

# Detección de la madurez del banano basado en la inteligencia artificial

Jorge Enríquez

Universidad de las Fuerzas Armadas ESPE

9 de agosto de 2022



CIT 2022 - Congreso Internacional de Ciencia y Tecnología

- 1 Introducción
- 2 Antecedentes
- 3 Diseño de la Investigación
- 4 Implementación
- 5 Resultados
- 6 Discusión
- 7 Conclusiones y Trabajos Futuros

- En la actualidad, con el inminente crecimiento de la población mundial, se estima que el sector agrícola necesitará producir un 50 % más de alimentos.
- Un tercio de las exportaciones mundiales de banano proceden de Ecuador. Por lo tanto, el cultivo del banano se considera **la actividad agrícola más importante para la economía del país.**
- Los métodos manuales que determinan el grado de madurez de los plátanos suelen ser de **baja precisión y poco fiables, ya que se basan en la experiencia del evaluador.** Por lo tanto, es necesaria la automatización de dicho proceso.

- **P1.** ¿Cómo diseñar e implementar un sistema inteligente para la detección de la madurez del banano?
- **P2.** ¿Es posible determinar el grado de madurez del banano según la escala de maduración de Von Loesecke utilizando técnicas de aprendizaje automático? **En particular, la máquina de vectores de soporte multiclase (MSVM), el bosque aleatorio (RF) y el aprendizaje profundo (DL).**
- **P3.** ¿Cómo se compara el rendimiento del modelo propuesto con las técnicas de referencia más avanzadas?

# Objetivos

- Identificar las técnicas de aprendizaje automático más eficaces para la clasificación de imágenes.
- Crear un modelo de aprendizaje automático que permita clasificar los bananos según el nivel de madurez.
- Evaluar los modelos de predicción configurados mediante la obtención de las métricas de rendimiento, con el fin de determinar la técnica que permite obtener el modelo de predicción más preciso.

- **Escala de Von Loesecke:**

- Esta escala se emplea para **clasificar la maduración del banano.**
- Tiene siete subclases de maduración que corresponden a las **etapas de color de la cáscara.**
- El análisis del color depende del **ángulo de observación, la iluminación del entorno y la precisión del observador.**



Figura 1: Cada racimo de bananos representa una etapa del proceso de maduración.

- **Support Vector Machine:**

- SVM es un **clasificador binario lineal** que identifica un único límite entre dos clases.
- Dispone de una amplia variedad de tipos de kernel, entre los que destacan **Lineal, RBF, Polinomial y Sigmoidal**, ya que se pueden utilizar independientemente del tipo de problema.
- Para resolver una tarea multiclase, debe ser **descompuesta en una serie de clasificadores SVM binarios simples**, utilizando métodos como **one-versus-one (OVO) y one-versus-all (OVA)**.

# Diseño de la Investigación (1/4)

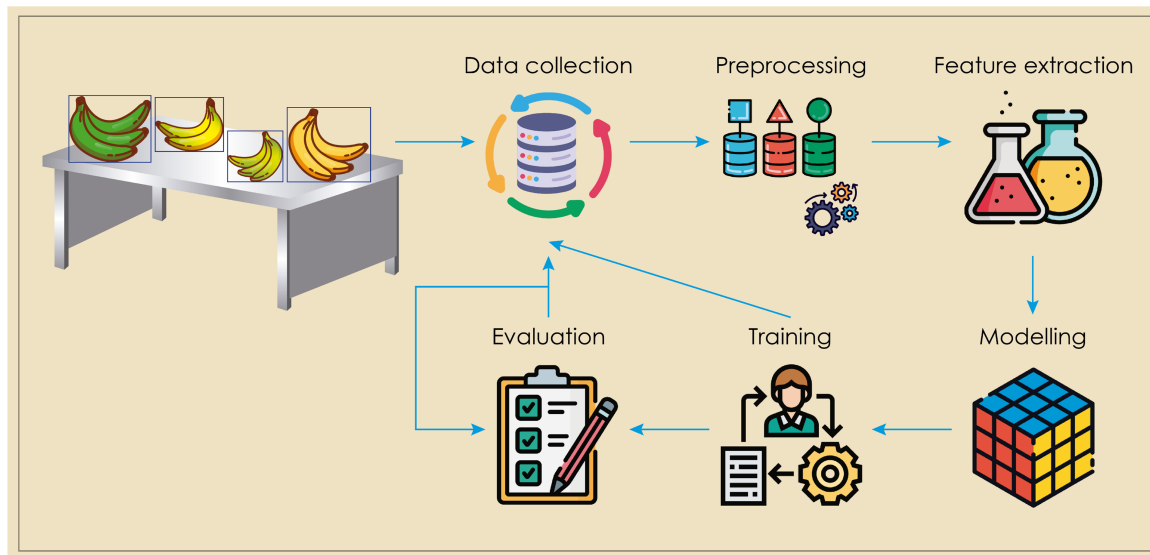


Figura 2: Visión general del marco propuesto.



