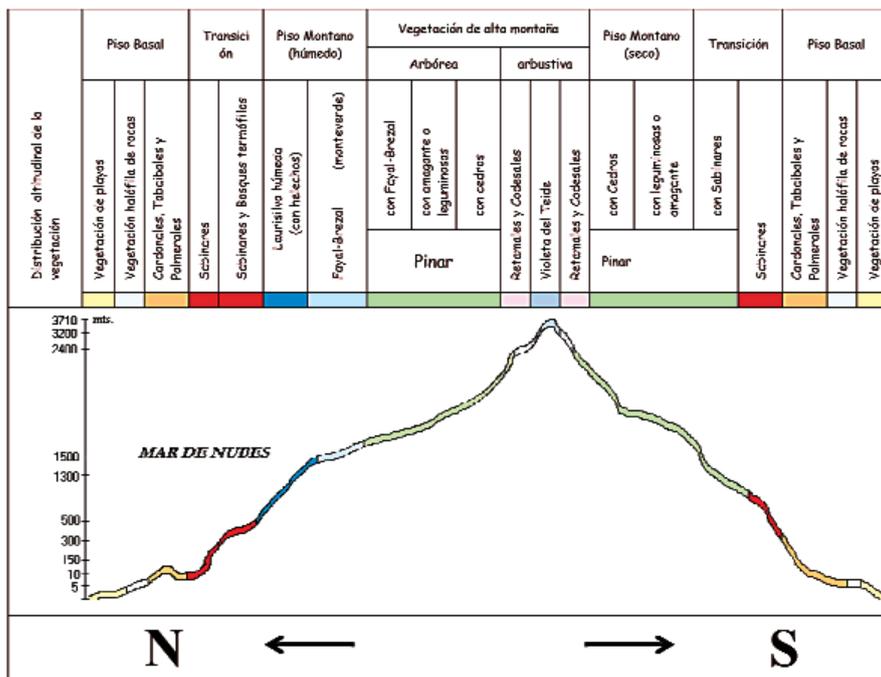


Figura 2. Distribución altitudinal de los tipos de vegetación en las Islas Canarias.

