

Evaluación

Adrián Matute, Pablo Martínez, Osvaldo Del Valle, Andres Callealta, Jorge Martínez
Tecnológico de Monterrey Campus Querétaro, México

Introducción

Este reporte detalla la evaluación de los modelos desarrollados para la detección y monitoreo de vacas en la fila de ordeño, centrándose en métricas clave, objetivos de negocio y descubrimientos significativos. La etapa de evaluación permite validar si el modelo cumple con los objetivos propuestos y analizar áreas de mejora antes de una posible implementación a mayor escala.

A. Evaluación de Resultados

Los resultados obtenidos de los modelos se analizaron con respecto a los objetivos de negocio definidos en las fases iniciales del proyecto. Este análisis permite verificar si el modelo cumple con los criterios establecidos para la detección de patrones y la optimización del proceso de ordeño.

Objetivos de negocio evaluados:

1. Identificar patrones significativos de aglomeraciones en la fila de espera para el ordeño.
 - a. El socio formador determinará si las aglomeraciones afectan la eficiencia del proceso de ordeño o el bienestar animal.
2. Identificar los cuellos de botella en la fila de ordeño.
 - a. El socio formador determinará si los hallazgos proporcionan información útil para optimizar la programación e identificar los cuellos de botella en la fila.

Resultados Clave

Iteración de Modelado I:

Precisión del modelo híbrido: 95.14%.

Detección de cuellos de botella:

- Se identificaron momentos críticos de congestión en horas tempranas de la mañana y tardías por la noche.
- Esta información sugiere ajustes en la programación para evitar aglomeraciones, optimizando tanto la logística como el bienestar animal.

Iteración de Modelado II:

Precisión del modelo nocturno y diurno:

- **Modelo para condiciones diurnas:** 96.39%.
- **Modelo para condiciones nocturnas:** 92.9% (91.8% en la primera iteración).

Segmentación por iluminación:

- Las diferencias entre métricas diurnas y nocturnas refuerzan la necesidad de segmentar los modelos o usar el modelo híbrido para una mayor generalización.

Primera Iteración de Modelado III:

Precisión del modelo híbrido Faster R-CNN: 20.50%.

Segunda Iteración de Modelado III:

Precisión del modelo híbrido Faster R-CNN: 45.50%.

El modelo de Faster R-CNN muestra dificultades para detectar vacas en áreas pequeñas y medianas, lo que impide realizar un análisis detallado sobre las aglomeraciones. A pesar de obtener un mAP@50 decente, el modelo aún no es suficientemente robusto para detectar con precisión aglomeraciones de vacas en diferentes condiciones de iluminación, lo que afecta su capacidad para identificar patrones significativos en la fila.

Impacto en los Objetivos de Negocio

Cumplimiento parcial de objetivos:

- **Modelo diurno:** Resultados sólidos en condiciones óptimas de iluminación.
- **Modelo nocturno:** Aunque menos preciso, ofrece información relevante bajo condiciones adversas.
- **Modelo híbrido:** Mayor generalización y cumplimiento más integral de los objetivos planteados.
- **Modelo Faster R-CNN:** Muestra dificultades para detectar vacas en áreas pequeñas y medianas, esto impide realizar un análisis detallado sobre las aglomeraciones.

Beneficios potenciales:

- Mejor comprensión de los patrones de comportamiento del ganado.
- Generación de datos accionables para optimizar procesos.
- Identificación de áreas críticas para futuras intervenciones.

Interpretación de los Resultados

1. Fortalezas:

- **Modelos diurno e híbrido:** Alta precisión en entornos controlados y mayor adaptabilidad.
- **Modelo nocturno:** Desempeño aceptable considerando las limitaciones de iluminación
- Modelo Faster R-CNN: Precisión promedio en términos de identificar objetos en áreas grandes.

2. Áreas de oportunidad:

- Mejorar la detección en condiciones nocturnas mediante la recolección de más datos específicos y el uso de técnicas avanzadas de aumento de datos.
- Incorporar un preprocesamiento especializado para condiciones de baja visibilidad.

Descubrimientos Significativos

1. La congestión detectada sugiere que los momentos de mayor actividad coinciden con las horas más críticas del ordeño.
2. Existen patrones de comportamiento que podrían correlacionarse con estrés animal o ineficiencias en la logística.
3. Los datos recolectados podrían usarse para entrenar modelos adicionales de predicción del flujo en la fila.

Modelos Aprobados

1. Modelo diurno: Recomendado para condiciones de luz natural.
2. Modelo híbrido: Más versátil, recomendado para una implementación inicial en diferentes condiciones.
3. Modelo nocturno: Sujeto a mejoras, pero utilizable como complemento.

Proceso de Revisión

Actividades para la revisión:

Dar una visión general del proceso usado para la minería de datos

Analizar el proceso de minería de datos. Para cada fase del proceso analizar:

En retrospectiva ¿realmente fue necesaria?

¿Se ejecutó de manera óptima?

¿De qué manera puede ser mejorado el proceso?

Actividades para la revisión (cont.):

Identificar fallas

Identificar pasos engañosos

Identificar posibles acciones alternativas y caminos inesperados en el proceso

Revisar los resultados de minería de datos con respecto a los criterios de éxito de negocio

- Visión general del proceso para la minería de datos: al haber recolectado las 7000 imágenes que componen nuestro actual dataset, tuvimos que pasar mucho tiempo haciendo anotaciones en las imágenes manualmente. Esa larga tarea nos fue además útil para conocer nuestros datos. De esa manera supimos que algunas noches se apagan todas las luces, que algunas

veces hay manchas de leche en el suelo lo cual complica las cosas para nuestro modelo, y varias otras cosas.