# Diseño de la Investigación (2/4)

- **Recolección de datos:** Las imágenes utilizadas consistían en 305 muestras de bananos, que se clasificaron manualmente en tres categorías: **totalmente verde, amarillo con puntas verdes y amarillo con manchas cafés**. Para equilibrar el conjunto de datos, se aplicó la técnica data augmentation para obtener aproximadamente 400 imágenes por clase con un total de 1242 muestras.
- **Preprocesamiento:** Se utilizó la técnica **NLM para la eliminación de ruido** y la **herramienta de código abierto Rembg para eliminar el fondo** de las imágenes.
- **Extracción de características:** Para la **extracción de características de color**, se realizó una conversión de la imagen del banano RGB al espacio de color HSV para obtener el histograma de color. Para la **extracción de características de la textura**, se empleó matrices de co-ocurrencia de nivel de gris (GLCM) basadas en la textura.

## ● Modelado:

- Se implementó la técnica MSVM, en la que se pueden utilizar muchos métodos para construir un **clasificador multiclase como OVO y OVA**.
- En OVA, se necesitan  **$N$  pruebas binarias**, mientras que para OVO, se necesitan  **$N(N-1)/2$  pruebas binarias** para la clasificación de  $N$  clases.
- **OVO implica menos vectores de soporte que OVA**, reduciendo así la complejidad y aumentando la eficiencia.
- Se consideraron dos tipos de MSVM basados en OVO: el **grafo acíclico dirigido MSVM (DAGMSVM)** y **basado en votación MSVM (VBMSVM)**.

## ¿Cómo funciona DAGMSVM?

En la fase de prueba, utiliza el grafo acíclico dirigido (DAG) binario arraigado, que consta de  $N(N - 1)/2$  nodos internos y  $N$  hojas, para combinar estas SVM binarias.

## ¿Cómo funciona VBMSVM?

Una SVM  $C_{ij}$ , donde  $i \neq j, i = 1, \dots, N; j = 1, \dots, N$ , está diseñada para distinguir la clase  $i^{ma}$  de la clase  $j^{ma}$ . Cada  $C_{ij}$  SVM se entrena sobre todos los ejemplos de la clase  $i^{ma}$  con etiquetas positivas y los ejemplos de la clase  $j^{ma}$  con etiquetas negativas.

- **Entrenamiento:**

- ⇒ 80 % para los datos de entrenamiento
- ⇒ 20 % para los datos de la prueba
- ⇒ validación cruzada de 8 veces

- **Evaluación:**

- **Exactitud:** Proporción de predicciones correctas sobre la entrada.
- **Precisión:** Porción de ejemplos asignados a la clase positiva que pertenecen a la clase positiva.
- **Recuperación:** Indica lo bien que se predijo la clase positiva.
- **Puntuación F1:** La precisión y la recuperación se combinan en un único valor, que trata de equilibrar ambos valores.
- **Curva ROC:** Es un gráfico que muestra el rendimiento de un modelo de clasificación en todos los umbrales de clasificación.

# Implementación (1/3)

- Se comparó MSVM con un **método de ensamble (Random Forest)** y un **modelo de red neuronal (TL basado en CNN)**.
- Random Forest (RF) incluye múltiples árboles de decisión y se **basa en el voto mayoritario para la clasificación**. Los hiperparámetros considerados para esta técnica fueron: **n\_estimators, criterion, max\_depth, y max\_features**.
- El modelo de red neuronal es **basado en CNN y aprovecha el aprendizaje por transferencia (TL)**. Específicamente, se utilizó el VGG16 como modelo preentrenado, **que consta de trece capas convolucionales y cinco capas de maxpooling**. Los hiperparámetros seleccionados del modelo propuesto fueron: **entropía cruzada categórica dispersa, el algoritmo de optimización *adam* y se evaluó mediante la métrica de precisión**.

