



TRABAJO FIN DE MASTER

SISTEMA DE SEPARACIÓN AUTOMÁTICA DE GRILLOS POR SEXO – SOLUCIÓN EMBEBIDA DE BAJO COSTE USANDO YOLOV8 Y RASPBERRY PI

TRABAJO FIN DE MASTER
PARA LA OBTENCIÓN DEL
TÍTULO DE MASTER EN
INGENIERÍA INDUSTRIAL

SEPTIEMBRE 2025

**Juan Manuel Cantarero
Angulo**

DIRECTOR DEL TRABAJO FIN DE MASTER:
Dr. Matthew Smith

Sistema de separación automática de grillos por sexo

Solución embebida de bajo coste usando YOLOv8 y Raspberry Pi

Juan Manuel Cantarero Angulo, *Estudiante del Master Habilitante en Ingeniería Industrial* y Dr. Matthew Smith, *Supervisor*

(jcantareroangulo@hawk.illinoistech.edu)

(msmith84@illinoistech.edu)

*Illinois Institute of Technology, Chicago, Illinois
Universidad Politécnica de Madrid, Madrid, Spain*

La creciente demanda global de fuentes de proteína sostenibles está impulsando un interés creciente en los insectos comestibles, identificándose al *Acheta domesticus* (grillo doméstico) como una de las especies más adecuadas para la producción industrial. Las prácticas de cría actuales suelen mantener poblaciones mixtas de grillos sin una separación automatizada por sexos, a pesar de los beneficios potenciales que esto conlleva, como la selección genética, la optimización de las proporciones de reproducción y la diferenciación nutricional. Este trabajo presenta un sistema de bajo coste y en tiempo real para la separación automática por sexo de *Acheta domesticus*, que combina visión por computador y actuación física. El dispositivo integra una Raspberry Pi 5 con la cámara oficial Raspberry Pi AI y un modelo de detección de objetos YOLOv8 nano entrenado específicamente, junto con un brazo de clasificación accionado por servomotor. El modelo alcanzó una precisión media ($mAP@0.5$) de 0.977 durante las pruebas, y los experimentos en condiciones reales con grupos de grillos lograron una precisión global de separación del 86.8%. Estos resultados demuestran la viabilidad de desplegar modelos ligeros de deep learning en dispositivos con recursos limitados para aplicaciones en la cría de insectos, ofreciendo una solución práctica para mejorar la eficiencia y la sostenibilidad en la producción de grillos.

Palabras clave: YOLO, Visión por computadora, Cría de insectos, Detección de objetos, Raspberry Pi, Clasificación en tiempo real, *Acheta domesticus*, Edge AI

0 INTRODUCCIÓN Y TRABAJOS RELACIONADOS

La demanda global de fuentes de proteína sostenibles está aumentando rápidamente, impulsada por el crecimiento poblacional, las preocupaciones medioambientales y las limitaciones de la producción ganadera convencional. Los insectos, en particular, han sido destacados por la FAO y numerosos investigadores como una alternativa altamente eficiente, ya que requieren significativamente menos tierra, agua y alimento, a la vez que producen menores emisiones de gases de efecto invernadero [1]. En este contexto, se proyecta que el mercado de proteína de insectos se expanda de aproximadamente 2.4 mil millones de USD en 2025 a alrededor de 28.54 mil millones de USD en 2035, lo que representa una tasa de crecimiento anual compuesta del 28.1% [2].

Entre los insectos comestibles, el *Acheta domesticus*, comúnmente conocido como grillo doméstico, ha surgido como uno de los principales candidatos para la cría a escala industrial debido a su rápido ciclo de crecimiento, sabor y

aceptación en diferentes culturas. En 2022, fue autorizado como nuevo alimento por la Unión Europea (2022/188) [3]. Además, el *Acheta domesticus* ha sido identificado como una de las siete especies de insectos aptas para la producción industrial y el consumo humano [4].

Actualmente, los grillos domésticos se crían en poblaciones mixtas bajo condiciones ambientales controladas [5]. Sin embargo, no se han realizado esfuerzos industriales significativos para separarlos por sexo. La separación por sexo podría aportar múltiples ventajas, incluyendo la selección genética, la optimización de las proporciones macho-hembra para la reproducción y mejoras nutricionales, dado que las hembras suelen ser de mayor tamaño [6] y más ricas en proteína [7] en comparación con los machos. El exceso de machos, por otro lado, puede generar competencia, estrés e incluso conductas de canibalismo dentro de las colonias [8]. Superar estas barreras resulta crítico para la escalabilidad de la cría industrial de insectos.

Se han llevado a cabo investigaciones previas sobre la identificación de sexo y el conteo de *Acheta domesti-*

cus. Por ejemplo, Giulietti et al. [9] desarrollaron un sistema de medición basado en visión artificial para estimar el número y la distribución por sexo de los grillos mediante un banco de pruebas. Hansen et al. [10] también aplicaron modelos de deep learning de última generación (YOLOv5 para la detección y VGG-16 para la clasificación por sexo) a grillos, logrando una alta precisión en la identificación del sexo, aunque sin integrar un mecanismo físico de separación. No obstante, ambos trabajos se centraron únicamente en las tareas de clasificación, sin proporcionar un flujo de trabajo de separación completamente automatizado y en tiempo real.

En este trabajo, vamos más allá de estos enfoques previos al presentar un prototipo de prueba de concepto que no solo clasifica, sino que también separa físicamente en tiempo real a los grillos según su sexo. El sistema integra un modelo de detección de objetos YOLOv8 nano [11] embebido en una Raspberry Pi 5 [12], que controla un brazo de separación accionado por servomotor. Este diseño introduce la contribución novedosa de combinar visión por computador embebida con un mecanismo de separación mecánico, lo que permite una operación autónoma y de bajo coste, adecuada para la cría de insectos a pequeña escala y que ofrece una base para aplicaciones industriales más avanzadas.

1 METODOLOGÍA

El sistema propuesto permite la clasificación y separación en tiempo real de *Acheta domesticus* mediante la combinación de visión por computador con control físico embebido. La configuración de hardware incluye una Raspberry Pi 5, la cámara oficial Raspberry Pi AI y un brazo robótico accionado por un servomotor SG90. El software se basa en un modelo ligero de detección de objetos (YOLOv8 nano), entrenado con un dataset personalizado e integrado con la lógica de control que activa el brazo según el sexo predicho del grillo.

Los grillos son inducidos a desplazarse desde una cámara inicial (BOX 1) hasta una cámara de destino (BOX 2) a través de un puente transparente de metacrilato. A medida que un grillo cruza el puente, la cámara captura de manera continua los fotogramas y los envía al modelo de detección de objetos, que los procesa en tiempo real para determinar tanto la posición como el sexo del grillo. Basándose en estas predicciones, el sistema, orquestado por la Raspberry Pi, ordena al servomotor activar un mecanismo de tipo compuerta, dirigiendo al individuo hacia el compartimento de machos o al de hembras.

1.1 Hardware y configuración física

Las paredes de BOX 1 y BOX 2 están construidas en acrílico oscuro, mientras que el puente de conexión está hecho de acrílico transparente (Figura 1); todos los componentes fueron diseñados en CAD y fabricados mediante una cortadora láser. El brazo de clasificación también fue diseñado con software CAD y fabricado con una impresora 3D.

Una única cámara (Raspberry Pi AI Camera) se encuentra montada en una posición fija sobre el puente transparente, como se muestra en la Figura 1, alineada con el brazo de clasificación y la entrada a BOX 2. Esta ubicación permite que el sistema detecte a los grillos justo antes de que lleguen al punto de clasificación, posibilitando la activación oportuna del brazo. Para mejorar el contraste visual y facilitar la clasificación, el suelo del puente acrílico se pintó de blanco, proporcionando un fondo uniforme para el procesamiento de imágenes.

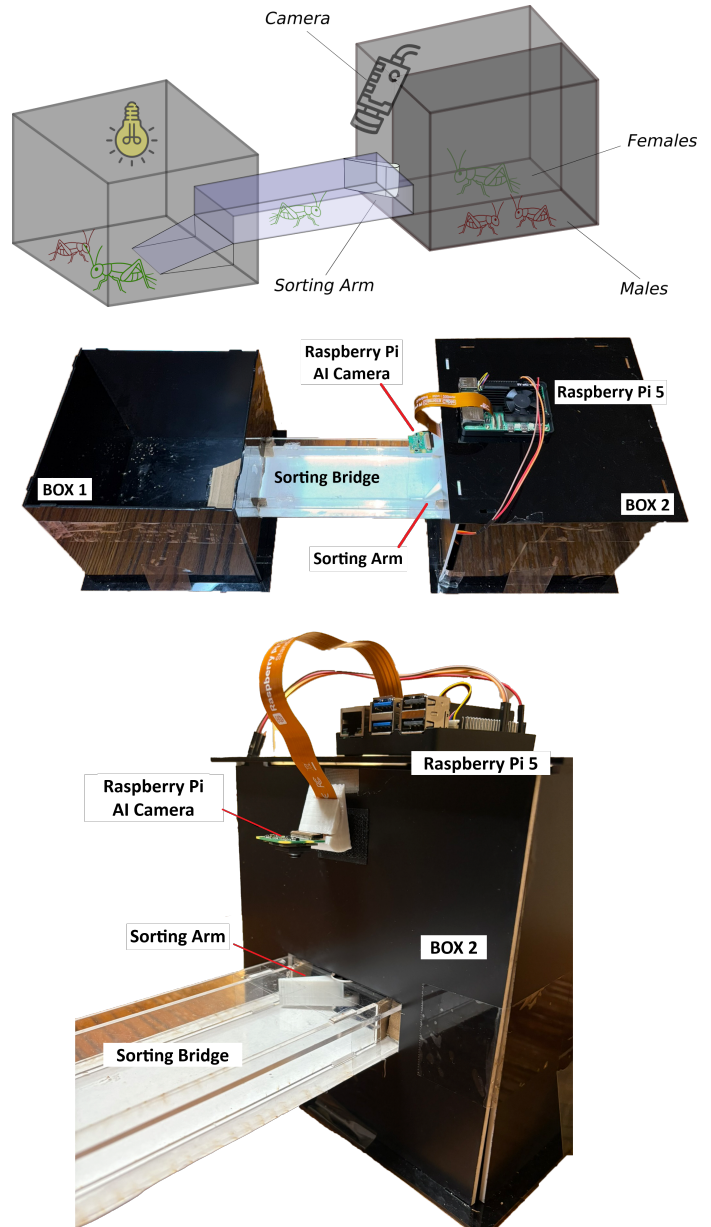


Fig. 1: Vista general del prototipo de separación de grillos: **arriba** — dibujo del dispositivo; **centro** — vista superior del dispositivo de clasificación; **abajo** — cámara Raspberry Pi AI y brazo de separación.