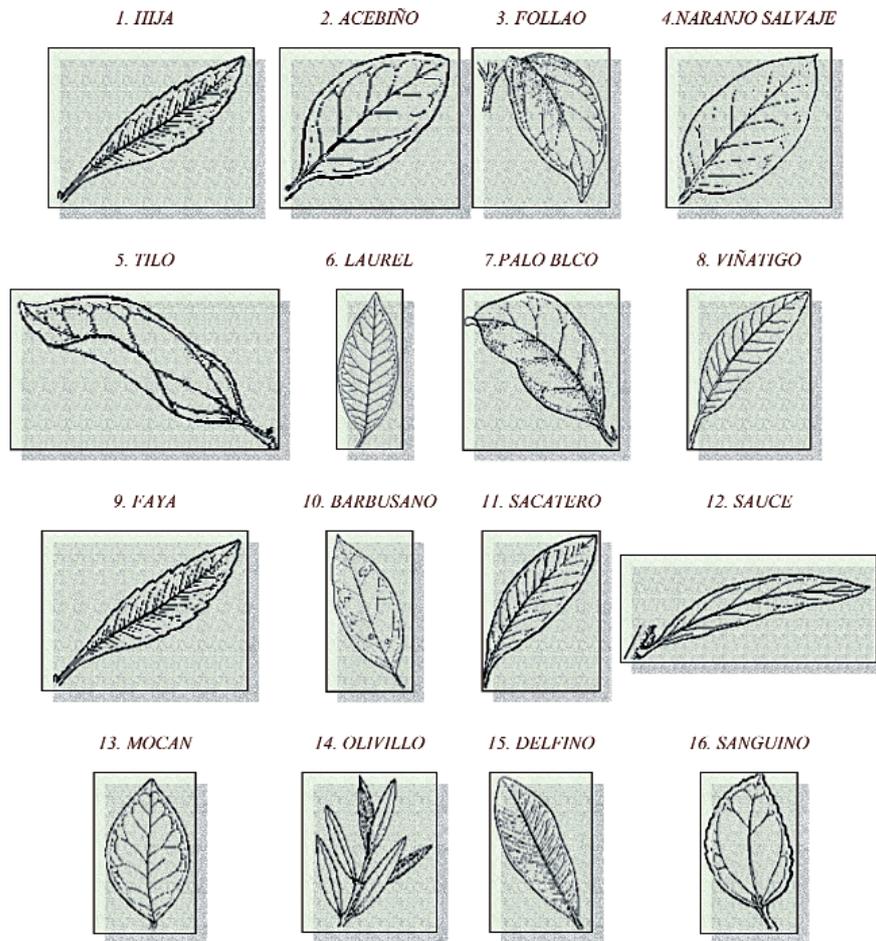


Tabla I. Descripción de las hojas de laurisilva Canaria reconocidas por el sistema de taxonomía foliar automática.

excelsa (Ait.) DC), *Especie 8*: Viñátigo (*Persea indica* (L.) Spreng.), *Especie 9*: Faya (*Myrica faya* Ait.), *Especie 10*: Barbusano (*Apollonias barbujana* (Cav.) Bornm), *Especie 11*: Sacatero o Aderno (*Heberdenia excelsa* (Ait.) & DC.), *Especie 12*: Sauce o Sao (*Salix canariensis* Chr.Sm. ex Link), *Especie 13*: Mocán (*Visnea mocanera* L. f.), *Especie 14*: Olivillo (*Phillyrea angustifolia* L.), *Especie 15*: Delfino (*Myrsine canariensis* (Willd.) A.DC), *Especie 16*: Sanguino (*Rhamnus glandulosa* Ait.). En la Tabla I se describen las hojas de dichas especies. (Font Quer, 1973)

Como se observará, hay especies que no aparecen y que se definen en el apartado anterior. La razón de que no estén el almácigo,

el lentisco, el brezo y el tejo, es que son especies de hoja compuesta y están divididas en foliolos. El objetivo de este trabajo es reconocer hojas y no foliolos, por lo que más adelante se dejará propuesta otra metodología posible para el reconocimiento de estas hojas compuestas (Torres, 2001).

Para poder desarrollar la base de datos se ha hecho necesario seguir una metodología de trabajo basada en la recogida y posterior conservación de las especies a similitud de la realizada por parte de los biólogos. Consiste en los siguientes pasos: Recogida selectiva de las hojas en los tiempos establecidos y zonas localizadas (ver Tabla II), limpieza de impurezas no deseadas como polvo o tierra, pren-

| <i>Zona de Recogida / Localidad / Isla</i> | <i>Coordenadas UTM</i> | <i>Fecha Recogida</i> |
|--|--|-----------------------|
| Pico del Inglés. Monte de Las Mercedes. (La Laguna) Tenerife | X = 376.280 Y = 3.156.840 Z = 903 m. | 15 de Abril de 2000 |
| | | 11 de Febrero de 2001 |
| Camino de los Laureles. Entrada a Tegueste. (La Laguna) Tenerife | X = 369.600 Y = 3.155.460 Z = 475 m. | 15 de Junio de 2000 |
| | | 11 de Febrero de 2001 |
| Roque de las Ánimas. (Taganana) Tenerife | X = 379.420 Y = 3.159.580 Z = 689 m. | 15 de Abril de 2000 |
| Campo de la Manzanilla (La Laguna) Tenerife | X = 370.900 Y = 3.152.760 Z = 552 m. | 9 de Febrero de 2001 |
| Cabezo de Zapata. Monte de Las Mercedes. (La Laguna) Tenerife | X = 374.505 Y = 3.156.780 Z = 982 m. | 15 de Abril de 2000 |
| | | 28 de Enero de 2001 |
| | | 8 de Febrero de 2001 |

TABLA II. Datos sobre la recolección.

sado de las especies, que elimina el agua que poseen las hojas para una adecuada conservación de las mismas, y adhesión de las hojas prensadas a las plantillas, que va a servir para escanearlas y al mismo tiempo para su conservación.