# Implementación (2/3)

- La fase experimental se desarrolló en el lenguaje de programación Python.
- Para equilibrar el conjunto de datos, se aplicaron varias técnicas de data augmentation, como **el volteo, el desplazamiento, la rotación, el corte, el acercamiento y el modo de relleno más cercano.**
- A continuación, se utilizó la técnica NLM para eliminar el ruido de las imágenes utilizando los parámetros mínimos recomendados, que son:  **$h = 10$ , `templateWindowSize = 7` y `searchWindowSize = 21`.** La herramienta Rembg se utilizó para **eliminar el fondo de las imágenes.**

# Implementación (3/3)

- La extracción de características se realizó **considerando el color y la textura de las imágenes**.
  - Para las características de color, las imágenes se transformaron al espacio de color HSV para obtener el histograma de color utilizando los siguientes parámetros: **image, channels = [0, 1, 2], mask = None, histSize = [8, 8, 8], ranges = [0, 256, 0, 256, 0, 256]**.
  - Para las características texturales, se empleó la técnica de **matrices de co-ocurrencia de nivel de gris (GLCMs)**, según la cual se transformó la imagen a escala de grises para obtener las **características texturales para cuatro tipos de adyacencia** y se calculó la media entre las mismas.

¿Cuál es la diferencia entre DAGMSVM y VBMSVM?

En VBMSVM, cada predicción realizada se pasa a cada clasificador y la clase predicha se almacena para asignar la clase que obtiene el máximo número de votos. Por otro lado, para DAGMSVM, en cada predicción se elimina la clase rechazada hasta que sólo queda una clase por asignar.

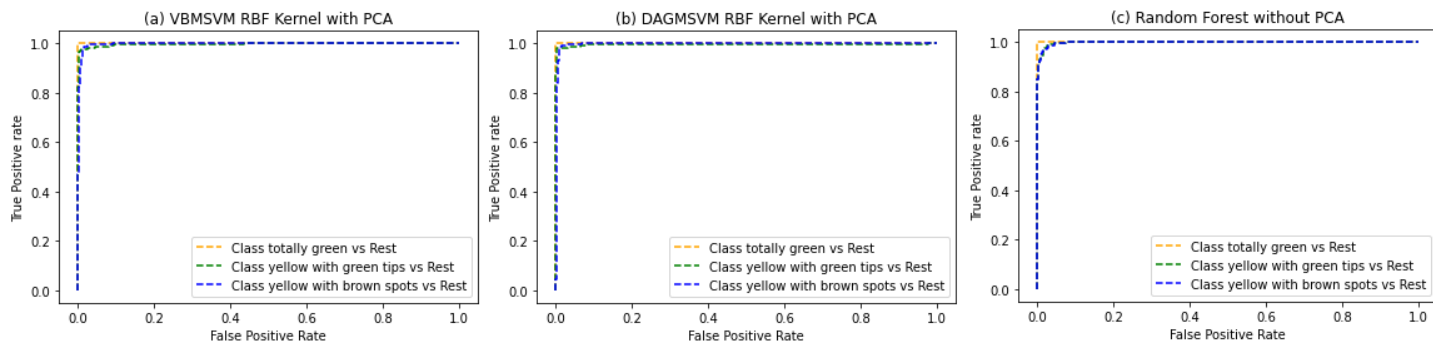
# Resultados (1/4)

Table 1. Resumen de los modelos Metrics

Experiment N°	Technique	Accuracy	F1	Precision	Recall
1st Experiment: Concatenated Color Features and Textural Features	VBMSVM Linear Kernel	98.19	98.19	98.22	98.19
	VBMSVM Polynomial Kernel	98.09	98.09	98.13	98.09
	VBMSVM RBF Kernel	98.39	98.39	98.42	98.39
	VBMSVM Sigmoid Kernel	97.88	97.88	97.95	97.88
	DAGMSVM Linear Kernel	98.09	98.09	98.10	98.09
	DAGMSVM Polynomial Kernel	98.19	98.19	98.24	98.19
	DAGMSVM RBF Kernel	98.09	98.09	98.15	98.09
	DAGMSVM Sigmoid Kernel	97.88	97.88	97.95	97.88
Random Forest	98.09	98.09	98.14	98.09	
2nd Experiment: Concatenated Color Features and Textural Features with PCA	VBMSVM Linear Kernel	97.48	97.47	97.51	97.48
	VBMSVM Polynomial Kernel	98.69	98.69	98.72	98.69
	VBMSVM RBF Kernel	98.39	98.39	98.44	98.39
	VBMSVM Sigmoid Kernel	96.37	96.37	96.49	96.37
	DAGMSVM Linear Kernel	97.07	97.07	97.14	97.07
	DAGMSVM Polynomial Kernel	98.69	98.69	98.72	98.69
	<b>DAGMSVM RBF Kernel</b>	<b>98.89</b>	<b>98.89</b>	<b>98.93</b>	<b>98.89</b>
	DAGMSVM Sigmoid Kernel	96.47	96.47	96.57	96.47
Random Forest	96.98	96.98	97.13	96.98	
3rd Experiment: Deep Learning	CNN-based TL	48.19	37.07	30.38	48.19

# Resultados (2/4)

- Ambas variantes de la técnica MSVM superaron a RF (véase la figura (c)).
- Entre las técnicas MSVM, se observa que DAGMSVM (véase la figura (a)) permite que **las clases amarillo con puntas verdes y amarillo con manchas cafés** sean reconocidas con mayor precisión que con VBMSVM (véase la figura (b)).