El mecanismo de clasificación está impulsado por un servomotor SG90 (Figura 2), programado para operar en dos posiciones fijas (Figura S1). Cuando el modelo de detección identifica una hembra, el brazo bloquea el acceso

al compartimento de machos, dirigiendo el grillo hacia la sección de hembras (Figura S1b). Por el contrario, si se detecta un macho, el brazo cambia a la posición opuesta (Figura S1c). Este mecanismo no invasivo permite la separación automática sin contacto directo, minimizando el riesgo de daño a los insectos.

Para incentivar el movimiento de los grillos a través del sistema, se crean dos entornos contrastantes: BOX 1 contiene una fuente de luz intensa y carece de alimento y agua, lo que la hace inhóspita, mientras que BOX 2 permanece oscura y contiene tanto alimento como humedad, atrayendo de forma natural a los grillos hacia la cámara de destino.

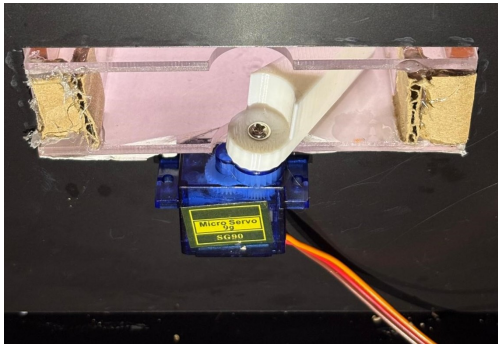


Fig. 2: Servomotor y brazo de clasificación impreso en 3D ubicado al final del puente.

1.2 dataset y anotación

Para entrenar y validar el modelo de detección de grillos, se desarrolló un dataset personalizado y se puso a disposición pública en Roboflow [13]. El dataset estuvo inicialmente compuesto por 335 imágenes con una resolución de 480×480 píxeles, capturando individuos de *Acheta domesticus* de ambos sexos mientras se aproximaban al brazo de clasificación. Todas las imágenes fueron adquiridas utilizando la cámara Raspberry Pi AI, montada en una posición fija superior idéntica a la empleada durante la fase experimental, garantizando la consistencia entre la recopilación de datos y las condiciones de despliegue.

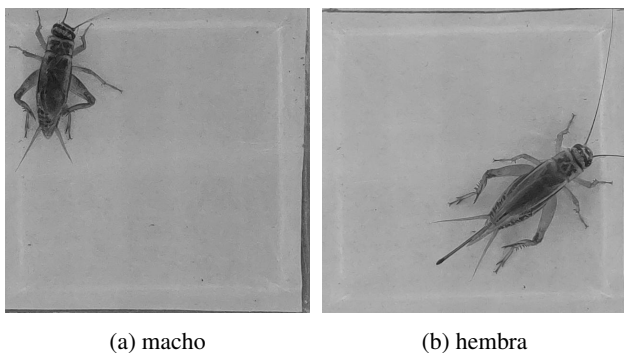


Fig. 4: Comparación entre macho y hembra de *Acheta domesticus*; obsérvese el ovopositor.

Para garantizar la robustez en condiciones reales, se utilizaron diversas fuentes de luz durante la adquisición de

imágenes, colocadas en distintos ángulos para compensar los reflejos producidos por la superficie acrílica. Esto introdujo una variabilidad controlada en la iluminación, ayudando al modelo a generalizar frente a distintos escenarios lumínicos. Además, se tomaron imágenes con uno, dos e incluso tres grillos de distinto sexo cruzando el puente simultáneamente (véase Figura 5), aumentando así la diversidad y complejidad del dataset.



Fig. 5: Imagen de entrenamiento con tres *Acheta domesticus*.

La Figura 6 ofrece una visión general del dataset. En primer lugar, se destaca que el conjunto de datos está balanceado, conteniendo aproximadamente el mismo número de individuos machos y hembras. También se ilustra la distribución espacial de los grillos dentro de las imágenes, mostrando que el *Acheta domesticus* tiende a desplazarse cerca de las paredes del puente, así como sus tamaños relativos (normalizados entre 0 y 1). Esta variabilidad en posición y escala contribuye a la robustez del conjunto y respalda la expectativa de resultados fiables durante el entrenamiento.

Cada imagen fue convertida a escala de grises para reducir el tensor de entrada a un único canal, optimizando así la velocidad de inferencia sin afectar de manera significativa la precisión de la detección. Posteriormente, todas las imágenes fueron anotadas manualmente utilizando Roboflow [14]. Para cada grillo se dibujó una caja delimitadora (bounding box) y se etiquetó según su sexo. La identificación del sexo resulta sencilla a simple vista, ya que las hembras son generalmente de mayor tamaño y presentan un ovopositor claramente visible en la parte posterior del abdomen, como se muestra en la Figura 4.

Para incrementar el tamaño y la variabilidad del conjunto, se aplicó un proceso de aumento de datos. Se introdujo un zoom aleatorio en el rango del 10% al 26% para simular oclusiones parciales, como en los casos en que un grillo es visible solo parcialmente debido al encuadre o a la superposición de individuos. Tras la aplicación de aumento, el conjunto final contenía 597 imágenes de entre-

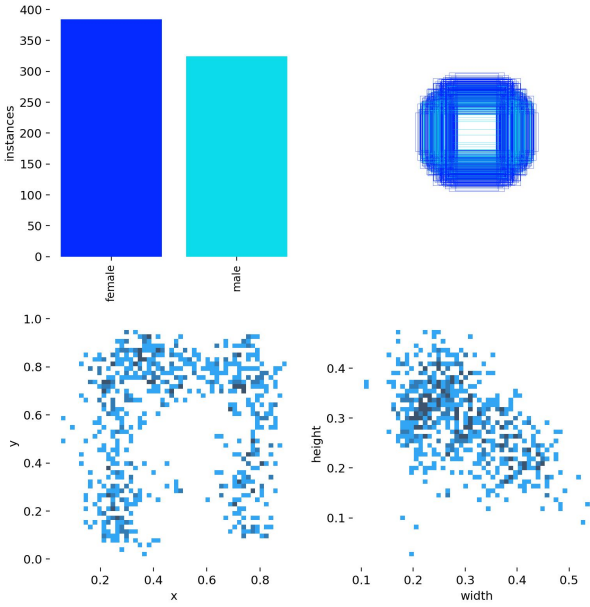


Fig. 6: Distribución espacial de los grillos en el dataset.

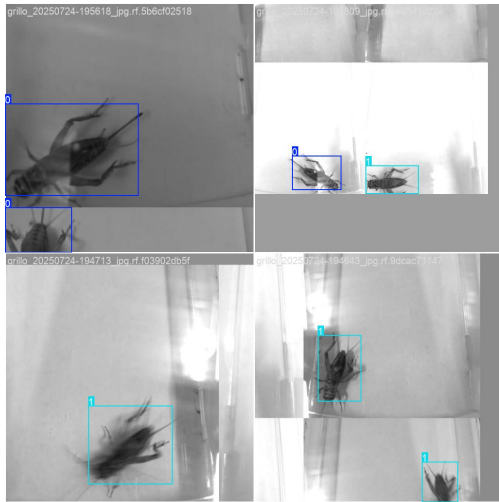


Fig. 7: Imágenes de entrenamiento.

namiento (81%), 63 de validación (9%) y 73 de prueba (10%), todas exportadas en formato YOLOv8 con coordenadas normalizadas de las cajas delimitadoras.

1.3 Modelo de detección de objetos: YOLOv8 Nano

El modelo de detección de objetos empleado en este trabajo es YOLOv8 nano, la variante más ligera de la familia YOLOv8 desarrollada por Ultralytics [11]. Los modelos YOLO son detectores de una sola etapa (SSDs), capaces de predecir cajas delimitadoras y sus clases asociadas en una sola pasada hacia adelante a través de la red neuronal. Este diseño elimina la necesidad de pasos intermedios de propuesta de regiones, lo que resulta en tiempos de inferencia significativamente reducidos en comparación con detectores de dos etapas como Faster R-CNN.

YOLOv8 introduce varias mejoras respecto a versiones anteriores, incluyendo cabeceras desacopladas para regresión y clasificación, y una nueva arquitectura de *backbone* basada en CSPDarknet [15] con conexiones parciales (*Cross Stage Partial*, CSP). La versión nano está específicamente optimizada para su despliegue en dispositivos con recursos limitados, reduciendo los multiplicadores de profundidad y ancho de la red y, con ello, tanto el número de parámetros como la carga computacional. Esto permite la ejecución en tiempo real en la CPU de la Raspberry Pi 5 sin necesidad de aceleración mediante GPU externa.

En este trabajo, YOLOv8 nano fue entrenado localmente sobre el conjunto de datos en escala de grises descrito anteriormente. Durante la inferencia, el modelo procesa las imágenes capturadas por la cámara Raspberry Pi AI y, para cada fotograma, devuelve las coordenadas de las cajas delimitadoras detectadas, la clase predicha (macho o hembra) para cada una y la correspondiente confianza de dicha clasificación. Si no se detecta ningún grillo en el fotograma, el modelo devuelve: “no detection”.

El rendimiento de los modelos de detección de objetos se evalúa habitualmente mediante la *Intersección sobre Unión* (*IoU*) y la *Precisión Media Promedio* (*mAP*). *IoU* cuantifica la superposición entre la caja predicha B_p y la caja de referencia B_{gt} :

$$IoU = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|} \quad (1)$$

Una detección se considera correcta cuando $IoU \geq 0.5$. La *Precisión Promedio* (*AP*) de una clase es el área bajo su curva *Precisión-Exhaustividad*:

$$AP_c = \int_0^1 P_c(R) dR \quad (2)$$

y la *mAP* corresponde al promedio de las *AP* de todas las clases. En este trabajo, los resultados se reportan como $mAP@0.5$, es decir, las detecciones se consideran correctas si $IoU \geq 0.5$.