A continuación se hace un recuento del material utilizado para la metodología de trabajo seguida. En la recogida es necesario unas bolsas plásticas para la clasificación y posterior transporte de las especies; para la limpieza, un trapo de algodón impregnado en agua y un cepillo de dientes para eliminar posibles pulgones que han sido detectados en ciertas especies; para el prensado, papel absorbente y un cartón separador entre capas, y una prensa de 20 a 30 kilos. La adhesión de las hojas a las plantillas se realiza mediante cinta con capa adherente de doble cara, y para el almacenamiento de cada especie debidamente clasificada y caracterizada se han utilizado carpetas de

cartón a similitud de la metodología seguida para la clasificación y conservación de especies en el herbario del Jardín de Aclimatación de la Orotava.

Las opciones para el escaneo son una resolución de 300 dpi, escala de grises a 8 bits (256 niveles de gris) con el ápice y pecíolo alienados en vertical (ver figura). Las imágenes escaneadas se guardan en formato 'bmp'. Posteriormente se transformarán a formato "jpeg" (Joint Photographic Experts Group). Las figuras 3 y 4 muestran el procedimiento de recogida y un ejemplo de hoja escaneada.

PARAMETRIZACIÓN

El primer paso que debe darse para poder realizar un sistema de reconocimiento consiste en la extracción de parámetros característicos. Cabe preguntarse ahora cuáles son los parámetros que defi-

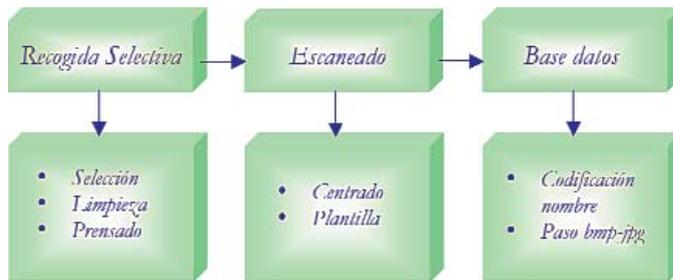


Figura 3. Esquema del procedimiento de recogida de hojas.

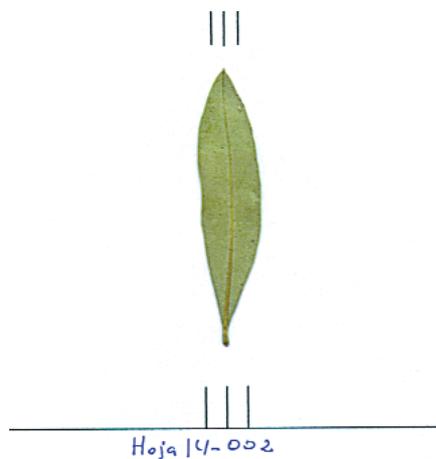


Figura 4. Ejemplo de hoja escaneada antes de su paso a escala de grises.

nen a todas y cada una de las especies y nos permiten diferenciarlas entre sí con un porcentaje de error mínimo. Pues bien, muchos de esos parámetros son intuitivos: las dimensiones (largo x ancho), la forma, la rugosidad, la longitud del peciolo, los colores, etc. Son parámetros que inconscientemente usamos cuando reconocemos las cosas a diario. Por lo tanto, serán de ese tipo los que se usarán para tratar de diferenciar las especies vegetales.

Para ello, se elaborarán diversos programas en el lenguaje de programación matemático MATLAB. Con esos algoritmos se podrá, en algunos casos, medir diversas magnitudes (como puede ser el ancho y el largo de la hoja), y en otros, verificar la presencia o ausencia de una determinada característica de la

hoja (por ejemplo, la existencia de glándulas en algunas especies).

Este proceso de extracción de parámetros se realizará en dos fases: en primer lugar, se hará un corte a grandes rasgos de la zona de estudio (se centra la atención en aquello que estamos estudiando) y, en segundo lugar, se obtendrán las peculiaridades de la zona bajo observación. Las zonas de estudio son tres: el cuerpo de la hoja, el peciolo y el ápice (ver figura 5), siendo los parámetros obtenidos de cada zona, los reflejados en la figura 6.

