

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 992 109**

51 Int. Cl.:

G06V 10/26 (2012.01)

G06V 10/44 (2012.01)

G06V 10/82 (2012.01)

G06V 20/10 (2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.03.2021** **PCT/EP2021/056503**

87 Fecha y número de publicación internacional: **23.09.2021** **WO21185745**

96 Fecha de presentación y número de la solicitud europea: **15.03.2021** **E 21711260 (6)**

97 Fecha y número de publicación de la concesión europea: **14.08.2024** **EP 4121894**

54 Título: **Cuantificación del daño biótico en hojas de plantas, por redes neuronales convolucionales**

30 Prioridad:

16.03.2020 EP 20163426

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.12.2024

73 Titular/es:

BASF SE (100.0%)
Carl-Bosch-Strasse 38
67056 Ludwigshafen am Rhein, DE

72 Inventor/es:

BERECIARTUA-PEREZ, ARANZAZU;
PICON RUIZ, ARTZAI;
SPANGLER, CORINNA MARIA;
KLUKAS, CHRISTIAN;
EGGERS, TILL;
ECHAZARRA HUGUET, JONE y
NAVARRA-MESTRE, RAMON

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 992 109 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Cuantificación del daño biótico en hojas de plantas, por redes neuronales convolucionales

5 Campo técnico

La divulgación se refiere en general al procesamiento de imágenes por ordenadores, y más en particular se refiere a técnicas para cuantificar el daño biótico en las hojas de plantas.

10 Antecedentes

Las plantas de cultivo son plantas que se van a consumir por seres humanos (o por animales). El daño a las plantas, especialmente el daño biótico, se debe minimizar. Esto es especialmente pertinente en tanto que la planta de cultivo todavía está creciendo en el campo agrícola. Cuantificar el daño es una condición previa para que el agricultor aplique una medida adecuada que limite el daño.

Más detalladamente, los expertos en patología vegetal están familiarizados con las enfermedades en las plantas provocadas por patógenos (organismos infecciosos) y condiciones ambientales (factores fisiológicos). Muy simplificado y para señalar sólo un aspecto de muchos aspectos: como las plantas tienen hojas, el daño biótico se puede estimar al inspeccionar visualmente las hojas de plantas.

El daño puede tener diferentes razones. Por ejemplo, el daño por alimentación es el daño provocado por animales plaga (entre ellos algunas especies de insectos) que comen partes de las hojas. El daño bacteriano o el daño por hongos son ejemplos de daño en la estructura celular de la hoja. Se acaba de mencionar el término general "necrosis" para el daño celular.

El daño cambia la apariencia visual de las hojas, usualmente en su color. El daño por alimentación es visible como insectos que se quedan en las hojas, o como agujeros o partes cortadas de las hojas. Los cambios estructurales pueden cambiar el color de las hojas, al menos parcialmente. Puede haber otras razones más complejas. Por ejemplo, un ratón puede comer de la raíz tal que las hojas pierdan su color "verde".

Los agricultores pueden aplicar una medida de control de plagas adecuada u otra contramedida: tal como la aplicación de productos químicos específicos (tal como, por ejemplo, insecticidas, bactericidas, fungicidas). La eficiencia de las medidas está con relación a la cantidad del daño, a la cantidad de los productos químicos, así como a la exactitud y el tiempo que se tarda en cuantificar el daño. Sin embargo, diferentes agricultores pueden cuantificar los daños de manera diferente.

Al cuantificar el daño, los agricultores pueden operar cámaras (en dispositivos móviles y similares) para tomar imágenes de hojas. Las herramientas de procesamiento de imágenes desempeñan una función importante en la asistencia a los agricultores.

Sin embargo, las condiciones del mundo real para tomar imágenes de hojas habitualmente están lejos de ser óptimas: en condiciones reales de campo, los agricultores pueden tomar las imágenes, por ejemplo, por teléfonos inteligentes no enfocados con iluminación deficiente.

En tanto que las condiciones son difíciles de cambiar, existe una tarea para que las herramientas de procesamiento de imágenes compensen estas y otras insuficiencias. En otras palabras, existe el deseo de estimar el daño lo más objetivo posible y lo más repetible posible.

Yang Wu et al: Segmentación de órganos de cultivos e identificación de enfermedades con base en una red neuronal profunda débilmente supervisada. AGRONOMY, vol. 9, no. 11, 1 de noviembre de 2019 (2019-11-01), página 737 explica el uso de una red neuronal convolucional para reconocer enfermedades en las hojas de tomate. Para preparar la entrada para la CNN, las hojas se identifican al adicionar manualmente anotaciones de imagen en forma de rectángulos, en combinación con la aplicación posterior de procesamiento de color dentro de los rectángulos. Entonces, la CNN clasifica la enfermedad. No hay cuantificación.

Eric Hitimana et al: Estimación automática de la infección de la hoja de café en vivo con base en técnicas de procesamiento de imágenes, COMPUTER SCIENCE & INFORMATION TECHNOLOGY (CS& IT), 21 de febrero de 2014 (21-02-2014), páginas 255-266 explica un enfoque de dos pasos para estimar los daños, con el primer paso para remover el fondo de la imagen y el segundo paso para estimar la gravedad del daño. Las técnicas clásicas de procesamiento de imágenes se utilizan para ambos pasos.

Sumario

De acuerdo con realizaciones de la presente invención, un ordenador se adapta para cuantificar el daño biótico en las hojas de plantas de cultivo. el ordenador realiza un método implementado por ordenador que comprende recibir

imágenes de hojas y proporcionar indicadores de cantidad de daños. el ordenador proporciona los indicadores como grados de daño. Puesto que el ordenador utiliza herramientas de aprendizaje automático, realiza el método durante la llamada fase de prueba (o la llamada fase de producción).

- 5 Por el uso de estos indicadores, el agricultor puede aplicar las medidas en consecuencia.

El ordenador cuantifica el daño para las hojas de las plantas como daño por hoja. El daño por hoja es un valor numérico que ayuda al agricultor a aplicar las medidas en consecuencia.

- 10 Puesto que la imagen de la hoja muestra una hoja en primer plano junto con otras hojas o con otros objetos, el ordenador no puede derivar el daño por hoja de la imagen de hoja directamente. Por lo tanto, el procesamiento se distingue por las secuencias de pasos realizadas posteriormente. En la primera secuencia, el ordenador segmenta la imagen de hoja a una imagen de hoja individual. En otras palabras, la imagen de hoja individual es la imagen de hoja en la que permanece la hoja en primer plano y de la que se han removido las otras hojas u objetos. En la
15 segunda secuencia, el ordenador obtiene el daño por hoja de la imagen de hoja individual.

- En general, el ordenador utiliza técnicas de aprendizaje automático (ML) y, más en particular, el ordenador realiza ambas secuencias de pasos por dos redes neuronales convolucionales (CNN) entrenadas por separado. Una primera CNN realiza la primera secuencia, y una segunda CNN realiza la segunda secuencia. Puesto que cada
20 una de las CNN tiene múltiples capas, el ML se puede considerar como "aprendizaje profundo".

- Puesto que las imágenes son conjuntos de píxeles, la primera CNN proporciona la imagen de hoja individual como un subconjunto contiguo de píxeles que muestra sólo una hoja en primer plano en particular. La primera CNN deriva la imagen de hoja individual de la imagen de hoja por una técnica de separación (o técnica de delimitación, o técnica de segmentación). La segunda CNN procesa la imagen de hoja individual para obtener el grado de daño.
25 Este indicador de daño por hoja se calcula como la relación entre la superficie de una región de daño en la hoja con relación a la superficie de la hoja. La superficie de la hoja es con relación al número de píxeles que representan la hoja (con la resolución de la cámara que es un factor). Por lo tanto, la relación también se puede obtener como una relación de píxeles (por ejemplo, el número de píxeles que corresponden al daño y el número de píxeles que corresponden a la hoja).
30

Además de la relación de daño, el daño se podría cuantificar de otra manera, tal como por el número de puntos de daño o por el número de agrupaciones de puntos de daño.

- 35 La segunda CNN puede funcionar de acuerdo con dos alternativas de la segunda secuencia.

- En la primera alternativa, la segunda CNN utiliza la regresión para obtener el grado de daño como el resultado. En la segunda alternativa, la segunda CNN proporciona una imagen binaria como resultado intermedio, en donde la imagen binaria tiene un número contable de píxeles que muestran daño, y en donde la imagen de hoja individual
40 tiene un número contable de píxeles que muestran la hoja en primer plano.

Como condición, el ordenador se ha entrenado con conjuntos de datos de entrenamiento. El entrenamiento adecuado está disponible para ambas CNN.

- 45 La primera CNN se entrena con imágenes anotadas por humanos que son imágenes de hojas anotadas en la hoja (es decir, imágenes que muestran la hoja en primer plano, con anotaciones que identifican la hoja en primer plano sobre el resto de la imagen).

- La segunda CNN se entrena por separado para las alternativas.
50

Para la primera alternativa, la segunda CNN se entrena con anotaciones de daño (es decir, los usuarios expertos humanos indican el daño en términos de la relación). Las anotaciones de daño no identifican qué píxeles muestran daño y qué píxeles no.

- 55 Para la segunda alternativa, las anotaciones de usuarios humanos no están disponibles. Hay una razón simple para eso: las regiones de daño pueden ser demasiado pequeñas para que el experto haga anotaciones correctas. En su lugar, para la segunda alternativa, la segunda CNN se entrena con (una pluralidad de) pares de imágenes como entrada. Un par de imágenes comprende una imagen de hoja individual y una imagen binaria. Para cada par, la imagen binaria se ha derivado por procesamiento de color a partir de la respectiva imagen de hoja individual.
60 En tanto que las imágenes de hojas individual son imágenes en color, una herramienta de procesamiento de color modifica la codificación por color tal que las imágenes binarias tienen píxeles que identifican regiones de daño y tienen píxeles que identifican regiones de hoja.

- 65 Durante el entrenamiento, la segunda CNN (segunda alternativa) recibe los pares y aprende cómo llegar de las imágenes de hojas individual codificadas por colores a las imágenes binarias. De esta manera, la segunda CNN obtiene pesos y otros parámetros para el modelo de red, pero la segunda CNN no aplica procesamiento de color.

Durante la fase de prueba, la segunda CNN en la primera alternativa utiliza la regresión mencionada para obtener el grado de daño. La segunda CNN en la segunda alternativa proporciona la imagen binaria como resultado intermedio, y un contador adicional obtiene el grado de daño al contar los píxeles.

Breve descripción de las figuras

La figura 1 ilustra una descripción general de un enfoque implementado por ordenador para enseñar, en una fase de entrenamiento, redes neuronales convolucionales (CNN) para cuantificar el daño biótico en hojas de plantas de cultivo;

La figura 2 ilustra una descripción general de un enfoque implementado por ordenador para cuantificar el daño biótico de las hojas de plantas en una fase de prueba;

La figura 3 ilustra una descripción general de los métodos implementados por ordenador;

La figura 4 ilustran una imagen de hoja, una imagen de hoja anotada en la hoja y una imagen de hoja individual;

La figura 5 ilustra el daño en una hoja a manera de ejemplo, con la hoja que es una hoja principal u hoja en primer plano de la planta;

La figura 6 ilustra una imagen de hoja individual con una región de daño, e ilustra una imagen de hoja anotada por daños (primera alternativa);

La figura 7 ilustra las CNN con capas, en una descripción general;

La figura 8 ilustra la transformación de color (como un proceso auxiliar para la segunda alternativa) con una imagen que tiene píxeles en dos colores y con la asignación de píxeles codificados por colores a valores binarios;

La figura 9 ilustra que el ordenador cambia la codificación por color de una imagen de hoja individual y llega a una imagen binaria;

La figura 10 ilustra varias fotografías en escala de grises con imágenes reales; y

La figura 11 ilustra un ejemplo de un dispositivo informático genérico y un dispositivo informático móvil genérico.

Descripción detallada

Estructura

La descripción inicia al explicar las convenciones de escritura e introduce secuencias de pasos desde una perspectiva de alto nivel al diferenciar una fase de entrenamiento y una fase de prueba (figuras 1-3). La descripción entonces analiza las plantas, hojas e imágenes con más detalle (figuras 4-6), analiza los detalles de las CNN (figura 7), e introducir el procesamiento de imágenes por transformación de color que es un proceso auxiliar para entrenar la segunda CNN en la segunda alternativa (figuras 8-10).

El término "imagen" representa la estructura de datos de una fotografía digital (es decir, una estructura de datos que utiliza un formato de archivo tal como JPEG, TIFF, BMP, RAW o similares). Frases como "la cámara toma una imagen" representan una cámara con un ángulo de visión para capturar un objeto (tal como una planta) y dejar que la cámara almacene la imagen.

La descripción utiliza el término "mostrar" cuando explica el contenido de las imágenes (es decir, la semántica), por ejemplo, en frases tal como "la imagen muestra una planta". Sin embargo, no es necesario que un usuario humano mire la imagen. La descripción expresa las interacciones del usuario de ordenador con el usuario que mira la imagen por el término "pantalla", tal como en "El ordenador muestra un valor numérico" o similares.