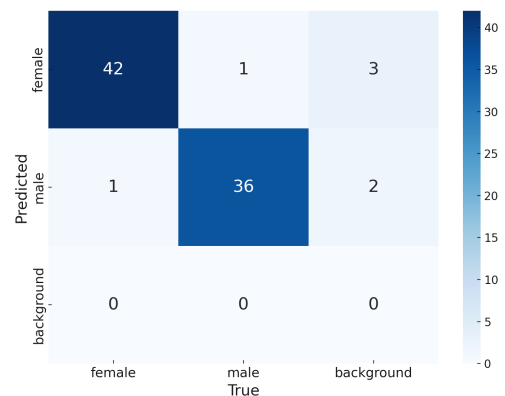


Fig. 8: Matriz de confusión obtenida a partir de imágenes de test.

Los resultados de entrenamiento son muy prometedores. La Figura 8 presenta la matriz de confusión obtenida durante la fase de prueba, mientras que la Figura 9 muestra

la curva Precisión–Exhaustividad utilizada para calcular el $mAP@0.5$, que alcanzó un valor de 0.977. La tasa de error fue mínima: solo 1 de 43 hembras (2.3%) fue clasificada como macho, y 1 de 37 machos (2.7%) fue clasificada como hembra, resultando en una precisión global del 97.5%. No obstante, el modelo generó cinco falsos positivos en regiones de fondo sin grillos presentes. Estos errores probablemente se debieron a ligeras variaciones en la iluminación y podrían mitigarse mediante un conjunto de datos más grande y diverso.

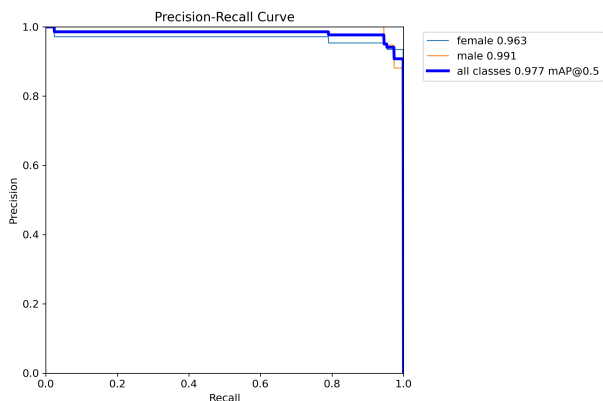


Fig. 9: Curva Precisión–Exhaustividad del modelo YOLO entrenado durante la fase de testing.

2 Lógica de control y actuación

El sistema no se limita a determinar el sexo de cada grillo; también incorpora una lógica de control para decidir cuándo activar el mecanismo de separación, reduciendo así la probabilidad de acciones de clasificación incorrectas.

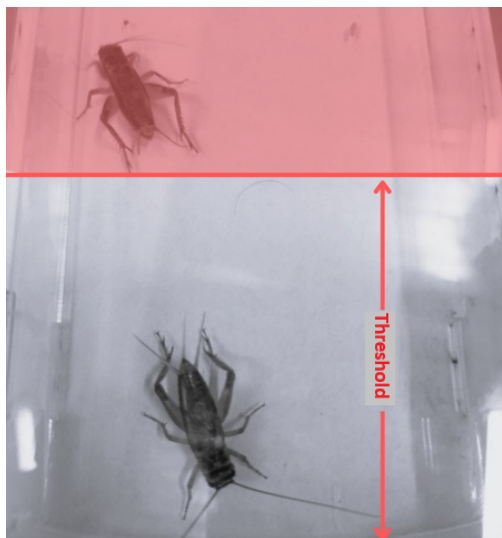


Fig. 10: Umbral de distancia. El grillo superior es ignorado.

El brazo de clasificación se activa únicamente cuando el grillo está lo suficientemente cerca del mecanismo. Como varios grillos pueden aparecer en el mismo fotograma, el

sistema prioriza al individuo más próximo al punto de separación. Para ello se define un umbral de proximidad fijo: si el grillo detectado se encuentra más allá de este límite, el brazo permanece inactivo. Este umbral se representa visualmente mediante una línea horizontal en la Figura 10, que indica la distancia mínima requerida (100 píxeles en este caso) para activar el actuador.

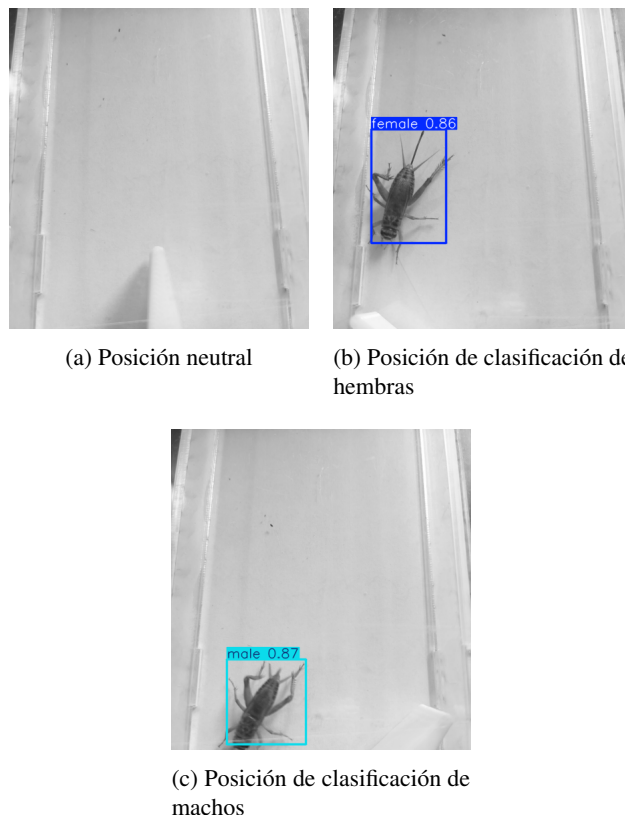


Fig. S1: Posiciones del brazo de clasificación.

Para evitar activaciones no deseadas causadas por clasificaciones erróneas ocasionales —un desafío inherente a los sistemas de detección de objetos— se aplica también una ventana deslizante de 10 fotogramas. Con una tasa de adquisición de aproximadamente un fotograma cada 300 ms, la ventana abarca unos tres segundos. El sistema calcula una predicción promedio sobre estos 10 fotogramas; el brazo solo se activa si la mayoría cumple simultáneamente con los criterios de clase y proximidad.

La Figura S1 ilustra las tres posiciones posibles del brazo de clasificación: neutral (inicial), posición de clasificación de hembras y posición de clasificación de machos.

3 EXPERIMENTOS Y RESULTADOS

Se llevaron a cabo cuatro experimentos, cada uno con aproximadamente 25 individuos de *Acheta domesticus*. El objetivo de estas pruebas de separación en tiempo real fue evaluar la eficiencia global de separación del dispositivo bajo diferentes condiciones de estrés. En todos los casos, BOX 1 se iluminó con una fuente de luz intensa y no contenía alimento ni agua, mientras que BOX 2 se mantuvo a

oscuras y disponía de alimento y agua, atrayendo así a los grillos.

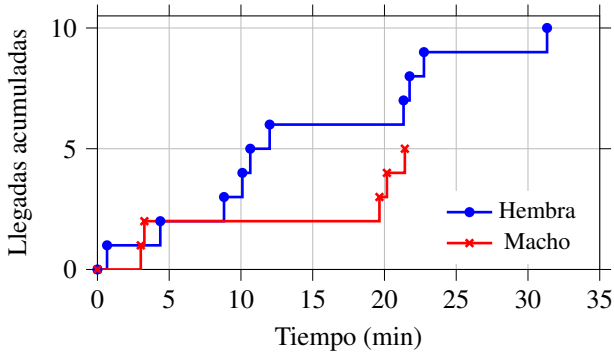


Fig. 12: Experimento 3: llegadas acumuladas por sexo.

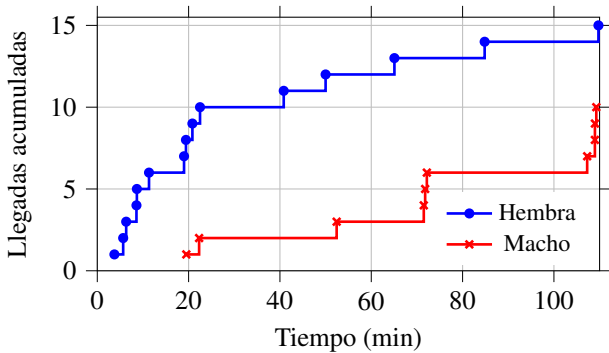


Fig. 13: Experimento 4: llegadas acumuladas por sexo.

En el Experimento 1, los grillos fueron forzados manualmente a cruzar el puente uno a uno. Este procedimiento generó un alto nivel de estrés, provocando desplazamientos nerviosos y a gran velocidad. Como resultado, en algunos casos el modelo YOLOv8 detectó correctamente al grillo, pero el brazo de clasificación no pudo reaccionar con suficiente rapidez, lo que condujo a errores de separación. En este contexto, definimos la *precisión de separación* como la precisión del sistema completo, que incluye tanto la clasificación basada en visión como la actuación mecánica del brazo. No debe confundirse con la precisión de detección del clasificador visual por sí solo. La precisión de separación se calcula como el cociente entre el número de grillos separados correctamente en el compartimento adecuado por el brazo y el número total de grillos procesados, esto es,

$$\text{Sorting Accuracy} = \frac{\text{Number of correctly sorted crickets}}{\text{Total number of sorted crickets}}$$

Para el Experimento 1, la precisión de separación alcanzada fue del 83%, considerablemente menor que el 94% obtenido en el Experimento 3. Aunque el desempeño fue inferior al de otros escenarios, la principal ventaja de este enfoque fue la rapidez: todos los grillos se separaron en solo 10 minutos. Sin embargo, su mayor desventaja fue la necesidad de intervención humana constante, lo que lo hace inadecuado para un sistema autónomo y escalable.

Table 1: Matriz de confusión para el Experimento 1.

	Pred. Macho	Pred. Hembra
True Macho	10	2
True Hembra	2	9

En el Experimento 2, en lugar de empujar individualmente a los grillos, se agitó BOX 1 para inducir movimiento colectivo. Esto provocó en *Acheta domestica* un nerviosismo muy similar al observado en el Experimento 1. Los resultados de clasificación (Tabla 2) fueron similares a los del Experimento 1, con una ventaja clave: la agitación de la caja puede automatizarse, evitando la interacción humana. La duración total del experimento fue de 15 minutos, muy similar a los 10 minutos del Experimento 1. Como se muestra en la Tabla 8, la precisión obtenida fue del 83%.