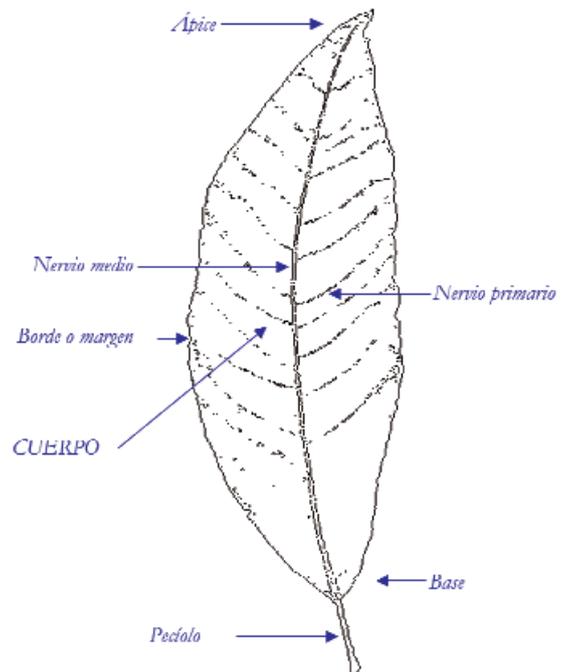


Figura 5. Partes de la hoja.

EL CUERPO

El cuerpo de la hoja es, sin duda, la zona que más información aporta al sistema de reconocimiento. Aquí es necesario distinguir dos partes en la extracción de parámetros: una primera que comprende los parámetros relacionados con el cuerpo en su totalidad, como son el largo, ancho, escala, color y lóbulos, y una

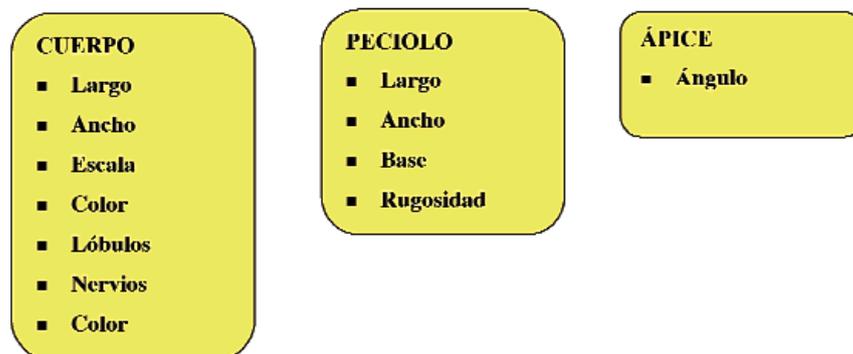


Figura 6. Parámetros obtenidos de cada parte de la hoja.

segunda parte, que tiene que ver con zonas más específicas del cuerpo de la hoja, como es la nerviación y la existencia de glándulas.

Para estudiar cada una de esas peculiaridades es imprescindible, ya que de esta forma se facilitan mucho las operaciones, realizar un corte de la imagen de la hoja en la zona de interés. Por ejemplo, para calcular la existencia de glándulas en la hoja, se recorta la imagen a la zona central inferior de la hoja, ya que es ahí donde es posible encontrar las glándulas. Así, las operaciones que han de aplicarse para determinar la existencia o no de esas glándulas se llevan a cabo en una zona específica del cuerpo de la hoja y se minimizan los errores que puedan producirse si se estudia el cuerpo en su totalidad.

Siguiendo con la misma línea, podemos establecer que para el estudio de la nervadura se eligió la zona central de los nervios de la hoja. Se optó por esta zona ya que es la que, en general, tiene una tonalidad más uniforme y, por lo tanto, es en ella donde más fácilmente se pueden llevar a cabo las diferentes operaciones. Por último, para medir el largo, ancho y determinar los lóbulos del borde del limbo, se seleccionó, como es natural, el cuerpo de la hoja en su totalidad.

EL PECÍOLO

Los parámetros obtenidos del pecíolo fueron largo, ancho, existencia / ausencia de la base y rugosidad. El proceso seguido para el cálculo de estas características fue

El cuerpo de la hoja es, sin duda, la zona que más información aporta al sistema de reconocimiento. Aquí es necesario distinguir dos partes en la extracción de parámetros: una primera que comprende los parámetros relacionados con el cuerpo en su totalidad, como son el largo, ancho, escala, color y lóbulos, y una segunda parte, que tiene que ver con zonas más específicas del cuerpo de la hoja, como es la nerviación y la existencia de glándulas.

