



MANUAL TÉCNICO: **ETIQUETADO DE IMÁGENES PARA ALGORITMOS DE DEEP LEARNING**

Centro de Electricidad y Automatización Industrial
Regional Valle



Esta obra está bajo una Licencia Creative Commons
Atribución – No comercial – Compartir igual - 4.0 Internacional



**MANUAL TÉCNICO:
ETIQUETADO DE IMÁGENES PARA ALGORITMOS DE DEEP LEARNING**

Sistema de Investigación, Desarrollo Tecnológico e Innovación
SENNOVA - SENA

Regional Valle
Centro de Electricidad y Automatización Industrial
Área de Innovación y Competitividad

Calle 52 No 2 Bis 15 Complejo Salomia. Tel 431 5800. Ext. 22581 - 22763
Cali - Colombia



SERVICIO NACIONAL DE APRENDIZAJE, SENA

Centro de Electricidad y Automatización Industrial, CEAI

Fanny Clemencia Montenegro Maya
Subdirectora de centro (E)

José Fernando Pérez Villa
Dinamizador Área de Innovación y Competitividad del CEAI

Grupo de Investigación: Unidad de Investigación Aplicada, Desarrollo Tecnológico e Innovación, UIADTI

Autores:

Cesar Augusto Agudelo

Instructor Investigador

Kelyn Botina

Investigador SENNOVA

Edinson Cortés

Investigador SENNOVA

Brayan Andrés Gutierrez Rojas

Investigador Junior

Comité editorial CEAI

Maggie Carolina Rafet Jaimes

César Augusto Agudelo Forero

Karen Andrea Saldarriaga Ramírez

José Fernando Pérez Villa

ISBN:

Primera edición

Diseño y diagramación:

Karen Andrea Saldarriaga Ramírez

© Servicio Nacional de Aprendizaje SENA.

Este libro salvo las excepciones previstas por la ley, no puede ser reproducido por ningún medio, sin previa autorización escrita del autor. Los textos publicados son de propiedad intelectual del autor y pueden utilizarse con propósitos educativos y académicos, siempre que se cite el autor y la publicación. Las opiniones aquí contenidas son responsabilidad de los autores y no reflejan necesariamente el pensamiento del editor del SENA.

Santiago de Cali, diciembre de 2024

Contenido

INTRODUCCIÓN	5
Objetivo del manual.....	7
Contextualización.....	8
Importancia del etiquetado de imágenes en algoritmos de Deep Learning.....	8
Desafíos en el etiquetado de imágenes	9
Aplicaciones del etiquetado de imágenes en el análisis de Cristales De Azúcar	11
Monitorización y Control de Calidad	11
Presentación de la herramienta VGG Annotator.....	13
Conceptualización	14
Fundamentos del Deep Learning y redes neuronales convolucionales	14
Introducción al Deep Learning.....	14
Redes Neuronales Convolucionales (CNN).....	14
Aplicación de CNN en el Etiquetado de Imágenes.....	15
Técnicas de Procesamiento de Imágenes Aplicadas al Etiquetado.....	16
Mejora de la Calidad de las Imágenes	16
Segmentación y Detección de Objetos	16
Extracción de Características Relevantes.....	17
Eliminación de Ruido	17
Normalización y Escalamiento.....	17
Descripción Del Desarrollo	18
Preparación de los Datos de Entrenamiento	19
Recopilación y Selección de Imágenes de Cristales de Azúcar	19
Limpieza y Mejora de la Calidad de las Imágenes.....	20
Etiquetado Manual de las Imágenes con VGG Annotator.....	21
Instalación y Configuración de VGG Annotator	22
Creación y Gestión de Proyectos de Etiquetado	23
Caso De Estudio: Etiquetado De Cristales De Azúcar.....	24
Contexto	24
Procedimiento de Etiquetado de Imágenes de Cristales de Azúcar	27
Descripción de las Categorías de Etiquetas Utilizadas en el Etiquetado Manual de las Imágenes	32
Estrategias de etiquetado específicas para Cristales de Azúcar.....	36

Aplicación del Modelo de Deep Learning.....	39
Resultados Obtenidos	42
Refinamiento de las regiones etiquetadas.....	42
Exclusión de imágenes borrosas o no relevantes.....	42
Identificación y clasificación de tipologías específicas	42
Mejora Continua A Lo Largo Del Proceso.....	42
Impacto Positivo En La Calidad Del Conjunto De Datos.....	43
Lecciones Aprendidas Y Futuras Iteraciones	43
Glosario	44
Bibliografía	46

Índice figuras

Figura 1. Pantalla inicial herramienta VGG Image Annotator (VIA).	13
Figura 2. Contenido del archivo comprimido via-2.0.12.zip e identificación del archivo ejecutable	23
Figura 3. Opción para cargar las imágenes que se desean anotar.....	28
Figura 4. Despliegue de las opciones de atributos del etiquetado.....	28
Figura 5. Nombre del atributo.	28
Figura 6. Definición del tipo de atributo creado.....	29
Figura 7. Descripción de id en el atributo “Forma” creado.	29
Figura 8. Tipo de forma empleada para definir la región de la etiqueta.....	30
Figura 9. Ejemplo de trazado de región sobre la imagen a etiquetar.....	30
Figura 10. Asignación del tipo de etiqueta a la región creada.	31
Figura 11. Marcación del tipo de cristal sobre la región creada.	31
Figura 12. Menú para el guardado del proyecto.....	32
Figura 13. Imagen con un cristal cúbico seleccionado.....	33
Figura 14. Imagen con un cristal alargado seleccionado.	33
Figura 15. Imagen con un cristal Tipo D seleccionado.	34
Figura 16. Imagen con cristales aglomerados seleccionados.....	35
Figura 17. Imagen con cristales tipo “Otros” seleccionados.	35
Figura 18. Resultados obtenidos con los modelos entrenados.....	41

INTRODUCCIÓN

En la era actual de rápido avance tecnológico, los algoritmos de Deep Learning han emergido como herramientas fundamentales para abordar problemas complejos y extraer información valiosa de datos visuales (Goodfellow, Bengio, & Courville, 2016). La eficacia de estos algoritmos está intrínsecamente ligada a la calidad y precisión de los datos de entrenamiento, especialmente en el análisis de imágenes (LeCun, Bengio, & Hinton, 2015). En este contexto, el presente “Manual Técnico: Etiquetado de Imágenes para Algoritmos de Deep Learning” se posiciona como una guía exhaustiva, diseñada para optimizar los resultados en aplicaciones que demandan un análisis visual riguroso y preciso.

Este manual abarca un espectro completo, desde los conceptos fundamentales hasta las técnicas avanzadas de etiquetado de imágenes, cruciales para el entrenamiento efectivo de algoritmos de Deep Learning. Se exploran meticulosamente las técnicas de procesamiento de imágenes y los métodos de anotación, haciendo especial énfasis en la utilización de la herramienta VGG Annotator para agilizar y mejorar el proceso de etiquetado (Dutta & Zisserman, 2019). Aunque el enfoque principal se centra en el análisis de imágenes, este manual se circunscribe a los casos donde el etiquetado manual es aplicable, sin adentrarse en técnicas automáticas de etiquetado o áreas ajenas al análisis de imágenes.

El público objetivo de este manual incluye investigadores, ingenieros de datos y profesionales en el campo del aprendizaje automático que trabajan en proyectos que involucran análisis de imágenes. Asimismo, resulta de gran utilidad para estudiantes avanzados y académicos interesados en comprender y aplicar las mejores prácticas en el etiquetado de datos visuales para el entrenamiento de algoritmos de Deep Learning (Krizhevsky, Sutskever, & Hinton, 2017).

La estructura del manual guía al lector a través de un recorrido lógico y progresivo. Comienza con una explicación detallada sobre la importancia crucial del etiquetado preciso, seguida por capítulos que profundizan en técnicas de procesamiento de imágenes, estrategias para mejorar la calidad de las anotaciones, y el uso práctico de la herramienta VGG Annotator. Como elemento central y distintivo, se incluye un análisis pormenorizado basado en el proyecto de investigación “Evaluación de técnicas de Visión Artificial basadas en Deep Learning para detección de morfologías en sustancias cristalinas de origen vegetal – SENNOVA 2023”. Este caso de estudio sirve como ejemplo práctico para ilustrar la aplicación real de las técnicas discutidas, proporcionando un valioso puente entre la teoría y la práctica.

El manual concluye con una sección dedicada a buenas prácticas y recomendaciones específicas, diseñadas para asegurar la consistencia y calidad de las anotaciones en diversos contextos de aplicación.

La metodología empleada en este manual combina una revisión exhaustiva de la literatura existente con la aplicación práctica de herramientas y técnicas de vanguardia en el etiquetado de imágenes. Se fundamenta en fuentes ampliamente reconocidas en el campo del Deep Learning (LeCun et al., 2015; Krizhevsky et al., 2017), complementadas con experiencias prácticas en la utilización de herramientas como VGG Annotator (Dutta & Zisserman, 2019).

Objetivo del manual

El objetivo principal de este manual técnico es proporcionar una guía exhaustiva y práctica para el etiquetado preciso y eficiente de imágenes, orientado específicamente al entrenamiento de algoritmos de Deep Learning. Este manual se fundamenta en técnicas avanzadas de procesamiento de imágenes y en la utilización de herramientas especializadas como VGG Annotator (Dutta & Zisserman, 2019), con el fin de capacitar a los lectores en todas las etapas del proceso de etiquetado, desde la optimización de la calidad visual de las imágenes hasta la implementación precisa de anotaciones.

Este manual está diseñado para satisfacer las necesidades específicas de investigadores, ingenieros de datos y estudiantes avanzados en el campo del aprendizaje automático y la visión por computador. Su propósito es dotar a estos profesionales con los conocimientos y habilidades necesarios para generar datos de entrenamiento de alta calidad, contribuyendo así a mejorar significativamente la precisión y el rendimiento de los algoritmos de Deep Learning en aplicaciones específicas (LeCun, Bengio, & Hinton, 2015).

Al concluir la lectura de este manual, los usuarios estarán capacitados para:

- 1.** Aplicar técnicas avanzadas de procesamiento de imágenes para mejorar la calidad de los datos de entrada.
- 2.** Utilizar eficientemente la herramienta VGG Annotator para el etiquetado preciso de imágenes.
- 3.** Implementar estrategias para garantizar la coherencia y utilidad de los conjuntos de datos generados.
- 4.** Evaluar la calidad de las anotaciones y su impacto en el rendimiento de los algoritmos de Deep Learning.

El éxito en la aplicación de los conocimientos adquiridos se medirá a través de la capacidad de los usuarios para crear conjuntos de datos coherentes y de alta calidad, que contribuyan a mejorar significativamente los resultados de sus proyectos de Deep Learning (Goodfellow, Bengio, & Courville,

2016). El contenido de este manual se ha seleccionado y estructurado cuidadosamente para ser relevante y realista, alineándose con las necesidades actuales de optimización de procesos de etiquetado en entornos que requieren precisión. Además, se ha diseñado para que pueda ser completado en un marco de tiempo razonable, ajustándose a los plazos habituales en proyectos técnicos de esta naturaleza.

Contextualización

Importancia del etiquetado de imágenes en algoritmos de Deep Learning.

En el ámbito de la visión artificial, el etiquetado preciso de imágenes se erige como un pilar fundamental para el éxito de los algoritmos de Deep Learning, una rama avanzada del aprendizaje automático que emplea redes neuronales profundas para el análisis y clasificación de vastos conjuntos de datos visuales. LeCun, Bengio y Hinton (2015) subrayan que la fortaleza del Deep Learning radica en la capacidad de estas redes para aprender representaciones jerárquicas a partir de datos no estructurados, lo cual resulta particularmente eficaz en tareas complejas de detección y clasificación de imágenes.

La eficacia y precisión de los modelos de Deep Learning están intrínsecamente ligadas a la calidad y exhaustividad de los datos de entrenamiento. En este contexto, el etiquetado meticuloso de imágenes se convierte en una tarea crítica para alcanzar un rendimiento óptimo del modelo (Goodfellow, Bengio, & Courville, 2016).