Table 2: Matriz de confusión para el Experimento 2.

	Pred. Macho	Pred. Hembra
True Macho	13	1
True Hembra	4	11

En el Experimento 3, solo se aplicaron estímulos ocasionales, consistentes en una ligera agitación de BOX 1 y suaves soplos de aire. Esto redujo el nivel de estrés de los grillos, permitiéndoles cruzar el puente con mayor calma. En consecuencia, la separación fue más fluida, como se refleja en la Tabla 3. No obstante, el menor ritmo incrementó significativamente la duración, requiriéndose 35 minutos para que cruzaran todos los individuos.

Table 3: Matriz de confusión para el Experimento 3.

	Pred. Macho	Pred. Hembra
True Macho	5	0
True Hembra	1	11

En el Experimento 4 no se aplicó ningún tipo de agitación, y los grillos cruzaron a su propio ritmo. Como se muestra en la Tabla 4, esta configuración de bajo estrés produjo las mejores métricas de clasificación, pero a costa de la eficiencia temporal. Tras 1 h y 50 minutos, solo el 72% de los grillos había cruzado y el experimento se dio por concluido. En estas condiciones no se puede garantizar que todos los individuos crucen el puente. Las llegadas acumuladas (Fig. 13) revelan una tasa de cruce decreciente: la mayor parte de los movimientos se produjo durante los primeros 30 minutos, mientras que el resto de individuos retrasaron o evitaron el cruce. Esto sugiere que, cuanto más grillos permanecen en la caja de partida, menor es su tendencia a moverse, posiblemente por confort, efectos de grupo o reticencia a explorar.

Para profundizar en las condiciones que afectaron al rendimiento de la separación, analizamos la velocidad individual de los grillos en relación con su sexo real y con si fueron clasificados correctamente o no por el sis-

Table 4: Matriz de confusión para el Experimento 4.

	Pred. Macho	Pred. Hembra
True Macho	8	1
True Hembra	1	13

Table 5: Resumen de métricas.

Exp	Acc	Prec	Rec	F1	Tiempo (min)
1	0.83	0.83	0.83	0.83	10
2	0.83	0.84	0.83	0.83	15
3	0.94	0.92	0.96	0.93	35
4	0.91	0.91	0.91	0.91	110

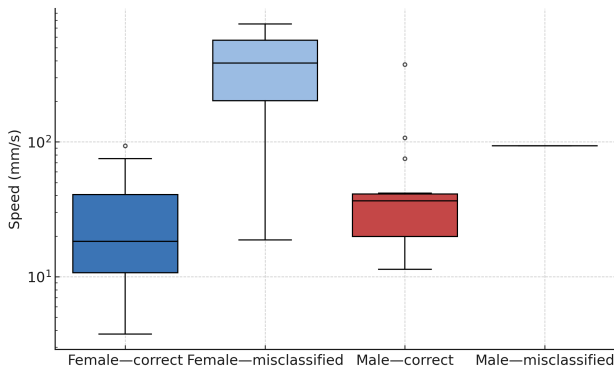


Fig. 14: Distribución de velocidades de los grillos (mm/s) por sexo y resultado de clasificación. Los mal clasificados, de ambos sexos, muestran medianas de velocidad mayores que los correctamente clasificados.

tema (Fig. 14). Las velocidades se obtuvieron midiendo el tiempo transcurrido desde el momento en que cada grillo apareció en la zona de detección del campo de visión de la cámara hasta que completó el cruce del puente, y dividiendo dicho tiempo entre la longitud del tramo recorrido (Tabla 6). Los resultados revelan dos tendencias principales. Primero, los machos fueron, en promedio, más rápidos que las hembras (mediana de machos: 39.5 mm/s; mediana de hembras: 18.8 mm/s). Segundo, los individuos mal clasificados, independientemente del sexo, tendieron a moverse a velocidades más altas que los correctamente separados (mediana de mal clasificados: 75.0 mm/s frente a 20.3 mm/s en correctamente clasificados). Estos hallazgos sugieren que una mayor velocidad de movimiento desafía la respuesta mecánica del brazo de separación, conduciendo a clasificaciones incorrectas.

La Tabla 7 presenta la matriz de confusión global obtenida al agregar los cuatro experimentos en tiempo real. Esta visión consolidada ofrece una imagen completa del rendimiento del dispositivo en distintas condiciones, resultando en una precisión global de separación del 86.8%.

Table 6: Velocidad medida de los grillos en los Experimentos 3 y 4.

Cricket	Sexo	Tiempo (s)	Clasif.	Velocidad (mm/s)
1	f	7.00	f	10.71
2	f	4.00	f	18.75
3	f	1.80	f	41.67
4	f	2.05	f	36.59
5	f	1.90	f	39.47
6	f	1.10	f	68.18
7	m	0.80	f	93.75
8	f	4.10	f	18.29
9	f	6.00	f	12.50
10	m	5.10	m	14.71
11	f	4.00	f	18.75
12	f	4.10	f	18.29
13	f	1.80	f	41.67
14	m	1.80	m	41.67
15	f	5.10	f	14.71
16	m	1.90	m	39.47
17	m	2.00	m	37.50
18	m	1.95	m	38.46
19	f	1.80	f	41.67
20	m	2.10	m	35.71
21	m	0.20	m	375.00
22	m	0.70	m	107.14
23	m	1.00	m	75.00
24	f	0.10	m	750.00
25	f	1.00	f	75.00
26	m	4.50	m	16.67
27	m	3.10	m	24.19
28	f	0.80	f	93.75
29	f	11.00	f	6.82
30	f	2.10	f	35.71
31	f	7.00	f	10.71
32	f	10.00	f	7.50
33	f	8.50	f	8.82
34	m	6.60	m	11.36
35	m	3.80	m	19.74
36	f	4.00	m	18.75
37	m	3.70	m	20.27
38	f	20.00	f	3.75
39	f	5.60	f	13.39
40	f	8.10	f	9.26

Table 7: Matriz de confusión global obtenida sumando los cuatro experimentos en tiempo real.

	Pred. Macho	Pred. Hembra
True Macho	36	6
True Hembra	6	43

4 CONCLUSIONES Y TRABAJO FUTURO

Este trabajo demuestra la viabilidad de un dispositivo autónomo y de bajo coste para la separación en tiempo real por sexo de *Acheta domesticus*, integrando visión por computador y control embebido. Los experimentos revelaron que, bajo condiciones de bajo estrés, el sistema alcanzó precisiones superiores al 90%, mientras que en condiciones de alto estrés se obtuvo un mayor ritmo de procesamiento a costa de un menor rendimiento en la clasificación.

Table 8: Métricas por clase derivadas de la matriz de confusión global.

	Macho	Hembra
Precisión	0.857	0.878
Exhaustividad	0.857	0.878
F1-score	0.857	0.878
Especificidad	0.878	0.857
Exactitud	0.868	0.868

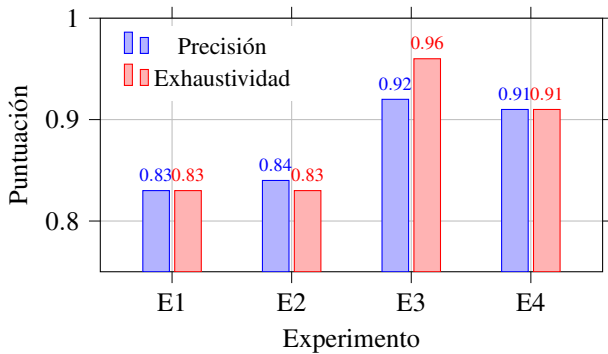


Fig. 15: Precisión y exhaustividad por experimento.

Hasta donde alcanza nuestro conocimiento, este es el primer prototipo capaz de separar físicamente grillos domésticos por sexo en tiempo real. Un sistema de este tipo tiene el potencial de contribuir significativamente a la industria de cría de insectos en rápida expansión, al permitir la selección genética, la optimización de las proporciones sexuales para la reproducción y el ajuste de las estrategias nutricionales.

Se identificaron limitaciones, incluyendo la influencia de factores de estrés ambiental, los reflejos de luz en la estructura acrílica y el ruido del servomotor, que en ocasiones desalentaba el cruce de los grillos. Una observación adicional de los Experimentos 3 y 4 es que, debido al estado de calma de los grillos, algunos se sobresaltaban con el ruido del brazo separador mientras cruzaban el puente, retrocediendo con frecuencia hacia BOX 1 y retrasando el proceso. Curiosamente, los machos parecieron ser más asustadizos que las hembras: en los dos últimos experimentos, donde los grillos pudieron circular libremente, los machos fueron de forma consistente los últimos en cruzar el puente, aparentemente más afectados por el sonido del brazo de clasificación.

El trabajo futuro se centrará en mejorar la robustez y escalabilidad del hardware, así como en probar actuadores alternativos. Además, el método de estimulación aplicado para inducir movimiento puede automatizarse aún más, por ejemplo mediante la introducción de paneles térmicos controlables en la base de BOX 1 o pequeños ventiladores, evitando la intervención humana. También serán necesarios experimentos adicionales con poblaciones más grandes de grillos para aumentar la fiabilidad y la significación estadística de los resultados. Finalmente, extender el enfoque a otras especies de insectos y evaluar métricas de eficiencia energética permitirá explorar con mayor profundidad el

potencial industrial de la separación automatizada de insectos.