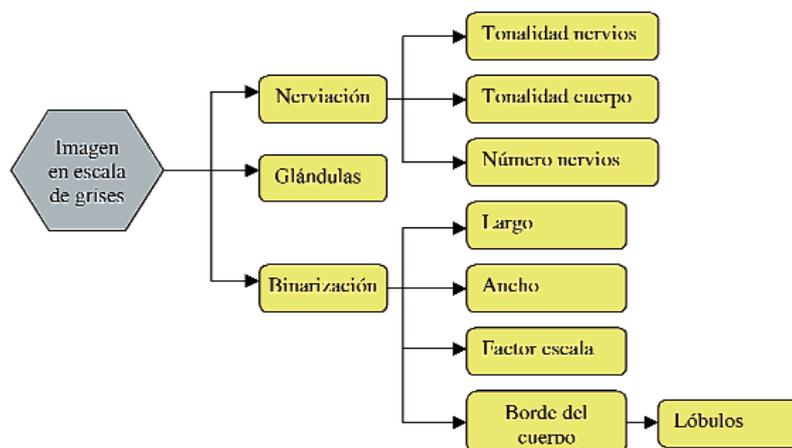


Figura 7. Obtención de los parámetros del cuerpo.

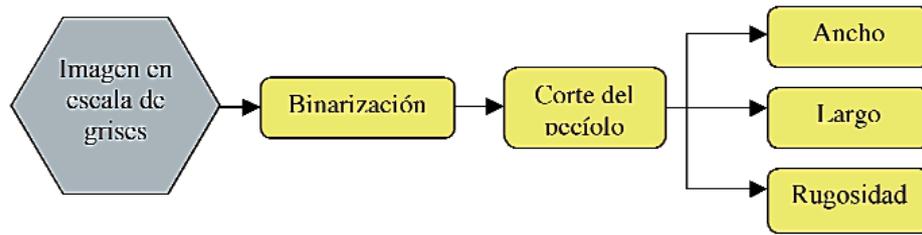


Figura 8. Obtención de los parámetros del pecíolo.

el siguiente: Recorte de la imagen a la zona de interés, cálculo del largo y del ancho del pecíolo, se determina también si la hoja presenta o no base, y estudio de la rugosidad del pecíolo.

El punto más conflictivo aquí, se encuentra a la hora de diferenciar lo que pertenece al cuerpo de la hoja de lo que pertenece al pecíolo. Una vez que se salva esta dificultad, el resto de las medidas es relativamente más sencillo.

EL ÁPICE

El ángulo que forma el ápice de la hoja se erige como un parámetro muy importante a la hora de caracterizar las distintas especies vegetales, y es exactamente la medida del ángulo del ápice lo que se busca en este apartado. Para ello, y al igual que con los anteriores parámetros, se comienza seleccionando la zona de la hoja de interés, se trazan las tangentes cerca del ápice y se calcula el ángulo.

CLASIFICACIÓN

Una vez que se ha conseguido caracterizar las distintas especies

vegetales mediante la obtención de una serie de parámetros, se deben construir las redes neuronales que permitan clasificar las clases de árboles que se están analizando.

Para ello, se realiza un estudio de las distintas posibilidades que ofrecen las redes neuronales (funcionamiento, topología, etc.). El tipo de red neuronal más adecuado para el problema que se está tratando no se sabe a priori. Únicamente después de probar varias opciones, se podrá optar por el que ofrezca unos resultados más acordes con los objetivos buscados.

A continuación se detallarán los pasos seguidos para elaborar el sistema más adecuado para clasificar las muestras de las distintas especies vegetales. (Bishop, 1997)

NORMALIZACIÓN

Para poder utilizar los datos obtenidos en el proceso de parametrización anterior, éstos deben sufrir algún tipo de modificación. Esto se debe a que existen grandes diferencias entre ellos, aún tratándose de datos referidos al mismo aspecto (por ejemplo, grandes diferencias

Una vez que se ha conseguido caracterizar las distintas especies vegetales mediante la obtención de una serie de parámetros, se deben construir las redes neuronales que permitan clasificar las clases de árboles que se están analizando.