Un etiquetado adecuado permite a los algoritmos identificar patrones distintivos y características específicas, facilitando así la clasificación y detección precisa en una amplia gama de aplicaciones, desde el reconocimiento facial hasta la identificación de objetos en entornos complejos.

En escenarios de visión artificial avanzada, la delimitación precisa de contornos y la identificación de características intrínsecas son elementos cruciales para garantizar la exactitud del modelo. Krizhevsky, Sutskever y Hinton (2012) demostraron la importancia de estos aspectos en su trabajo

pionero con redes neuronales convolucionales, estableciendo un nuevo estándar en el reconocimiento de imágenes a gran escala. La precisión en el etiquetado impacta directamente en la confiabilidad y aplicabilidad de las soluciones de visión artificial desarrolladas, determinando en gran medida su éxito en entornos del mundo real (He, Zhang, Ren, & Sun, 2016).

Además, el etiquetado de alta calidad contribuye significativamente a la robustez y generalización de los modelos de Deep Learning. Russakovsky et al. (2015) enfatizan que la diversidad y representatividad de los datos etiquetados son fundamentales para crear modelos que puedan funcionar eficazmente en una variedad de condiciones y escenarios.

Esto es particularmente relevante en aplicaciones críticas como la conducción autónoma o el diagnóstico médico asistido por IA, donde la precisión y fiabilidad son imperativas.

En resumen, el etiquetado preciso de imágenes no es simplemente un paso preliminar en el desarrollo de algoritmos de Deep Learning para visión artificial, sino un componente integral que determina en gran medida el rendimiento, la fiabilidad y la aplicabilidad práctica de estos sistemas avanzados.

Desafíos en el etiquetado de imágenes

El etiquetado de imágenes, fundamental para el entrenamiento de algoritmos de Deep Learning, presenta diversos desafíos que pueden impactar significativamente en la precisión y eficacia de los modelos resultantes. A continuación, se analizan los principales retos identificados en la literatura científica y en la práctica profesional.

Ambigüedad y Subjetividad en las Anotaciones

La interpretación subjetiva de las imágenes constituye uno de los desafíos más complejos en el proceso de etiquetado. Torralba et al. (2011) señalan que la variabilidad en las interpretaciones entre diferentes anotadores puede generar inconsistencias significativas en los datos de entrenamiento. Esta divergencia interpretativa no solo dificulta la creación

de un conjunto de datos coherente, sino que también puede comprometer la capacidad del modelo para generalizar eficazmente a nuevos datos (Zhang et al., 2018).

Volumen de Datos y Costo de Anotación

El etiquetado de imágenes es un proceso que demanda una considerable inversión de tiempo y recursos, especialmente cuando se requieren grandes volúmenes de datos para entrenar modelos de Deep Learning robustos. Sun et al. (2017) enfatizan que la anotación manual de extensos conjuntos de datos puede resultar prohibitivamente costosa y laboriosa. Este desafío se magnifica en proyectos que requieren anotaciones de alta calidad en grandes escalas, lo que puede imponer restricciones significativas en términos de presupuesto y plazos (Russakovsky et al., 2015).

Balanceo de Clases

El desequilibrio en la representación de clases dentro de los conjuntos de datos etiquetados es un problema recurrente que puede inducir sesgos en los modelos de aprendizaje. He y Garcia (2009) advierten que este desequilibrio puede llevar a modelos que favorecen las clases mayoritarias en detrimento de la precisión en la detección de clases minoritarias. Abordar este reto requiere la implementación de estrategias sofisticadas de muestreo o técnicas de aumento de datos para garantizar una representación equitativa de todas las clases en el conjunto de entrenamiento (Chawla et al., 2002).

Calidad y Variabilidad de las Imágenes

La diversidad en la calidad de las imágenes y las condiciones de captura representan un desafío significativo para el etiquetado preciso. Krizhevsky, Sutskever y Hinton (2012) destacan que factores como la baja resolución, el ruido y la iluminación deficiente pueden obstaculizar la identificación precisa de objetos y características relevantes. Además, la variabilidad en poses, ángulos y contextos de los objetos exige un alto nivel de consistencia por parte de los anotadores, un estándar difícil de mantener uniformemente (Everingham et al., 2015).

Etiquetado de Detalles Finos y Complejos

En aplicaciones que demandan un alto nivel de precisión, como la segmentación de imágenes médicas o el análisis de materiales, el etiquetado se torna particularmente desafiante. Bejnordi et al. (2017) subrayan la dificultad de lograr una precisión óptima en la segmentación de imágenes médicas, donde incluso las variaciones más sutiles pueden ser críticas. La anotación precisa de estos detalles finos requiere no solo un conocimiento especializado sino también herramientas avanzadas de anotación (Litjens et al., 2017).

Escalabilidad y Automatización

A medida que los proyectos crecen en escala, la necesidad de etiquetar volúmenes masivos de datos se convierte en un reto formidable. Aunque herramientas como VGG Annotator (Dutta & Zisserman, 2019) ofrecen soluciones semiautomáticas para el etiquetado, Everingham et al. (2015) señalan que la escalabilidad de estos sistemas aún enfrenta limitaciones significativas. La transición hacia procesos más automatizados, si bien promete eficiencia, también introduce nuevos desafíos relacionados con la precisión y la adaptabilidad a diversos tipos de imágenes (Papadopoulos et al., 2017).

Aplicaciones del etiquetado de imágenes en el análisis de Cristales De Azúcar

El etiquetado de imágenes se ha convertido en un proceso fundamental para el entrenamiento de algoritmos de Deep Learning, especialmente en aplicaciones que demandan una alta precisión en la detección y clasificación de estructuras complejas. En el contexto específico del análisis de cristales de azúcar, esta técnica ofrece beneficios significativos tanto para la industria alimentaria como para la investigación científica, mejorando la calidad del producto y profundizando la comprensión de los procesos de cristalización.

Monitorización y Control de Calidad

En la industria azucarera, el etiquetado de imágenes desempeña un papel crucial en la monitorización y el control de calidad durante la producción de azúcar. La identificación precisa de patrones específicos relacionados con la pureza,

tamaño y forma de los cristales es esencial para garantizar la consistencia en la producción. Singh et al. (2019) demostraron que el análisis automatizado de imágenes etiquetadas puede detectar eficazmente impurezas y variaciones morfológicas que podrían comprometer la calidad del azúcar. Este nivel de control no solo previene problemas de calidad, sino que también asegura que el producto final cumpla con los estándares requeridos por la industria y los consumidores.

Investigación Científica en Morfología Cristalina

En el ámbito de la investigación científica, el etiquetado de imágenes se ha convertido en una herramienta invaluable para el análisis detallado de la morfología de los cristales de azúcar. Esta técnica permite a los investigadores identificar características específicas de los cristales que indican la presencia de impurezas, defectos o variaciones en la estructura cristalina. Ramírez et al. (2020) utilizaron el etiquetado de imágenes para estudiar las diferencias en la formación cristalina bajo diversas condiciones de temperatura y concentración. Sus hallazgos no solo profundizaron el entendimiento de los procesos de cristalización, sino que también abrieron nuevas posibilidades para mejorar las técnicas de producción.

Optimización de la Producción

La optimización del proceso de fabricación representa otra aplicación clave del etiquetado de imágenes en el análisis de cristales de azúcar. Al etiquetar imágenes de cristales en diferentes etapas del proceso de producción, se pueden identificar áreas de mejora y ajustar los parámetros de producción para obtener cristales con las propiedades deseadas. Zhang et al. (2018) demostraron que el uso de imágenes etiquetadas permitió ajustar con precisión las variables de temperatura y tiempo de cristalización, maximizando así el rendimiento y la calidad de los cristales.

Esta optimización no solo mejora la eficiencia del proceso, sino que también reduce significativamente los costos asociados con la producción de azúcar al minimizar el desperdicio y optimizar el uso de recursos.

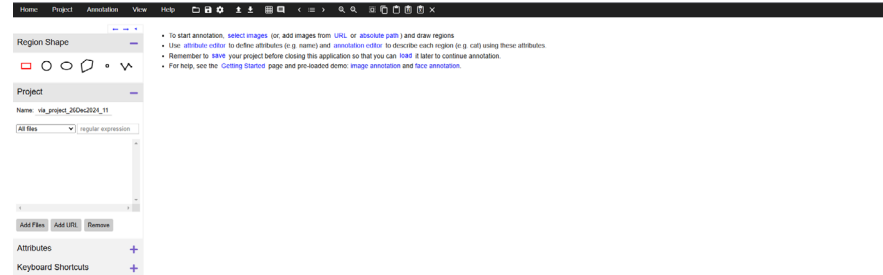
La aplicación del etiquetado de imágenes en el análisis de cristales de azúcar ha demostrado ser una herramienta poderosa que impulsa la innovación y la eficiencia en la industria azucarera. Desde el control de calidad hasta la investigación científica y la optimización de procesos, esta técnica continúa abriendo nuevas posibilidades para mejorar la producción y comprensión de los cristales de azúcar.

Presentación de la herramienta VGG Annotator

El VGG Image Annotator (VIA) es una herramienta de anotación de imágenes de código abierto, diseñada para definir regiones dentro de una imagen y proporcionar descripciones textuales detalladas de esas regiones. Desarrollada por el Visual Geometry Group (VGG) de la Universidad de Oxford, VIA se ha convertido en una herramienta ampliamente utilizada en la comunidad de visión por computadora debido a su simplicidad, versatilidad y disponibilidad bajo la licencia BSD-2 (Dutta & Zisserman, 2019).

Figura 1.

Pantalla inicial herramienta VGG Image Annotator (VIA).



Dentro de los principales investigadores y desarrolladores del aplicativo se encuentran Andrew Zisserman, profesor de ingeniería de visión por computadora en Oxford y profesor de investigación de la Royal Society; el profesor Andrea Vedaldi, profesor de Visión por Computador y Aprendizaje Automático y codirector del grupo VGG en el departamento de Ciencias de la Ingeniería de la Universidad de Oxford; el Dr João Henriques, investigador de la Real Academia de Ingeniería VGG, Universidad de Oxford; y el Dr Christian Rupprecht, profesor departamental de visión por computadora en VGG en Oxford.

Conceptualización

Fundamentos del Deep Learning y redes neuronales convolucionales

Introducción al Deep Learning

El Deep Learning, una rama avanzada del aprendizaje automático ha transformado significativamente la capacidad de las máquinas para procesar y comprender grandes volúmenes de datos complejos. Esta disciplina, fundamentada en el uso de redes neuronales profundas, permite a los sistemas aprender representaciones jerárquicas de datos, lo cual es particularmente valioso en tareas complejas como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la identificación de patrones en datos no estructurados (LeCun et al., 2015).

A diferencia de los métodos tradicionales de aprendizaje automático, que dependen de características diseñadas manualmente, los modelos de Deep Learning destacan por su capacidad de extraer y combinar características de forma autónoma a través de múltiples capas de procesamiento. Esta habilidad les confiere una ventaja significativa en el análisis de datos complejos y multidimensionales.

Redes Neuronales Convolucionales (CNN)

Dentro del ecosistema del Deep Learning, las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés) han emergido como una arquitectura particularmente eficaz para el análisis de imágenes. Las CNN están diseñadas específicamente para procesar datos con una estructura de cuadrícula, como las imágenes digitales, y son capaces de capturar características espaciales y temporales mediante la aplicación de filtros convolucionales (Krizhevsky et al., 2017).

La arquitectura de una CNN típica incluye varias capas fundamentales:

- **Capa Convolucional:** Es el núcleo de la CNN, donde se aplican filtros o kernels a la imagen de entrada para generar mapas de características. Estos filtros pueden detectar patrones como bordes, texturas y formas, ajustándose durante el entrenamiento para optimizar la extracción de características relevantes (He et al., 2016).