5 REFERENCES

- [1] M. Montowska, P. Łukasz Kowalczewski, I. Rybicka, E. Fornal, “Nutritional value, protein and peptide composition of edible cricket powders,” *Food Chemistry*, vol. 289, pp. 130–138 (2019), doi:<https://doi.org/10.1016/j.foodchem.2019.03.062>, URL <https://www.sciencedirect.com/science/article/pii/S0308814619305394>.
- [2] Meticulous Market Research Private Limited, “Edible Insects Market to Reach \$28.54 Billion by 2035, Growing at an Exceptional CAGR of 28.1% from 2025,” (2025), press release, June 29, 2025.
- [3] “Commission Implementing Regulation (EU) 2022/188 of 10 February 2022 authorising the placing on the market of frozen, dried and powder forms of house cricket (*Acheta domesticus*) as a novel food,” https://eur-lex.europa.eu/eli/reg_imp/2022/188/oj/eng (2022), official Journal of the European Union, L 31/181, 11 February 2022.
- [4] A. van Huis, *Insects As Food and Feed: From Production to Consumption* (Wageningen Academic Publishers) (2017), URL <https://doi.org/10.3920/978-90-8686-296-2>.
- [5] R. L. Patton, “Growth and Development Parameters for *Acheta domesticus*1,2,” *Annals of the Entomological Society of America*, vol. 71, no. 1, pp. 40–42 (1978 01), doi:10.1093/aesa/71.1.40, URL <https://doi.org/10.1093/aesa/71.1.40>.
- [6] R. C. del Castillo, “The quantitative genetic basis of female and male body size and their implications on the evolution of body size dimorphism in the house cricket *Acheta domesticus* (Gryllidae),” *Genetics and Molecular Biology*, vol. 28, no. 4, pp. 843–848 (2005), doi:10.1590/S1415-47572005000500030, URL <https://www.scielo.br/j/gmb/a/zVQwZ33LX8CzcDXskSphfSQ/?lang=en>.
- [7] Y. Gutiérrez, M. Fresch, D. Ott, J. Brockmeyer, C. Scherber, “Diet composition and social environment determine food consumption, phenotype and fecundity in an omnivorous insect,” *Royal Society Open Science*, vol. 7, no. 4, p. 200100 (2020).
- [8] J. Stångberg, E. Immonen, P. P. Moreno, E. Bolund, “Experimentally induced intrasexual mating competition and sex-specific evolution in female and male nematodes,” *Journal of Evolutionary Biology*, vol. 33, no. 12, pp. 1677–1688 (2020), doi:10.1111/jeb.13706, URL <https://doi.org/10.1111/jeb.13706>, epub 2020 Oct 1.
- [9] N. Giulietti, P. Castellini, C. Truzzi, B. Ajdini, M. Martarelli, “Vision Measurement System for Gender-Based Counting of *Acheta domesticus*,” *Sensors*, vol. 24, no. 15, p. 4936 (2024).
- [10] M. F. Hansen, A. Oparaeke, R. Gallagher, A. Karimi, F. Tariq, M. L. Smith, “Towards Machine Vision for Insect Welfare Monitoring and Behavioural

Insights,” *Frontiers in Veterinary Science*, vol. Volume 9 - 2022 (2022), doi:10.3389/fvets.2022.835529, URL <https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2022.835529>.

[11] G. Jocher, J. Qiu, A. Chaurasia, “Ultralytics YOLO,” (2023 Jan.), URL <https://github.com/ultralytics/ultralytics>.

[12] Raspberry Pi Foundation, “Raspberry Pi Documentation,” <https://www.raspberrypi.org/documentation/> (2025), accessed: 2025-08-23.

[13] C. S. Classification, “auto-cricket Dataset,” (2025 jul), URL https://universe.roboflow.com/cricket-sex-classification/auto_cricket_s-cf8wj.

[14] B. Dwyer, J. Nelson, T. Hansen, *et al.*, “Roboflow (Version 1.0) [Software],” <https://roboflow.com> (2025), computer vision.

[15] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” (2020).

[16] J. Cantarero, “Cricket-Pipeline [Source Code],” <https://github.com/JuanCantarero/cricket-pipeline> (2025).

APÉNDICE

El conjunto de datos y el código fuente que respaldan los hallazgos de este trabajo están disponibles públicamente en Roboflow y GitHub [13, 16].

NOMENCLATURE

AD = *Acheta domesticus*

YOLO = You Only Look Once

IoU = Intersection over Union

mAP = mean Average Precision

R-CNN = Region-based Convolutional Neural Network

B_p = Predicted bounding box

B_{gt} = Ground truth bounding box

AP = Average Precision

VGG = Visual Geometry Group

SSD = Single-Shot Detector

Real-time Cricket Sorting By Sex

A low-cost embedded solution using YOLOv8 and Raspberry Pi

Juan Manuel Cantarero Angulo, *Artificial Intelligence Master's Student* and Dr. Matthew Smith, *Supervisor*
(jcantareroangulo@hawk.illinoistech.edu) (msmith84@illinoistech.edu)

Illinois Institute of Technology, Chicago, Illinois
Universidad Politécnica de Madrid, Madrid, Spain

The global demand for sustainable protein sources is driving increasing interest in edible insects, with *Acheta domesticus* (house cricket) identified as one of the most suitable species for industrial production. Current farming practices typically rear crickets in mixed-sex populations without automated sex sorting, despite potential benefits such as selective breeding, optimized reproduction ratios, and nutritional differentiation. This work presents a low-cost, real-time system for automated sex-based sorting of *Acheta domesticus*, combining computer vision and physical actuation. The device integrates a Raspberry Pi 5 with the official Raspberry AI Camera and a custom YOLOv8 nano object detection model, together with a servo-actuated sorting arm. The model reached a mean Average Precision at IoU 0.5 ($mAP@0.5$) of 0.977 during testing, and real-world experiments with groups of crickets achieved an overall sorting accuracy of 86.8%. These results demonstrate the feasibility of deploying lightweight deep learning models on resource-constrained devices for insect farming applications, offering a practical solution to improve efficiency and sustainability in cricket production.

Keywords: YOLO, Embedded computer vision, Insect farming, Object detection, Raspberry Pi, Real-time sorting, *Acheta domesticus*, Edge AI

0 INTRODUCTION AND RELATED WORK

The global demand for sustainable protein sources is rapidly increasing, driven by population growth, environmental concerns, and the limitations of conventional livestock production. Insects, in particular, have been highlighted by the FAO and numerous researchers as a highly efficient alternative, requiring significantly less land, water, and feed while producing lower greenhouse gas emissions [1]. Within this context, the insect protein market is projected to expand from approximately USD 2.4 billion in 2025 to about USD 28.54 billion by 2035, representing a compound annual growth rate of 28.1% [2].

Among edible insects, *Acheta domesticus*, commonly known as the house cricket, has emerged as one of the leading candidates for industrial-scale farming due to its fast growth cycle, palatability, and acceptance across different cultures. In 2022, it was authorized as a novel food by the European Union (2022/188) [3]. Furthermore, *Acheta domesticus* has been identified as one of the seven insect species suitable for industrial-scale production and human consumption [4].

Currently, house crickets are reared in mixed-sex populations under controlled environmental conditions [5]. However, no significant industrial efforts have been made

to sort them by sex. Sex-based sorting could provide multiple advantages, including selective breeding, optimization of male-to-female ratios for reproduction, and nutritional improvements, as female crickets are generally larger [6] and richer in protein [7] compared to males. Excess males, on the other hand, can generate competition, stress, and even cannibalistic behavior within colonies [8]. Bridging this gap is critical to support the scalability of insect farming.

Previous research has been conducted on *Acheta domesticus* gender identification and counting. For instance, Giulietti et al. [9] developed a vision-based measurement system to estimate the number and gender distribution of crickets using a test bench setup. Hansen et al. [10] also applied state-of-the-art deep learning models (YOLOv5 for detection and VGG-16 for sex classification) to crickets, achieving high accuracy in sex identification but without integrating a physical sorting mechanism. However, both works focused solely on classification tasks, without providing a fully automated, real-time sorting pipeline.

In this paper, we go beyond these previous approaches by presenting a proof-of-concept prototype that not only classifies but also physically separates house crickets by sex in real time. The system integrates a YOLOv8 nano object detection model [11] embedded on a Raspberry Pi

5 [12], which controls a servo-actuated sorting arm. This design introduces the novel contribution of combining embedded computer vision with a mechanical sorting feature, enabling low-cost, autonomous operation suitable for small-scale insect farming and providing a foundation for more advanced industrial applications.

1 METHODOLOGY

The proposed system enables real-time classification and separation of AD by combining computer vision with embedded physical control. The hardware setup includes a Raspberry Pi 5, the official Raspberry Pi AI Camera, and a robotic arm actuated by a servo motor SG90. The software component is based on a lightweight object detection model (YOLOv8 nano), trained on a custom dataset, and integrated with control logic that activates the arm according to the predicted sex of the cricket.

The crickets are encouraged to move from an initial chamber (BOX 1) to a target chamber (BOX 2) via a transparent acrylic bridge. As a cricket crosses the bridge, the camera continuously captures its frames and sends them to the object detection model, which processes them in real time to determine both the cricket’s position and sex. Based on these prediction, the system commands the servo motor to activate a flipper-like mechanism, directing the individual to either the male or female compartment within BOX 2.

1.1 Hardware and Physical Setup

The walls of BOX 1 and BOX 2 are made of dark acrylic, while the connecting bridge is made of transparent acrylic (Figure 1); all components were designed in CAD and fabricated using a laser cutter. The sorting arm was also designed using CAD software and manufactured with a 3D printer.

A single camera (Raspberry Pi AI Camera) is mounted in a fixed overhead position above the transparent bridge, as shown in Fig. 1, aligned with the sorting arm and the entrance to BOX 2. This placement allows the system to detect crickets just before they reach the sorting point, enabling timely activation of the arm. To enhance visual contrast and facilitate classification, the floor of the acrylic bridge has been painted white, providing a uniform background for image processing.

The sorting mechanism is driven by an SG90 servomotor (Fig. 2), programmed to operate in two fixed positions (Fig. S1). When the detection model identifies a female, the sorting arm blocks access to the male compartment, directing the cricket into the female section (Fig. S1b). Conversely, if a male is detected, the arm switches to the opposite position (Fig. S1c). This non-invasive mechanism enables automated sorting without direct contact, minimizing the risk of injury to the insects.