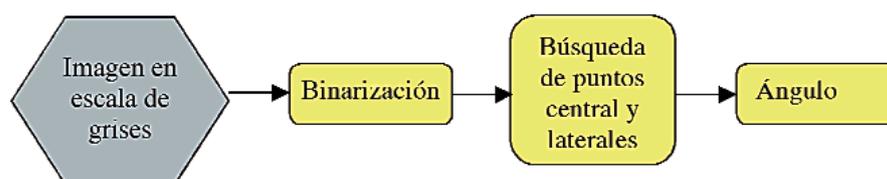


Figura 9. Obtención de los parámetros del ápice.

de magnitud entre los resultados que se obtuvieron en el proceso de medida del largo de pecíolo).

Este proceso de normalización es muy sencillo, y consta de los siguientes pasos: primero, se calcula el máximo de cada parámetro para cada una de las especies; a continuación, se calcula el máximo de los máximos anteriores; y por último, se divide cada uno de los valores de los diferentes parámetros y de las distintas especies por el correspondiente máximo.

RED NEURONAL DE ALIMENTACIÓN PROGRESIVA

Como ya se ha advertido, la clasificación taxonómica de especies de árboles a partir de la hoja presenta varios problemas. Uno de ellos resulta de las grandes diferencias existentes entre las distintas hojas de una misma especie. Esto conduce a pensar que un modelo jerárquico que vaya estudiando las distintas características y discrimine directamente a partir de la presencia o ausencia de una de ellas, no sea la mejor solución.

Por lo tanto, parece que un modelo en el que se tengan en cuenta todos los parámetros, y en el que se pondere adecuadamente el peso específico de cada uno de ellos en la decisión final, sí es el que pudiera ofrecer el mejor resultado. Es esta la razón fundamental por la que se ha optado por la red neuronal de alimentación progresiva.

De manera general, una red neuronal se define como un grafo compuesto por varios nodos conectados entre sí, donde esos nodos poseen una determinada función de transferencia, y donde los enlaces entre ellos están ponderados por diversos valores. Estos valores tienen la capacidad de ir variando dinámicamente durante el proceso de aprendizaje de la red. Además, la red en sí, se

dice que tiene un procedimiento de funcionamiento concreto.

Existen varios tipos de redes neuronales, muchos de ellos diseñados para resolver problemas específicos. Las características que presentaba el problema de la clasificación de diversas especies arbóreas a partir de su hoja, hicieron pensar que la clase de red neuronal más adecuada para dar respuesta a esta dificultad era la red neuronal de alimentación progresiva con una sola capa oculta y con funciones de transferencia sigmoideas. (Bishop, 1997)

ALGORITMO DE APRENDIZAJE

Una vez que se ha elegido una determinada arquitectura (un tipo de red neuronal) se debe optar por un algoritmo de aprendizaje para la red; es decir, un método que ayude a calcular los pesos que deben asignarse a cada uno de los nodos. El objetivo que se persigue con esto es tratar de conseguir que la red asigne unos valores de salida considerados como correctos para unos valores de entrada determinados.

El método que se eligió para el entrenamiento y el aprendizaje de la red fue el algoritmo de *backpropagation*, combinado con el LMS (menor error cuadrático medio). Éste se utiliza para conseguir los pesos óptimos para una sola neurona, pero sin contar con la función de transferencia. Con el algoritmo de *backpropagation* se obtiene un algoritmo de aprendizaje válido para la red de alimentación progresiva al completo. (Briceño, 2002a)

RESULTADOS

Para finalizar la presentación del trabajo realizado, sólo resta presentar los resultados conseguidos y las conclusiones que de ellos se derivan.

PORCENTAJES DE PRUEBA

La red neuronal elegida necesita de un periodo de entrenamiento o aprendizaje. Para que pueda llevarse a cabo dicho entrenamiento, se proporciona un determinado porcentaje de muestras de la cantidad total disponible. Las restantes muestras se utilizarán para testar la red, estudiando si dicha red ofrece unos resultados adecuados.