- **Capa de Pooling:** Situada generalmente después de la capa convolucional, la capa de pooling reduce la dimensionalidad de los mapas de características, preservando la información más relevante. Este proceso no solo disminuye la carga computacional, sino que también contribuye a hacer el modelo más robusto frente a variaciones de escala y traslación (Scherer et al., 2010).
- **Capa de Activación:** Las funciones de activación, como la Unidad Lineal Rectificada (ReLU), se aplican para introducir no linealidad en el modelo. Esto permite a la red aprender y modelar relaciones complejas en los datos (Nair & Hinton, 2010).
- **Capa Fully Connected:** En las etapas finales de la red, una o más capas completamente conectadas integran las características extraídas para realizar tareas de clasificación o regresión (Simonyan & Zisserman, 2014).

Aplicación de CNN en el Etiquetado de Imágenes

Las CNN han demostrado ser excepcionalmente efectivas en tareas de etiquetado de imágenes para algoritmos de Deep Learning. Su capacidad para aprender y reconocer patrones complejos en datos visuales las hace ideales para identificar objetos, segmentar regiones específicas y clasificar elementos dentro de una imagen con alta precisión.

El proceso de entrenamiento de una CNN para el etiquetado de imágenes implica el uso de grandes conjuntos de datos de imágenes etiquetadas. Estas etiquetas permiten a la red aprender las características visuales asociadas con cada clase o categoría.

Durante el entrenamiento, la red ajusta iterativamente sus filtros y pesos para minimizar la diferencia entre sus predicciones y las etiquetas reales, mejorando así su precisión en tareas de clasificación (Goodfellow et al., 2016).

La aplicación de CNN en el etiquetado de imágenes ha revolucionado diversos campos, incluyendo:

- **Análisis de imágenes médicas:** Mejorando la precisión en el diagnóstico y la detección temprana de enfermedades (Litjens et al., 2017).

- Visión por computadora en vehículos autónomos: Facilitando la detección de objetos y la navegación en tiempo real (Bojarski et al., 2016).
- Inspección de calidad en manufactura: Automatizando y mejorando la precisión en la detección de defectos (Ferguson et al., 2018).

Técnicas de Procesamiento de Imágenes Aplicadas al Etiquetado

En el contexto de la detección de morfologías cristalinas, el procesamiento de imágenes desempeña un papel crucial en la optimización de los datos de entrenamiento para modelos de Deep Learning. Estas técnicas permiten refinar las imágenes, reducir el ruido y resaltar características esenciales, facilitando un etiquetado más preciso y mejorando la capacidad del modelo para identificar patrones relevantes. A continuación, se describen las técnicas más significativas aplicadas al etiquetado de imágenes en este ámbito:

Mejora de la Calidad de las Imágenes

- a. Normalización de Intensidad:** Este proceso ajusta el rango de intensidades de los píxeles para mejorar la consistencia en la iluminación y mitigar el impacto de las variaciones de luz. La normalización es fundamental para asegurar que las diferencias en la iluminación no comprometan la capacidad del modelo para aprender características clave (Gonzalez & Woods, 2018).
- b. Corrección de Color:** Esta técnica se emplea para eliminar desviaciones no deseadas en la reproducción del color, asegurando la coherencia en la percepción visual. La corrección de color facilita la identificación de características relevantes en las morfologías cristalinas al garantizar que los colores representen fielmente las propiedades físicas de los objetos (Burger & Burge, 2016).

Segmentación y Detección de Objetos

- a. Umbralización:** La umbralización es una técnica que divide la imagen en regiones binarias basadas en un valor umbral. Esta segmentación binaria es útil para identificar áreas de interés, como los contornos de las morfologías cristalinas, facilitando la posterior etiquetación (Otsu, 1979).

b. Segmentación Semántica: Este método asigna etiquetas a cada píxel de la imagen, permitiendo distinguir entre diferentes clases de objetos dentro de la misma imagen. La segmentación semántica es crucial en la delimitación precisa de morfologías cristalinas, ya que permite una clasificación detallada a nivel de píxel (Long et al., 2015).

Extracción de Características Relevantes

a. Filtros Convolucionales: Estos filtros aplican operaciones de convolución sobre la imagen para resaltar características específicas, como bordes o texturas. La aplicación de estos filtros mejora la capacidad del modelo para identificar detalles distintivos en las morfologías cristalinas, ayudando a diferenciar entre diferentes tipos de cristales (Goodfellow et al., 2016).

b. Transformada de Fourier: Esta técnica permite analizar la frecuencia de patrones dentro de una imagen, siendo particularmente útil para identificar características repetitivas en los cristales. La Transformada de Fourier es valiosa en la detección de patrones periódicos y en la identificación de anomalías en la estructura cristalina (Brigham & Morrow, 1967).

Eliminación de Ruido

a. Filtrado Gaussiano o Mediano: Estos métodos se emplean para reducir el ruido presente en las imágenes, lo cual es esencial para evitar distracciones en el proceso de etiquetado. La eliminación de ruido asegura que el modelo se enfoque en las características relevantes y no en artefactos aleatorios que puedan estar presentes en la imagen (Gonzalez & Woods, 2018).

Normalización y Escalamiento

a. Normalización Z-score: Este proceso ajusta los valores de los píxeles para que tengan una media de cero y una desviación estándar de uno, lo que mejora la convergencia del modelo durante el entrenamiento. La normalización Z-score es particularmente útil

para asegurar que las variaciones de escala en los valores de píxeles no afecten negativamente el aprendizaje del modelo (Bishop, 2006).

b. Escalamiento de Imágenes: El escalamiento ajusta las dimensiones de las imágenes para garantizar la consistencia en el tamaño, lo que facilita la comparación y la detección de patrones independientemente de las dimensiones originales. Esto es crucial cuando se trabaja con conjuntos de datos que contienen imágenes de diferentes resoluciones y tamaños (Szeliski, 2010). La implementación adecuada de estas técnicas de procesamiento de imágenes es fundamental para optimizar la calidad y la información contenida en las imágenes etiquetadas. Su correcta aplicación contribuye a un etiquetado más preciso y, por consiguiente, a un mejor rendimiento de los modelos de Deep Learning en la detección de morfologías cristalinas. Asegurar la implementación apropiada de estas técnicas es esencial para obtener resultados confiables y robustos en proyectos de visión por computadora enfocados en el análisis de estructuras cristalinas.

Descripción Del Desarrollo

A continuación, se ofrece una guía detallada sobre los pasos y procesos fundamentales involucrados en la implementación del etiquetado de imágenes para algoritmos de Deep Learning, con un enfoque particular en el análisis de cristales de azúcar. Este capítulo es esencial para comprender cómo se estructuran y ejecutan las distintas fases de un proyecto de etiquetado, desde la preparación de los datos de entrenamiento hasta la creación y gestión de proyectos de anotación manual.

El desarrollo de un proyecto de Deep Learning exitoso depende en gran medida de la calidad de los datos y la precisión del etiquetado. Por ello, este capítulo explora cada aspecto crítico del proceso, comenzando con la preparación y limpieza de las imágenes, que son la base del conjunto de datos utilizado para entrenar el modelo.

A continuación, se describe el uso de la herramienta VGG Annotator para el etiquetado manual de imágenes, proporcionando instrucciones paso a paso para asegurar que cada imagen sea anotada de manera consistente y precisa.

Además, este capítulo incluye un estudio de caso específico centrado en la detección de morfologías cristalinas en el contexto de la industria azucarera. A través de este estudio de caso, se ejemplifica la aplicación práctica de las técnicas y metodologías discutidas, demostrando su eficacia en un entorno industrial real.

Preparación de los Datos de Entrenamiento

La preparación de los datos de entrenamiento es un paso crítico para garantizar el éxito de cualquier modelo de Deep Learning, ya que la calidad y estructura de los datos influyen directamente el rendimiento del modelo. Este proceso involucra la adquisición, selección, limpieza y mejora de las imágenes que serán utilizadas para entrenar algoritmos de visión artificial, asegurando que el conjunto de datos sea representativo, coherente y óptimo para el aprendizaje automático.

Recopilación y Selección de Imágenes de Cristales de Azúcar

La base de un modelo robusto es la calidad de los datos que lo alimentan. En el contexto de la detección de morfologías cristalinas de azúcar, es crucial desarrollar una estrategia eficaz para la recopilación de imágenes. Esta estrategia debe involucrar múltiples fuentes, como imágenes capturadas en laboratorio bajo diferentes condiciones de iluminación y microscopía, así como imágenes obtenidas de bases de datos especializadas. La selección de estas imágenes debe enfocarse en capturar la diversidad dentro del dominio de los cristales de azúcar, asegurando que se abarquen distintas morfologías, tamaños y formaciones cristalinas (Krizhevsky, Sutskever, & Hinton, 2012). La inclusión de una amplia variabilidad en las imágenes ayuda a evitar el sesgo en el entrenamiento y mejora la capacidad del modelo para generalizar a nuevos datos.

Limpieza y Mejora de la Calidad de las Imágenes

Las imágenes en bruto a menudo requieren un procesamiento adicional para maximizar su utilidad en el entrenamiento de modelos de Deep Learning. Este proceso incluye la limpieza de las imágenes para eliminar artefactos y ruido, así como la corrección de problemas de iluminación que podrían introducir sesgos en el modelo.

- **Limpieza de Imágenes:** La limpieza de imágenes es esencial para asegurar que los datos de entrenamiento estén libres de interferencias que puedan afectar negativamente el rendimiento del modelo. Las siguientes técnicas son comúnmente utilizadas:
 - **Eliminación de Ruido:** Se aplican filtros para reducir el ruido presente en las imágenes. Entre los métodos más comunes se encuentran el filtro de mediana, que suaviza el ruido mientras preserva los bordes, y el filtrado bilateral, que elimina el ruido mientras mantiene los detalles de la imagen (González & Woods, 2018).
 - **Eliminación de Artefactos:** Los artefactos, como manchas o distorsiones, pueden ser eliminados mediante técnicas de interpolación y restauración de imágenes. Estos métodos ayudan a mejorar la claridad de la imagen y a preservar las características relevantes para el análisis de los cristales de azúcar (Jain, 1989).
 - **Normalización de Color:** La normalización del color ajusta la intensidad y el equilibrio de color de las imágenes para lograr una representación uniforme. Esto es importante para garantizar que las variaciones en la captura de imágenes no afecten el entrenamiento del modelo (Huang et al., 2016).
 - **Corrección de Iluminación.** La corrección de la iluminación es un paso crucial para asegurar que las variaciones en la iluminación no introduzcan sesgos en el modelo de Deep Learning. Las siguientes técnicas son clave para este propósito:
 - **Ecualización de Histograma:** Esta técnica ajusta el histograma de la imagen para distribuir de manera uniforme las intensidades de los píxeles, lo que mejora el contraste y la visibilidad de los detalles importantes en los cristales de azúcar (Szeliski, 2010).

- Corrección Gamma: La corrección gamma modifica la curva de gamma de la imagen, ajustando la relación de contraste. Esta técnica es especialmente útil para imágenes que presentan variaciones significativas en la iluminación, ayudando a mejorar la consistencia visual en el conjunto de datos (Gonzalez & Woods, 2018).
- Compensación de Iluminación Local: Técnicas como el filtrado de mediana adaptativo se utilizan para compensar las variaciones locales en la iluminación, asegurando que los detalles importantes se preserven en todas las regiones de la imagen (Jain, 1989).

Etiquetado Manual de las Imágenes con VGG Annotator

El VGG Image Annotator (VIA) es una herramienta ligera y versátil para la anotación manual de imágenes, audio y video. Desarrollado por el Visual Geometry Group (VGG) de la Universidad de Oxford, VIA se ha convertido en una herramienta esencial para la creación de conjuntos de datos etiquetados que son fundamentales en el entrenamiento de modelos de Deep Learning. Su simplicidad y eficacia radican en su diseño independiente, ya que se ejecuta completamente en un navegador web sin necesidad de instalación o configuración adicional, y su tamaño es inferior a 400 kilobytes (Dutta & Zisserman, 2019).

El desarrollo del proceso de etiquetado manual de imágenes para el análisis de cristales de azúcar involucró una colaboración multidisciplinaria y educativa que integró a aprendices de diferentes programas de formación del Centro de Electricidad y Automatización Industrial.

