

Numero Usp: 10655837

Área de proyecto: Aprendizado de Características

Contextualización del Problema : Cada año el mercado de producción de la papa sufre pérdida debido a infestación de plagas, una estimación de la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), se alcanza un porcentaje del 20% al 40% de pérdida en la producción . Si se detecta estas plagas a tiempo se puede tomar medidas preventivas, mitigar las pérdidas de producción y económicas. Tradicionalmente la detección de estas plagas es por un especialista humano , pero no todos los productores pueden cubrir estos gastos. Así el objetivo de este proyecto es la detección y Clasificación de las enfermedades de la papa causadas por plagas .

El tipo de imágenes a ser utilizadas están separadas en 3 grupos: Imágenes de la hoja de papa en buen estado, Imágenes de la hoja de papa afectadas por *Phytophthora infestans*(plaga avanzada) , Imágenes de la hoja de papa afectadas por *Alternaria solani*(plaga leve) imágenes fueron obtenidas de Plant_village_Dataset [link aqui](#) las imagenes estan en formato RGB con dimensiones de 256x256.



Objetivo del proyecto: Detección y clasificación de enfermedades de la papa

Etapas:

- Segmentación:
 - Segmentación de la hoja del fondo:
 - Usar una máscara para separar el fondo de la Hoja, usando un algoritmo automático que calcule el threshold como [Otsu](#)
 - Segmentación de las áreas afectadas de la hoja:
 - Eliminar las regiones verdes de la imagen tomando como principio que representan zonas en buen estado. Posibles métodos: Trabajar con el canal verde de la imagen , separar las regiones verdes por medio de una umbralización .Igualmente es posible trabajar en la escala $L^* a^* b^*$.
- Extracción de Características:
 - Se calcularán descriptores de la imagen resultante tanto asi de color como de textura(Ej: Euclidean Distance ,Logarithmic Distance, Mean , Kurtosis ,standard deviation , Entropy , etc)
 - Reducir la dimensionalidad de las características (Opcional): Evaluar diferentes conjuntos de features, o aplicar principal component analysis(PCA) , para encontrar la mejor combinatoria de los descriptores
- Separación de la data(Posible métodos):

- Hold - Out
 - K-fold Cross Validation
- Clasificación:
 - Se utiliza un algoritmo de clasificación que recibirá como entrada los features extraídas y datos rotulados y dará como salida un clasificador. Posible algoritmos a ser testados:
 - KNN
 - MLP
- Testeo:
 - Se medirá la eficacia con una medida de error cuadrático para cada clasificador con el conjunto de testeo