1.- Modelado 3D de plantas de fresa por medio de segmentación semántica

2.- Resumen. La inteligencia artificial aplicada a la Agroindustria es un área de investigación que ha ido incrementando su uso. Se ha trabajado con la detección de Biomasa utilizando algoritmos que utilizan imagenes 2D y han tenido buenos resultados. Sin embargo, existen problemas en la detección de caracteristicas como la altura y follaje. En este proyecto, utilización la segmentación por medio de YOLOv7 y los resultados los combinamos con información de profundidad proporcionada por una cámara Intel RealSense D435. La propuesta la dividimos en tres fases. La primera es la captura de fotos RGB y de profundidad. En la segunda, el modelo YOLOv7 se encarga de llevar a cabo la segmentación de las plantas en las imágenes RGB. Por último, realizamos la combinación de la información 2D y de profundidad por medio del modelo para realizar la reconstrucción 3D. Este proyecto representa una aplicación técnica sumamente interesante en el campo de la visión por computadora y la generación de modelos tridimensionales.

3.- Introducción.

La agroindustria es un componente esencial de la economía global y desempeña un papel crucial en la producción de alimentos para satisfacer las crecientes demandas de una población en constante crecimiento (Carlos A. da Silva, 2013). Sin embargo, esta industria se enfrenta a una serie de desafíos críticos que amenazan su sostenibilidad y la seguridad alimentaria a nivel mundial (FDA, 2021). En este contexto, el cambio climático emerge como una problemática central que afecta directamente a la agroindustria, impactando la producción agrícola y la disponibilidad de recursos naturales (Carlos A. da Silva, Doyle Baker. 2013).

Según las estadísticas proporcionadas por el Instituto Nacional de Estadística y Geografía (INEGI, 2022), se evidencia un aumento preocupante en la variabilidad climática y eventos climáticos extremos que afectan de manera significativa la producción agrícola en México y en todo el mundo (INEGI, 2019). Estos datos respaldan la urgente necesidad de abordar los problemas alimentarios relacionados con el cambio climático, como la pérdida de cultivos, la disminución de la calidad de los alimentos y la amenaza a la seguridad alimentaria (INEGI, 2019).

En este contexto, la agricultura de precisión, apoyada por tecnologías emergentes como la visión artificial y algoritmos de Deep Learning, como YOLOv7 (Y Wang, H Wang, 2022), se presenta como una solución prometedora para abordar los desafíos planteados por la agroindustria (Pablo Esteban Villota Neira, 2019). Estas tecnologías avanzadas permiten una gestión más eficiente de los recursos agrícolas, la detección temprana de problemas en los cultivos y la optimización de la producción, contribuyendo así a una mayor resiliencia y sostenibilidad de la agroindustria en un contexto de cambio climático (JPT Soto, JSS Suárez, 2019).

la combinación de la agricultura de precisión, la visión artificial y los algoritmos de Deep Learning, como YOLOv7, pueden desempeñar un papel crucial en la transformación de la agroindustria y en la búsqueda de soluciones innovadoras para los problemas alimentarios en un mundo afectado por el cambio climático (Isabel Cisternas, Ignacio Velásquez, 2020).

4.- Metodología.

En este proyecto, se propone una metodología (ver Figura 1) donde se usa una cámara RGB-D Intel RealSense D435 junto con Python y la herramienta YOLO v7 para adquirir imágenes en formato RGB y capturar datos de profundidad. Este enfoque permite llevar a cabo la segmentación de hojas de plantas de fresa mediante el uso de YOLO V7. El etiquetado de estas imágenes se realiza a través de la plataforma Roboflow. Una vez que las imágenes están etiquetadas, el modelo YOLO v7 se encarga de llevar a cabo la segmentación de las plantas en las imágenes. Posteriormente, utilizando Python, se procede a la reconstrucción 3D de las plantas a partir de las imágenes 2D capturadas. Este proyecto representa una aplicación técnica en el ámbito de la visión por computadora y la generación de modelos tridimensionales.

Imagen RGB. Una imagen RGB, siglas de "Red, Green, Blue" (rojo, verde y azul en español), es una representación digital que se basa en la combinación de tres canales de color. Cada canal representa la intensidad de uno de estos tres colores primarios, lo que permite una amplia gama de colores al variar las intensidades en cada canal. La representación de imágenes en formato RGB es esencial en la fotografía. Cada píxel en una imagen RGB se define por sus valores en estos tres canales, lo que posibilita una representación fiel y detallada.