El término "anotación" se refiere a los metadatos que recibe un ordenador cuando un usuario experto mira la pantalla de una imagen e interactúa con el ordenador. El término "imagen anotada" indica la disponibilidad de estos metadatos para una imagen (o para una subregión de esa imagen), pero no hay necesidad de almacenar los metadatos y la imagen en la misma estructura de datos. Las anotaciones son metadatos y no hay necesidad de incrustarlas en la estructura de datos de la imagen.

El término "área" identifica la cantidad de una superficie de una planta (en términos de metros cuadrados, o píxeles cuadrados). La descripción ocasionalmente escribe colores entre comillas, tal como para hojas "verdes", lo que indica que la naturaleza ofrece muchos tonos de color diferentes.

Por conveniencia, la descripción se refiere a componentes de hardware (tal como ordenadores, cámaras, dispositivos móviles, redes de comunicación) en términos singulares. Sin embargo, las implementaciones pueden usar múltiples componentes. Por ejemplo, "la cámara que toma una pluralidad de imágenes" comprende escenarios en los que participan múltiples cámaras tal que algunas imágenes se toman de una primera cámara, algunas imágenes se toman de una segunda cámara y así sucesivamente.

Descripción general de las fases de entrenamiento y prueba

Las figuras 1-3 ilustran descripciones generales de enfoques implementados por ordenador, para entrenar redes neuronales convolucionales (CNN) para cuantificar el daño biótico en hojas de plantas de cultivo en una fase de entrenamiento **1, y para usar las CNN entrenadas para cuantificar el daño durante una fase de prueba (posterior) **2.

De principio a fin de la mayor parte de esta descripción, las referencias señaladas como * 1/* *2 representan elementos que son similares pero que tienen un uso diferente en ambas fases.

De izquierda a derecha, las figuras 1-2 ilustran plantas 111/112 (con hojas y con regiones de daño 181/182), cámaras 311/312 para tomar imágenes de hojas 411/412 y ordenadores 201/202 con CNN. Las figuras también ilustran a los usuarios humanos 191/192. Las figuras 1-2 ilustran los ordenadores 201/202 por rectángulos con marcos en negrita.

Las figuras 1-2 también ilustran el ordenador 301 y el dispositivo móvil 302, realizando actividades auxiliares (o participando en ellas), tal como tomar imágenes, transmitir imágenes, recibir anotaciones, reenviar resultados a otros ordenadores, mostrar grados de daño a los usuarios, etc.

Como en la figura 3, los ordenadores 201/202 implementan las secuencias de pasos 601B, 602B, 701B y 702B. La secuencia 701B (entrenamiento para obtener el grado de daño) tiene dos alternativas

- (1) entrenar a CNN 271 para realizar la regresión, y
- (2) entrenar a CNN 271 para realizar la segmentación de daños.

Algunas de las actividades auxiliares son actividades de preprocesamiento que preparan ejecuciones de métodos. La figura 3 ilustra las actividades de preprocesamiento por las referencias 601A, 602A y 701A.

Los ordenadores 201/202 utilizan CNN y otros módulos (tal como interfaces de usuario). En tanto que las figuras 1-2 sólo introduzca las CNN, las otras figuras proporcionan detalles para el preprocesamiento de imágenes y para establecer parámetros para las CNN. Las CNN 261 y 271 se están entrenando en la fase de entrenamiento **1 para convertirse en las CNN 262 y 272 entrenadas, respectivamente. En otras palabras, la diferencia entre las CNN no entrenadas y entrenadas es la disponibilidad de los parámetros obtenidos a través del entrenamiento.

La figura 3 ilustra un diagrama de flujo de actividad en notación de matriz. La figura continúa usando líneas en negrita para ilustrar las secuencias de pasos que se realizan por las CNN.

La descripción explica las actividades por separado para la fase de entrenamiento **1 (en la fila anterior) y para la fase de prueba **2 (en la fila a continuación). La figura 3 diferencia las actividades de preprocesamiento 601A, 602A y 701A (tal como tomar / recibir imágenes y anotar imágenes) y el procesamiento intermedio 401 de las secuencias de pasos implementadas por ordenador 601B, 602B, 701B y 702B con técnicas de aprendizaje automático.

Las secuencias 601B y 602B se realizan con las CNN 261/262, y las secuencias 701B y 702B se realizan con las CNN 271/272. Las CNN 261/262 y las CNN 271/272 se diferencian entre sí por los parámetros y por la función.

Fase de entrenamiento con alternativa (1)

Al igual que en el cuadro 601A, la fase de entrenamiento **1 alternativa (1) inicia por el ordenador 301 que recibe imágenes de hojas 411 (de la cámara 311). En interacción con el usuario experto 191, el ordenador 301 recibe además anotaciones que indican los bordes de las hojas en las imágenes (denominadas anotaciones de hoja). el ordenador 301 recibe imágenes de hojas 411 y recibe anotaciones 451 (ver figura 4) para una pluralidad de imágenes, para obtener una pluralidad de imágenes de hojas anotadas en la hoja 461. Esta primera pluralidad es el conjunto de entrenamiento de segmentación.

Como en el cuadro 601B, la alternativa de la fase de entrenamiento **1 (1) continúa por el ordenador 201 que realiza el método implementado por ordenador 601B para usar este conjunto de entrenamiento (imágenes de hojas anotadas en la hoja 461) para entrenar a CNN 261 para separar las imágenes de hojas en imágenes de hojas individual. De esta manera, CNN 261 se convierte en CNN 262. En otras palabras, CNN 261 es "entrada" y CNN 262 es "salida" de este método.

Como en el cuadro 701A, la alternativa de la fase de entrenamiento **1 (1) continúa por el ordenador 301 que recibe imágenes de hojas 411 (de la cámara 311, opcionalmente las mismas imágenes que en 601A). En interacción con el usuario experto 191, el ordenador 301 recibe anotaciones que indican valores de daño en las hojas, por ejemplo, por porcentajes de daño. Puede ser ventajoso permitir que el ordenador 301 reciba imágenes porque la hoja ya está separada del resto de la imagen, tal como por la recepción de imágenes de hojas anotadas en la hoja 461 o la recepción de imágenes de hojas (salida de CNN 262). Recibir las anotaciones conduce a una pluralidad de imágenes de hojas anotadas por daños 471. Esta segunda pluralidad es el conjunto de entrenamiento de estimación de daños.

Como en el cuadro 701B, la alternativa de la fase de entrenamiento **1 (1) continúa por el ordenador 201 que realiza el método implementado por ordenador 701B para usar imágenes de hojas anotadas por daño 471 para entrenar a CNN 271 para estimar los valores de daño. De esta manera, CNN 271 se convierte en CNN 272. En otras palabras, CNN 271 es "entrada" y CNN 272 es "salida" de este método.

Fase de prueba con alternativa (1)

Como la fase de prueba **2 conduce a un grado de daño para una hoja particular, la descripción utiliza palabras en singular para imágenes particulares, valores de daño particulares, etc.

Como en el cuadro 602A, la fase de prueba **2 inicia con el ordenador 202 recibiendo la imagen de hoja (particular) 412.

Como en el cuadro 602B, el ordenador 202 utiliza CNN 262 para procesar la imagen de hoja 412 para derivar una imagen de hoja individual (particular) 422 (es decir, una imagen de hoja segmentada). La imagen 422 es un conjunto contiguo de píxeles que muestra la hoja principal 122-1 de la planta (particular) 112 completamente (ver figura 2). En otras palabras, la imagen 422 muestra toda la hoja principal.

Como en el cuadro 702B, el ordenador 202 utiliza CNN 272 para procesar la imagen de hoja individual 422 por regresión para obtener un grado de daño (particular). El grado de daño corresponde a los valores estimados que los expertos habían anotado anteriormente en la fase de entrenamiento. Puesto que los expertos hacen sus estimaciones al comparar las superficies dañadas con las superficies sanas, el grado corresponde aproximadamente a la relación de superficie (superficie de las regiones de daño sobre la superficie de hoja).

Fase de entrenamiento con alternativa (2)

Como en el cuadro 601A, la alternativa de la fase de entrenamiento **1 (2) también inicia por el ordenador 301 que recibe imágenes de hojas 411 (de la cámara 311), pero la alternativa (2) no aplica la secuencia 601B. el ordenador realiza la segmentación para obtener imágenes de hojas individual 421. La segmentación se puede realizar manualmente (con un experto que corta la hoja en primer plano), pero también se puede realizar al utilizar la secuencia 602B. La figura ilustra esto por líneas punteadas desde el cuadro 601A al cuadro 602B. En ese caso, el entrenamiento de la CNN para realizar la separación se realizó anteriormente. La salida sería imágenes de hojas individual 421 (es decir, una pluralidad de imágenes).

Como en el cuadro 401, el ordenador 201 procesa imágenes de hojas individual 421 al convertirlas en imágenes binarias 431. Con detalles que se van a explicar con relación a la figura 8 (TRANSFORMACIÓN DE COLOR), las imágenes binarias 431 tienen píxeles que representan daño y tienen píxeles que representan las partes no dañadas de las hojas (es decir, regiones de daño versus regiones sin daño).

Como en el cuadro 701B, la fase de entrenamiento **1 alternativa (2) continúa por el ordenador 201 que realiza el método implementado por ordenador 701B. el ordenador 201 recibe pares de imágenes con imágenes de hojas individual 421 y con imágenes binarias correspondientes 431. La CNN ahora aprende cómo obtener las imágenes binarias por múltiples convoluciones, pero sin realizar la transformación de color. De nuevo, de esta manera, CNN 271 se convierte en CNN 272. En otras palabras, CNN 271 es "entrada" y CNN 272 es "salida" de este método.

Fase de prueba con alternativa (2)

Como la fase de prueba **2 conduce a un grado de daño para una hoja particular, la descripción utiliza nuevamente palabras en singular para imágenes particulares, etc.

Como en el cuadro 602A, la fase de prueba **2 inicia con el ordenador 202 que recibe la imagen de hoja (particular) 412 y que proporciona la imagen de hoja individual (particular) 422. No hay una diferencia sustancial a la alternativa (1), excepto que el ordenador también obtiene N_HOJA (es decir, el número de píxeles de hoja).

Como en el cuadro 702B, el ordenador 202 utiliza la alternativa CNN 272 (2) para procesar la imagen de hoja individual 422 por segmentación de daño para obtener una imagen binaria (particular) 432. Puesto que la imagen

binaria utiliza una codificación diferente para las regiones de daño y para las regiones sin daño en la hoja, el número de píxeles de daño N_DAÑO está disponible.

El ordenador calcula el grado de daño como la relación entre el número de píxeles (que corresponde al daño) sobre (es decir, dividido por) el número de píxeles (que corresponde a la hoja), es decir

$$\text{RELACIÓN} = \text{N\_DAÑO} / \text{N\_HOJA}.$$

Exactitud

Opcionalmente, el ordenador 202 reenvía la RELACIÓN de grado de daño (ambas alternativas (1) y (2)) al usuario de campo 192 (ver la figura 2).

El usuario de campo 192 también podría obtener el grado de daño al enviar plantas (u hojas) a un laboratorio en el que los expertos inspeccionan visualmente las hojas. Los expertos de laboratorio podrían utilizar herramientas de medición (tal como microscopios) para obtener datos con una exactitud relativamente alta. La duración del enfoque de laboratorio contradice la exactitud: el retardo de laboratorio (en términos de horas o días) impediría contramedidas eficientes. El resultado puede llegar cuando el daño se ha extendido por toda la planta y ha destruido la planta.

En contraste, el retardo al transmitir electrónicamente imágenes de hojas 412 desde el campo agrícola a el ordenador 202 (ver la red 342 en la figura 2), el procesamiento de las imágenes por ambas CNN 262/272 y al retornar el grado de daño es insignificante. Idealmente, el agricultor hace la foto y ve los resultados en un período de segundos. Durante ese tiempo, el daño a las plantas no cambiará significativamente.

El grado de daño RELACIÓN no es tan exacto como la verdad fundamental del laboratorio, pero la RELACIÓN es lo suficientemente exacta como para que el agricultor identifique (y cuantifique) las contramedidas en consecuencia.

Implicación del usuario

En la fase de entrenamiento **1, el usuario experto 191 puede tener diferentes funciones. En la primera función, el usuario 191 sabe exactamente dónde tienen sus bordes las hojas principales, y en la segunda función, el usuario 191 estima un grado de daño.

Planta y cámara en situaciones del mundo real

Un ordenador cuantificaría el daño con la mayor exactitud si las imágenes (tanto para las fases de entrenamiento como de prueba) se tomaran en condiciones repetibles con equipos de cámara de alto rendimiento. Sin embargo, no hay necesidad de visitar un estudio fotográfico para tomar fotos de alta calidad de plantas bien formadas y sanas seleccionadas.

En contraste, el usuario 192 en su función de operador de cámara toma las fotos (es decir, las imágenes) en ubicaciones que son habituales para las plantas, tal como en el campo, en el invernadero o similares, justo en el lugar de trabajo.