En principio, no se conoce cuál es el porcentaje de muestras idóneo para el entrenamiento de la red. No todas las especies presentan el mismo grado de variabilidad en la forma y estructura de sus hojas, por lo que no se sabe que número de muestras es el adecuado para lograr cubrir todas las posibilidades. Hay especies en las que sus hojas tienen un alto grado de homogeneidad entre sí, lo que permite reducir el número de muestras necesarias para el entrenamiento. Pero también existen otras clases de árboles con hojas que exhiben claras diferencias entre sí, lo que hace que sea muy importante aumentar la cantidad de muestras utilizadas para el entrenamiento, de forma que puedan tenerse en cuenta la mayor cantidad de casos posibles.

Finalmente se optó por escoger un 30% de las muestras para realizar el entrenamiento, mientras que el 70% restante se usará para testar la red.

ARQUITECTURAS UTILIZADAS

Dentro de la red neuronal de alimentación progresiva elegida, se estudiaron diversas variantes. Las características comunes a todas ellas son las siguientes: Tienen 13 entradas diferentes, una por cada uno de los parámetros estudiados en las hojas, Cada una cuenta con 16 salidas, una por cada una de las especies en estudio. Para cada una de las estructuras diseñadas se realizaron diez pruebas. Se ha obtenido un porcentaje del grado de reconocimiento para cada una de las redes propuestas, así como la varianza y la desviación típica.

Los experimentos realizados son los siguientes: Experimento 1, con red neuronal de 26 nodos en la capa oculta; Experimento 2, con red neuronal de 2 capas ocultas de 10 y 30 neuronas; Experimento 3, con red neuronal de 13 nodos en la única capa oculta; Experimento 4, con red neuronal de 16 nodos en la capa oculta; Experimento 5, con red neuronal igual al experimento 4 pero entrenando con un 50% de la base de datos y testeado con un 50%. Los resultados se muestran en la Tabla III (Briceño, 2002b).

MÉTODO DEL MODELO MÁS VOTADO

Este modelo consiste en coger varias redes y aplicarles a todas el mismo test en paralelo. A continua-

| | EXPERIMENTO | | | | |
|----------|-------------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| MEDIA | 85,7267 | 84,7093 | 84,0262 | 76,8459 | 83,4884 |
| VARIANZA | 1,0834 | 1,0697 | 1,3153 | 5,447 | 0,9613 |

Tabla III. Tasa de Acierto del Sistema de Taxonomía Foliar Automática.

| Nº redes | EXPERIMENTO | | | | |
|----------|-------------|---------|---------|---------|---------|
| | 1 | 2 | 3 | 4 | 5 |
| 4 | 87,936 | 87,3547 | 84,3023 | 85,3198 | 86,6279 |
| 6 | 87,7907 | 87,7909 | 85,7558 | 86,3372 | 86,1919 |
| 8 | 88,0814 | 88,2267 | 86,7733 | 87,2093 | 86,1919 |
| 10 | 88,3721 | 88,0814 | 86,0465 | 88,0814 | 86,1919 |

Tabla IV. Tasa de Acierto Mediante el Modelo Más Votado.

ción, sus resultados serán comparados y se seleccionará el resultado por el que más redes hayan votado. Los resultados obtenidos, utilizando este procedimiento, se muestran en la tabla IV resaltando en amarillo los óptimos.

ENCUESTA

Para finalizar, se realizó una encuesta con el fin de comparar los resultados obtenidos por el clasificador con la capacidad de percepción del ser humano para realizar la clasificación de las diversas especies. A los encuestados se les explicó detenidamente el proceso y el fin que se pretendía conseguir. Se les dejó un tiempo para que se familiarizaran con las distintas especies, con sus características, sus formas, etc. Además, para llevar a cabo la clasificación, tenían una muestra, a modo de ejemplo, de cada una de las especies para que pudieran comparar y decidir cuál de ellas les parecía la más acertada. La media de aciertos que se logró con estos encuestados fue del 51,7%, resultado claramente inferior al 88,8% conseguido con el sistema desarrollado.

