

Título del Informe

Miguel Alejandro Yáñez Martínez C411

María Fernanda Suárez C411

Adrián Navarro Foya C411

Daniel Polanco Pérez C441

Daniel Ángel Arró Moreno C412

Pedro Pablo Alvarez Portelles C412

Abel Llerena Domingo C412

1 Definición del problema

El problema asignado consistió en desarrollar un sistema para detectar y cuantificar esporas de los géneros *Cladosporium* y *Curvularia* en muestras recolectadas mediante un dispositivo ubicado en la azotea de la Facultad de Biología. El sistema debe ser capaz de analizar imágenes microscópicas de las muestras y clasificar las esporas en función de su género, además de cuantificar la cantidad de esporas presentes en cada imagen. El objetivo es automatizar el proceso de análisis de esporas fúngicas, facilitando la identificación y clasificación de las mismas en un entorno de laboratorio.

2 Estado del Arte

2.1 Pre-procesamiento de Imágenes

El pre-procesamiento de imágenes es una etapa crucial que mejora la calidad de las imágenes antes de su análisis, facilitando la identificación y cuantificación precisas de las esporas fúngicas.

2.1.1 Filtrado

- **Filtro Gaussiano:** Este filtro aplica una convolución gaussiana a la imagen, suavizando las variaciones bruscas de intensidad y reduciendo el ruido, mientras mantiene los bordes importantes. Es especialmente útil en imágenes microscópicas donde las esporas pueden estar rodeadas de artefactos o partículas no deseadas que podrían interferir en el análisis.
- **Filtro de Mediana:** Reemplaza el valor de cada píxel por la mediana de los valores de los píxeles en su vecindario inmediato, eliminando eficazmente el ruido impulsivo conocido como "sal y pimienta". Ideal para eliminar píxeles

aislados extremadamente brillantes o oscuros que pueden dificultar la segmentación precisa de las esporas.

2.1.2 Segmentación

- **Umbralización (Otsu):** Este método determina automáticamente un umbral óptimo que separa la imagen en dos clases: fondo y objeto, maximizando la varianza entre ambas. Es eficaz en imágenes donde existe un contraste claro entre las esporas y el fondo, permitiendo una segmentación sencilla.
- **Detección de Bordes:** Algoritmos como Canny, Sobel y Prewitt se utilizan para identificar cambios abruptos en la intensidad de la imagen, que suelen corresponder a los bordes de las esporas. Resalta los contornos de las esporas, facilitando su delimitación y posterior análisis.
- **Clustering (K-means):** Agrupa píxeles en clusters basados en características como intensidad o color, asignando cada píxel al cluster con el centroide más cercano. Eficaz para separar esporas de fondos complejos o con texturas variadas.

2.1.3 Mejora de Contraste

- **Ajuste de Brillo y Contraste:** Modifica la intensidad de los píxeles para resaltar características de interés, como las esporas. Mejora la visibilidad de esporas en imágenes subexpuestas o con iluminación deficiente, facilitando su análisis.
- **Ecualización del Histograma:** Redistribuye las intensidades de los píxeles para obtener un histograma más uniforme, mejorando el contraste global de la imagen. Útil en imágenes donde las esporas y el fondo tienen contrastes similares, dificultando su diferenciación.

2.2 Extracción de Características

Una vez pre-procesadas las imágenes, se procede a extraer características relevantes que permitan identificar y clasificar las esporas fúngicas.

2.2.1 Características Manuales

Forma

- **Área:** Número total de píxeles que componen la espora, proporcionando una medida de su tamaño.
- **Perímetro:** Longitud del contorno de la espora, útil para evaluar su complejidad.
- **Excentricidad:** Indica cuánto se desvía la forma de la espora de una circunferencia perfecta; valores cercanos a 0 indican formas más circulares, mientras que valores cercanos a 1 indican formas más alargadas.

- **Relación de Aspecto:** Proporción entre el ancho y el alto de la espora, útil para distinguir entre esporas de formas similares.

Textura

- **Histograma de Gradientes Orientados (HOG):** Captura la distribución de las direcciones de los bordes en la imagen, proporcionando información sobre la textura y estructura de la espora.
- **Matriz de Co-ocurrencia de Niveles de Gris (GLCM):** Analiza la frecuencia con la que pares de píxeles con valores específicos aparecen en una relación espacial determinada, ofreciendo información detallada sobre la textura.

Color

- **Histogramas RGB/HSV:** Representan la distribución de colores en la imagen, permitiendo diferenciar esporas basadas en su coloración.

2.2.2 Características Profundas

Redes Neuronales Convolucionales (CNN): Estas redes aprenden automáticamente características jerárquicas de las imágenes a través de múltiples capas convolucionales, capturando desde bordes simples hasta patrones complejos. Son especialmente efectivas en imágenes complejas donde las características manuales pueden no ser suficientes para una clasificación precisa.