Este entorno de no estudio crea varias restricciones, tal como las siguientes:

- La iluminación (de la planta 112 y/o el fondo) puede no estar optimizada para tener imágenes que muestren detalles de las hojas. Las imágenes de hojas 412 pueden padecer de sobreexposición o subexposición. Conforme el sol avanza a lo largo de su camino, la iluminación cambia durante el día.
- La distancia de la hoja a la cámara puede ser impredecible. El usuario retiene la cámara en su mano, pero diferentes usuarios pueden tomar imágenes a diferentes distancias.
- La resolución no es necesariamente una resolución de 24 M o similar de las cámaras de estudio, pero puede ser más pequeña. Por ejemplo, muchas cámaras tienen opciones para ajustar la resolución de imagen a valores más bajos (por ejemplo, con la opción L para resolución completa, las opciones M y S para resoluciones reducidas en píxeles)
- Aunque algunas cámaras utilizan enfoque automático, no hay garantía de que todas las hojas estén enfocadas. Es necesario esperar imágenes desenfocadas.
- Cámaras (especialmente esas de dispositivos móviles, teléfonos inteligentes, ver, dispositivo 302 en la figura 2) puede ajustar la configuración de cámara y el procesamiento de imágenes para optimizar la imagen de las personas, pero la cámara no está bien ajustada para tomar imágenes de hojas.
- Las cámaras usualmente no se montan en un trípode o similar. Se espera agitación.
- Las hojas pueden oscilar en el aire.
- Las hojas pueden contener otros elementos que son diferentes de las regiones de daño. Estos elementos son, por ejemplo, insectos, gotas de agua, partículas de suelo, o agujeros.

Segmentación de hojas

La figura 4 ilustran la imagen de hoja 411/412, la imagen de hoja anotada en la hoja 461 y la imagen de hoja 421/422. La notación de **1/**2 indica la aplicabilidad a las fases. Las coordenadas de píxeles son (I, J).

En ambas fases **1 y **2, los operadores de cámara (por ejemplo, los usuarios 191/192) dirigen la cámara 311/312 a la planta 111/112 tal que las imágenes de hojas 411/412 son imágenes con la mayor parte de los píxeles que muestran al menos una hoja completamente (es decir, tal como la imagen de hoja muestra al menos una hoja entera con todos los bordes). Es ventajoso si la hoja principal está representada por el 50% o más de los píxeles de la imagen de hoja.

Puesto que las hojas de las plantas se traslapan entre sí, la imagen de hoja 411/412 muestra, por lo tanto, la hoja principal 421-1/421-2 en primer plano (es decir, hoja/hojas en primer plano) y muestra otra hoja 421-2/422-2 (u hoja/hojas de fondo) en el fondo. En el ejemplo, la imagen 411/412 muestra la hoja principal 421-1/422-1 como la hoja completa (es decir, toda la hoja), y muestra la hoja adyacente 421-2/422-2 sólo parcialmente (debido al traslape).

Para simplificar, la figura ilustra imágenes de hojas 411/412 para una planta ideal sin daños, pero se podrían presentar daños en todas las hojas. Puesto que el valor de daño es un valor por hoja, la hoja principal se separa.

En la fase de entrenamiento **1, las imágenes de hojas 411 son imágenes de entrada a el ordenador 301, y en la fase de prueba **2, las imágenes de hojas 412 son imágenes de entrada a la CNN 262 del ordenador 202.

La imagen de hoja anotada en la hoja 461 se ha anotado por el usuario experto 191 (alternativa (1) solamente). La anotación en la hoja identifica el límite de hoja (o borde de hoja) de la hoja principal 421-1 en diferencia con la hoja adyacente 421-2 y en diferencia al fondo (también con el suelo, si es visible en la imagen). En implementaciones, el usuario 191 puede dibujar el polígono 451 (línea discontinua) alrededor de esa parte de la imagen de hoja 411 que muestra la hoja completa (es decir, la hoja principal). ver figura 5. Es conveniente mostrar el polígono 451 al usuario experto 191, pero esto no se requiere. el ordenador 301 puede cerrar el polígono 451 automáticamente.

En lugar de polígonos, el experto en la técnica puede usar otras interfaces de usuario, por ejemplo, herramientas de procesamiento de imágenes para manipular imágenes, por ejemplo, al "borrar" los píxeles que circundan la hoja principal o de otra manera.

La anotación de hoja permite que el ordenador 201 (ver figura 1) para cada píxel de la imagen de hoja 411 para diferenciar si el píxel pertenece a la hoja principal o no. Esta diferenciación es pertinente para realizar el método 601B (ver figura 3, segmentación de hoja).

La anotación en la hoja permite que la CNN se entrene para diferenciar las regiones de imagen que muestran dos tipos de "límites": entre hoja y hoja, y entre la hoja y el fondo (tal como el suelo). Una vez entrenada, la CNN no puede más que la hoja a lo largo de estos límites. En otras palabras, los límites (o márgenes) representan una línea de corte.

Para la anotación de hoja, no importa si la hoja muestra daño o no.

La imagen de hoja 421 es la salida de CNN 271 que realiza la segmentación y muestra sólo la hoja principal. Los píxeles fuera de la hoja se van a ignorar para su procesamiento (por CNN 272).

Daños (en las hojas)

Habiendo descrito la segmentación de la hoja principal de la planta, tanto para las plantas en la vida real, como para las imágenes, la descripción ahora analiza el daño que se presenta en las hojas.

La figura 5 ilustra el daño en la hoja 121/122 a manera de ejemplo. La hoja 121/122 es la hoja de primer plano o principal (ver figura 4). Por simplicidad, esta hoja se ilustra por un rectángulo de forma redonda. Las líneas negras simbolizan pliegues y nervios (de la hoja). La figura tiene las versiones (A), (B) y (C).

El daño es visible en la superficie de la hoja, como variaciones en forma de punto en el color de superficie de la hoja. Muy simplificado, una hoja sin daño tiene una superficie que es sustancialmente "verde" en todas partes, una hoja con daño tiene una superficie que todavía es "verde", pero que muestra manchas (o puntos) "no verdes".

En el ejemplo simplificado de la figura 5 versión (A), la hoja 121/122 es una hoja dañada con dos puntos de daño. Para simbolizar las variaciones de tamaño, ambos puntos se ilustran por diferentes tamaños. En realidad, la mayoría de los puntos no son necesariamente negros (como en la figura), sino más bien de un color "no verde". La región de daño 181/182 puede ser un grupo de puntos (es una región en la hoja que tiene una concentración

relativamente alta de puntos de daño)

Las razones del daño incluyen daños en la alimentación (los insectos se han comido parte de la hoja, potencialmente los insectos aún se están quedando en la hoja), daños por enfermedades, y otros. Diferentes razones conducen a diferentes tipos de daño.

Las hojas y las regiones de daño tienen áreas (medidas, por ejemplo, en milímetros cuadrados o centímetros cuadrados). Para cuantificar los daños, las áreas con daños pueden estar con relación al área de superficie general de la hoja:

$$\text{RELACIÓN} = \text{ÁREA\_DAÑO} / \text{ÁREA\_HOJA}$$

La figura 5 (C) ilustra la hoja de nuevo, pero con la sombra 183. La sombra NO es un daño. Estas sombras (u oclusiones, etc.) se presentan de acuerdo con las direcciones relativas entre la luz y (el eje óptico de) la cámara.

La descripción explicará ahora cómo el ordenador 202 procesa las imágenes para identificar los píxeles que muestran daño. Calcular la RELACIÓN al relacionar números de píxeles conduce al grado de daño.

Anotación del daño en la alternativa (1)

La figura 6 ilustra la imagen de hoja 421 con la región de daño 481 en el lado izquierdo, e ilustra la imagen de hoja anotada por daños 471 en el lado derecho. La figura 6 corresponde al cuadro 601A en la figura 3. La imagen de hoja individual 421 se puede obtener al operar CNN 262, o puede ser la imagen de hoja 411 que se está anotando (ver figura 4). La anotación de imágenes para obtener el conjunto de entrenamiento de estimación de daños comprende la asignación de daños estimados visualmente.

En la fase de entrenamiento **1, el usuario experto 191 inspecciona la imagen de hoja 421 y estima los porcentajes de área. La región de daño 481 se extiende sobre ÁREA_DAÑO y la hoja se extiende sobre ÁREA_HOJA. La relación ÁREA_DAÑO / ÁREA_HOJA es el valor de anotación (es decir, los metadatos en la imagen 471).

En el ejemplo, aproximadamente el 20% de la superficie de hoja muestra daño. No hay necesidad de contar píxeles.

Como ya se menciona, se obtienen anotaciones para pluralidades de hojas (conjunto de entrenamiento de estimación de daños).

Variaciones del enfoque son posibles. Por ejemplo, el usuario experto 191 no tiene que mirar una imagen. Él o ella puede mirar la hoja directamente, y puede ingresar la estimación en el ordenador 301 directamente. (Esto es diferente de anotar los bordes de las hojas como en la figura 5, el valor de daño puede ser sólo un valor numérico individual).

Además, el usuario experto 191 podría anotar el área de daño por polígonos o similares (similar al borde de hoja) y el ordenador podría calcular la relación de daño al relacionar las áreas obtenidas por las anotaciones (ver la anotación 451 para la hoja).

Indicadores de cantidad para el daño en las hojas (o plantas)

El daño se cuantifica por grados de daño (o indicadores de cantidad de daño). El píxel en la imagen de hoja 422-2 (ver figura 4) se puede diferenciar en

- píxeles que corresponden al daño
- píxeles que corresponden a la hoja (con o sin daño)
- píxeles que se ignoran

El ordenador con CNN 272 (alternativa (2)) puede contar los píxeles (N) por separado para los píxeles de daño y para los píxeles de hoja. La obtención del grado de daño puede seguir la RELACIÓN = $N_{\text{DAÑO}} / N_{\text{HOJA}}$ (u otras fórmulas modificadas).

En otras palabras, el daño corresponde a la parte de un área dañada sobre la superficie general de la hoja. Las relaciones se proporcionan convenientemente como porcentajes.

Retornando brevemente a la figura 2, la RELACIÓN se puede retornar al operador de cámara 192 (en la función de ser el agricultor). Para el agricultor puede marcar la diferencia entender, por ejemplo, que una planta en particular muestra un daño del 50 % o 20 %.

El experto en la técnica puede calcular la RELACIÓN por diferentes fórmulas (por ejemplo, $N_{\text{DAÑO}} /$

$N\_SIN\_DAÑO$, con $N\_DAÑO + N\_SIN\_DAÑO = N\_HOJA$)

Además, los puntos de daño son contables. En el ejemplo simplificado hay sólo $N = 2$ puntos (figura 5, en la región 181), pero en realidad puede haber cientos de puntos en una región individual (o en una hoja individual). Las regiones de daño también son contables. En el ejemplo, hay una región individual solamente. Pero las regiones también son medibles por sus áreas de superficie. Es conveniente medir la región de daño con relación a la superficie general de la hoja. En el ejemplo de la figura 6, la región 481 ocupa aproximadamente una cuarta parte del área de superficie. El daño puede no aparecer como puntos contables, sino más bien como regiones relativamente grandes (con diferentes colores).

Lado frontal/posterior

En tanto que las figuras ilustran la hoja con los lados frontales, se observa que también se puede presentar el daño en los lados posteriores de las hojas. el ordenador que procesa las imágenes no tiene necesariamente datos a priori que diferencien el lado frontal del lado posterior.

En el lado frontal o anverso, el contraste entre el "verde" de la hoja y el "otro color" del daño usualmente es mayor que en el lado posterior o reverso.

Durante el entrenamiento **2, la secuencia 601A (recibir imágenes de hojas 411) se realiza con imágenes que muestran el lado frontal y con imágenes que muestran el lado posterior.

Capas de CNN

La figura 7 ilustra las CNN 261/262/271/272 con capas, en una descripción general. Las CNN se implementan por colecciones de rutinas de programa que se ejecutan por un ordenador, tal como por el ordenador 201/202. La figura 7 ilustra las CNN con la entrada a una capa de entrada y con la salida de una capa de salida. La figura 7 también ilustra (al menos simbólicamente) capas intermedias. Las CNN 261/262/271/272 son redes neuronales profundas debido a que tienen múltiples capas intermedias. Las capas intermedias están ocultas. En otras palabras, el aprendizaje profundo se aplica aquí.

La figura 7 también ilustra algunos parámetros. Puesto que las CNN son bien conocidas en la técnica, la descripción se enfoca en los parámetros que se aplican especialmente para la segmentación por las CNN 261/262 y para estimar el daño por las CNN 271/272.

En la fase de entrenamiento **1, las CNN 261/271 reciben imágenes anotadas 461,471 y convierten la CNN 261 no entrenada en CNN 262 entrenada (usando las imágenes de hojas anotadas en hojas) y convierten la CNN 271 no entrenada en CNN 272 entrenada (usando imágenes de hojas anotadas por daños).

En la fase de prueba, las CNN 262 y 272 reciben la imagen de hoja 412 y proporcionan la salida como el indicador de cantidad de daño. Por ejemplo, el indicador puede ser la RELACIÓN (alternativa (1)). O, como en la alternativa (2), la salida es imágenes binarias 432 con $N\_DAÑO$, y se puede calcular la RELACIÓN (debido a que se conoce N_HOJA).

Tipos de red

Las redes están disponibles públicamente en una variedad de implementaciones, y las redes se configuran por parámetros de configuración.

La descripción se refiere brevemente a los parámetros de entrada / salida en general, así como a los parámetros de configuración y entonces especifica los parámetros. Ocasionalmente, se modifica una red existente.