Estos aprendices, vinculados al semillero de investigación de Inteligencia Artificial, desempeñaron un papel crucial en la ejecución del proyecto, aportando tanto a la realización técnica como al aprendizaje práctico en un entorno de investigación aplicada.

Antes de iniciar el proceso de etiquetado, los aprendices recibieron capacitación específica en el uso de VGG Image Annotator (VIA), utilizada para la anotación manual de imágenes. La capacitación incluyó:

- **Introducción a VGG Annotator:** Los aprendices se familiarizaron con la interfaz de usuario de VIA, aprendiendo a manejar las diferentes funciones y herramientas disponibles para realizar un etiquetado preciso y consistente.
- **Definición de Atributos y Parámetros de Anotación:** Se instruyó a los aprendices en la configuración de atributos relevantes para el análisis de cristales de azúcar, como la forma y el tipo de cristal, asegurando que las etiquetas fueran coherentes y útiles para el entrenamiento del modelo.
- **Práctica de Etiquetado:** Los aprendices realizaron sesiones prácticas de etiquetado, donde aprendieron a dibujar regiones utilizando la herramienta de polígono, asignar etiquetas a los cristales según su morfología, y guardar los proyectos de manera segura.

Durante este proceso, los aprendices adquirieron habilidades clave en la gestión de proyectos de etiquetado, comprendiendo la importancia de la precisión y consistencia en el etiquetado manual para la calidad del modelo de Deep Learning.

El proceso de etiquetado fue supervisado por instructores y coordinadores del semillero de investigación, quienes aseguraron que las prácticas de etiquetado se alinearan con los estándares del proyecto. Se implementaron revisiones periódicas para evaluar la precisión de las etiquetas y la consistencia entre los diferentes anotadores. Esto incluyó la validación de las etiquetas y la retroalimentación continua para mejorar el proceso.

Instalación y Configuración de VGG Annotator

Para comenzar a utilizar VIA, se puede acceder directamente a través del enlace web proporcionado por los desarrolladores: VGG Annotator en línea. Alternativamente, los usuarios pueden optar por descargar la versión sin conexión desde VGG Annotator descarga. Este archivo ZIP contiene todo lo necesario para ejecutar VIA en un entorno local, descomprimiendo el archivo y ejecutando el archivo `via.html` en cualquier navegador web moderno como Firefox, Chrome o Safari.

Figura 2.

Contenido del archivo comprimido via-2.0.12.zip e identificación del archivo ejecutable

Nombre	Fecha de modificación	Tipo	Tamaño
CHANGELOG	26/09/2022 6:52 a. m.	Archivo	12 KB
Contributors.md	26/09/2022 6:52 a. m.	Archivo MD	6 KB
LICENSE	26/09/2022 6:52 a. m.	Archivo	2 KB
README.md	26/09/2022 6:52 a. m.	Archivo MD	8 KB
via.html	26/09/2022 6:52 a. m.	Chrome HTML Do...	415 KB
via_demo.html	26/09/2022 6:52 a. m.	Chrome HTML Do...	613 KB

VIA es totalmente independiente, basado únicamente en HTML, CSS y JavaScript, sin la necesidad de bibliotecas externas. Soporta varias formas de regiones para la anotación, incluyendo rectángulos, círculos, elipses, polígonos, puntos y polilíneas, y permite importar y exportar datos de anotación en formatos CSV y JSON.

Creación y Gestión de Proyectos de Etiquetado

El proceso de creación y gestión de proyectos de etiquetado con VIA es esencial para estructurar y optimizar la eficiencia del etiquetado manual de imágenes. A continuación, se detalla el proceso paso a paso:

- **Inicio de un Nuevo Proyecto:** Para iniciar un proyecto de etiquetado, seleccione “Project” en el menú principal de VIA y elija “New Project”. Defina los parámetros del proyecto, incluyendo la ubicación de los archivos y la selección de imágenes base. Es importante crear categorías de etiquetas específicas que sean relevantes para el análisis de cristales de azúcar, lo que facilitará el etiquetado coherente a lo largo del proyecto (Henriques et al., 2018).
- **Interfaz de Usuario de VGG Annotator:** VIA presenta una interfaz intuitiva que facilita el marcado preciso y consistente de objetos en las imágenes. La barra de herramientas ofrece acceso rápido a las funciones de anotación, mientras que la ventana principal permite una visualización clara de las imágenes y las regiones anotadas.

- **Establecimiento de Parámetros de Anotación:** Configure los parámetros de anotación según las necesidades del proyecto. Esto incluye seleccionar el tipo de anotación, como cajas delimitadoras o polígonos, y establecer cómo se registrarán las coordenadas de las etiquetas. Por ejemplo, al etiquetar cristales de azúcar, se puede optar por utilizar la herramienta de polígono para trazar con precisión los contornos de cada cristal.
- **Gestión Eficiente de Conjuntos de Datos:** Es fundamental organizar y gestionar de manera eficiente los conjuntos de datos a medida que se etiquetan. VIA permite la exportación de datos de regiones en formatos CSV y JSON, lo que facilita la integración con otros sistemas y el manejo de grandes volúmenes de datos para futuras iteraciones de etiquetado (Dutta, Gupta & Zisserman, 2016).

Caso De Estudio: Etiquetado De Cristales De Azúcar

Contexto

En la industria azucarera, la calidad y consistencia del producto final dependen en gran medida del control preciso de los procesos de cristalización. Durante la producción de azúcar, las características morfológicas de los cristales, como su tamaño y forma, son indicadores críticos que afectan tanto la eficiencia del proceso como la calidad del azúcar producido. Sin embargo, la identificación y clasificación precisas de estas morfologías representan un desafío significativo, especialmente cuando se requiere un análisis rápido y consistente a gran escala.

El proyecto de investigación “Evaluación de técnicas de Visión Artificial basadas en Deep Learning para detección de morfologías en sustancias cristalinas de origen vegetal – SENNOVA 2023” fue desarrollado para abordar este desafío, aplicando técnicas avanzadas de visión artificial y Deep Learning en la detección y análisis de cristales de azúcar. Este proyecto, en colaboración con dos ingenios azucareros del Valle del Cauca y el Centro de Investigación de la Caña de Azúcar de Colombia (CENICAÑA), se centró en la creación de un conjunto de datos robusto y representativo, así como en el etiquetado manual de imágenes de cristales de azúcar,

con el objetivo de entrenar modelos de Deep Learning capaces de mejorar la precisión y eficiencia en la clasificación de estos cristales.

Propósito del Caso de Estudio

El propósito de este caso de estudio es demostrar la aplicación práctica de las técnicas de etiquetado de imágenes descritas en este manual, específicamente en el contexto del análisis de cristales de azúcar. A través de este caso de estudio, se pretende:

- **Validar la Eficacia del Etiquetado Manual:** Mostrar cómo un etiquetado preciso y bien estructurado puede mejorar significativamente la capacidad de los modelos de Deep Learning para identificar y clasificar morfologías cristalinas en entornos industriales.
- **Ilustrar el Proceso de Implementación:** Proporcionar una guía detallada sobre cómo se llevó a cabo el proceso de etiquetado manual utilizando VGG Annotator, desde la preparación de los datos hasta la configuración de los parámetros de anotación, con el fin de que otros profesionales puedan replicar este enfoque en sus propios proyectos.
- **Evaluar el Impacto en la Industria:** Analizar cómo los resultados obtenidos mediante este enfoque pueden aplicarse para mejorar la eficiencia y calidad en la producción de azúcar, ofreciendo un ejemplo claro de cómo las técnicas de Deep Learning pueden integrarse en procesos industriales.
- **Proveer un Modelo de Referencia:** Establecer un modelo de referencia que pueda ser utilizado para futuras investigaciones o proyectos que involucren la detección de morfologías cristalinas, no solo en la industria azucarera, sino también en otros campos donde el control de calidad y la precisión en la clasificación de partículas son esenciales.

Descripción del conjunto de datos

El conjunto de datos utilizado en este manual es fundamental para garantizar la efectividad y robustez de los

algoritmos de Deep Learning en el análisis de cristales de azúcar. Este conjunto de datos se compone de imágenes proporcionadas por dos ingenios azucareros del Valle del Cauca, junto con el Centro de Investigación de la Caña de Azúcar de Colombia (CENICAÑA). Las imágenes fueron capturadas durante las etapas críticas del proceso de cristalización en las masas A y B, momentos clave en la producción de azúcar donde la formación y evolución de los cristales son monitoreadas de cerca.

El proyecto de investigación “Evaluación de técnicas de Visión Artificial basadas en Deep Learning para detección de morfologías en sustancias cristalinas de origen vegetal – SENNOVA 2023” utilizó este conjunto de datos para desarrollar y entrenar modelos que pudieran identificar y clasificar diferentes morfologías de cristales de azúcar con alta precisión. Las imágenes capturadas reflejan una amplia variedad de condiciones de cristalización, lo que garantiza que el conjunto de datos sea representativo y permita al modelo generalizar de manera efectiva a nuevas imágenes (Ramírez et al., 2020).

El conjunto de datos incluye imágenes de alta resolución que capturan detalles finos en la estructura de los cristales, lo cual es crucial para el entrenamiento de modelos de Deep Learning en tareas de clasificación y detección. Además, la diversidad de imágenes en condiciones de iluminación, tamaño de cristales y etapas de desarrollo asegura que los modelos pueden manejar variaciones reales en un entorno de producción industrial.

Este enfoque colaborativo, que involucra tanto a la industria azucarera como a centros de investigación, resalta la importancia de contar con datos de alta calidad y representativos para el desarrollo de soluciones tecnológicas avanzadas en la industria azucarera. La combinación de expertise industrial y científico en la creación de este conjunto de datos establece un estándar elevado para futuros proyectos en el campo de la visión artificial aplicada a la agricultura y la manufactura (Singh et al., 2019).

Procedimiento de Etiquetado de Imágenes de Cristales de Azúcar

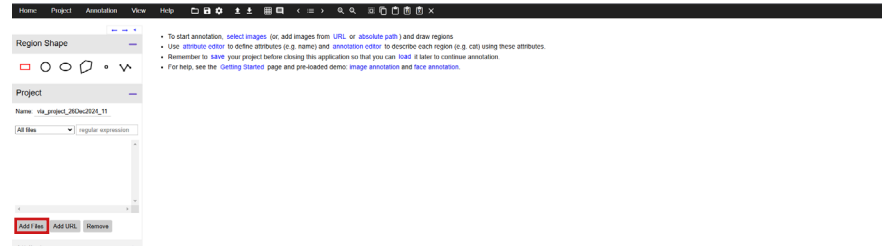
El procedimiento de etiquetado es una etapa fundamental en el desarrollo de modelos de Deep Learning, especialmente en proyectos que involucran la detección y clasificación de morfologías cristalinas en imágenes. Un etiquetado preciso y coherente de las imágenes es crucial para entrenar modelos que puedan identificar patrones complejos y realizar segmentaciones exactas en entornos industriales.

En el contexto del análisis de cristales de azúcar, el procedimiento de etiquetado no solo define cómo se estructuran los datos de entrada, sino que también impacta directamente en la eficacia y precisión del modelo final.

Este apartado detalla el proceso sistemático de etiquetado manual llevado a cabo utilizando la herramienta VGG Annotator, abarcando desde la carga inicial de imágenes hasta la configuración de parámetros específicos y la aplicación de etiquetas. Cada paso ha sido diseñado para maximizar la calidad del conjunto de datos y garantizar que los modelos de Deep Learning sean entrenados con la mayor precisión posible. Además, se describen las mejores prácticas y consideraciones clave que se deben tener en cuenta durante el etiquetado, asegurando la consistencia y fiabilidad de los resultados obtenidos.

El etiquetado de imágenes en VIA se realiza en varios pasos clave:

Paso 1. Cargar Imágenes: Comience cargando todas las imágenes que se desean anotar. Esto se puede hacer seleccionando archivos locales a través del menú “Project” y eligiendo “Add local files”. Seleccione las imágenes deseadas y haga clic en “Abrir” para cargarlas en el proyecto.

