

# Informe Técnico: Modelado de Mosquitos con Knowledge Distillation

## Datos del Autor

Nombre: Matteo Martinez

Carrera: Ingeniería en Mecatrónica

Institución: Facultad de Ingeniería UNA

Fecha: Junio de 2025

## 1. Introducción

En el marco del desarrollo de tecnologías para el control vectorial de enfermedades, se diseñó una trampa inteligente capaz de detectar y contar mosquitos en tiempo real utilizando inteligencia artificial. Esta trampa busca integrarse en redes de monitoreo ambiental, aportando datos clave para la toma de decisiones en salud pública.

Uno de los mayores desafíos en este tipo de aplicaciones es lograr modelos eficientes que puedan ejecutarse en dispositivos con recursos limitados, como microcomputadoras o sistemas embebidos. Para abordar esto, se aplicó la técnica de Knowledge Distillation, la cual permite entrenar modelos más livianos sin sacrificar precisión.

## 2. Descripción del Dataset

El conjunto de datos utilizado fue proporcionado por el Instituto de Investigaciones en Ciencias de la Salud (IICS), y consistió en 1.199 imágenes reales de mosquitos. Estas imágenes fueron cuidadosamente etiquetadas para permitir la detección de cinco clases:

- Female
- Male
- Incomplete
- Out of Focus
- Undefined

La distribución fue la siguiente:

- Entrenamiento: 719 imágenes
- Validación: 240 imágenes
- Pruebas (test): 240 imágenes

# Informe Técnico: Modelado de Mosquitos con Knowledge Distillation

Para mejorar la capacidad de generalización, se aplicaron técnicas de aumento de datos. Esto permitió elevar el total de imágenes de entrenamiento a 2.157. Las técnicas incluyeron rotaciones, ajustes de brillo y saturación, desenfoque y adición de ruido controlado.

## 3. Arquitectura del Modelo y Distillation

Se trabajó con la arquitectura YOLOv8, reconocida por su velocidad y precisión en tareas de detección de objetos. El modelo YOLOv8-small se utilizó como 'teacher', mientras que YOLOv8-nano, de menor tamaño y complejidad, actuó como 'student'.

La técnica de Knowledge Distillation consiste en transferir el conocimiento aprendido por un modelo grande (profesor) a uno pequeño (estudiante), mediante la supervisión de sus salidas y representación intermedia.

Se llevaron a cabo tres experimentos:

1. Entrenamiento estándar del modelo nano por 100 épocas.
2. Entrenamiento del modelo nano guiado por el teacher mediante distillation.
3. Reentrenamiento del student, utilizando predicciones con confianza mayor al 40% para filtrar ejemplos útiles.

Todos los entrenamientos se realizaron con imágenes de tamaño 224x224, batch size 16, y tasa de aprendizaje (learning rate) de 0.01.

## 4. Resultados Obtenidos

Los modelos fueron evaluados mediante métricas de precisión (mAP50 y mAP50-95) y tiempo de inferencia en una GPU RTX 3090. Los resultados fueron los siguientes:

- Modelo Nano Normal:
  - mAP50 Hembras: 87.1%
  - mAP50-95 Hembras: 44.9%
  - mAP50 Machos: 93.9%
  - mAP50-95 Machos: 52.4%
- Tiempo de Inferencia: 4.5 ms

# Informe Técnico: Modelado de Mosquitos con Knowledge Distillation

- Modelo Nano con Distillation:
  - Mismo desempeño en precisión
  - Tiempo de Inferencia: 4.4 ms
- Modelo Nano Reentrenado con Distillation:
  - mAP50-95 Hembras: 47.6%
  - mAP50-95 Machos: 54.6%
  - Tiempo de Inferencia: 4.6 ms

Se observó que la técnica de distillation mejoró la precisión de forma significativa en el reentrenamiento, sin comprometer el tiempo de inferencia, manteniéndose en el orden de los 4-5 ms por imagen.

## 5. Discusión

La implementación de modelos ligeros con buena precisión es crucial para aplicaciones en campo, donde la disponibilidad energética y de cómputo es limitada. El uso de Knowledge Distillation permitió refinar el desempeño del modelo sin incurrir en penalizaciones significativas de tiempo. Esta técnica, aunque no es novedosa, se vuelve altamente relevante en escenarios de Edge AI.

El modelo reentrenado no solo superó a sus pares en precisión, sino que también mantuvo un perfil computacional muy eficiente, siendo ideal para ejecutarse en plataformas como Raspberry Pi o NVIDIA Jetson Nano.

## 6. Conclusiones y Trabajo Futuro

- La distillation mejoró el desempeño del modelo nano sin afectar el tiempo de respuesta.
- Es posible implementar modelos con mAP cercanos al 50% en dispositivos de bajo costo.
- El modelo es capaz de diferenciar clases biológicas relevantes para la vigilancia vectorial.

Como trabajo futuro se contempla:

1. Despliegue del modelo en campo.
2. Integración con sensores adicionales (temperatura, humedad).
3. Validación de desempeño con imágenes reales captadas en tiempo real.

# **Informe Técnico: Modelado de Mosquitos con Knowledge Distillation**

4. Publicación científica de los resultados en el Congreso CLEI 2025.