Las redes de ejemplo comprenden los siguientes tipos de red (o "arquitecturas"):

- El tipo UNet se divulga por Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Redes convolucionales para segmentación de imágenes biomédicas, en: Lecture Notes in Computer Science (Including Sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). págs. 234-241. doi:10.1007/978-3-319-24574-4_28
- El tipo DenseNet se divulga por Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Redes convolucionales densamente conectadas, en: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. págs. 2261-2269. doi:10.1109/CVPR.2017.243

Las CNN tienen las siguientes propiedades:

- Puede ser un modelo totalmente convolucional (como para DenseNet), que es capaz de proporcionar segmentación semántica (es decir, segmentación de hojas, segmentación de daños).

- La diferencia es la función de pérdida. La regresión se conoce en la técnica.
- La CNN para segmentación de hojas se ha probado para implementaciones para los tipos UNet y DenseNet.

En las implementaciones, CNN 261/262 es una CNN del tipo DenseNet. Con este propósito, los siguientes parámetros de implementación son convenientes:

- La función de pérdida puede ser una función "binary_crossentropy".
- La activación de la última capa puede utilizar una función "softmax".
- Las dimensiones de la imagen de entrada y salida pueden ser de 224 x 224 píxeles.

Para la alternativa (1) - regresión - CNN 271/272 puede ser con base en ResNet18 (He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition. doi: 10.1109/CVPR.2016.90, publicado para la Conferencia IEEE 2016 sobre Visión por Ordenador y Reconocimiento de Patrones (CVPR), 27-30 de junio de 2016). En otras palabras, la CNN tiene una estructura de capas como se divulga en el documento, pero se ha entrenado como se describe en la presente. Desde 2015, otras versiones de ResNet están disponibles en los países de seguimiento, y el experto en la técnica también puede usar versiones más nuevas.

Una red neuronal residual (ResNet) es una red neuronal artificial (ANN) que es con base en construcciones conocidas de células piramidales en la corteza cerebral. Las redes neuronales residuales hacen esto al utilizar conexiones de salto, o accesos directos para saltar sobre algunas capas. Los modelos habituales de Res-Net se implementan con saltos de doble o triple capa que contienen no linealidades (ReLU) y normalización de lotes entremedio.

Para la alternativa (2) - segmentación de daños - CNN 271/272 es con base en el tipo DenseNet.

Procesamiento de imágenes en la segunda alternativa

La descripción ahora analiza la transformación de color (ver 401) que es un proceso auxiliar.

Como la cámara proporciona imágenes como imágenes en color (usualmente en el espacio de color RGB), las imágenes mantienen el color a través del procesamiento. La descripción ahora explica cómo el procesamiento de color (transformación del espacio de color y filtración) puede adaptar las imágenes para que su contenido sea más adecuado para las CNN.

El procesamiento de imágenes implementado por ordenador puede tomar ventaja de las siguientes observaciones (con respecto a la planta):

- El daño en las hojas cambia el color de la hoja (ver figura 5), pero este cambio de color no es uniforme para la planta en su conjunto.
- Las imágenes de hojas pueden mostrar hojas en el fondo, con colores diferentes (por ejemplo, hoja "verde" sobre fondo "no verde").
- Las imágenes de hojas muestran píxeles de hojas "verdes" y píxeles dañados "no verdes" en diferentes colores.

En estas situaciones, las diferencias pueden ser diferencias binarias (VERDADERO/FALSO, VERDE/NO VERDE, etc.). Potenciar el contraste de color en las imágenes haría que estas diferencias fueran más visibles para los espectadores humanos. Sin embargo, las imágenes no se muestran necesariamente a usuarios humanos (excepto algunas interacciones de anotación, ver 451 en la figura 4).

La figura 8 ilustra la transformación de color 401 (ver figura 3) con una imagen que tiene píxeles en dos colores y con la asignación de píxeles codificados por colores a valores binarios

En la codificación de RGB (es decir, la codificación original de las imágenes), cada píxel se codifica por un triplete de números (por ejemplo, números reales en el intervalo cerrado $[0; 1]$, o números enteros sobre un intervalo numérico diferente). Por ejemplo, un píxel en verde ideal sería (R, G, B) codificado como (0, 1, 0).

El ejemplo de la figura 8 ilustra la región sin daño 477 (o región sana) en, por ejemplo, "verde" (0,1, 0,8, 0,1) e ilustra la región de daño 478 en, por ejemplo, "blanco" (0,3, 0,3, 0,3). La codificación de RGB es un ejemplo para la primera codificación por color 400-1.

El color biológico "verde" no es ideal como (0, 1, 0) y no es ideal como en el ejemplo. En las imágenes de plantas u hojas, los píxeles "verdes" se codifican por una mezcla de diferentes componentes que incluso varían de píxel a píxel. Los artefactos (ver las sombras) adicionarían una variación adicional en la codificación.

Para diferenciar los colores, el ordenador tendría que aplicar operaciones multidimensionales. El espacio de color RGB es un espacio 3D, y cualquier umbral que separe el color en dos grupos (binarios) sería un plano 2D en ese espacio. La implementación requeriría múltiples cálculos computacionales.

La transformación de color proporciona una solución a esta restricción. La imagen se transforma en una segunda codificación por color 400-2 en el llamado espacio de color $L^*a^*b^*$. Este espacio también se conoce como CIELAB. Para implementar esta transformación del espacio de color, el experto en la técnica puede usar bibliotecas con instrucciones de ordenador para la transformación de imágenes. Estas bibliotecas están disponibles, por ejemplo, en Matlab (mathworks.com), skimage (scikit-image.org), openCV (opencv.org), etc. En las realizaciones, la transformación se implementó por skimage para python.

La codificación por color es bien conocida en la técnica, y el espacio $L^*a^*b^*$ está estandarizado por la CIE (Commission Internationale de l'Eclairage, Comisión Internacional de Iluminación, ISO/CIE 11664-4:2019 Colorimetría). En resumen, los componentes de color definen BLANCO, NEGRO, AMARILLO, AZUL, ROJO y VERDE. El componente a^* indica VERDE (si es negativo) e indica ROJO (si es cero o positivo).

Puesto que, para muchas plantas, "verde" es el color dominante de la planta, el uso de "verde" es conveniente para las ilustraciones. Como nota lateral, la reducción del esfuerzo de cálculo puede reducir el tiempo de espera por lo que el agricultor de la figura 2 espera a que se muestre el grado de daño (ahorrando tiempo de cálculo).

En la transformación de color citada por la figura 8, la asignación de los píxeles codificados por colores a la codificación binaria 400-3 depende del componente a^* , con "verde a verdadero" (para $a^* < 0$) y "no verde a falso" (para $a^* > 0$). La asignación es arbitraria y la ilustración en negro para VERDADERO y blanco para FALSO se utiliza para la ilustración. Los nombres de los valores binarios "(no) -verde..." no importan.

El uso del componente a^* (de CIELAB) como el diferenciador de color es ventajoso como se explica a continuación. Ya no hay necesidad de una diferenciación 2D.

El experto en la técnica puede aplicar otras transformaciones de color de manera similar (tal como RGB a XYZ, estandarizado por CIE también).

En general, la transformación se selecciona tal que uno de los componentes (en la segunda codificación por color) tenga su variación máxima de los colores que se van a diferenciar. En el ejemplo de la figura 8, el color "verde" se codifica por a^* que eventualmente cambia su signo. En este caso, el cambio de signo es la variación máxima. Se observa que una CNN puede aprender de manera más eficiente cuando los signos cambian.

En otro ejemplo, un componente de código particular tendría su variación máxima por un cambio de "0" a "1".

Por una simple operación (una comparación para sustancialmente cada píxel), el ordenador puede clasificar cada píxel en valores binarios para obtener una imagen en codificación binaria 400-3. En el ejemplo, los valores binarios se llaman "verde a verde" o "no verde a falso". Pero, de nuevo, el nombre no importa.

La figura 9 ilustra que el ordenador cambia la codificación por color de la imagen de hoja individual 421 y llega a la imagen binaria 431. En el ejemplo, la imagen de hoja individual 421 tiene 5 regiones de daño. La imagen binaria 431 muestra las regiones sin daño en un primer código binario, tal como "verde a verdadero" (ilustrado en negro) y muestra las regiones de daño en un segundo código binario, tal como "no verde a falso" (ilustrado en blanco).

Durante la fase de entrenamiento **1, se realiza un cambio de color para una pluralidad de imágenes de hojas individual, tal como para 1000 imágenes.

Es posible que el cambio de color no detecte regiones de daño en algunas situaciones, pero las CNN pueden detectar regiones de daño en más situaciones. El hecho de proporcionar muchos ejemplos "débiles" (durante el entrenamiento) a las CNN los fuerza a aprender qué es el daño. Durante la fase de prueba, la CNN entonces puede ser capaz de detectar regiones dañadas con mayor exactitud.

Dos escenarios de uso de ejemplo

El enfoque (es decir, las secuencias 601B, 602B, 701B, 702B ver figura 3) es aplicable a al menos dos escenarios experimentales de diferentes combinaciones de causa de daño y planta dañada.

- En el primer escenario, la planta 112 es una berenjena, y el daño es el daño de alimentación por tisanóptaros (FRANOC). En el primer escenario, se esperan daños en ambos lados de la hoja (anverso y reverso).
- En el segundo escenario, la planta 112 es un tomate, y el daño es el daño de alimentación por tuta (GNORAB).

Las siglas en mayúscula son los llamados códigos EPOP (Organización Europea y Mediterránea de Protección Fitosanitaria). Los escenarios son sólo ejemplos, el enfoque también se podría aplicar a otros cultivos con hojas.

Imágenes reales

La figura 10 ilustra varias fotografías en escala de grises con imágenes reales de berenjenas. La figura ilustra las imágenes junto con los números de referencia.

La imagen de hoja 411/412 (es decir, antes de la segmentación) se ilustra incluso con la mano del operador de cámara. La imagen de hoja individual 421/422 es la imagen de hoja (por ejemplo, después de la segmentación). El fondo en "negro" sólo indica que los píxeles de fondo no transmiten información. Estas imágenes son aplicables tanto a la fase de entrenamiento **1 como a la fase de prueba **2. Las imágenes 411/412 muestran una hoja con pequeños puntos de bajo contraste, casi como para una hoja sana.

Las siguientes imágenes muestran un ejemplo con daños, que pertenece a la alternativa (2) en la fase de entrenamiento **1.

La imagen de hoja individual 421 muestra la región de daño 481 (ver la flecha) y muestra el resto de la hoja sin daño.

La imagen binaria 431 se ha derivado de la imagen 421 por procesamiento de color como se describe. El color del daño se ha asignado al BLANCO. En otras palabras, la imagen 421/431 es un par de imágenes.

La imagen binaria 431 se ha derivado de la imagen 421.

Durante la fase de prueba **2, la imagen de hoja individual 421 sería la entrada a la alternativa (2) de CNN 272 y la imagen binaria 432 sería la salida. La RELACIÓN se puede calcular, debido a que (i) el número de píxeles BLANCOS se puede derivar de la imagen binaria 432 y (ii) el número de píxeles en la hoja se puede derivar de la imagen 422.

Sistema de ordenador

La figura 11 ilustra un ejemplo de un dispositivo informático genérico 900 y un dispositivo informático móvil genérico 950, que se puede utilizar con las técnicas descritas en la presente. Se propone que el dispositivo informático 900 represente diversas formas de ordenadores digitales, tal como ordenadores portátiles, ordenadores de escritorio, estaciones de trabajo, asistentes digitales personales, servidores, servidores blade, ordenadores centrales, y otras ordenadores apropiadas. El dispositivo informático genérico puede 900 corresponder a los ordenadores 201/202 de las figuras 1-2. Se propone que el dispositivo informático 950 represente diversas formas de dispositivos móviles, tal como asistentes digitales personales, teléfonos celulares, teléfonos inteligentes, y otros dispositivos informáticos similares. Por ejemplo, el dispositivo informático 950 puede incluir los componentes de almacenamiento de datos y/o componentes de procesamiento de dispositivos como se muestra en la figura 1. Los componentes mostrados aquí, sus conexiones y relaciones, y sus funciones, se propone que sean sólo a manera de ejemplo, y no se propone que limiten las implementaciones de las invenciones descritas y/o reivindicadas en este documento.

El dispositivo informático 900 incluye un procesador 902, una memoria 904, un dispositivo de almacenamiento 906, una interfaz de alta velocidad 908 que se conecta a la memoria 904 y a los puertos de expansión de alta velocidad 910, y una interfaz de baja velocidad 912 que se conecta al bus de baja velocidad 914 y al dispositivo de almacenamiento 906. Cada uno de los componentes 902, 904, 906, 908, 910 y 912, se interconectan usando varios buses, y se pueden montar en una placa madre común o de otras maneras, según sea apropiado. El procesador 902 puede procesar instrucciones para la ejecución dentro del dispositivo informático 900, que incluyen instrucciones almacenadas en la memoria 904 o en el dispositivo de almacenamiento 906 para mostrar información gráfica para una GUI en un dispositivo de entrada/salida externo, tal como la pantalla 916 acoplada a la interfaz de alta velocidad 908. En otras implementaciones, se pueden utilizar múltiples procesadores y/o múltiples buses, según sea apropiado, junto con múltiples memorias y tipos de memoria. Además, se pueden conectar múltiples dispositivos informáticos 900, y cada dispositivo proporciona porciones de las operaciones necesarias (por ejemplo, como un banco de servidores, un grupo de servidores blade o un sistema de múltiples procesadores).