Figura 3.*Opción para cargar las imágenes que se desean anotar.*

Paso 2. Configuración de Atributos: Defina los atributos para las anotaciones, como la “Forma” del cristal. Agregue este atributo en la interfaz de VIA utilizando el botón “+” (figuras 4 y 5) y configure el tipo de atributo como “checkbox” (figura 6). Esto es esencial para categorizar adecuadamente los diferentes tipos de cristales (Henriques et al., 2018).

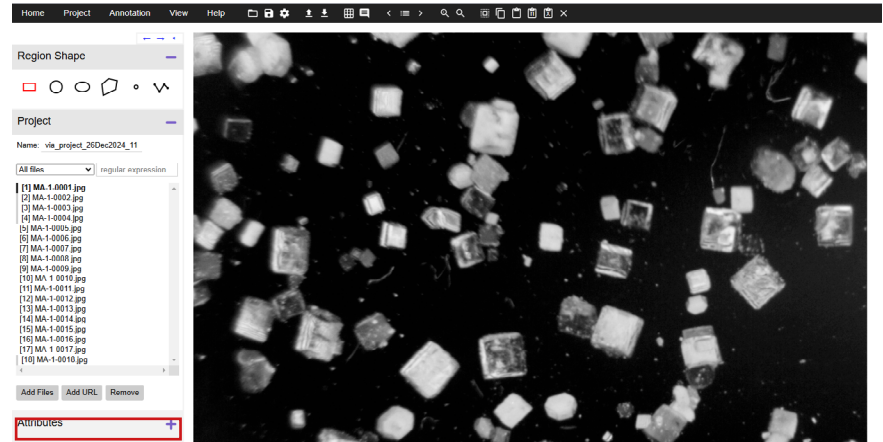
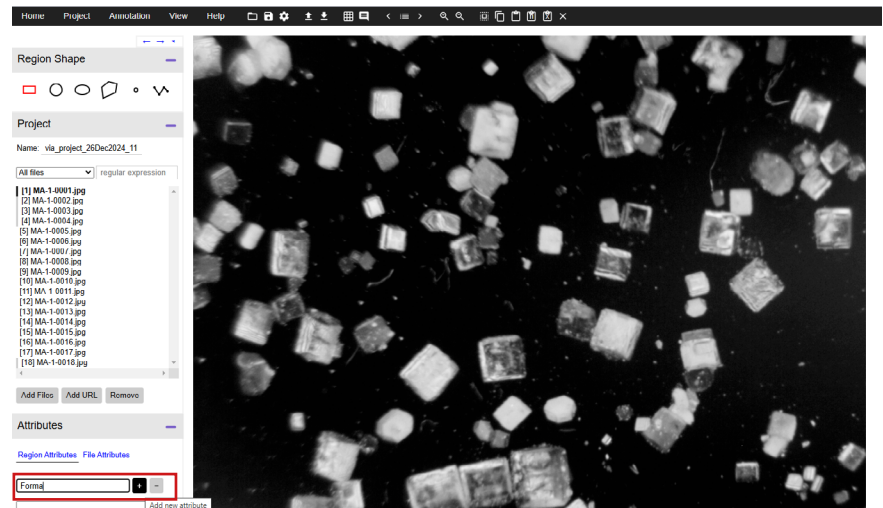
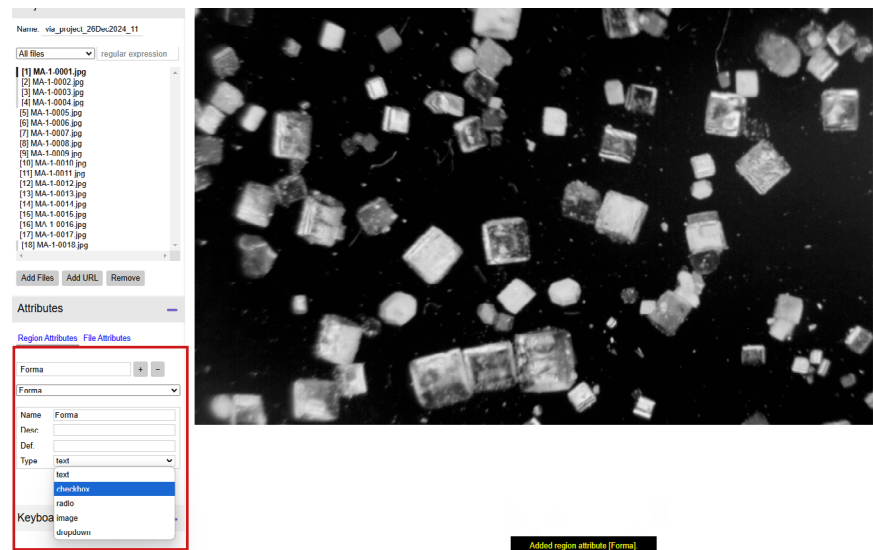
Figura 4.*Despliegue de las opciones de atributos del etiquetado.***Figura 5.***Nombre del atributo.*

Figura 6.

Definición del tipo de atributo creado.

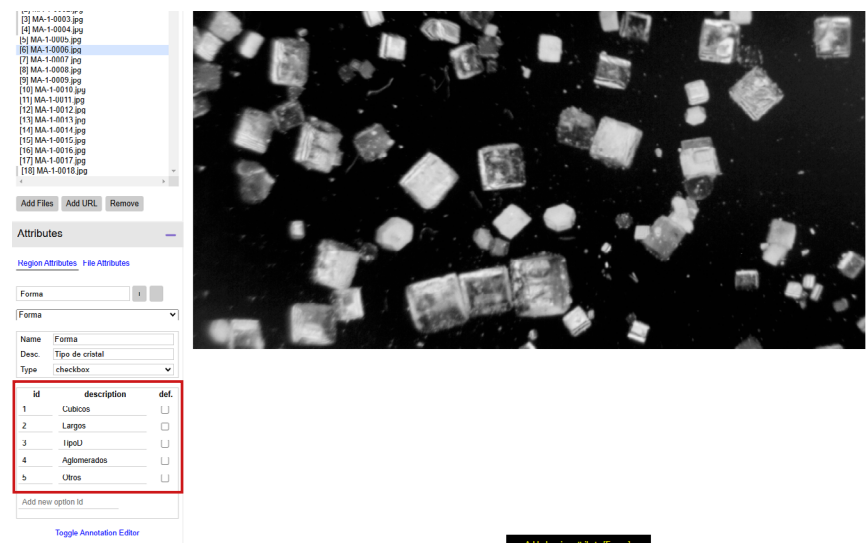


Adicione los id 1, 2, 3, 4 y 5 escribiendo los números en el campo denominado “id” y posteriormente presionando la tecla Enter. Seguidamente podrá adicionar las descripciones de cada uno de los “id”, como se muestra en la figura 7.

En el caso de estudio expuesto en este manual técnico se emplearon las siguientes descripciones para los tipos de cristales identificados: Cúbicos, Largos, TipoD, Aglomerados y Otros.

Figura 7.

Descripción de id en el atributo “Forma” creado.



Paso 3. Dibujar Regiones: Amplíe la imagen según sea necesario para identificar el cristal a etiquetar. Utilice la herramienta “Polygon region shape” (Figura 8), para trazar el contorno del cristal, asegurándose de realizar un trazado preciso que capture las características relevantes (Figura 9). Finalice el dibujo de la región con un doble clic o presionando “Enter”.

Figura 8.

Tipo de forma empleada para definir la región de la etiqueta.

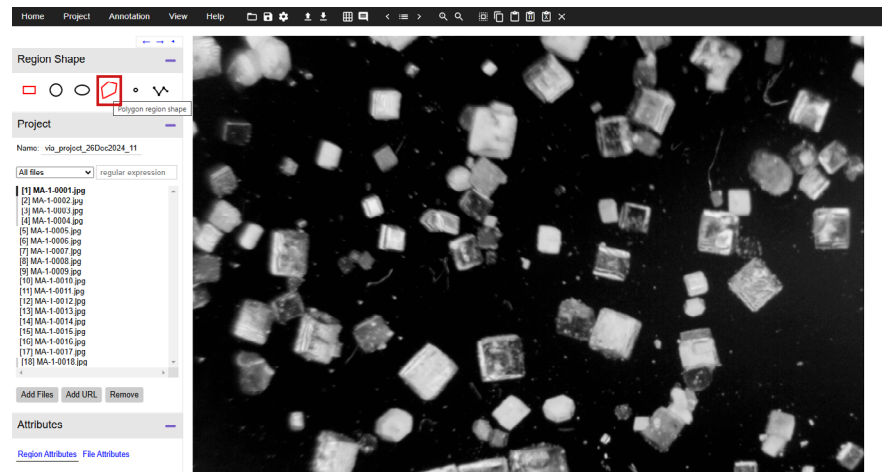
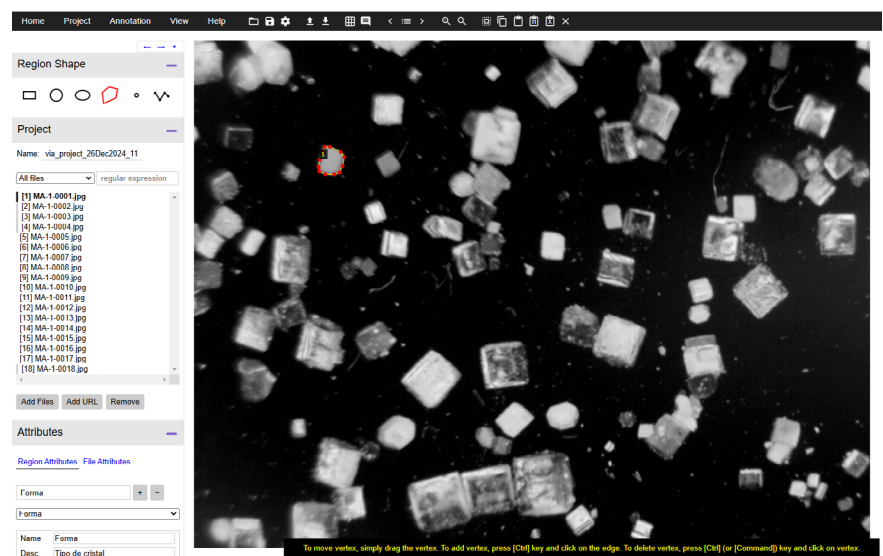


Figura 9.

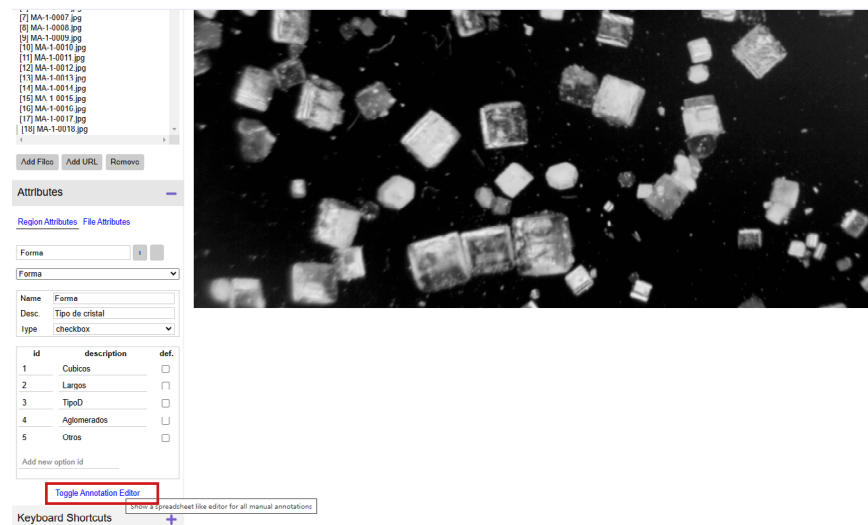
Ejemplo de trazado de región sobre la imagen a etiquetar.