To encourage the crickets to move through the system, two contrasting environments are created: BOX 1 contains an intense light source and lacks food and water, making it inhospitable, while BOX 2 remains dark and contains both

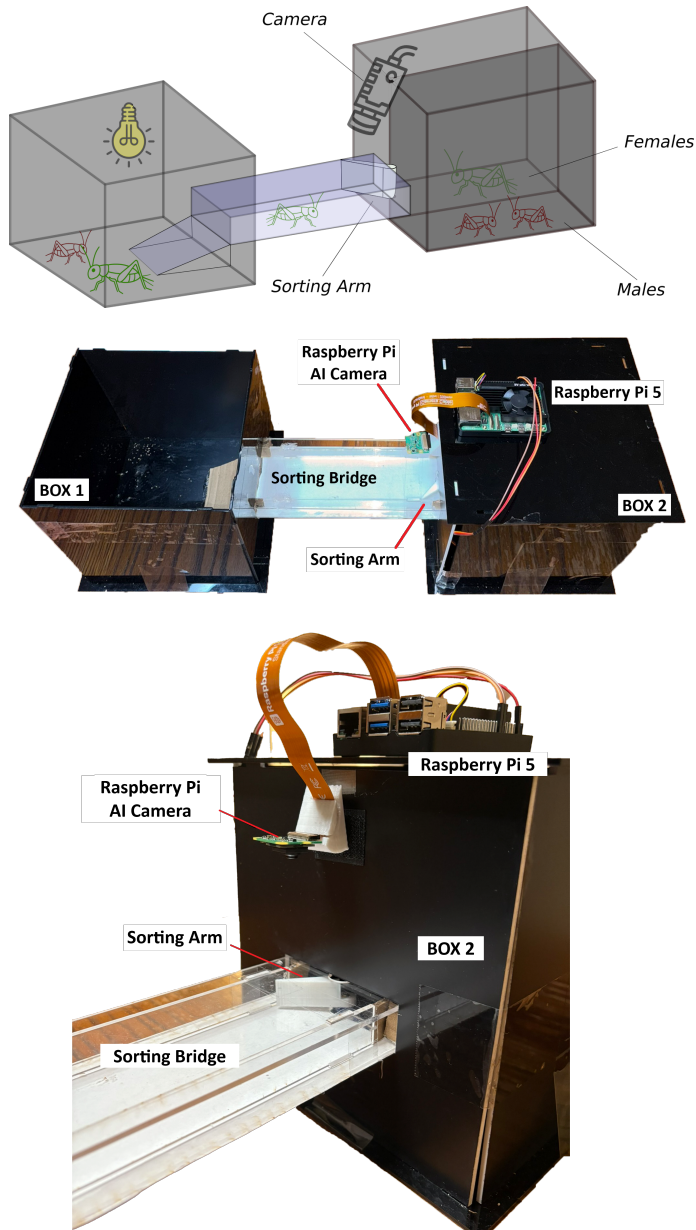


Fig. 1: Overview of the cricket sorting prototype: **top** — drawing of the device; **center** — top view of the sorting device; **bottom** — Raspberry Pi AI camera and sorting arm.

food and moisture, thereby attracting the crickets naturally toward the destination chamber.

1.2 Dataset and Annotation

To train and validate the cricket detection model, a custom dataset was developed and made publicly available on Roboflow [13]. The dataset initially comprised 335 images with a resolution of 480×480 pixels, capturing *Acheta domesticus* individuals of both sexes as they approached the sorting arm. All images were acquired using the Raspberry Pi AI Camera, mounted in a fixed overhead position identical to that used during the experimental phase, ensuring consistency between data collection and deployment conditions.

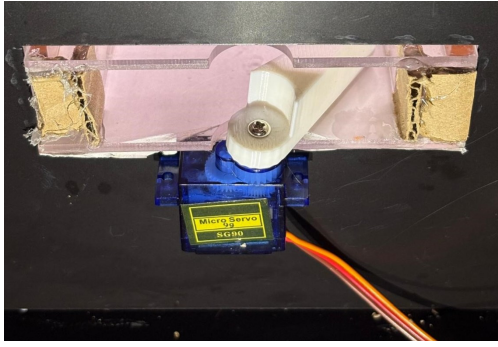


Fig. 2: Servomotor and 3D-printed sorting arm located at the end of the bridge.

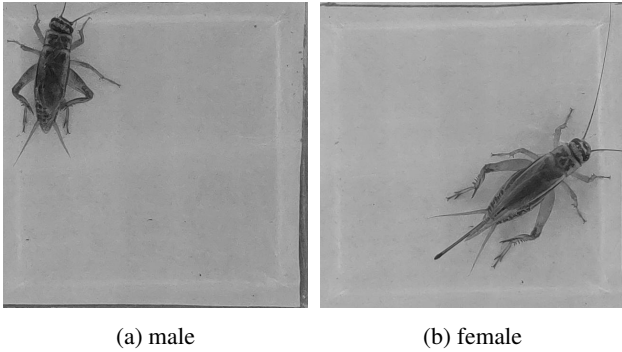


Fig. 4: Comparison between male and female *Acheta domestica*; note the ovipositor.

To ensure robustness under real-world conditions, various light sources were used during image acquisition, placed at different angles to compensate for reflections caused by the acrylic surface. This introduced controlled variability in illumination, helping the model generalize to different lighting scenarios. Additionally, images were taken with one, two, or even three (see figure 5) crickets of different sexes crossing the bridge simultaneously, increasing the diversity and complexity of the dataset.



Fig. 5: Training image containing three *Acheta domestica*.

Figure 6 provides an overview of the dataset. It first highlights that the dataset is balanced, containing approximately the same number of male and female individuals. It also illustrates the spatial distribution of crickets within the images, showing that *Acheta domestica* often tend to walk close to the bridge walls, as well as their relative sizes (normalized between 0 and 1). This variability in both position and scale contributes to the robustness of the dataset and supports the expectation of reliable training outcomes.

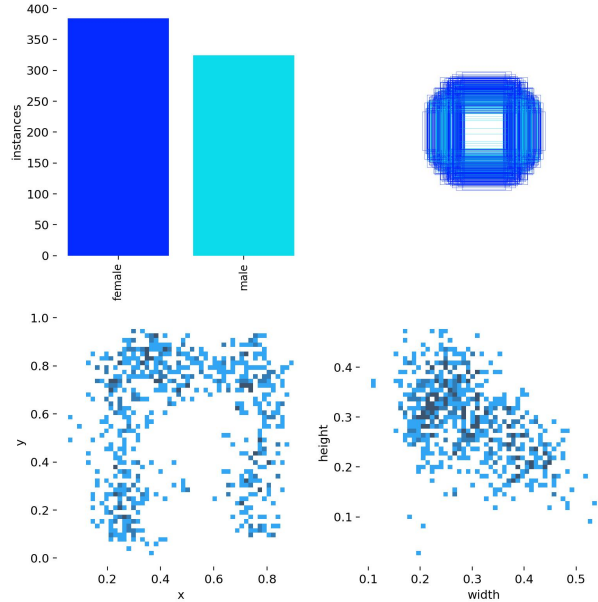


Fig. 6: Spatial distribution of crickets in the dataset.

Each image was converted to grayscale to reduce the input tensor to a single channel, thereby optimizing inference speed without significantly impacting detection accuracy. Following that, all images were annotated manually using Roboflow [14]. For each cricket, a bounding box was drawn and labeled according to its sex. Sex identification is straightforward to the human eye, with females being generally larger and exhibiting a clearly visible ovipositor at the rear end of the abdomen, as shown in figure 4.

To increase the size and variability of the dataset, a data augmentation process was applied. A random zoom was introduced in the range of 10% to 26% to simulate partial occlusions, such as cases where a cricket is only partially visible due to the frame or overlapping individuals. After augmentation, the final dataset contained 597 training images (81%), 63 validation images (9%), and 73 test images (10%), all exported in YOLOv8 format with normalized bounding box coordinates.

1.3 Object Detection Model: YOLOv8 Nano

The object detection model employed in this work is YOLOv8 nano, the lightest variant of the YOLOv8 family developed by Ultralytics [11]. YOLO models are single-shot detectors (SSDs) capable of predicting bounding boxes and their associated classes in a single forward pass through the neural network. This design eliminates

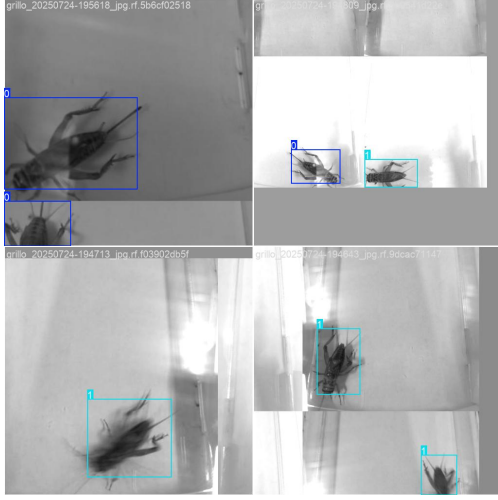


Fig. 7: Training images.

the need for intermediate region proposal steps, resulting in significantly reduced inference times compared to two-stage detectors such as Faster R-CNN.

YOLOv8 introduces several enhancements over previous versions, including decoupled heads for regression and classification, and a redesigned backbone based on CSP-Darknet [15] with Cross Stage Partial (CSP) connections. The nano version is specifically optimized for deployment on resource-constrained devices by reducing the depth and width multipliers of the network, thereby decreasing both the number of parameters and the computational load. This enables real-time execution on the Raspberry Pi 5 CPU without the need for external GPU acceleration.

In this work, YOLOv8 nano was trained locally on the grayscale dataset described before. During inference, the model processes images captured by the Raspberry Pi AI Camera and, for each frame, outputs the coordinates of the detected bounding box(es), the predicted class (male or female) for each box, and the corresponding confidence score. If there is no cricket detected on the frame, then the model returns: "no detection".

Object detection model performance is commonly evaluated using *Intersection over Union (IoU)* and *mean Average Precision (mAP)*. IoU quantifies the overlap between the predicted box B_p and the ground truth box B_{gt} :

$$IoU = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|} \quad (1)$$

A detection is considered correct when $IoU \geq 0.5$. The *Average Precision (AP)* of a class is the area under its Precision–Recall curve:

$$AP_c = \int_0^1 P_c(R) dR \quad (2)$$

and the *mAP* is the mean of all AP classes. In this work, results are reported as *mAP@0.5*, i.e., detections are correct if $IoU \geq 0.5$.