La memoria 904 almacena información dentro del dispositivo informático 900. En una implementación, la memoria 904 es una unidad o unidades de memoria volátil. En otra implementación, la memoria 904 es una unidad o unidades de memoria no volátil. La memoria 904 también puede ser otra forma de medio leíble por ordenador, tal como un disco magnético u óptico.

El dispositivo de almacenamiento 906 es capaz de proporcionar almacenamiento masivo para el dispositivo informático 900. En una implementación, el dispositivo de almacenamiento 906 puede ser o contener un medio leíble por ordenador, tal como un dispositivo de disquete, un dispositivo de disco duro, un dispositivo de disco óptico, o un dispositivo de cinta, una memoria flash u otro dispositivo de memoria de estado sólido similar, o una matriz de dispositivos, incluyendo dispositivos en una red de área de almacenamiento u otras configuraciones. Un producto de programa de ordenador se puede incorporar de manera tangible en un soporte de información. El producto de programa de ordenador también puede contener instrucciones que, cuando se ejecutan, realizan uno o más métodos, tal como aquellos descritos anteriormente. El portador de información es un medio leíble por

ordenador o máquina, tal como la memoria 904, el dispositivo de almacenamiento 906, o la memoria en el procesador 902.

El controlador de alta velocidad 908 gestiona operaciones intensivas en ancho de banda para el dispositivo informático 900, en tanto que el controlador de baja velocidad 912 gestiona operaciones intensivas en ancho de banda más bajas. Esta asignación de funciones es sólo a manera de ejemplo. En una implementación, el controlador de alta velocidad 908 se acopla a la memoria 904, la pantalla 916 (por ejemplo, a través de un procesador gráfico o acelerador) y a los puertos de expansión de alta velocidad 910, que pueden aceptar varias tarjetas de expansión (no mostradas). En la implementación, el controlador de baja velocidad 912 se acopla al dispositivo de almacenamiento 906 y al puerto de expansión de baja velocidad 914. El puerto de expansión de baja velocidad, que puede incluir varios puertos de comunicación (por ejemplo, USB, Bluetooth, Ethernet, Ethernet inalámbrica) se puede acoplar a uno o más dispositivos de entrada/salida, tal como un teclado, un dispositivo señalador, un escáner, o un dispositivo de red tal como un conmutador o encaminador, por ejemplo, a través de un adaptador de red.

El dispositivo informático 900 se puede implementar en una cantidad de formas diferentes, como se muestra en la figura. Por ejemplo, se puede implementar como un servidor estándar 920, o varias veces en un grupo de estos servidores. También se puede implementar como parte de un sistema de servidor en bastidor 924. Además, se puede implementar en un ordenador personal tal como un ordenador portátil 922. De manera alternativa, los componentes del dispositivo informático 900 se pueden combinar con otros componentes en un dispositivo móvil (no mostrado), tal como el dispositivo 950. Cada uno de dichos dispositivos puede contener uno o más de los dispositivos informáticos 900, 950, y un sistema completo se puede componer de múltiples dispositivos informáticos 900, 950 que se comunican entre sí.

El dispositivo informático 950 incluye un procesador 952, memoria 964, un dispositivo de entrada/salida tal como una pantalla 954, una interfaz de comunicación 966, y un transceptor 968, entre otros componentes. El dispositivo 950 también se puede proporcionar con un dispositivo de almacenamiento, tal como un microaccionador u otro dispositivo, para proporcionar almacenamiento adicional. Cada uno de los componentes 950, 952, 964, 954, 966 y 968, se interconectan usando varios buses, y varios de los componentes se pueden montar en una placa madre común o de otras maneras, según sea apropiado.

El procesador 952 puede ejecutar instrucciones dentro del dispositivo informático 950, que incluyen instrucciones almacenadas en la memoria 964. El procesador se puede implementar como un conjunto de chips que incluye procesadores analógicos y digitales separados y múltiples. El procesador puede proporcionar, por ejemplo, la coordinación de los otros componentes del dispositivo 950, tal como el control de las interfaces de usuario, las aplicaciones ejecutadas por el dispositivo 950 y la comunicación inalámbrica por el dispositivo 950.

El procesador 952 se puede comunicar con un usuario a través de la interfaz de control 958 y la interfaz de pantalla 956 acoplada a una pantalla 954. La pantalla 954 puede ser, por ejemplo, una pantalla TFT LCD (Pantalla de cristal líquido con transistores de película delgada) o una pantalla OLED (Diodo orgánico emisor de luz), u otra tecnología de pantalla apropiada. La interfaz de pantalla 956 puede comprender circuitos apropiados para accionar la pantalla 954 para presentar información gráfica y otra información a un usuario. La interfaz de control 958 puede recibir comandos de un usuario y convertirlos para su envío al procesador 952. Además, se puede proporcionar una interfaz externa 962 en comunicación con el procesador 952, a fin de permitir la comunicación de área cercana del dispositivo 950 con otros dispositivos. La interfaz externa 962 puede proporcionar, por ejemplo, comunicación cableada en algunas implementaciones, o para comunicación inalámbrica en otras implementaciones, y también se pueden utilizar múltiples interfaces.

La memoria 964 almacena información dentro del dispositivo informático 950. La memoria 964 se puede implementar como uno o más de un medio o medios leíbles por ordenador, una unidad o unidades de memoria volátil, o una unidad o unidades de memoria no volátil. La memoria de expansión 984 también se puede proporcionar y conectar al dispositivo 950 a través de la interfaz de expansión 982, que puede incluir, por ejemplo, una interfaz de tarjeta SIMM (Módulo de Memoria en Línea Único). Esta memoria de expansión 984 puede proporcionar espacio de almacenamiento adicional para el dispositivo 950, o también puede almacenar aplicaciones u otra información para el dispositivo 950. Específicamente, la memoria de expansión 984 puede incluir instrucciones para llevar a cabo o complementar los procesos descritos anteriormente, y también puede incluir información segura. Por lo tanto, por ejemplo, la memoria de expansión 984 puede actuar como un módulo de seguridad para el dispositivo 950, y se puede programar con instrucciones que permitan el uso seguro del dispositivo 950. Además, se pueden proporcionar aplicaciones seguras mediante las tarjetas SIMM, junto con información adicional, tal como colocar la información de identificación en la tarjeta SIMM de una manera no hackeable.

La memoria puede incluir, por ejemplo, memoria flash y/o memoria NVRAM, como se analiza a continuación. En una implementación, un producto de programa de ordenador se incorpora de manera tangible en un soporte de información. El producto de programa de ordenador contiene instrucciones que, cuando se ejecutan, realizan uno o más métodos, tal como aquellos descritos anteriormente. El portador de información es un medio leíble por

ordenador o máquina, tal como la memoria 964, la memoria de expansión 984, o la memoria en el procesador 952, que se puede recibir, por ejemplo, a través del transceptor 968 o la interfaz externa 962.

El dispositivo 950 puede comunicarse de manera inalámbrica a través de la interfaz de comunicación 966, que puede incluir circuitería de procesamiento de señal digital donde sea necesario. La interfaz de comunicación 966 puede proporcionar comunicaciones bajo diversos modos o protocolos, tal como llamadas de voz GSM, SMS, EMS, o mensajería MMS, CDMA, TD-MA, PDC, WCDMA, CDMA2000, o GPRS, entre otros. Esta comunicación se puede presentar, por ejemplo, a través del transceptor de radiofrecuencia 968. Además, se puede presentar comunicación de corto alcance, tal como usando un Bluetooth, Wi-Fi, u otro tal transceptor (no mostrado). Además, el módulo de receptor GPS (Sistema de Posicionamiento Global) 980 puede proporcionar datos inalámbricos adicionales relacionados con la navegación y la ubicación al dispositivo 950, que se pueden utilizar según sea apropiado por las aplicaciones que se ejecutan en el dispositivo 950.

El dispositivo 950 también se puede comunicar de manera audible usando el codificador-decodificador de audio 960, que puede recibir información hablada de un usuario y convertirla en información digital utilizable. El codificador-decodificador de audio 960 también puede generar sonido audible para un usuario, tal como a través de un altavoz, por ejemplo, en un teléfono del dispositivo 950. Este sonido puede incluir sonido de llamadas telefónicas de voz, puede incluir sonido grabado (por ejemplo, mensajes de voz, archivos de música, etc.) y también puede incluir sonido generado por aplicaciones que operan en el dispositivo 950.

El dispositivo informático 950 se puede implementar en una cantidad de formas diferentes, como se muestra en la figura. Por ejemplo, se puede implementar como un teléfono celular 980. También se puede implementar como parte de un teléfono inteligente 982, asistente digital personal, u otro dispositivo móvil similar.

Se pueden realizar varias implementaciones de los sistemas y técnicas descritos en la presente en circuitería electrónica digital, circuitería integrada, ASIC (circuitos integrados de aplicación específica) especialmente diseñados, hardware de ordenador, firmware, software, y/o combinaciones de los mismos. Estas diversas implementaciones pueden incluir la implementación en uno o más programas de ordenador que son ejecutables y/o interpretables en un sistema programable que incluye al menos un procesador programable, que puede ser de propósito especial o general, acoplado para recibir datos e instrucciones de, y para transmitir datos e instrucciones a, un sistema de almacenamiento, al menos un dispositivo de entrada, y al menos un dispositivo de salida.

Estos programas de ordenador (también conocidos como programas, software, aplicaciones de software o código) incluyen instrucciones de máquina para un procesador programable, y se pueden implementar en un lenguaje de programación de procedimiento y/u orientado a objetos de alto nivel, y/o en lenguaje ensamblador/máquina. Como se utiliza en la presente, los términos "medio leíble por máquina" y "medio leíble por ordenador" se refieren a cualquier producto, aparato y/o dispositivo de programa de ordenador (por ejemplo, discos magnéticos, discos ópticos, memoria, Dispositivos Lógicos Programables (PLD)) utilizados para proporcionar instrucciones de máquina y/o datos a un procesador programable, que incluye un medio leíble por máquina que recibe instrucciones de máquina como una señal leíble por máquina. El término "señal leíble por máquina" se refiere a cualquier señal utilizada para proporcionar instrucciones y/o datos de la máquina a un procesador programable.

Para proporcionar interacción con un usuario, los sistemas y técnicas descritos aquí se pueden implementar en un ordenador que tiene un dispositivo de visualización (por ejemplo, un monitor de CRT (tubo de rayos catódicos) o LCD (pantalla de cristal líquido)) para mostrar información al usuario y un teclado y un dispositivo señalador (por ejemplo, un ratón o una bola de seguimiento) por lo cual el usuario puede proporcionar entrada a el ordenador. Se pueden usar otros tipos de dispositivos para proporcionar interacción con un usuario también; por ejemplo, la retroalimentación proporcionada al usuario puede ser cualquier forma de retroalimentación sensorial (por ejemplo, retroalimentación visual, retroalimentación auditiva, o retroalimentación táctil); y la entrada del usuario se puede recibir en cualquier forma, incluyendo la entrada acústica, de habla, o táctil.

Los sistemas y técnicas descritos en la presente se pueden implementar en un dispositivo informático que incluye un componente de motor (por ejemplo, como un servidor de datos), o que incluye un componente de middleware (por ejemplo, un servidor de aplicaciones), o que incluye un componente de interfaz (por ejemplo, un ordenador de cliente que tiene una interfaz gráfica de usuario o un navegador web a través del cual un usuario puede interactuar con una implementación de los sistemas y técnicas descritos en la presente), o cualquier combinación de estos componentes de motor, middleware o interfaz. Los componentes del sistema se pueden interconectar por cualquier forma o medio de comunicación de datos digitales (por ejemplo, una red de comunicación). Los ejemplos de redes de comunicación incluyen una red de área local ("LAN"), una red de área amplia ("WAN") e Internet.

El dispositivo informático puede incluir clientes y servidores. Un cliente y servidor están en general remotos entre sí y habitualmente interactúan a través de una red de comunicación. La relación de cliente y servidor surge en virtud de programas de ordenador que se ejecutan en los respectivos ordenadores y que tienen una relación de cliente-servidor entre sí.

Referencias

	111, 112 planta
	221, 122 hoja
	181, 182 daño
5	183 sombra
	191, 192 usuario
	201, 202 ordenador
	301 ordenador
	311, 312 cámara
10	342 red
	392 interfaz de usuario
	401 transformación de color
	400-1 primer codificación por color
	400-2 segunda codificación por color
15	400-3 codificación binaria
	411/412 imagen de hoja
	421-1, 422-1 hoja principal
	421-2, 422-2 hoja adyacente
	422 imagen de hoja
20	451 anotación por polígono
	461 imagen de hoja anotada en la hoja
	471 imagen de hoja anotada por daños
	477, 478 regiones sin daños/daño
	481, 482 región de daño
25	261, 262 CNN (segmentación)
	271, 272 CNN (estimación de daños)
	601A, 602AB secuencias de pasos
	701AB, 702B secuencias de pasos
	9xx sistema de ordenador, con componentes

REIVINDICACIONES

1. Un método implementado por ordenador (602B, 702B) para cuantificar el daño biótico en hojas (122) de plantas de cultivo (112), el método (602B, 702B) que comprende:

recibir (701A) una imagen de hoja (412) tomada de una planta de cultivo particular (122), la imagen de hoja que muestra al menos una de las hojas (122) de la planta de cultivo particular (112);
usar una primera red neuronal convolucional (262) para procesar la imagen de hoja (412) para derivar una imagen de hoja segmentada (422), la imagen de hoja segmentada en lo sucesivo conocida como la imagen de hoja individual, que es un conjunto contiguo de píxeles que muestra una hoja entera en primer plano (422-1) de la planta particular (112), la primera red neuronal convolucional (262) se ha entrenado por una pluralidad de imágenes de hojas anotada en la hoja (461), en donde las imágenes de hojas (411) en un conjunto de entrenamiento se anotan para identificar hojas en primer plano (451), con la anotación de hoja que identifica el límite de hoja de hojas en primer plano (421-1) en diferencia a hojas adyacentes (421-2) y en diferencia al fondo;
y caracterizado por
usar una segunda red neuronal convolucional (272) para procesar la imagen de hoja individual (422) para obtener un grado de daño, que es la relación del área de superficie de las regiones de daño sobre el área de superficie de la hoja en primer plano (422-1).