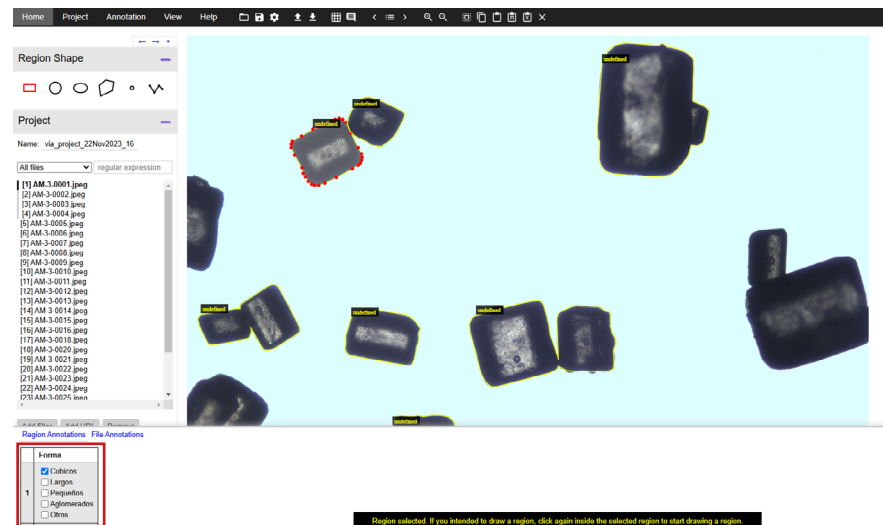
Paso 4. Asignación de Etiquetas: Asigne la etiqueta correspondiente al cristal utilizando la ventana de “Attributes”. Marque el tipo de cristal para cada región creada en la imagen, siguiendo las indicaciones establecidas para el proyecto.

Figura 10.

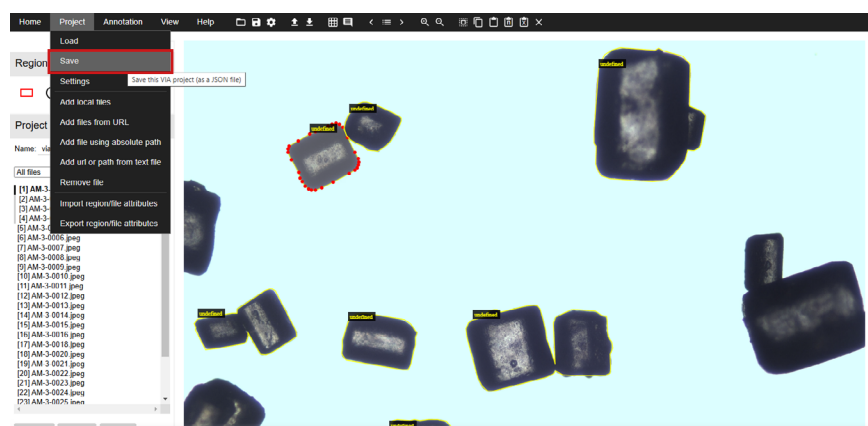
Asignación del tipo de etiqueta a la región creada.

**Figura 11.**

Marcación del tipo de cristal sobre la región creada.



Paso 5. Guardar el Proyecto: Guarde el proyecto regularmente para evitar la pérdida de datos. El proyecto se puede guardar como un archivo .json, eligiendo la opción “Save” en el menú “Project” y verificando que el nombre del proyecto sea correcto.

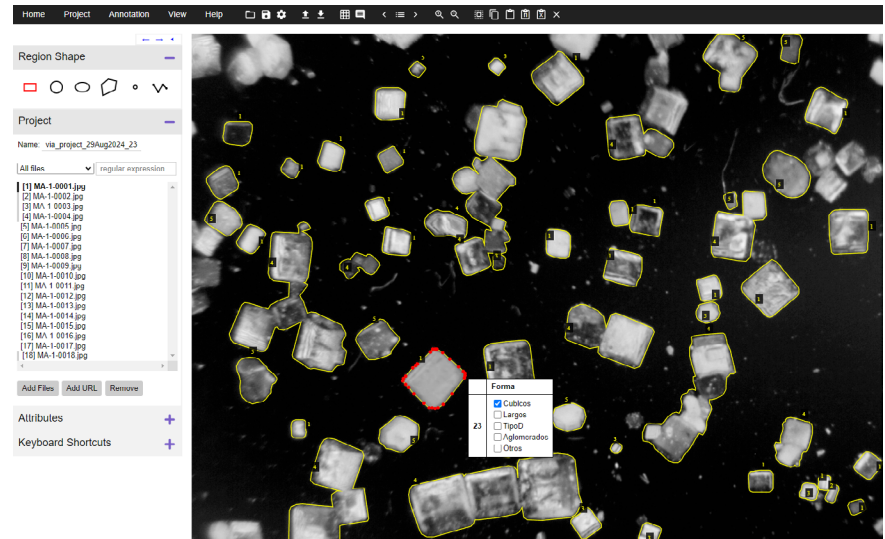
Figura 12.*Menú para el guardado del proyecto.*

Descripción de las Categorías de Etiquetas Utilizadas en el Etiquetado Manual de las Imágenes

El proceso de etiquetado manual de imágenes para el análisis de cristales de azúcar requiere la definición de categorías de etiquetas que representen las diversas morfologías cristalinas observadas durante la cristalización. Estas categorías son esenciales para entrenar algoritmos de Deep Learning que puedan identificar y clasificar con precisión diferentes tipos de cristales. A continuación, se describen las principales categorías de etiquetas utilizadas en este proceso:

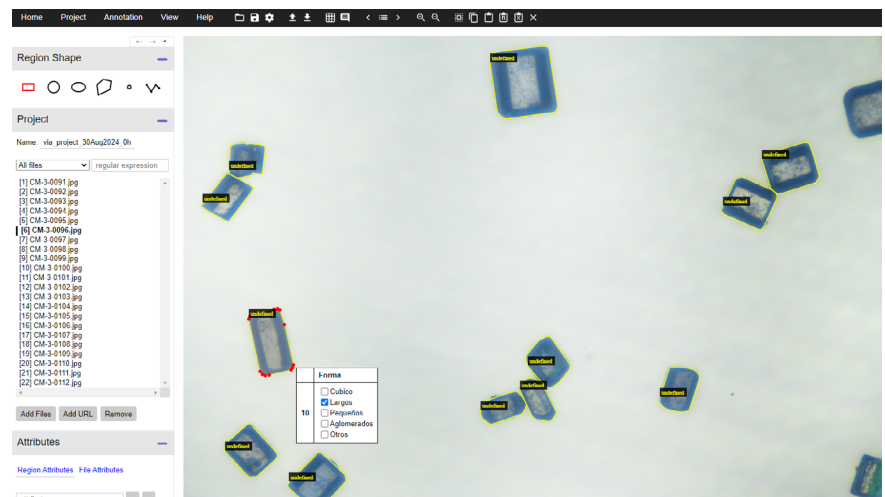
- 1. Cristales Cúbicos.** Los cristales cúbicos son cristales que presentan una forma casi perfecta de cubo, con caras planas y ángulos rectos. Estos cristales son típicos en ciertos procesos de cristalización controlada y son altamente deseados en la producción de azúcar de alta calidad.

La identificación precisa de cristales cúbicos es crucial, ya que su forma regular es un indicador de un proceso de cristalización óptimo. Etiquetar estos cristales permite al modelo de Deep Learning aprender a distinguirlos de otras morfologías menos regulares.

Figura 13.*Imagen con un cristal cúbico seleccionado.*

2. Cristales Largos. Los cristales largos son aquellos que tienen una forma alargada o rectangular, con una longitud significativamente mayor que su ancho. Este tipo de cristal puede formarse debido a variaciones en las condiciones de cristalización, como la concentración de la solución o la temperatura.

Etiquetar cristales largos es importante para analizar las condiciones de cristalización que los producen y ajustar el proceso para evitar su formación en exceso, ya que pueden afectar la uniformidad del producto final.

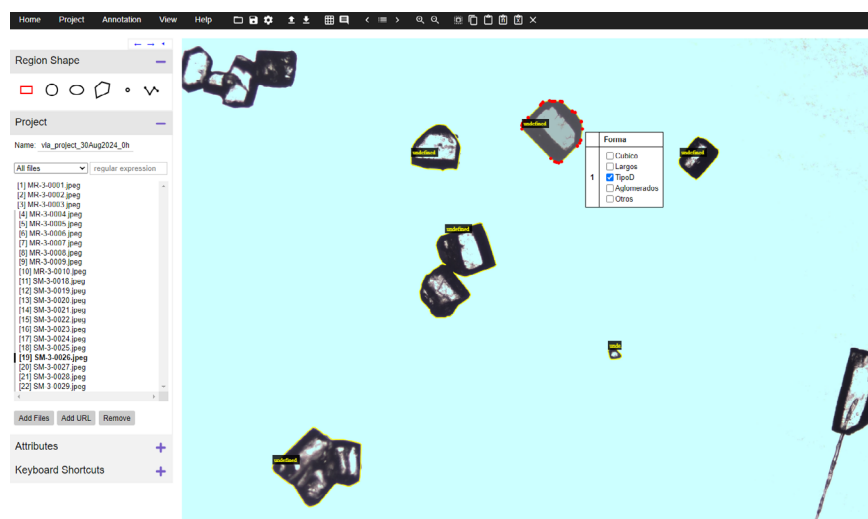
Figura 14.*Imagen con un cristal alargado seleccionado.*

3. Cristales Tipo D. Los cristales Tipo D tienen una morfología más compleja, con formas irregulares y bordes menos definidos. Este tipo de cristal puede ser indicativo de inestabilidades en el proceso de cristalización.

La identificación de cristales Tipo D es esencial para diagnosticar y corregir posibles problemas en el proceso de cristalización. Su presencia en grandes cantidades puede señalar la necesidad de ajustes en los parámetros de producción.

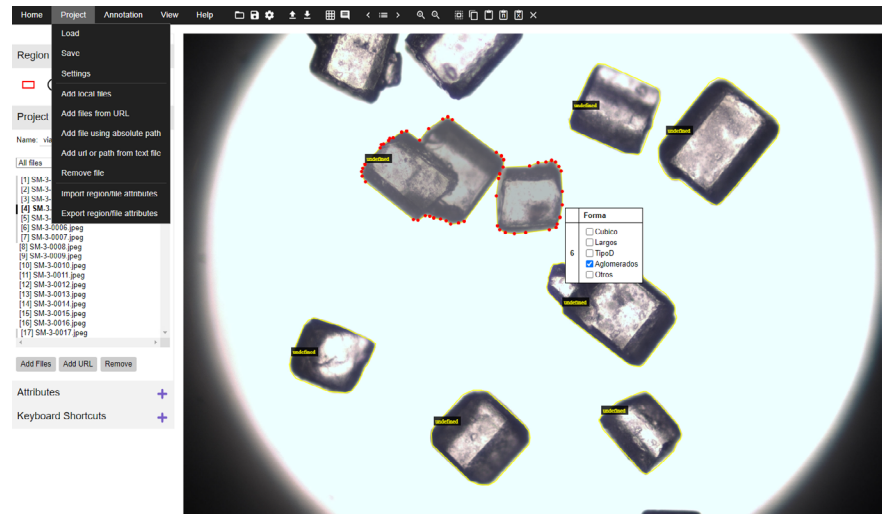
Figura 15.

Imagen con un cristal Tipo D seleccionado.