Ventajas:

- **Extracción Automática de Características:** Las CNN eliminan la necesidad de diseñar manualmente las características, ya que aprenden directamente de los datos relevantes para la tarea específica.
- **Invariancia Espacial:** Gracias a operaciones como el *pooling*, las CNN pueden reconocer patrones independientemente de su posición en la imagen.
- **Profundidad y Complejidad:** La estructura profunda de las CNN les permite modelar relaciones complejas y capturar patrones de alto nivel en las imágenes.
- **Transferencia de Aprendizaje:** Las CNN preentrenadas en grandes conjuntos de datos pueden adaptarse a nuevas tareas con datos limitados, mejorando la eficiencia y precisión del modelo.

Transfer Learning: Utiliza modelos preentrenados en grandes conjuntos de datos (e.g., ImageNet) y los adapta a tareas específicas mediante un proceso de fine-tuning. Reduce el tiempo de entrenamiento y mejora el rendimiento en conjuntos de datos pequeños o cuando se dispone de recursos computacionales limitados. Algunos de los modelos que se encontraron en los documentos analizados fueron VGG16, ResNet50, InceptionV3.

2.3 Clasificación y Cuantificación

Una vez extraídas las características relevantes, se procede a la clasificación y cuantificación de las esporas fúngicas utilizando diversos métodos de aprendizaje automático.

2.3.1 Métodos Tradicionales de Aprendizaje Automático

Máquinas de Vectores de Soporte (SVM): Encuentra un hiperplano óptimo que separa las diferentes clases en el espacio de características. Eficaz para la clasificación binaria y multiclase de esporas.

K-Vecinos más Cercanos (K-NN): Clasifica una espora basándose en la mayoría de sus k vecinos más cercanos en el espacio de características. Útil cuando las clases están bien separadas en el espacio de características pero es computacionalmente costoso para grandes conjuntos de datos.

Random Forest (RF): Combina múltiples árboles de decisión para mejorar la precisión y reducir el sobreajuste. Este es efectivo para clasificación y selección de características importantes y maneja bien datos no lineales y es robusto frente al ruido.

2.3.2 Métodos de Aprendizaje Profundo

- **Redes Neuronales Convolucionales (CNN):** Se utilizó para clasificación de esporas en imágenes complejas con alta precisión. Algunos ejemplos de los modelos utilizados fueron GoogleNet, ResNet, AlexNet.
- **Transfer Learning:** Se utilizaron los mencionados anteriormente.

2.4 Técnicas de Aumento de Datos

Para mejorar la generalización de los modelos y abordar problemas de desbalance de clases, se emplean técnicas de aumento de datos.

- **SMOTE (Synthetic Minority Over-sampling Technique):** Genera muestras sintéticas de clases minoritarias para equilibrar el conjunto de datos. Es útil cuando algunas especies de esporas abundan demasiado en una muestra. Puede generar muestras no realistas si no se aplica correctamente.
- **Rotación:** Gira la imagen en diferentes ángulos.
- **Volteo:** Invierte la imagen horizontal o verticalmente.
- **Escalado:** Cambia el tamaño de la imagen.

3 Soluciones implementadas

Debido a la poca cantidad de imágenes proporcionadas por la facultad de biología se decidió buscar varios dataset de investigaciones sobre detección y conteo de esporas fúngicas y unirlos para posteriormente realizar data augmentation.

3.1 Data Aumentation

Para ampliar el conjunto de datos, aplicamos transformaciones a las imágenes del dataset original, incluyendo rotaciones entre -25° y 25° , escalado entre 0.8x y 1.2x, y traslaciones en los ejes x e y de hasta un 20% del tamaño de la imagen. Además, recalculamos las anotaciones correspondientes a cada imagen. Estas modificaciones permitieron aumentar el dataset de 3 073 a 15 365 muestras.

3.2 Solución CBR

3.2.1 Entrenamiento

Dado que CBR analiza solo una imagen a la vez, se emplean las coordenadas proporcionadas por las etiquetas para extraer los fragmentos que contienen esporas y proceder a su análisis. Posteriormente, se extrajeron las siguientes características para generar vectores representativos de las esporas, los cuales fueron almacenados para su uso futuro: **Características de Color**

Estas características representan la información cromática de la imagen en distintos espacios de color:

- **Media y desviación estándar en RGB:**
 - `mean_rgb`: Promedio de cada canal de color (Rojo, Verde, Azul).
 - `std_rgb`: Desviación estándar en cada canal.
- **Media y desviación estándar en HSV:**
 - `mean_hsv`: Promedio de cada canal en el espacio HSV (Tono, Saturación, Valor).
 - `std_hsv`: Desviación estándar de cada canal en HSV.
- **Histogramas de color:**
 - `hist_rgb`: Distribución de intensidad en cada canal de color RGB.
 - `hist_hsv`: Distribución de intensidad en HSV, normalizada para que su suma sea 1.