2. El método de acuerdo con la reivindicación 1, donde la imagen de hoja (412) es una imagen en la que la mayor parte de los píxeles muestra al menos una hoja entera de la planta.

3. El método de acuerdo con cualquiera de las reivindicaciones 1-2, donde el uso de una segunda red neuronal convolucional (272) para procesar la imagen de hoja individual (422) para obtener un grado de daño, se realiza por regresión.

4. El método de acuerdo con la reivindicación 3, donde la segunda red neuronal convolucional (272) es del tipo ResNet.

5. Método de acuerdo con la reivindicación 4, donde la segunda red neuronal convolucional (272) se ha entrenado por una pluralidad de imágenes de hojas individual anotada por daños (421), con las anotaciones que son valores numéricos que corresponden al daño.

6. El método de acuerdo con cualquiera de las reivindicaciones 1-2, donde el uso de una segunda red neuronal convolucional (272) para procesar la imagen de hoja individual (422) para obtener un grado de daño, se realiza al procesar imágenes de hojas individual (421) y al convertirlas en imágenes binarias (431).

7. El método de acuerdo con la reivindicación 6, donde la segunda red neuronal convolucional (272) es del tipo DENSENET.

8. El método de acuerdo con la reivindicación 7, donde el grado de daño se calcula como la relación entre el número de píxeles que corresponden al daño en la imagen binaria, sobre el número de píxeles que corresponden a la hoja en la imagen de hoja individual.

9. El método de acuerdo con la reivindicación 8, donde la segunda red neuronal convolucional (272) se ha entrenado por imágenes obtenidas al convertir (401) una pluralidad de imágenes de hojas individual (421) de una primera codificación por color (400-1) a una segunda codificación por color (400-2), en donde el segundo código de color (400-1) tiene un componente de color individual que diferencia el color de las regiones sanas (477) sobre el color de las regiones de daño (478), y al asignar valores binarios a los píxeles de la imagen binaria de acuerdo con la magnitud del componente de color individual.

10. El método de acuerdo con la reivindicación 9, donde la primera codificación por color (400-1) es RGB, la segunda codificación por color (400-2) se codifica en el espacio de color $L^*a^*b^*$, en donde el componente de color individual es el componente a^* .

11. El método de acuerdo con la reivindicación 10, donde el componente de color individual es la magnitud del componente a^* que se diferencia entre ser negativo para las regiones sanas (477) y ser no negativo para las regiones de daño (478).

12. El método de acuerdo con la reivindicación 1, donde la segunda red neuronal convolucional es del tipo DENSENET para procesar la imagen de hoja individual para obtener el grado de daño, realizado al procesar imágenes de hojas individual y al convertirlas en imágenes binarias, en donde el grado de daño se calcula como la relación entre el número de píxeles que corresponden al daño en la imagen binaria sobre el número de píxeles que corresponden a la hoja en la imagen de hoja individual.

13. Un producto de programa de ordenador que, cuando se carga en una memoria de un ordenador y que se

ejecuta por al menos un procesador del ordenador, provoca que el ordenador realice los pasos del método implementado por ordenador de acuerdo con cualquiera de las reivindicaciones 1-12.

- 5 14. Un sistema de ordenador (201/202) adaptado para realizar cualquiera de los métodos (602B/702B) de acuerdo con las reivindicaciones 1-12.