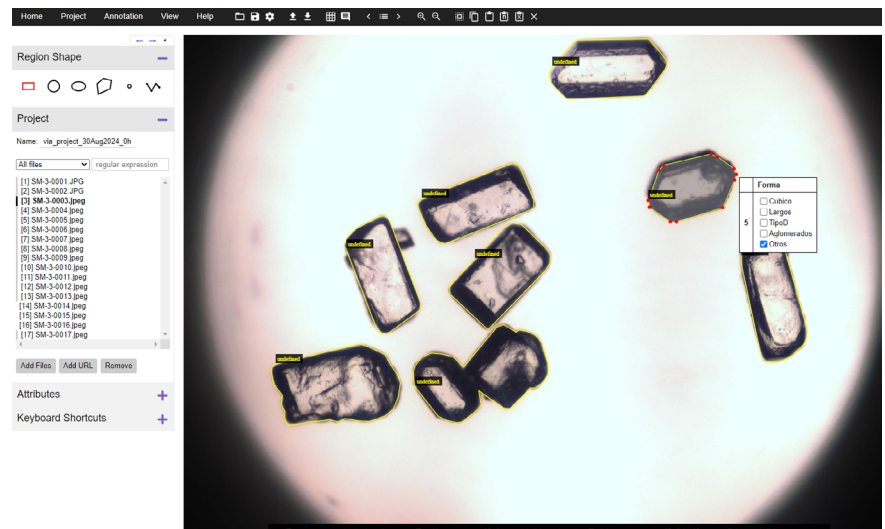
4. Cristales Aglomerados. Los cristales aglomerados son grupos de cristales que se han adherido entre sí, formando estructuras más grandes y complejas. Esta aglomeración puede ocurrir cuando los cristales se forman demasiado rápido o en condiciones de alta saturación.

Etiquetar cristales aglomerados ayuda a identificar fases del proceso donde la cristalización no está controlada adecuadamente. La formación de estos aglomerados puede afectar negativamente la calidad del azúcar y requiere atención especial.

Figura 16.*Imagen con cristales aglomerados seleccionados.*

- 5. Otros.** Esta categoría incluye cristales que no se ajustan a ninguna de las mencionadas. Puede abarcar morfologías atípicas o cristales que presentan características mixtas o indeterminadas.

La categoría “Otros” es útil para capturar la diversidad completa de morfologías cristalinas presentes en las imágenes. Permite al modelo de Deep Learning manejar casos atípicos y aprender a clasificar cristales que no encajan claramente en las demás categorías.

Figura 17.*Imagen con cristales tipo “Otros” seleccionados.*

Estrategias de etiquetado específicas para Cristales de Azúcar

El proceso de etiquetado para las imágenes de cristales de azúcar se llevó a cabo de manera meticulosa y estructurada, abordando las particularidades del proceso de etiquetado. Este apartado detalla las estrategias implementadas, comenzando con la estandarización del nombre de las imágenes y avanzando hasta la consolidación del banco de datos etiquetado necesario para el entrenamiento del algoritmo de segmentación.

Estandarización del nombre de imágenes: Se inició el proceso con la estandarización del nombre de las imágenes suministradas por los aliados del proyecto. Un código único fue creado, incorporando dos letras iniciales (MG, BL, AM, MR, MA, MB) para identificar el tipo de cristal al que correspondía la imagen, proporcionando así un marco de referencia claro y consistente.

Asignación de paquetes de imágenes: Las imágenes estandarizadas se organizaron en paquetes y se asignaron a 10 aprendices participantes en el proyecto. Cada aprendiz asumió la responsabilidad de etiquetar manualmente las regiones de interés en las imágenes utilizando la herramienta VGG Annotator. Esta estrategia permitió distribuir eficientemente la carga de trabajo y acelerar el proceso de etiquetado.

Formato de etiquetado en Archivos .JSON: Cada aprendiz generó un archivo .JSON que contenía la información de las regiones etiquetadas y sus correspondientes etiquetas. Este formato estructurado facilitó la recopilación y organización de los datos de manera coherente, permitiendo una fácil integración en el conjunto de datos general.

Revisión detallada de la calidad del etiquetado: La calidad del etiquetado fue sometida a una revisión detallada. Se implementaron revisiones periódicas para identificar posibles inconsistencias y asegurar que las anotaciones se ajustaran a los estándares de calidad requeridos. Este enfoque de control de calidad garantizó que las etiquetas fueran precisas y coherentes, un requisito esencial para el entrenamiento efectivo de los algoritmos de segmentación de imágenes.

Seguimiento periódico del proceso: A lo largo del proceso de etiquetado, se estableció un seguimiento periódico del proceso de etiquetado. Este seguimiento permitió identificar y corregir problemas de manera oportuna, asegurando una mejora continua en la calidad del etiquetado. La implementación de un sistema de seguimiento regular también facilitó la construcción progresiva de un banco de datos robusto y bien estructurado.

Consolidación del banco de datos etiquetado: La consolidación del banco de datos etiquetado se realizó de manera iterativa, incorporando los resultados de las revisiones y ajustes realizados durante el proceso. Este banco de datos etiquetado constituye la base esencial para el entrenamiento de los algoritmos de segmentación de Deep Learning, garantizando que el modelo pueda aprender de datos de alta calidad y representar con precisión las morfologías cristalinas presentes en las imágenes.

Alternativas para conformar el banco de datos: Para garantizar que el algoritmo de segmentación de imágenes de cristales de azúcar funcione de manera óptima, es fundamental disponer de un banco de datos completo y bien estructurado. Existen varias alternativas para conformar este banco de datos, dependiendo de cómo se almacenen las imágenes y las anotaciones generadas durante el proceso de etiquetado. A continuación, se presentan dos enfoques principales, cada uno con sus ventajas y consideraciones particulares, que permiten optimizar tanto el proceso de almacenamiento como el de entrenamiento del modelo.

- 1. Fotografías y Archivos con Coordenadas JSON Anotadas:** La primera alternativa consiste en utilizar las imágenes originales de los cristales de azúcar, acompañadas de sus respectivos archivos JSON, que contienen las coordenadas y etiquetas de las regiones de interés. Estos archivos JSON proveen la información necesaria para que el algoritmo de segmentación identifique con precisión las morfologías cristalinas presentes en las imágenes. La principal ventaja de esta alternativa es su flexibilidad, ya que las anotaciones pueden ser fácilmente ajustadas o reutilizadas sin modificar las imágenes originales. Al mantener la separación entre

las imágenes y las etiquetas, es más sencillo gestionar la revisión y actualización de las anotaciones. Además, esta opción permite un proceso de administración más eficiente, ya que los archivos JSON son más ligeros y fáciles de manejar que las máscaras en color.

2. Fotografías Originales con Máscaras en Color Generadas a partir de los Archivos JSON Anotado: La segunda alternativa involucra la creación de máscaras en color a partir de las anotaciones realizadas en los archivos JSON. Estas máscaras en color son imágenes donde las distintas regiones de interés (tipos de cristales) se representan mediante colores específicos. Por ejemplo, un tipo de cristal podría estar marcado en rojo, mientras que otro podría estar en azul o verde. Este enfoque es particularmente útil para el entrenamiento de algoritmos de segmentación basados en Deep Learning, ya que estos modelos pueden utilizar las máscaras en color para diferenciar con mayor precisión las distintas morfologías cristalinas dentro de las imágenes. Además, las máscaras en color facilitan la representación visual clara de múltiples clases en una sola imagen, haciendo el proceso de segmentación más eficiente.

Al utilizar máscaras en color, cada píxel de la imagen de salida está etiquetado con un color correspondiente a una clase, lo que permite al algoritmo identificar y aprender de las variaciones morfológicas de los diferentes cristales de azúcar. Esta alternativa también es ventajosa cuando se requiere una segmentación precisa de varias clases dentro de una misma imagen, ya que el color codificado permite una mejor separación de las regiones etiquetadas.

Consideraciones Adicionales:

- **Calidad y Precisión de las Máscaras en Color:** Es fundamental que las máscaras en color estén perfectamente alineadas con las imágenes originales. Cualquier desajuste entre la máscara y la imagen podría introducir errores en el entrenamiento del modelo. Además, es necesario asegurarse de que los colores seleccionados para representar cada tipo de cristal sean claramente distinguibles para el algoritmo y no presenten ambigüedades en la segmentación.

- **Gestión del Almacenamiento:** Las máscaras en color ocupan significativamente más espacio que los archivos JSON, ya que cada imagen debe almacenar información en múltiples canales de color (por ejemplo, en formato RGB). Por tanto, es importante prever un incremento en los requerimientos de almacenamiento y considerar el uso de formatos de compresión eficientes para gestionar el tamaño del banco de datos sin perder la calidad necesaria para el entrenamiento.
- **Flexibilidad y Uso Futuro:** Aunque las máscaras en color pueden facilitar el proceso de segmentación, también es recomendable conservar las imágenes originales junto con los archivos .JSON anotados. Esto permitirá realizar ajustes futuros en las anotaciones o probar nuevas técnicas de segmentación sin la necesidad de volver a generar las máscaras, manteniendo así la flexibilidad para optimizar el modelo en futuras iteraciones.

Aplicación del Modelo de Deep Learning

En el desarrollo del proyecto “Evaluación de técnicas de Visión Artificial basadas en Deep Learning para detección de morfologías en sustancias cristalinas de origen vegetal”, se implementaron dos modelos avanzados de Deep Learning: U-Net y Mask R-CNN. Estos modelos fueron seleccionados debido a su eficacia probada en tareas de segmentación y detección de objetos en imágenes complejas, siendo fundamentales para mejorar la precisión en la identificación y clasificación de cristales de azúcar.

Modelo U-Net: El modelo U-Net, propuesto originalmente por Ronneberger et al. (2015), se utilizó para la segmentación de cristales de azúcar debido a su capacidad para capturar tanto características locales como globales de las imágenes, lo que es crucial para delinear con precisión los contornos de los cristales, incluso en morfologías irregulares. U-Net fue entrenado utilizando un conjunto de 1200 imágenes de cristales de azúcar, con una división del conjunto de datos en 70% para entrenamiento, 10% para validación y 20% para pruebas.

Durante el proceso de entrenamiento, que se llevó a

cabo durante 20 épocas, se observó una mejora constante en las métricas de rendimiento del modelo. El U-Net alcanzó una precisión del 98.74% en el conjunto de entrenamiento y del 98.64% en el conjunto de validación, demostrando una capacidad sobresaliente para segmentar correctamente las áreas de interés en las imágenes. Además, la métrica de Intersección sobre Unión (IoU), que mide la superposición entre las áreas segmentadas y las áreas de referencia, se incrementó significativamente hasta alcanzar un valor de 0.8240, lo que refleja la precisión del modelo en la segmentación de cristales de azúcar.

Modelo Mask R-CNN: Mask R-CNN, desarrollado por He et al. (2017), fue implementado para la detección y clasificación de cinco morfologías distintas de cristales de azúcar, aprovechando su capacidad para realizar tanto la segmentación de instancias como la clasificación en un solo modelo. Inicialmente, el modelo fue entrenado con 244 imágenes etiquetadas manualmente. No obstante, durante las primeras 100 épocas de entrenamiento, se observó un sobreajuste, con una pérdida de 0.1684 en el conjunto de entrenamiento y de 0.6534 en el conjunto de validación, lo que indicaba la necesidad de un conjunto de datos más grande y diversificado para mejorar la capacidad de generalización del modelo.

Para abordar este problema, se generaron 976 imágenes sintéticas utilizando técnicas de aumento de datos, lo que permitió reducir la pérdida de validación a 0.5342 y mejorar la capacidad del modelo para generalizar a nuevas imágenes. A pesar de estas mejoras, el modelo Mask R-CNN mostró limitaciones cuando se aplicó a cristales distintos a los de azúcar, como los cristales de CBD, lo que subraya la importancia de adaptar el modelo a las características específicas del conjunto de datos.