Imagen que contiene gabinete, interior, cocina, pequeño

Descripción generada automáticamente

Imagen de profundidad. Una imagen de profundidad es una representación visual que proporciona información sobre la distancia o la ubicación tridimensional de los objetos. Esta representación visual ofrece una visión de la profundidad espacial en una imagen, permitiendo a los observadores percibir la disposición de los objetos en relación con su distancia unos de otros y del punto de vista del espectador. La imagen de profundidad es una herramienta valiosa en campos como la visión por computadora, la realidad virtual y la robótica, donde la percepción de la profundidad es esencial para una comprensión precisa del entorno.

Un dibujo de una persona

Descripción generada automáticamente con confianza baja

Imagen que contiene Escala de tiempo

Descripción generada automáticamente

Cámara RGB-D

Un control color blanco

Descripción generada automáticamente con confianza media

Imagen de la pantalla de un celular

Descripción generada automáticamente con confianza media

Para este proyecto hemos decidido utilizar la cámara Intel realsense d435 por las siguientes características:

* Resolución de Imagen: La cámara D435 ofrece una resolución de color de 1280 x 720 píxeles (HD) para imágenes RGB.
* Sensor de Profundidad: Utiliza tecnología de sensores de profundidad activos para capturar datos 3D con alta precisión.
* Campo de Visión: Tiene un campo de visión amplio, que permite capturar un área considerable en una sola toma.
* Tasa de Cuadros: La cámara puede capturar video a diferentes tasas de cuadros, generalmente hasta 30 fps.

**Método para obtención del Modelo 3D de una Planta de Fresa**

En el proceso de recrear un modelo 3D de una planta de fresa, se emplea la cámara Intel RealSense D435 junto con un software personalizado desarrollado en Python. El proceso de adquisición de datos e imágenes se divide en dos etapas, Captura de Imágenes RGB y de Profundidad.

Un hombre sentado en el piso

Descripción generada automáticamente con confianza baja

**Método para el etiquetado de imágenes por medio de Roboflow**

Avanzando en el proceso de etiquetado en Roboflow, estamos asignando etiquetas a las imágenes de las hojas de plantas de fresa que hemos adquirido previamente. En este proceso, nos esforzamos por ser precisos al etiquetar cada hoja de la planta con el fin de obtener resultados de alta calidad. Una vez que hemos etiquetado todas las hojas de la planta, la herramienta de Roboflow generará automáticamente tres grupos de imágenes: entrenamiento, prueba y validación. Estos grupos se utilizarán para el entrenamiento, la evaluación y la validación de las imágenes, lo que nos permitirá crear un conjunto de datos. Este conjunto de datos se exportará para su uso en nuestro modelo de entrenamiento a través de YOLO V7.

Imagen que contiene gabinete, edificio, texto, cocina

Descripción generada automáticamente

**Método para la segmentación por medio de YOLO V7**

El conjunto de datos generado por Roboflow se importa en Google Colab, donde se integra con el modelo YOLO V7. Configuramos los parámetros de entrenamiento para un total de 600 iteraciones, lo que nos permite llevar a cabo el proceso de entrenamiento y validación del conjunto de datos que utilizaremos en esta etapa. Google Colab se encarga de organizar y almacenar los archivos en Google Drive.

Gráfico de rectángulos

Descripción generada automáticamente con confianza baja

Una vez que todo el proceso ha concluido, el modelo YOLO V7 nos proporcionará un conjunto de datos que constituye estadísticas que validan la eficacia de nuestro entrenamiento. Como se muestra en la figura anterior, los resultados de una prueba muestran que la matriz de confusión indica una precisión de detección del 95% con un margen de error del 5%. Este puntaje es altamente favorable y respalda nuestros objetivos de manera significativa.

Interfaz de usuario gráfica, Aplicación, Word

Descripción generada automáticamente

Con estas graficas se demuestra todos los datos estadísticos de nuestras pruebas realizadas en todo el proceso del entrenamiento de nuestro modelo de Deep Learning por medio del modelo YOLO V7

**Desarrollo para Modelado 3D de plantas de fresa**

Captura de pantalla de computadora

Descripción generada automáticamente

Imagen que contiene Gráfico

Descripción generada automáticamente

Imagen que contiene Gráfico

Descripción generada automáticamente

El código presentado es una implementación en Python que tiene como objetivo la visualización de datos de profundidad en un espacio tridimensional utilizando dos imágenes 2D preprocesadas: una imagen de profundidad y una imagen en color (RGB). Esta representación tridimensional resultante puede ser valiosa en aplicaciones relacionadas con la visión por computadora y la percepción espacial, como en el campo de la robótica.