DIBUJOS

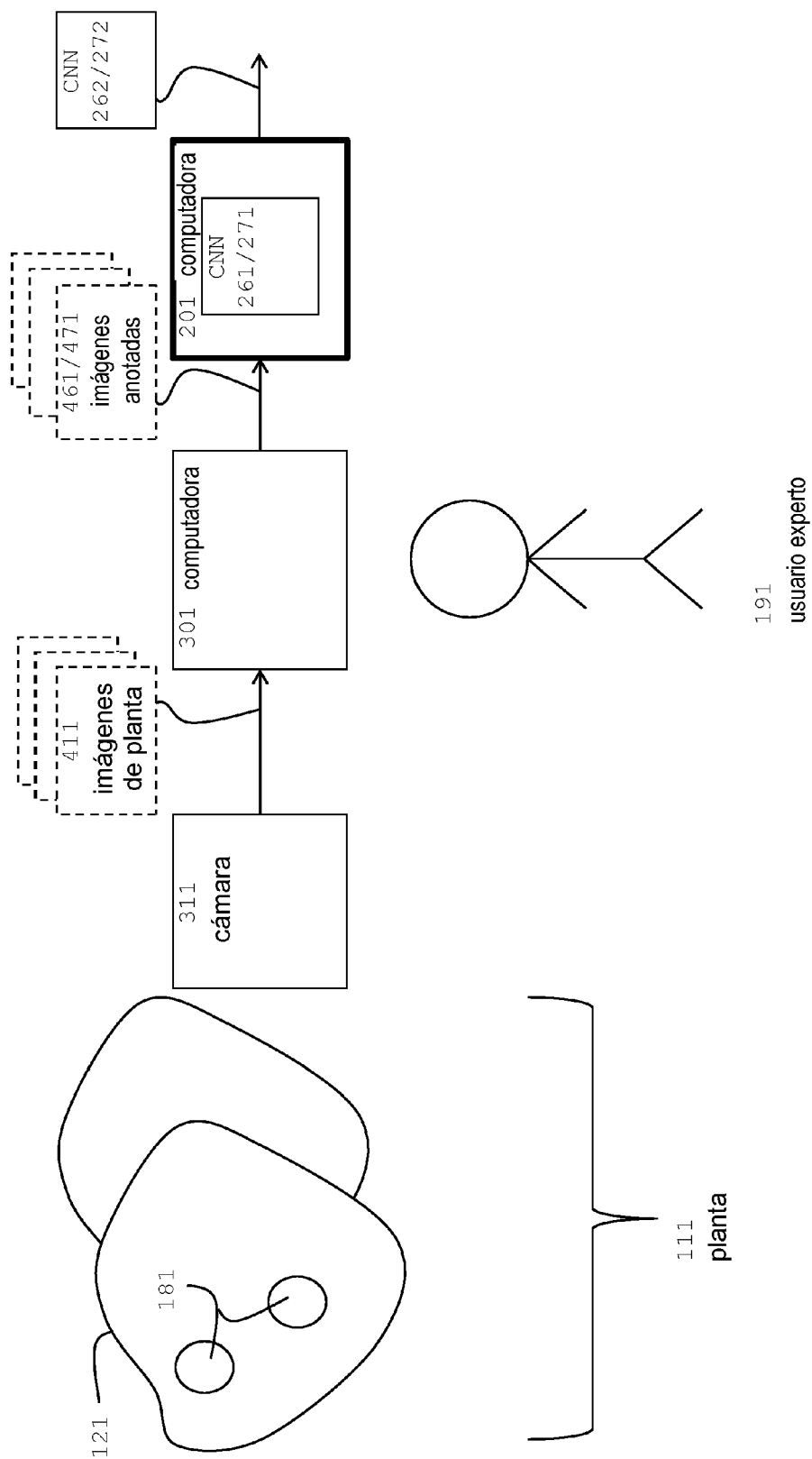


FIG. 1

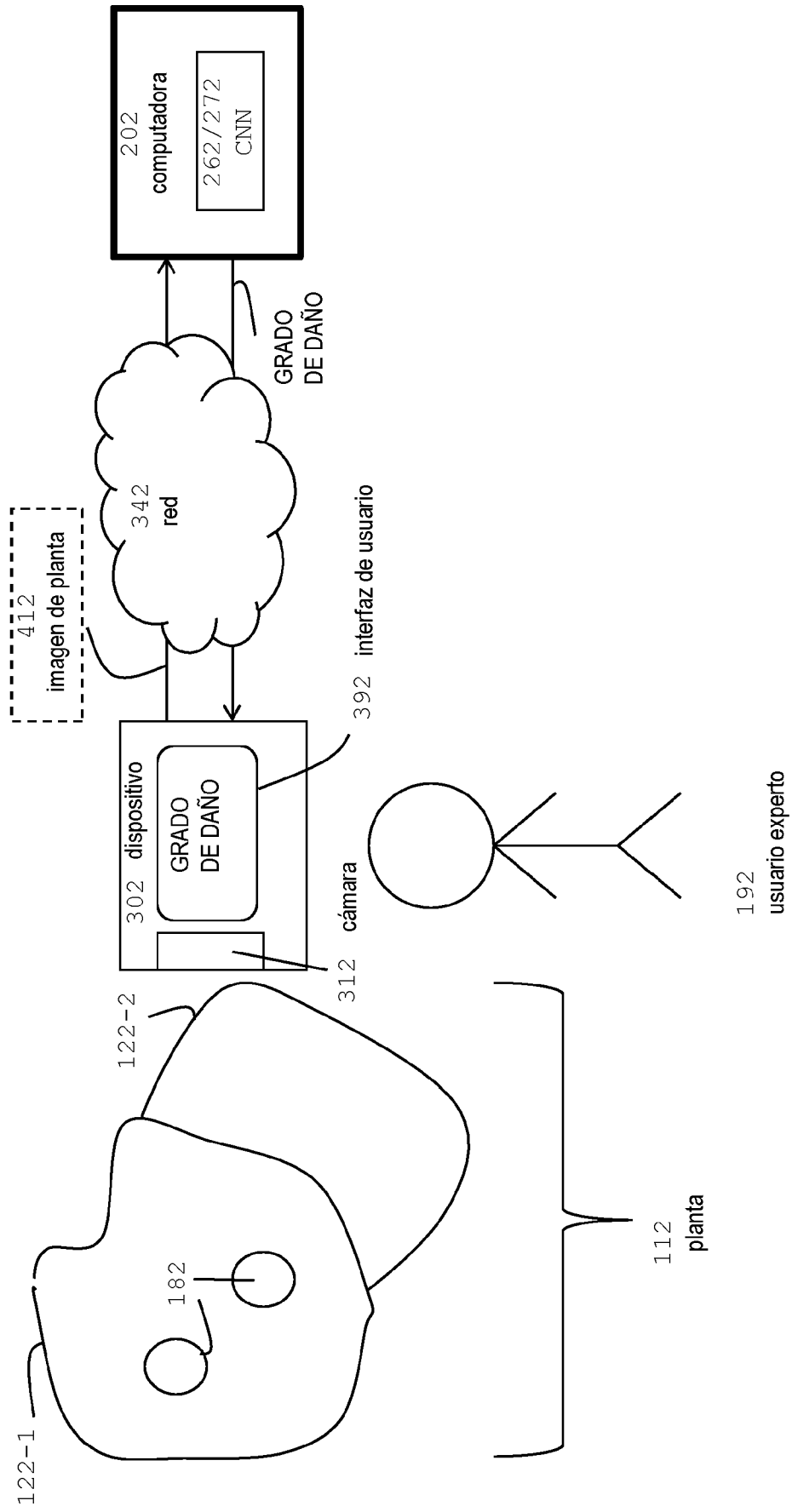


FIG. 2

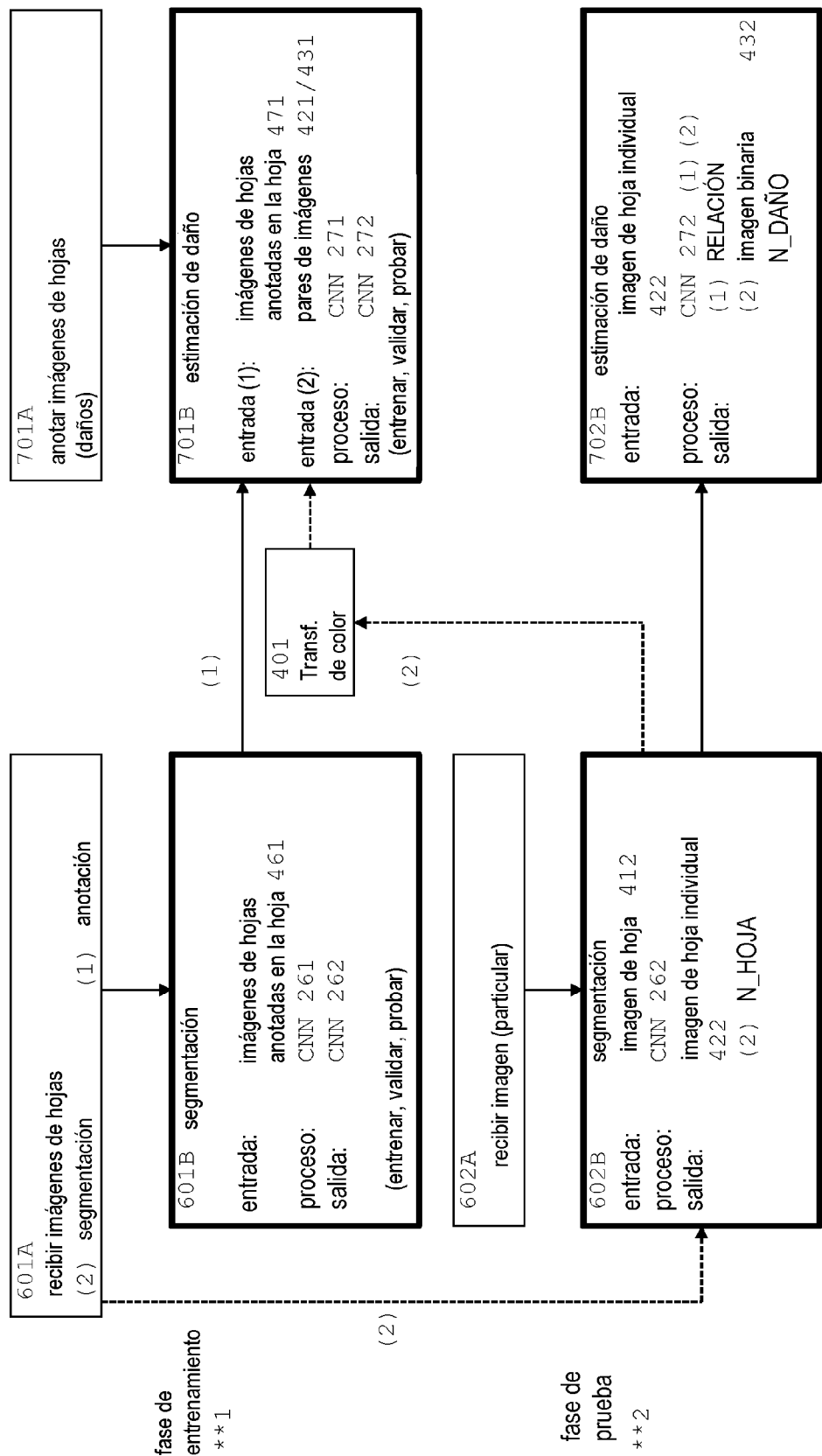


FIG. 3

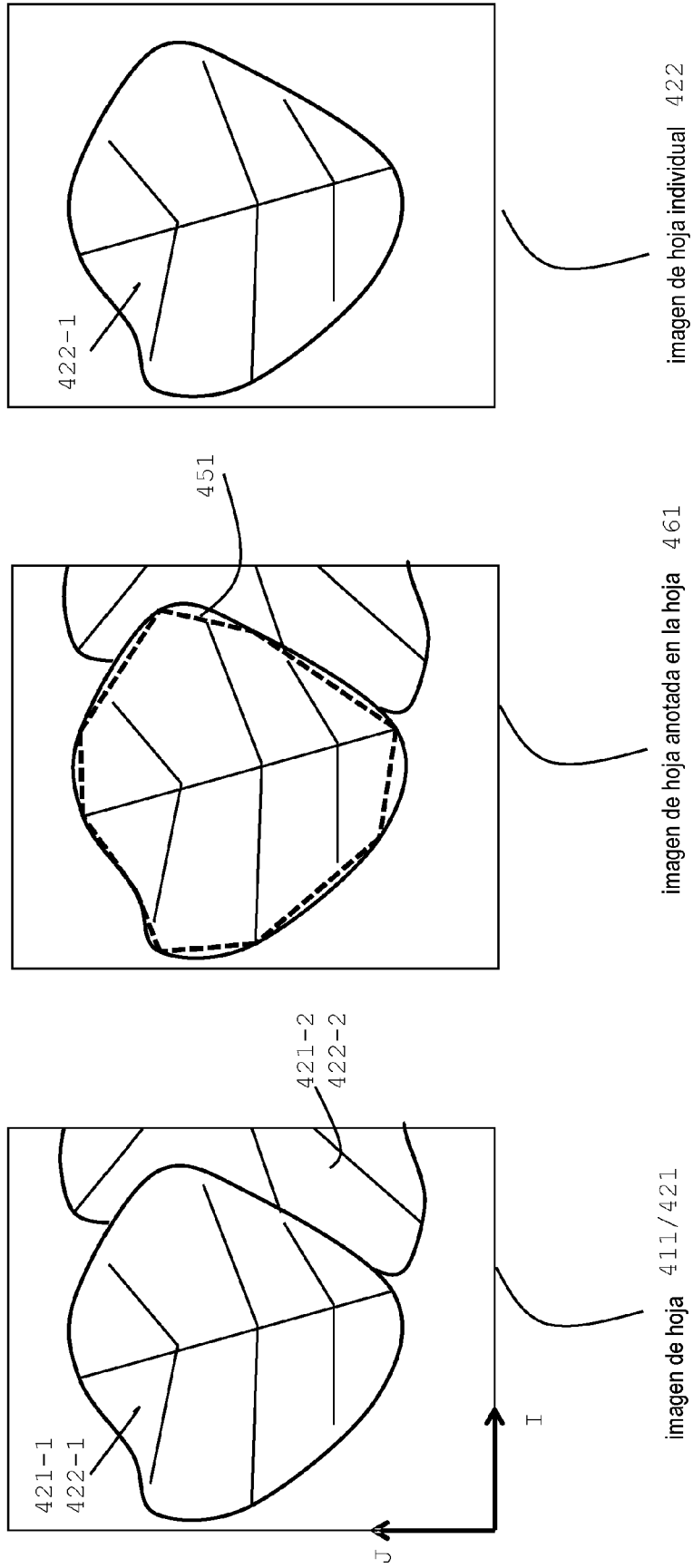
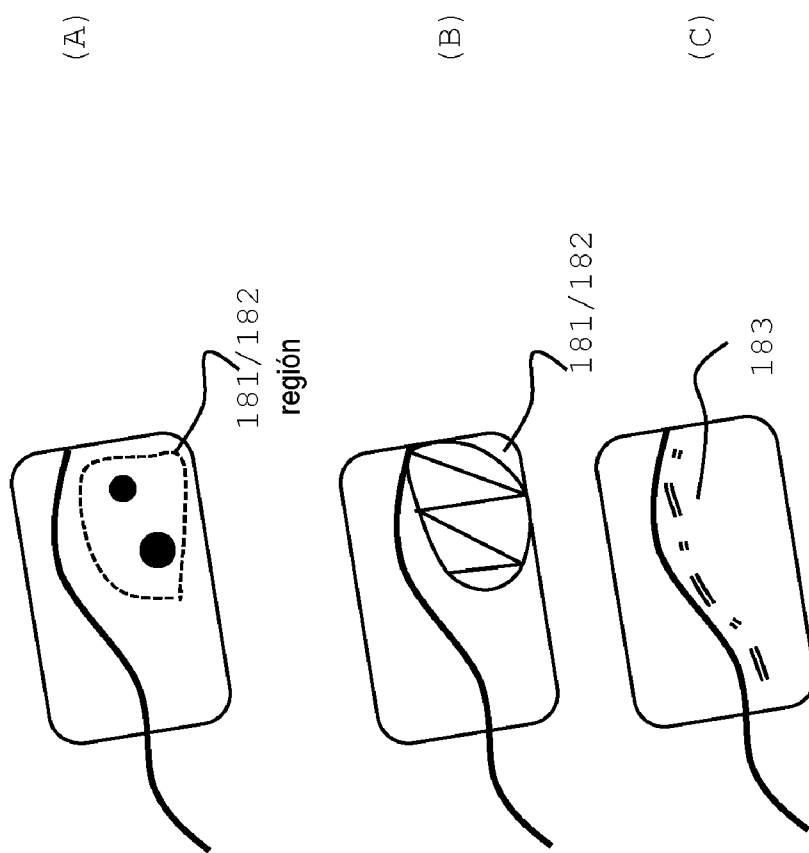