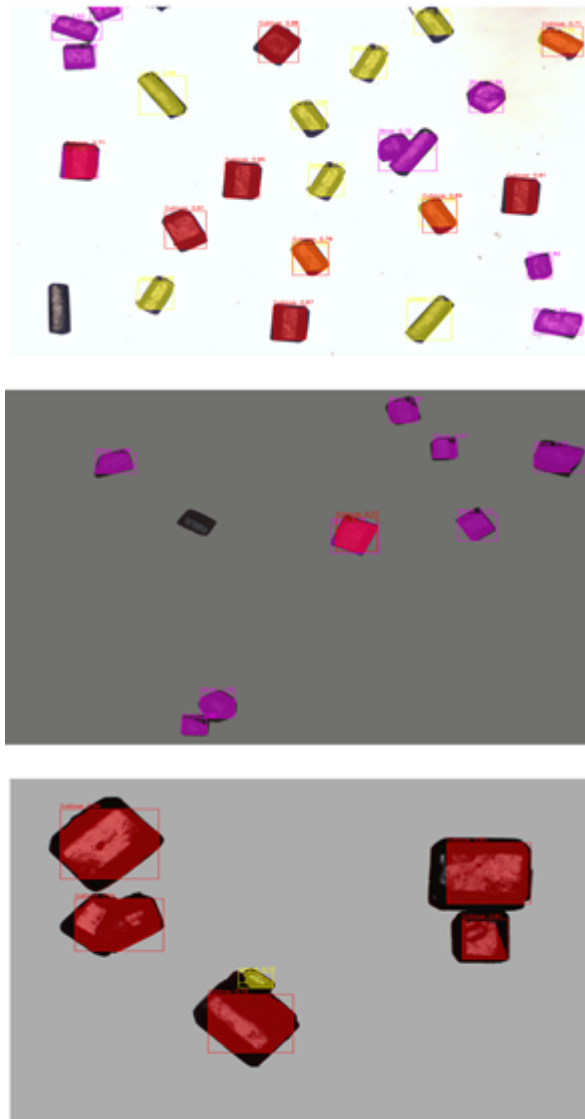
Evaluación y Análisis de Resultados. La evaluación de ambos modelos se realizó utilizando métricas cuantitativas, como la precisión y el IoU, complementadas con un análisis visual de las segmentaciones generadas. El modelo U-Net demostró una alta correspondencia entre las áreas segmentadas y las áreas reales de cristalización en las imágenes, lo que sugiere que este modelo es altamente

efectivo para su aplicación en entornos industriales donde la monitorización precisa del proceso de cristalización es esencial.

En contraste, aunque Mask R-CNN logró diferenciar entre diversas morfologías cristalinas, su rendimiento fue menos consistente, especialmente al enfrentarse a cristales con características morfológicas distintas a las del azúcar. Este resultado resalta la importancia de un conjunto de datos robusto y diversificado para entrenar modelos de Deep Learning, así como la necesidad de personalizar los modelos a las particularidades del problema que se está abordando.

Figura 18.

Resultados obtenidos con los modelos entrenados.



Resultados Obtenidos

El proceso de etiquetado y revisión de las imágenes de cristales de azúcar proporcionó valiosos insights y condujo a ajustes significativos para mejorar la precisión y relevancia de las anotaciones. A continuación, se detallan los resultados obtenidos y las adaptaciones realizadas durante el desarrollo del proceso de etiquetado.

Refinamiento de las regiones etiquetadas

Inicialmente, se observó la necesidad de ajustar las regiones etiquetadas para capturar de manera más precisa la morfología diversa de los cristales. Las regiones rectangulares iniciales se consideraron insuficientes para diferenciar entre las tipologías específicas deseadas, como cúbicos, alargados, aglomerados, forma D, entre otros. Como respuesta a esta observación, se implementaron regiones más ajustadas y específicas para cada tipo de cristal, mejorando la representación visual y la capacidad de discernir entre las diversas formas.

Exclusión de imágenes borrosas o no relevantes

Durante el proceso de revisión, se identificaron imágenes suministradas por los aliados que presentaban borrosidad significativa o contenían elementos dispersos que no correspondían a cristales. Estas imágenes fueron excluidas del conjunto de datos, optimizando la calidad y coherencia del conjunto de datos etiquetado.

Identificación y clasificación de tipologías específicas

La adaptación en la generación de regiones permitió una identificación más precisa y una clasificación más detallada de las tipologías específicas de cristales. Este refinamiento en la anotación sienta las bases para el entrenamiento de algoritmos capaces de discriminar entre diferentes formas y características de cristales de azúcar.

Mejora Continua A Lo Largo Del Proceso

El proceso de revisión y adaptación no fue estático; más bien, se implementó una estrategia de mejora continua. Se realizaron ajustes a medida que se identificaban áreas de oportunidad, asegurando que el conjunto de datos evolucionara y se perfeccionara a lo largo del tiempo.

Impacto Positivo En La Calidad Del Conjunto De Datos

Estas adaptaciones y refinamientos tuvieron un impacto positivo en la calidad del conjunto de datos etiquetado. El conjunto de datos mejorado se convierte en un recurso más robusto y preciso, fundamental para el éxito en la siguiente etapa del proyecto: el entrenamiento del algoritmo de segmentación.

Lecciones Aprendidas Y Futuras Iteraciones

Este proceso no solo brindó resultados tangibles sino también lecciones valiosas. La identificación temprana de necesidades de ajuste subraya la importancia de una revisión constante. Las lecciones aprendidas se incorporarán en futuras iteraciones, asegurando que el proceso de etiquetado y adaptación siga siendo ágil y sensible a las demandas cambiantes del proyecto.

En resumen, los resultados obtenidos no solo son reflejo de un proceso de etiquetado detallado y meticuloso, sino también de la capacidad de adaptación y mejora continua que impulsa la calidad y relevancia del conjunto de datos. Estos resultados positivos sientan las bases para avanzar con confianza en la implementación de algoritmos de Deep Learning específicos para el análisis avanzado de cristales de azúcar.

Glosario

Algoritmo de Deep Learning: Conjunto de instrucciones programadas que utilizan redes neuronales profundas para aprender patrones complejos en grandes volúmenes de datos. Son fundamentales para tareas como clasificación, segmentación y detección de objetos en imágenes (Goodfellow et al., 2016).

Anotación/ Etiquetado de Imágenes: Proceso de marcar regiones específicas en una imagen y asociarlas con etiquetas que describen sus características, lo cual es esencial para entrenar algoritmos de Deep Learning en tareas de reconocimiento y clasificación. Herramienta Usada: VGG Image Annotator (VIA).

Conjunto de Datos (Dataset),: Colección de datos estructurados, en este caso, imágenes etiquetadas, utilizadas para entrenar y validar modelos de Deep Learning. La calidad y la representatividad del conjunto de datos son cruciales para el éxito de los modelos (Krizhevsky, et al., 2012).

Cristales de Azúcar: Estructuras cristalinas formadas durante el proceso de cristalización en la producción de azúcar. La identificación de sus morfologías es crítica para asegurar la calidad y consistencia del producto final.

Deep Learning: Subcampo del aprendizaje automático que utiliza redes neuronales profundas para aprender representaciones de datos a través de múltiples capas de procesamiento. Es especialmente útil en el análisis de imágenes y reconocimiento de patrones (LeCun, et al., 2015).

Intersección sobre Unión (IoU): Métrica utilizada para evaluar la precisión de un modelo de segmentación de imágenes, calculando la superposición entre las áreas segmentadas por el modelo y las áreas de referencia. Un valor más alto indica una segmentación más precisa.

Mask R-CNN: Modelo de Deep Learning utilizado para la detección y segmentación de instancias, capaz de clasificar y delinear objetos dentro de una imagen. Es una extensión del

modelo Faster R-CNN, añadiendo una rama de segmentación a la red neuronal. (He, et al. 2017)

Morfología Cristalina: Forma y estructura de los cristales que se forman durante el proceso de cristalización. La identificación precisa de estas morfologías es fundamental en la industria azucarera para controlar la calidad del producto.

Red Neuronal Convolucional (CNN): Tipo de red neuronal utilizada principalmente en el análisis de imágenes. Emplea capas convolucionales para extraer características de las imágenes, como bordes y texturas, que son esenciales para la clasificación y segmentación. (Krizhevsky, et al.,2012)

Segmentación Semántica: Proceso de etiquetar cada píxel de una imagen para clasificar diferentes partes de la imagen en categorías específicas. Es una tarea clave en la visión por computadora, particularmente en la identificación de estructuras complejas como cristales.

Transfer Learning: Técnica en la que un modelo previamente entrenado en una tarea se adapta para una tarea relacionada. En el contexto del etiquetado de imágenes, permite reutilizar un modelo entrenado en un conjunto de datos diferente para mejorar la precisión en nuevas tareas (Pan & Yang, 2010)

U-Net: Arquitectura de red neuronal convolucional utilizada principalmente para la segmentación de imágenes biomédicas. Su diseño permite captar tanto las características locales como globales de una imagen, facilitando la segmentación precisa (Ronneberger, et al.,2015).

VGG Image Annotator (VIA): Herramienta de código abierto para la anotación manual de imágenes, ampliamente utilizada en la preparación de conjuntos de datos para el entrenamiento de algoritmos de Deep Learning. Es simple, independiente y se ejecuta en un navegador web (Dutta & Zisserman, 2019)

Z-Score: Método de normalización de datos que ajusta los valores de una imagen para que tengan una media de cero y una desviación estándar de uno, mejorando así la convergencia del modelo durante el entrenamiento (Bishop, C. M., 2006).

Bibliografía

Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., & Van Der Laak, J. A. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*, 318(22), 2199-2210.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zieba, K. (2016). End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316.

Brigham, E. O. (1988). The Fast Fourier Transform and Its Applications. Prentice-Hall.

Brigham, E. O., & Morrow, R. E. (1967). The fast Fourier transform. *IEEE Spectrum*, 4(12), 63-70.

Burger, W., & Burge, M. J. (2016). Digital Image Processing: An Algorithmic Introduction Using Java. Springer.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.

Dutta, A., & Zisserman, A. (2019). The VIA Annotation Software for Images, Audio and Video. Proceedings of the 27th ACM International Conference on Multimedia, 2276-2279. <https://doi.org/10.1145/3343031.3350535>

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2015). The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 88(2), 303-338.

Ferguson, M. K., Ronay, A., Lee, Y. T. T., & Law, K. H. (2018). Detection and segmentation of manufacturing defects with

convolutional neural networks and transfer learning. Smart and Sustainable Manufacturing Systems, 2(1), 133-144.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Gonzalez, R. C., & Woods, R. E. (2018). Digital Image Processing (4th ed.). Pearson.

He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263-1284.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2961-2969.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Huang, Z., Wang, W., Wang, L., & Tan, T. (2016). A Benchmark and Comparative Study of Color Correction Methods under Various Conditions. IEEE Transactions on Image Processing, 25(3), 1141-1156.

Jain, A. K. (1989). Fundamentals of Digital Image Processing. Prentice-Hall.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436-444.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3431-3440.

MIT Press. (2016). Deep Learning. MIT Press

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (pp. 807-814).

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62-66.

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359. <https://doi.org/10.1109/TKDE.2009.191>

Papadopoulos, D. P., Uijlings, J. R., Keller, F., & Ferrari, V. (2017). Extreme clicking for efficient object annotation. In Proceedings of the IEEE international conference on computer vision (pp. 4930-4939).

Ramírez, M., González, A., & López, J. (2020). Influence of temperature and concentration on sugar crystal formation: An image-based analysis. Journal of Crystal Growth, 532, 125411.

Ramírez, M., Hernández, J., & García, F. (2020). Estudio de la cristalización del azúcar mediante técnicas de imagen digital. Journal of Food Engineering, 245, 126-135.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252.

Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. In International conference on artificial neural networks (pp. 92-101). Springer, Berlin, Heidelberg.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Singh, R., Kumar, A., & Sharma, M. (2019). Automated detection of impurities in sugar crystals using machine learning and image analysis. Food Quality and Safety, 3(2), 107-115.

Singh, R., Verma, P., & Sharma, K. (2019). Automated Quality Control in Sugar Production using Image Processing. Food Quality and Preference, 73, 18-25.

Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting Unreasonable Effectiveness of Data in Deep Learning Era. IEEE International Conference on Computer Vision (ICCV), 843-852.

Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer.

Torralba, A., Russell, B. C., & Yuen, J. (2011). LabelMe: Online Image Annotation and Applications. Proceedings of the IEEE, 98(8), 1467-1484.

Zhang, Y., Li, X., & Chen, W. (2018). Optimization of Sugar Crystallization Process through Image Analysis and Machine Learning. Chemical Engineering Research and Design, 136, 98-106.