**Figura 3:** (a) VBMSVM Kernel RBF con PCA, (b) DAGMSVM Kernel RBF con PCA, (c) Random Forest sin PCA.



# Resultados (3/4)

- Esta figura demuestra la eficacia de las técnicas VBMSVM y DAGMSVM en comparación con el modelo TL basado en CNN. También podemos ver que **la técnica DAGMSVM RBF Kernel superó tanto al VBMSVM RBF Kernel como a TL basado en CNN.**

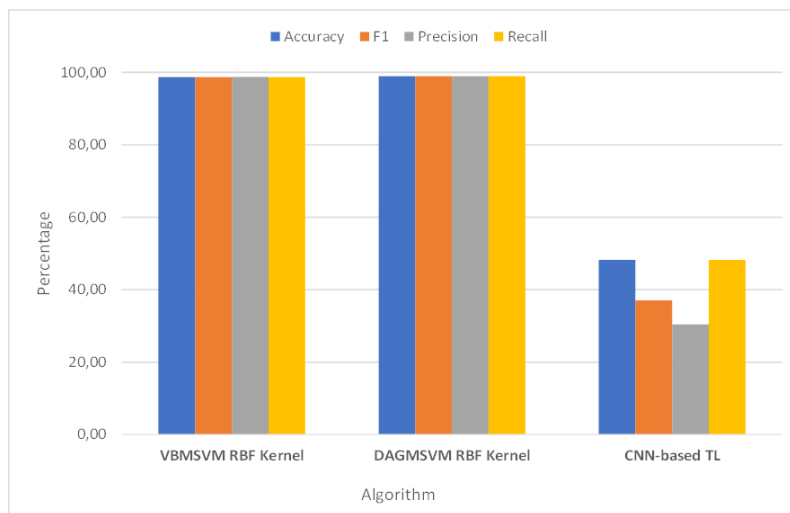


Figura 4: Comparación del tercer experimento con los mejores modelos de los experimentos anteriores.

- **Para DAGMSVM, VBMSVM y RF**, se consideraron los siguientes escenarios:
  - ① Características de color y características de textura concatenadas, donde se utilizó un espacio de características de 525 componentes.
  - ② Características de color y características de textura concatenadas con PCA, donde el espacio de características consistió en 19 componentes obtenidos aplicando PCA.
- El modelo TL basado en CNN obtuvo los siguientes resultados: 48,19 % de exactitud, 37,07 % de puntuación F1, 30,38 % de precisión y 48,19 % de recuperación. Es el modelo con menos precisión en este estudio.

- Los mejores resultados en **las técnicas MSVM** se obtuvieron **aplicando PCA con un espacio de características de 19 componentes**, mientras que **Random Forest fue mejor sin aplicar PCA**, por lo que se utilizó un espacio de características de 525 componentes.
- **Las técnicas de aprendizaje profundo son superiores a otras técnicas** para resolver problemas de clasificación de imágenes, sin embargo, **en este estudio** se puede determinar que su uso no es la mejor opción, **debido al número de muestras del conjunto de datos.**

# Conclusiones y Trabajos Futuros

- En este estudio, se desarrolló un modelo de inteligencia artificial para **detectar el nivel de madurez de los bananos** según tres categorías de la escala de Von Loesecke: **totalmente verde, amarillo con puntas verdes y amarillo con manchas cafés.**
- El modelo construido utilizó la **técnica de clasificación DAGMSVM y el kernel RBF**, haciendo uso de una validación cruzada de 8 veces, lo que permitió obtener **una exactitud del 98,89 %.**
- Para el trabajo futuro, se planea **recoger muestras de banano de diferentes estados de maduración en Ecuador**, preparando un ambiente totalmente controlado para evitar la interferencia de condiciones ambientales desfavorables.
- Además, se tiene previsto examinar más **técnicas de aprendizaje profundo, como las CNNs**, dada su reputación como modelos ideales para resolver problemas de visión por ordenador, gracias a su **capacidad para extraer representaciones detalladas del contenido de las imágenes.**

# Detección de la madurez del banano basado en la inteligencia artificial

Jorge Enríquez

Universidad de las Fuerzas Armadas ESPE

9 de agosto de 2022



CIT 2022 - Congreso Internacional de Ciencia y Tecnología

**Gracias, ¿alguna pregunta?**