Características de Textura

Estas características describen la estructura de la imagen en términos de patrones de repetición y variabilidad:

- **GLCM (Gray Level Co-occurrence Matrix)**
 - `contrast`: Mide la variación de intensidad entre píxeles vecinos.
 - `dissimilarity`: Indica la diferencia entre intensidades cercanas.
 - `homogeneity`: Evalúa qué tan homogénea es la textura.
 - `energy`: Mide la uniformidad de la imagen.
 - `correlation`: Mide la relación entre los valores de píxeles adyacentes.

- **LBP (Local Binary Pattern)**

- `lbp_histogram`: Histograma de patrones binarios locales, útil para detectar texturas repetitivas.

Estadísticas de Intensidad y Forma

Estas características describen la distribución de intensidad y la forma de los objetos en la imagen:

- **Estadísticas básicas de intensidad (escala de grises)**

- `mean_gray`: Media de los valores de intensidad en escala de grises.
- `std_gray`: Desviación estándar de la intensidad.
- `min_gray`: Valor mínimo de intensidad.
- `max_gray`: Valor máximo de intensidad.

- **Momentos de Hu (Forma y geometría)**

- `hu_moments`: Siete valores invariantes geométricos que describen la forma de los objetos detectados.

Y luego se calculan los percentiles. Los percentiles son valores que dividen un conjunto de datos ordenados en cien partes iguales, indicando la posición relativa de un dato dentro de la distribución. Esto nos permite clasificar un nuevo dato según unos criterios; en nuestro caso si la similitud con el mejor caso es menor que el percentil 80, la espora detectada podría considerarse no identificada o descartada, entre 80 y 95% la espora podría asignarse a una categoría con menor confianza y si la similitud es mayor que el percentil 95, la espora puede considerarse altamente similar a un caso existente y asignarse con alta confianza.

3.2.2 Predicción

Primero, se aplica segmentación de imágenes para identificar las regiones de interés mediante el siguiente proceso: la imagen se convierte a escala de grises para eliminar la información de color y reducir la complejidad computacional. Luego, se aplica un filtro Gaussiano para suavizar la imagen y eliminar el ruido. Posteriormente, se emplea la umbralización de Otsu para transformar la imagen en binaria, destacando las regiones de interés. A continuación, se detectan los contornos en la imagen binarizada, lo que permite identificar las esporas presentes. Finalmente, se calculan los rectángulos delimitadores (bounding boxes) que encierran las posibles esporas para su análisis.

Después, se extraen las características de cada imagen con el objetivo de predecir a qué grupo pertenece. En el proceso de predicción, el nuevo caso se compara con los casos existentes utilizando una función de similitud y seleccionando los k más similares. A partir de esta comparación, se clasifica en una de tres zonas de confianza: si la similitud es alta, se asigna automáticamente la clase más común y se

actualiza la base de datos; si se encuentra en una zona de incertidumbre, se realiza una búsqueda recursiva con más casos para refinar la decisión; y si la similitud es baja, se solicita la intervención de un usuario mediante aprendizaje activo. En caso de que el usuario proporcione una clasificación, el sistema la almacena para futuras referencias. De este modo, el modelo mejora con el tiempo, integrando nuevos conocimientos y refinando su capacidad de clasificación.

3.3 Modelo YOLOv11

YOLOv11 es la última versión de la serie "You Only Look Once" (YOLO) una familia de modelos. Desarrollado por Ultralytics, YOLO11 ofrece mejoras significativas en precisión, velocidad y versatilidad en comparación con sus predecesores. Soporta tareas como la detección de objetos, segmentación de instancias, estimación de pose, detección orientada y clasificación de imágenes. Ofrece una mayor precisión con menos parámetros, lo que resulta en modelos más rápidos y eficientes.

Es capaz de abordar una amplia gama de aplicaciones en visión por computadora, desde seguridad y vehículos autónomos hasta análisis de imágenes médicas.

3.3.1 Nuestra implementación

Primero, se carga un modelo YOLO preentrenado . Luego, se realiza un ajuste de hiperparámetros el cual viene integrado con YOLO, utilizando el optimizador AdamW y generando gráficos y reportes de rendimiento.

Finalmente, el modelo se entrena con los mejores hiperparámetros encontrados. Durante esta fase, se utiliza un tamaño de imagen predefinido y el optimizador AdamW. Se configuran 10 épocas de entrenamiento y el modelo resultante se guarda con un nombre específico. Además, se generan gráficos y reportes de validación para evaluar su desempeño. El modelo resultante del entrenamiento se encuentra en Google Drive