Training results are highly promising. Figure 8 presents the confusion matrix obtained during the test phase, while Figure 9 shows the Precision–Recall curve used to calcu-

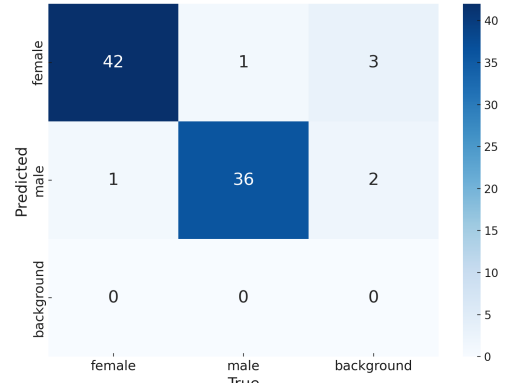


Fig. 8: Confusion matrix obtained from test images.

late the model’s *mAP@0.5*, which reached 0.977. Misclassification was minimal: only 1 out of 43 females (2.3%) was classified as male, and 1 out of 37 males (2.7%) as female, resulting in an overall accuracy of 97.5%. Nevertheless, the model produced five false positives in background regions with no crickets present. These errors are likely due to subtle variations in lighting conditions and could be mitigated by expanding the dataset with a larger and more diverse set of images.

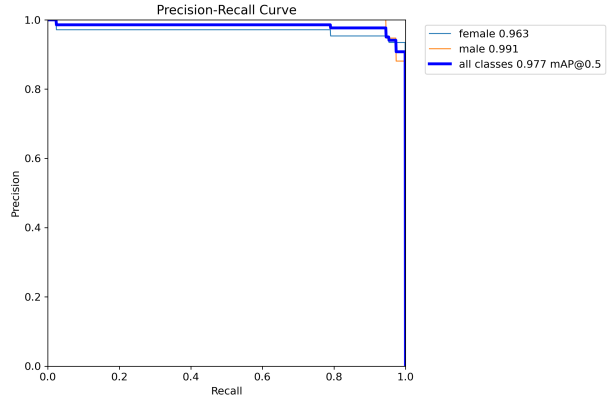


Fig. 9: Precision–Recall curve of the trained YOLO model during test.

2 Control Logic and Actuation

The system is not limited to determining the sex of each cricket; it also incorporates control logic to decide when to activate the sorting mechanism, thereby reducing the likelihood of incorrect sorting actions.

The sorting arm is triggered only when the cricket is sufficiently close to the mechanism. As multiple crickets may appear in the same frame, the system prioritizes the individual closest to the sorting point. For this purpose, a fixed proximity threshold is defined. If the detected cricket is located beyond this threshold, the arm remains inactive. This threshold is visually represented by a horizontal reference line in Figure 10, which indicates the minimum detection

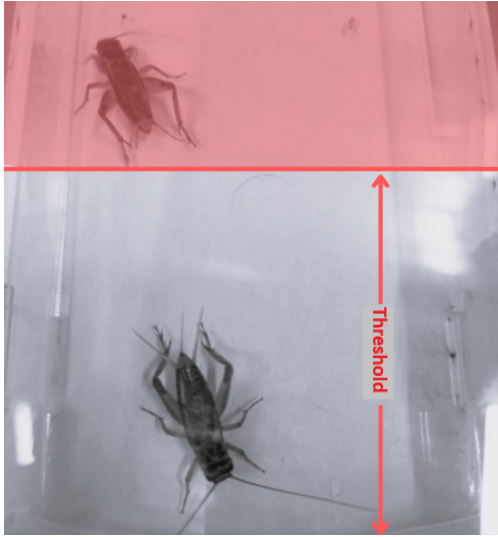
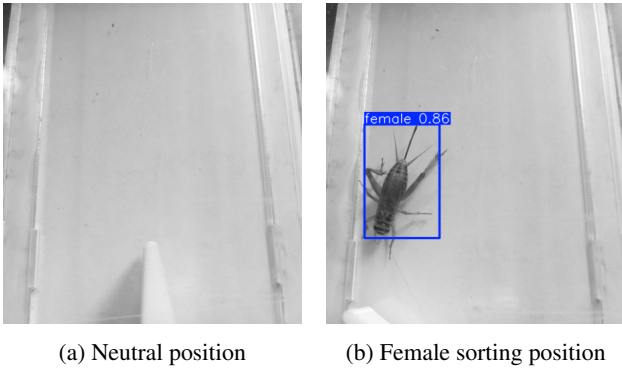


Fig. 10: Distance threshold. The upper cricket is ignored.

distance required to activate the actuator (100 pixels in this case).



(a) Neutral position

(b) Female sorting position



(c) Male sorting position

Fig. S1: Sorting arm positions.

To prevent unintended activations caused by occasional misclassifications—an inherent challenge in object detection systems—a sliding window of 10 frames is also applied. With an acquisition rate of approximately one frame every 300 ms, the window spans roughly three seconds. The system computes an average prediction over these 10 frames; the arm is activated only if the majority of them meet both the class and proximity criteria.

Figure S1 illustrates the three possible positions of the sorting arm: neutral (initial), female sorting position, and male sorting position.

3 EXPERIMENTS AND RESULTS

Four experiments were conducted, each involving approximately 25 *Acheta domesticus*. The objective of these real-time sorting trials was to assess the overall efficiency of the device under different stress conditions. In all cases, BOX 1 was illuminated with an intense light source and contained no food or water, while BOX 2 was kept dark and provided with food and water, thus attracting the crickets.

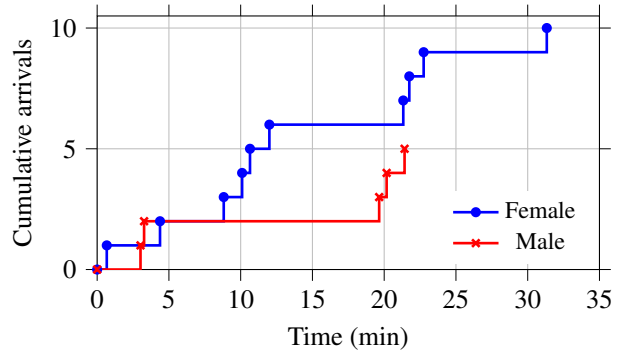


Fig. 12: Experiment 3: cumulative arrivals by sex.

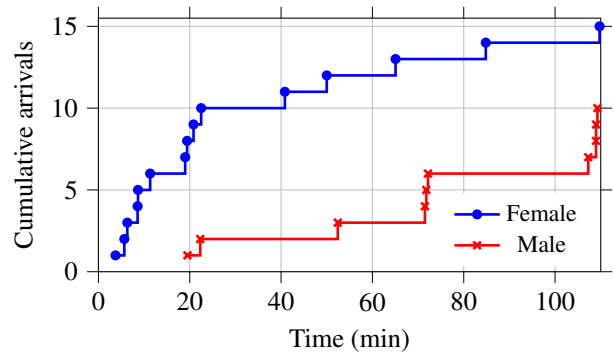


Fig. 13: Experiment 4: cumulative arrivals by sex.

In Experiment 1, crickets were manually forced to cross the bridge one by one. This procedure generated high stress, causing them to move nervously and at high speed. As a result, in some cases the YOLOv8 model detected the cricket correctly, but the sorting arm could not react fast enough, leading to misclassifications. In this context, we define the *sorting accuracy* as the accuracy of the entire system, which includes both the vision-based classification and the mechanical actuation of the sorting arm. It should not be confused with the object detection accuracy of the classifier alone. Sorting accuracy is calculated as the ratio between the number of crickets that were successfully separated into the correct compartment by the arm and the total number of crickets sorted, i.e.,

$$\text{Sorting Accuracy} = \frac{\text{Number of correctly sorted crickets}}{\text{Total number of sorted crickets}}$$

For Experiment 1, the sorting accuracy achieved was 83%, considerably lower than the 94% obtained in Experiment 3. Although performance was lower compared to other scenarios, the main advantage of this approach was speed: all crickets were sorted in only 10 minutes. However, its major drawback was the requirement of constant human intervention, which makes the method unsuitable for an autonomous and scalable system.

Table 1: Confusion matrix for Experiment 1.

	Pred. Male	Pred. Female
True Male	10	2
True Female	2	9

In Experiment 2, instead of pushing the crickets individually, BOX 1 was shaken to induce collective movement. This induced in *Acheta domesticus* a nervousness very similar to that obtained in Experiment 1. The classification results (Table 2) were similar to those in Experiment 1, but with a key advantage: box agitation can be automated, thereby avoiding the need for human interaction. The total duration of the experiment was 15 minutes, very similar to the 10 minutes of Experiment 1. As shown in table 8, the accuracy obtained was 83%

Table 2: Confusion matrix for Experiment 2.

	Pred. Male	Pred. Female
True Male	13	1
True Female	4	11

In Experiment 3, only occasional stimuli were applied, consisting of light agitation of BOX 1 and gentle air blows. This reduced the stress level of the crickets, allowing them to cross the bridge more calmly. Consequently, sorting was smoother, as reflected in Table 3. However, the reduced pace significantly increased the duration, requiring 35 minutes for all individuals to cross.

Table 3: Confusion matrix for Experiment 3.

	Pred. Male	Pred. Female
True Male	5	0
True Female	1	11

In Experiment 4, no agitation was applied, and the crickets were allowed to cross at their own pace. As shown in Table 4, this low-stress configuration produced the best classification metrics, but at the expense of time efficiency. After 1h 50 minutes, only 72% of the crickets had crossed, and the experiment was concluded. Under these conditions not all individuals could be guaranteed to cross the bridge. The cumulative arrivals (Fig. 13) reveal a progressively declining crossing rate: most movements occurred during the first 30 minutes, while the remaining individuals delayed or entirely avoided crossing. This suggests that the more crickets remain in the starting box, the lower their tendency

to move, possibly due to comfort, group effects, or reluctance to explore.

Table 4: Confusion matrix for Experiment 4.

	Pred. Male	Pred. Female
True Male	8	1
True Female	1	13

Table 5: Metrics summary.

Exp	Acc	Prec	Rec	F1	Time (min)
1	0.83	0.83	0.83	0.83	10
2	0.83	0.84	0.83	0.83	15
3	0.94	0.92	0.96	0.93	35
4	0.91	0.91	0.91	0.91	110

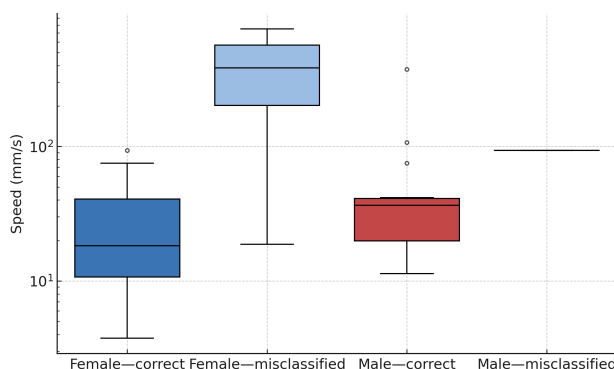


Fig. 14: Distribution of cricket speeds (mm/s) by sex and classification outcome. Misclassified individuals of both sexes exhibit higher median speeds compared to correctly classified ones.