CONCLUSIONES

Se ha expuesto aquí el trabajo realizado sobre la taxonomía de algunas especies vegetales usando un programa informático. Se comenzó presentado el proceso de construcción de la base de datos. Posteriormente, se vieron las

diversas técnicas aplicadas para extraer los parámetros y el por qué de su utilización. A continuación, se expusieron los motivos que llevaron a elegir la red neuronal de alimentación progresiva como la tipología más adecuada para este problema. Por último, se vio que los resultados conseguidos son satisfactorios, teniendo en cuenta los diversos problemas presentes. A partir de este punto, se ahondará aún más en la clasificación de estas mismas especies vegetales. La intención es tratar de obtener más parámetros significativos de las plantas. De esta forma se mejorará, entre otras cosas, la tasa media de reconocimiento, y se dotará al programa de mayor robustez frente a variaciones en la forma y la estructura de las hojas.

BIOGRAFÍA

JUAN CARLOS BRICEÑO LOBO

Graduado en matemáticas y matemáticas aplicadas por L' Université de Dijon, Francia, actualmente trabaja bajo la dirección del Dr. Miguel Ángel Ferrer Ballester en su tesis doctoral en el Departamento de Señales y Comunicaciones de la Universidad de Las Palmas de Gran Canaria. Ha trabajado en proyectos de investigación bajo la dirección del: Dr. Manuel María Murillo *Establecimiento de Modelos de Administración Pesquera en el Golfo*

Dulce (gobierno de Costa Rica), la Dra. Rafaela Sierra *Determinación de Parámetros de Predicción de Cáncer Gástrico* (JAICA (Japan International Cooperation Agency)/ Ministerio de Salud Pública (Gobierno de Costa Rica)) y el Dr. Ian Gould Director del Departamento de Entomología del Museo Británico *DASY A Classifying Digital Automatic System for Etchiumunae Wasp* (Darwin Foundation). Es profesor de la Universidad de Costa Rica, en el Departamento de Ciencias de la Computación e Informática.

Dirección de Juan Carlos Briceño Lobo:

Escuela de Ciencias de la Computación e Informática, Universidad de Costa Rica. Ciudad Universitaria Rodrigo Facio, San Pedro Montes de Oca, San José Costa Rica.

Correo electrónico:

juanca@ecci.ucr.ac.cr

Dirección actual del grupo de investigación

Miguel Ángel Ferrer Ballester
Departamento de Señales y Comunicaciones,
Universidad de Las Palmas de Gran Canaria
Campus de Tafira
35017 Las Palmas
Teléfono: 928451269
Fax: 928451243
Web: www.gpds.ulpgc.es
Mail: mferrer@dsc.ulpgc.es

BIBLIOGRAFÍA

Bishop C.M. (1997), "Neural Network for Pattern Recognition", Clarendon Press, Oxford.

Briceño J.C, Travieso C. y Ferrer M. A.(2002): "Automatic Recognition of Simple Laurisilvae Canariensis Leaves, by Perimeter Characterization". IASTED International Conference on Signal Processing, Pattern Recognition, and Applications, Crete SPPRA 2002.

Briceño J. C., Travieso C., Ferrer M. A.(2002): "Un modelo simple de probabilidades, para evaluar tamaños de contorno en la clasificación de objetos" XVII Simposium Nacional de la Unión Científica Internacional de Radio, URSI Alcalá de Henares, 2002

Font Quer Pío (1973): "Diccionario de Botánica". Barcelona. Editorial Labos, S.A.

Santos Arnoldo (1979): "Árboles de Canarias, Enciclopedia temática Canaria". Sta. Cruz de Tenerife. Editorial interinsular Canaria.

Silvey V.(1984) , "The Contribution of New Varieties to Cereals Yields in England and Wales between 1947 and 1983". Journal of the National Institute of Agricultural Botany, 17, 155-168.

Torres M. A., Briceño J.C, Alonso Itziar, Ferrer M.A.(2001): "Taxonomía de imágenes Foliares usando una red neuronal MLP". XVI Simposium Nacional de la Unión Científica Internacional de Radio, URSI 2001.

Warren David (1997): "Automated leaf shape description for variety testing in Chrysanthemums". IEEE Sixth International Conference on Image Processing and its Applications.

Patrocinadores de esta investigación:

BELEYMA, S.L.
UNIÓN ELÉCTRICA DE CANARIAS, S.A.
(UNELCO)