FIG. 4



121/122

FIG. 5

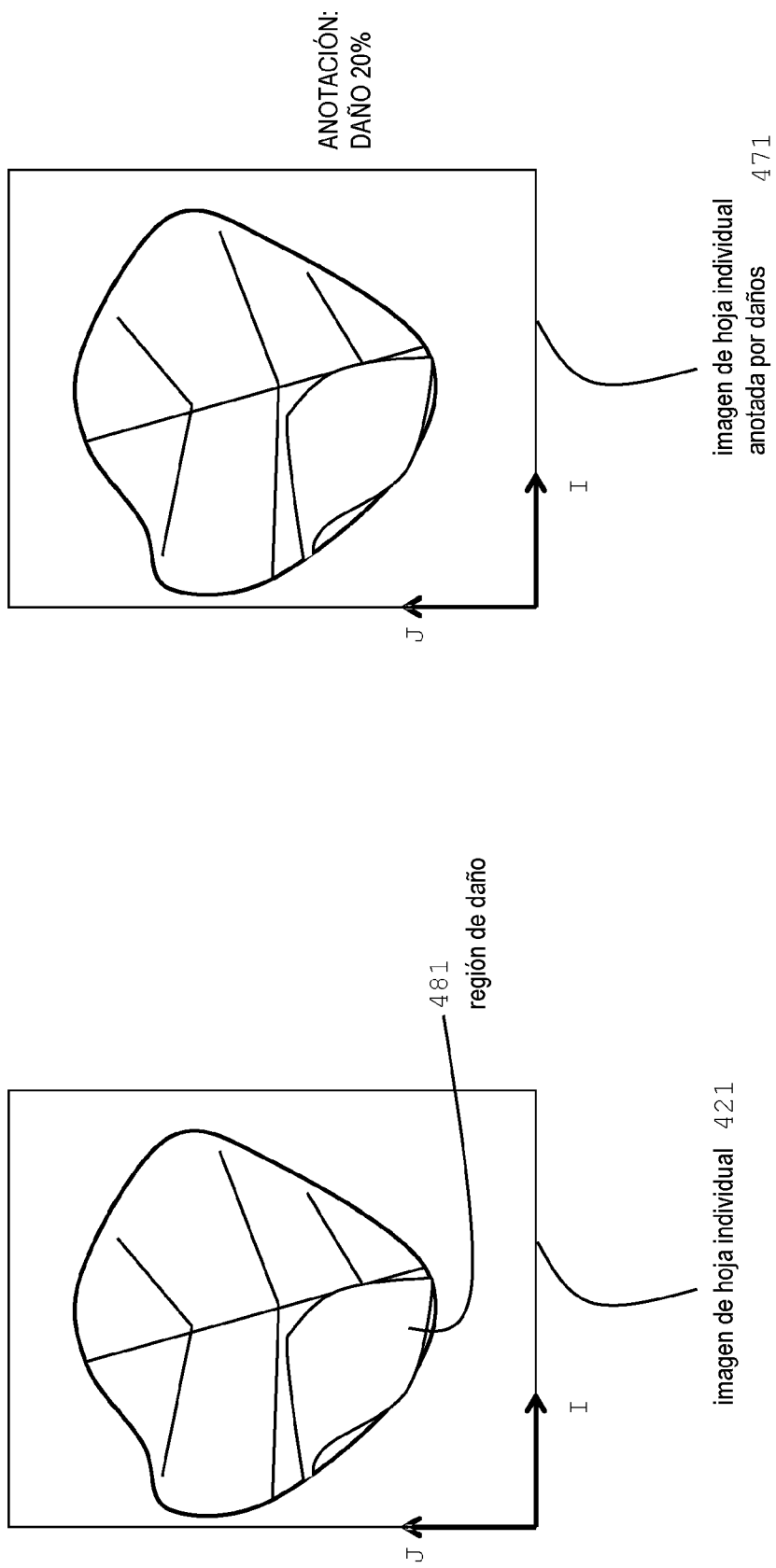


FIG. 6

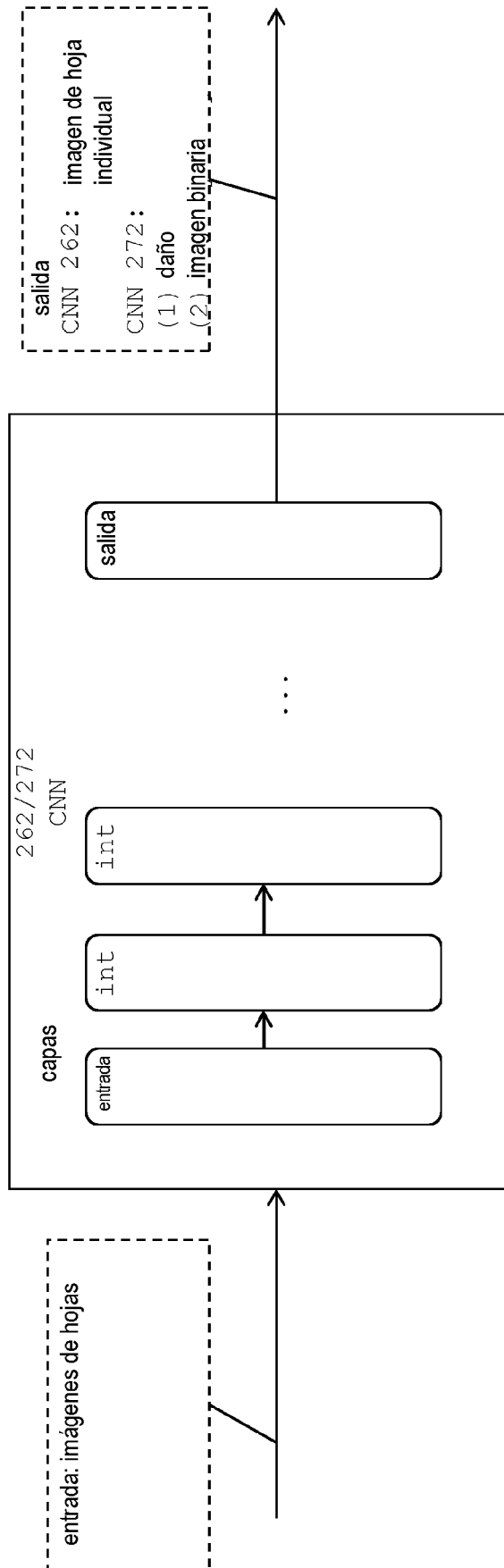
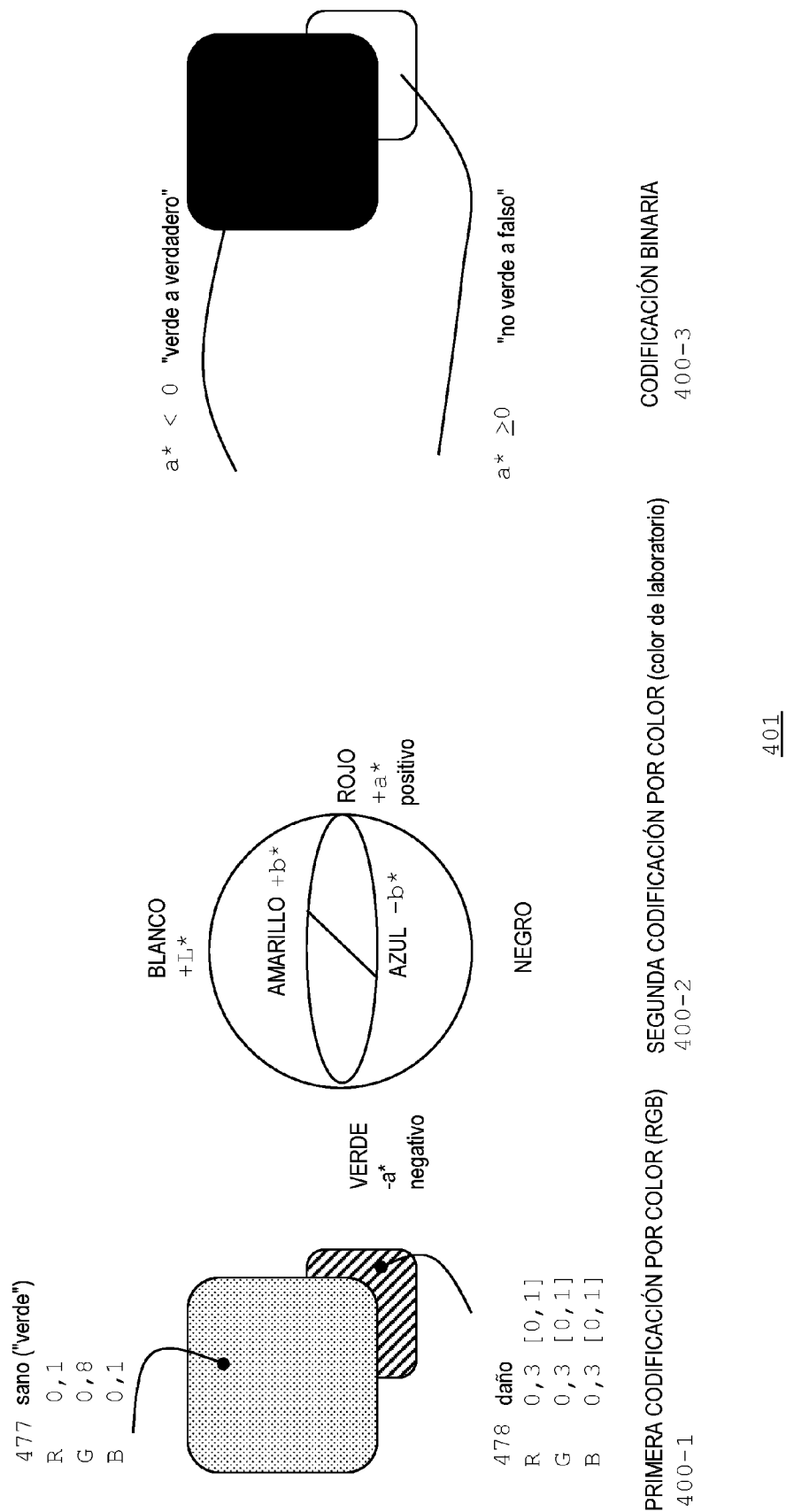


FIG. 7



CODIFICACIÓN BINARIA

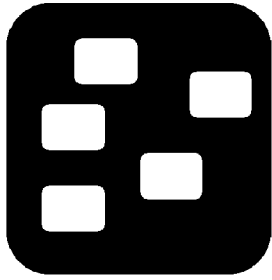
400-3

SEGUNDA CODIFICACIÓN POR COLOR (color de laboratorio)

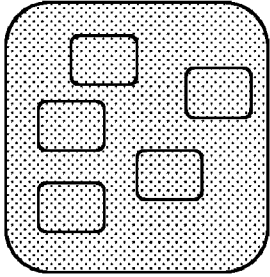
400-2

401

FIG. 8



431



421

FIG. 9

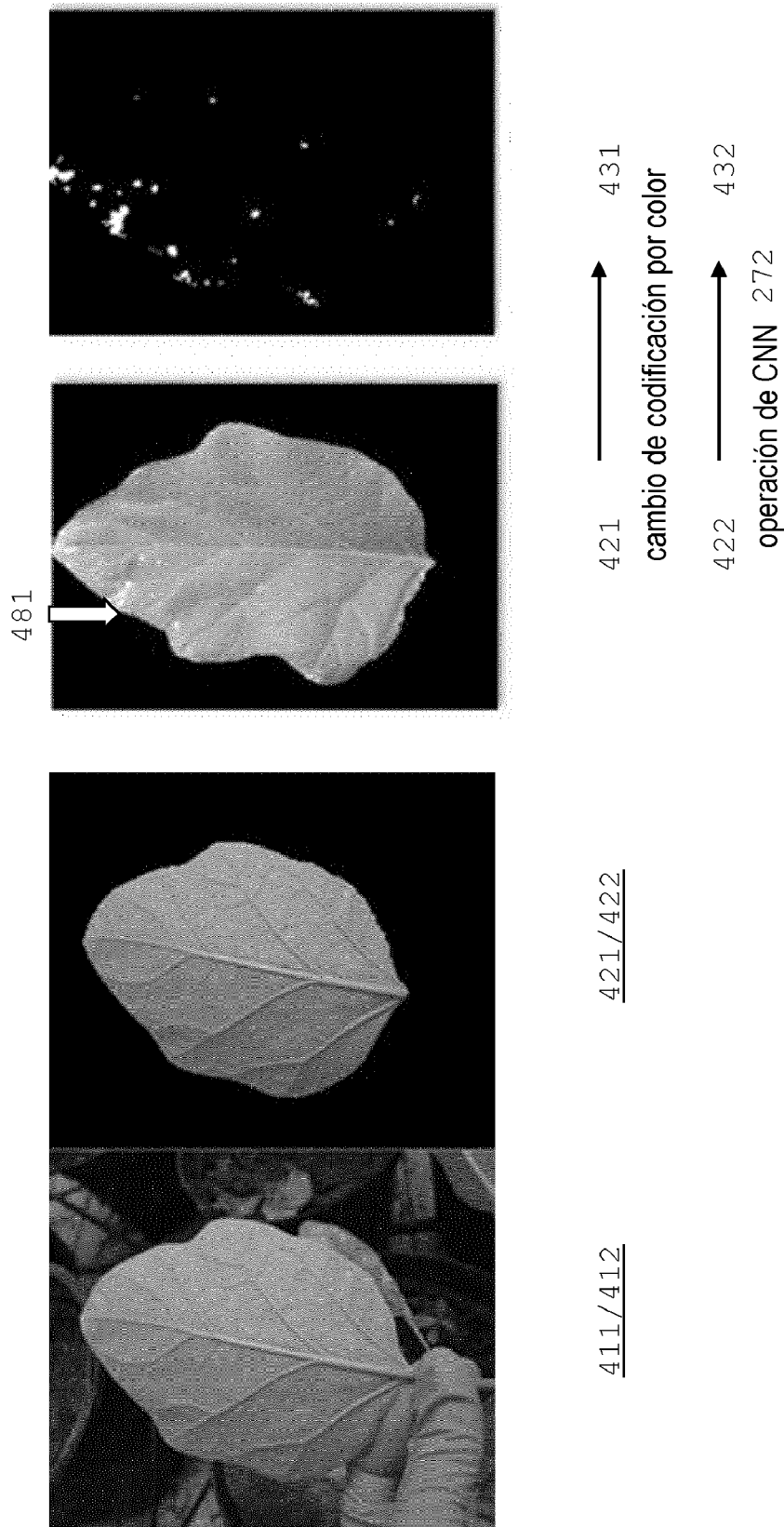


FIG. 10

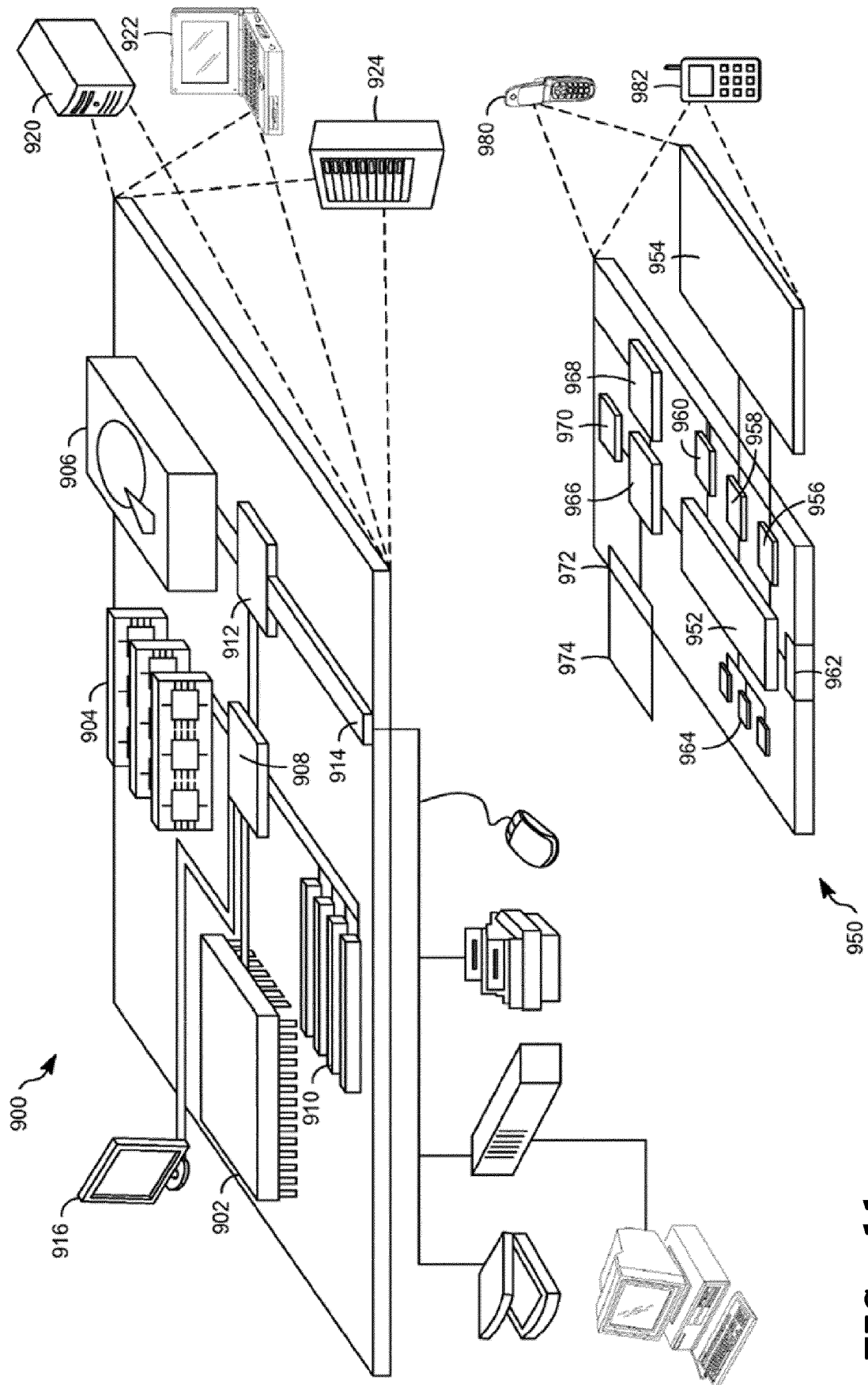


FIG. 11