To further investigate the conditions that affected sorting performance, we analyzed the speed of individual crickets in relation to their true sex and whether they were correctly or incorrectly classified by the system (Fig. 14). The speeds were obtained by measuring the time elapsed from the moment each cricket appeared in the detection zone of the camera’s field of view until it completely crossed the bridge, and dividing that time by the length of the traversed section (Table 6). The results reveal two main trends. First, males were on average faster than females (median male speed: 39.5 mm/s; median female speed: 18.8 mm/s). Second, misclassified individuals, regardless of sex, tended to move at higher speeds than those correctly sorted (median misclassified speed: 75.0 mm/s vs. 20.3 mm/s for correctly classified). These findings suggest that higher movement speed challenged the mechanical response of the sorting arm, leading to incorrect classifications.

Table 7 presents the global confusion matrix obtained by aggregating the four real-time experiments. This consolidated view provides a comprehensive picture of the device’s overall performance across different conditions, resulting in an overall sorting accuracy of 86.8%

Table 6: Measured speed of crickets in Experiments 3 and 4.

Cricket	Sex	Time (s)	Classified	Speed (mm/s)
1	f	7.00	f	10.71
2	f	4.00	f	18.75
3	f	1.80	f	41.67
4	f	2.05	f	36.59
5	f	1.90	f	39.47
6	f	1.10	f	68.18
7	m	0.80	f	93.75
8	f	4.10	f	18.29
9	f	6.00	f	12.50
10	m	5.10	m	14.71
11	f	4.00	f	18.75
12	f	4.10	f	18.29
13	f	1.80	f	41.67
14	m	1.80	m	41.67
15	f	5.10	f	14.71
16	m	1.90	m	39.47
17	m	2.00	m	37.50
18	m	1.95	m	38.46
19	f	1.80	f	41.67
20	m	2.10	m	35.71
21	m	0.20	m	375.00
22	m	0.70	m	107.14
23	m	1.00	m	75.00
24	f	0.10	m	750.00
25	f	1.00	f	75.00
26	m	4.50	m	16.67
27	m	3.10	m	24.19
28	f	0.80	f	93.75
29	f	11.00	f	6.82
30	f	2.10	f	35.71
31	f	7.00	f	10.71
32	f	10.00	f	7.50
33	f	8.50	f	8.82
34	m	6.60	m	11.36
35	m	3.80	m	19.74
36	f	4.00	m	18.75
37	m	3.70	m	20.27
38	f	20.00	f	3.75
39	f	5.60	f	13.39
40	f	8.10	f	9.26

Table 7: Global confusion matrix obtained by summing the four real-time sorting experiments.

	Pred. Male	Pred. Female
True Male	36	6
True Female	6	43

4 CONCLUSION AND FUTURE WORK

This work demonstrates the feasibility of a low-cost and autonomous device for real-time sex sorting of *Acheta domesticus*, integrating computer vision and embedded control. The experiments revealed that, under low-stress conditions, the system achieved accuracies above 90%, while

Table 8: Class-wise performance metrics derived from the global confusion matrix.

	Male	Female
Precision	0.857	0.878
Recall	0.857	0.878
F1-score	0.857	0.878
Specificity	0.878	0.857
Accuracy	0.868	0.868

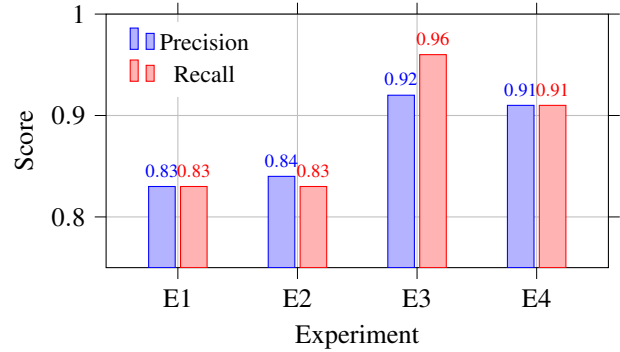


Fig. 15: Precision and recall per experiment.

high-stress conditions allowed faster throughput at the cost of lower classification performance.

To the best of our knowledge, this is the first prototype capable of physically separating house crickets by sex in real time. Such a system has the potential to contribute significantly to the rapidly expanding insect farming industry by enabling selective breeding, optimizing sex ratios for reproduction, and adjusting nutritional strategies.

Limitations were identified, including the influence of environmental stressors, light reflections from the acrylic setup, and the noise of the servo motor, which occasionally discouraged crickets from crossing. A further observation from Experiments 3 and 4 is that, due to the calm state of the crickets, some were startled by the noise of the sorting arm while crossing the bridge, often retreating back into BOX 1 and delaying the process. Interestingly, males appeared to be more skittish than females: in the last two experiments, where crickets were allowed to circulate freely, males were consistently the last to cross the bridge, seemingly more affected by the sound of the sorting arm.

Future work will focus on improving hardware robustness and scalability and testing alternative actuators. Moreover, the disturbing method applied to stimulate movement can be further automated, for example by introducing controllable heat pads in the base of BOX 1 or small fans, avoiding human interventions. Additional experiments with larger populations of crickets will also be necessary to increase the reliability and statistical significance of the results. Finally, extending the approach to other insect species and evaluating energy efficiency metrics will further explore the industrial potential of automated insect sorting.

5 REFERENCES

- [1] M. Montowska, P. Łukasz Kowalczewski, I. Rybicka, E. Fornal, “Nutritional value, protein and peptide composition of edible cricket powders,” *Food Chemistry*, vol. 289, pp. 130–138 (2019), doi:<https://doi.org/10.1016/j.foodchem.2019.03.062>, URL <https://www.sciencedirect.com/science/article/pii/S0308814619305394>.
- [2] Meticulous Market Research Private Limited, “Edible Insects Market to Reach \$28.54 Billion by 2035, Growing at an Exceptional CAGR of 28.1% from 2025,” (2025), press release, June 29, 2025.
- [3] “Commission Implementing Regulation (EU) 2022/188 of 10 February 2022 authorising the placing on the market of frozen, dried and powder forms of house cricket (*Acheta domesticus*) as a novel food,” https://eur-lex.europa.eu/eli/reg_imp/2022/188/oj/eng (2022), official Journal of the European Union, L 31/181, 11 February 2022.
- [4] A. van Huis, *Insects As Food and Feed: From Production to Consumption* (Wageningen Academic Publishers) (2017), URL <https://doi.org/10.3920/978-90-8686-296-2>.
- [5] R. L. Patton, “Growth and Development Parameters for *Acheta domesticus*1,2,” *Annals of the Entomological Society of America*, vol. 71, no. 1, pp. 40–42 (1978 01), doi:10.1093/aesa/71.1.40, URL <https://doi.org/10.1093/aesa/71.1.40>.
- [6] R. C. del Castillo, “The quantitative genetic basis of female and male body size and their implications on the evolution of body size dimorphism in the house cricket *Acheta domesticus* (Gryllidae),” *Genetics and Molecular Biology*, vol. 28, no. 4, pp. 843–848 (2005), doi:10.1590/S1415-47572005000500030, URL <https://www.scielo.br/j/gmb/a/zVQwZ33LX8CzcDXskSphfS/Q/?lang=en>.
- [7] Y. Gutiérrez, M. Fresch, D. Ott, J. Brockmeyer, C. Scherber, “Diet composition and social environment determine food consumption, phenotype and fecundity in an omnivorous insect,” *Royal Society Open Science*, vol. 7, no. 4, p. 200100 (2020).
- [8] J. Stångberg, E. Immonen, P. P. Moreno, E. Bolund, “Experimentally induced intrasexual mating competition and sex-specific evolution in female and male nematodes,” *Journal of Evolutionary Biology*, vol. 33, no. 12, pp. 1677–1688 (2020), doi:10.1111/jeb.13706, URL <https://doi.org/10.1111/jeb.13706>, epub 2020 Oct 1.
- [9] N. Giulietti, P. Castellini, C. Truzzi, B. Ajdini, M. Martarelli, “Vision Measurement System for Gender-Based Counting of *Acheta domesticus*,” *Sensors*, vol. 24, no. 15, p. 4936 (2024).
- [10] M. F. Hansen, A. Oparaeke, R. Gallagher, A. Karimi, F. Tariq, M. L. Smith, “Towards Machine Vision for Insect Welfare Monitoring and Behavioural Insights,” *Frontiers in Veterinary Science*, vol. Volume 9 - 2022 (2022), doi:10.3389/fvets.2022.835529, URL [\[veterinary-science/articles/10.3389/fvets.2022.835529\]\(https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2022.835529\).](https://www.frontiersin.org/journals/v</p></div><div data-bbox=)

- [11] G. Jocher, J. Qiu, A. Chaurasia, “YOLO,” (2023 Jan.), URL <https://github.com/ultralytics/ultralytics>.
- [12] Raspberry Pi Foundation, “Raspberry Pi Documentation,” <https://www.raspberrypi.org/documentation/> (2025), accessed: 2025-08-23.
- [13] C. S. Classification, “auto-cricket Dataset,” (2025 jul), URL https://universe.roboflow.com/cricket-sex-classification/auto_cricket_s-cf8wj.
- [14] B. Dwyer, J. Nelson, T. Hansen, *et al.*, “Roboflow (Version 1.0) [Software],” <https://roboflow.com> (2025), computer vision.
- [15] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” (2020).
- [16] J. Cantarero, “Cricket-Pipeline [Source Code],” <https://github.com/JuanCantarero/cricket-pipeline> (2025).

APPENDIX

The dataset and source code supporting the findings of this work are publicly available on Roboflow and GitHub [13, 16].

NOMENCLATURE

- AD = *Acheta domesticus*
YOLO = You Only Look Once
IoU = Intersection over Union
mAP = mean Average Precision
R-CNN = Region-based Convolutional Neural Network
 B_p = Predicted bounding box
 B_{gt} = Ground truth bounding box
AP = Average Precision
VGG = Visual Geometry Group
SSD = Single-Shot Detector