1. Importación de Bibliotecas: Comenzamos importando las bibliotecas esenciales, como `numpy`, `scikit-image` y `matplotlib`. Estas son ampliamente utilizadas para el procesamiento de imágenes y la creación de gráficos.

2. Carga de Imágenes: Luego, procedemos a cargar dos imágenes desde archivos: una imagen de profundidad y una imagen RGB. Ambas deben estar en formato numpy y corresponder al mismo objeto o escena.

3. Determinación de Dimensiones: En esta etapa, obtenemos las dimensiones (alto y ancho) de la imagen de profundidad, lo cual es crucial para mapear los datos en un espacio tridimensional de manera precisa.

4. Creación de la Figura 3D: Creamos una figura tridimensional que actúa como el lienzo en el cual representaremos los datos de profundidad en un espacio 3D.

5. Transformación de Datos: Mediante un ciclo, recorremos los píxeles de la imagen de profundidad. Para cada píxel, calculamos las coordenadas tridimensionales (X, Y, Z) basándonos en los valores de profundidad y ajustándolos según las dimensiones de la imagen y una constante de escala (k).

6. Almacenamiento de Datos: Los valores resultantes de las coordenadas 3D y los valores RGB correspondientes se almacenan en listas separadas (X, Y, Z y rgb).

7. Creación de una Visualización 3D: Representamos los datos 3D en la figura tridimensional. Cada punto 3D se colorea de acuerdo con el valor RGB correspondiente de la imagen.

8. Visualización de la Figura: Finalmente, mostramos la figura que representa los datos en un espacio tridimensional, lo que permite observar cómo se ven los objetos en función de la información de profundidad y color.

**Conclusiones**

La obtención del modelo 3D de las plantas de fresa se realizó de manera satisfactoria de acuerdo a la metodología propuesta. En primer lugar, se capturo y etiqueto con precisión las imágenes de hojas de plantas de fresa utilizando la cámara Intel RealSense D435 y la plataforma Roboflow. Este proceso de etiquetado requiere de precisión para asegurar la eficacia de nuestro modelo.

Posteriormente, se utilizo el dataset de las imágenes etiquetadas en Google Colab y los se empleo para entrenar un modelo YOLO V7, configurando parámetros de entrenamiento específicos que han logrado una detección de alta precisión.

Además, desarrollamos una herramienta en Python que nos permite visualizar datos de profundidad en un espacio tridimensional. Esto proporciona una comprensión más completa y precisa de la disposición de objetos en la escena, lo que es de gran utilidad en campos como la visión por computadora y la robótica.

En resumen, conseguimos la obtención de un modelo 3D detallado de una planta de fresa, cumpliendo nuestros objetivos de captura de datos precisos, entrenamiento de un modelo de alta calidad y representación visual efectiva en un espacio tridimensional. Estos logros respaldan aplicaciones diversas y prometen un impacto positivo en proyectos futuros relacionados con la percepción espacial y la representación visual de objetos.

Fuentes

La agroindustria <http://www.lamolina.edu.pe/postgrado/pmdas/cursos/innovacion/lecturas/Adicional/20%20-%20da%20Silva%20et%20al.pdf>

La agroindustria <https://www.fda.gov/news-events/press-announcements/actualizacion-sobre-el-covid-19-el-usda-y-la-fda-enfatizan-que-la-informacion-epidemiologica-y>

Datos de ingei agroindustria <https://camjol.info/index.php/NEXO/article/view/13122>

<https://www.inegi.org.mx/programas/ena/2019/>

el cambio climático en mexico por inegi <https://rde.inegi.org.mx/index.php/2015/05/10/politica-del-cambio-climatico-en-mexico-avances-obstaculos-y-retos/#:~:text=El%20cambio%20clim%C3%A1tico%20(CC)%20tendr%C3%A1,y%205%25%20en%20el%20verano>.

Visión artificial en la agroindustria <https://core.ac.uk/download/pdf/288582212.pdf>

Yolo v7 <https://ieeexplore.ieee.org/abstract/document/9994733/>

Agricultura de precisión <https://www.funlam.edu.co/revistas/index.php/lampsakos/article/view/3253>

<https://www.sciencedirect.com/science/article/pii/S0168169920312357>