Analizar cada fase del proceso:

1. Comenzamos con una fase de entendimiento del negocio, en este caso el CAETEC. Esta etapa fue crucial para alinear las metas del proyecto con los objetivos del negocio, definiendo como prioridades la optimización del proceso de ordeño y el monitoreo del bienestar animal. Estos objetivos, formulados bajo la metodología CRISP-DM, aseguraron que las soluciones desarrolladas fueran prácticas y alineadas con las necesidades específicas de CAETEC. Se logró una definición clara de metas específicas, como el conteo preciso de vacas y la identificación de cuellos de botella en las filas de espera, acompañada de una planeación efectiva ajustada a los recursos disponibles y a las expectativas del proyecto.

Sin embargo, se identificaron algunas fallas y oportunidades de mejora durante el proceso. La baja calidad de las imágenes nocturnas y las interrupciones por apagones de luz resultaron en una pérdida significativa de datos clave, lo que limitó la capacidad del modelo para generalizar. Además, se asumió erróneamente que la infraestructura tecnológica existente era suficiente para manejar el volumen de datos y las demandas computacionales del proyecto sin necesidad de ajustes.

Como acciones recomendadas, es fundamental realizar un análisis más profundo de los riesgos legales asociados con el uso de datos en entornos operativos, así como explorar mejoras tecnológicas que permitan una mayor resiliencia y adaptabilidad frente a condiciones adversas.

2. En la fase de entendimiento de los datos, el análisis resultó fundamental para identificar problemas críticos en el dataset, como la calidad inconsistente entre imágenes diurnas y nocturnas, así como las interrupciones en la captura de datos. Estos hallazgos permitieron establecer un panorama claro de los retos técnicos y operativos que impactan el desempeño de los modelos. Entre los logros más destacados se encuentra la identificación de problemas técnicos como imágenes oscuras o sobreexpuestas, que afectan significativamente la capacidad de los modelos para generalizar. Asimismo, se logró reconocer patrones específicos en los datos, como la dificultad para distinguir vacas en condiciones de poca luz.

Sin embargo, también se evidenciaron fallas y oportunidades de mejora. La baja calidad de las imágenes nocturnas, combinada con los apagones de luz, derivó en una pérdida de datos clave que limitó el análisis. Además, tratar las imágenes diurnas y nocturnas bajo un único enfoque sin considerar adecuadamente las diferencias en las condiciones lumínicas fue un paso engañoso que complicó el desarrollo de soluciones efectivas.

Como acciones recomendadas, se propuso utilizar modelos especializados para condiciones de día y noche, estrategia que fue implementada en iteraciones posteriores con éxito. También se sugirió aplicar técnicas de aumentación de datos y mejora de imágenes para mitigar los efectos de la baja calidad lumínica, mejorando así la estandarización y la efectividad del dataset.

3. Tras completar el entendimiento de los datos, se procedió a la fase de preparación de imágenes para garantizar que estuvieran en condiciones óptimas para ser utilizadas en el entrenamiento de los modelos. En esta etapa, se definieron estrategias claras para manejar datos problemáticos, como imágenes extremadamente oscuras o sobreexpuestas, con el objetivo de mejorar la calidad del dataset.

Entre los logros más destacados de esta fase se encuentra el manejo eficiente de valores atípicos, que incluyó la eliminación de imágenes de baja calidad que comprometían el entrenamiento del modelo. Además, la comprensión previa de los problemas potenciales permitió realizar una reparación rápida y efectiva del dataset, optimizando el tiempo dedicado a esta tarea.

No obstante, preparar las imágenes no estuvo exento de desafíos. Dado el estado inicial del dataset y las herramientas disponibles, se identificaron áreas de oportunidad para mejorar. Por ejemplo, un análisis más detallado de los valores atípicos mediante visualizaciones, como gráficos de distribución de luz, podría haber permitido tomar decisiones más informadas sobre qué datos eliminar.

Para proyectos futuros con problemáticas similares, se recomienda implementar análisis exploratorios visuales que faciliten la identificación de datos problemáticos de manera más precisa. Asimismo, sería valioso documentar de manera exhaustiva el impacto de la eliminación de imágenes en la calidad general del dataset, para evaluar cómo estas decisiones afectan los resultados finales de los modelos.

4. Modelado I: En esta primera fase de modelado, hemos tratado con un modelo bastante básico el cual es YOLOv8 nano. Probar diferentes modelos y arquitecturas es esencial en este tipo de proyecto, ya que todos son en teoría muy eficientes y no se puede saber con seguridad cuál es el mejor modelo para que problema.

No se puede realmente decir que hubo un fallo o un elemento destacadamente bueno en nuestro proceso, ya que hemos elegido uno de los modelos más comunes con los hyperparametros de base, lo que sí se puede decir es que esta fase fue necesaria.

5. Modelado II: Aquí, hemos decidido entrenar dos modelos: uno con fotos de día y otro con fotos de noche, oscuras. Esta fase si tuvo su importancia, ya que nos pudimos dar cuenta que la precisión del modelo es mucho más baja cuando son fotos con poca luz, y que es casi perfecto cuando está en buenas condiciones (de día y con pocas vacas, poca intersección). Pero aquí viene un punto negativo: si hay que hacer dos modelos para un mismo set de datos, es que el modelo no es lo suficientemente complejo para capturar patrones, pero en ese caso hubiéramos mejor entrenado la versión small, con tres veces más de parámetros que la versión nano. el mayor problema del proyecto se encuentra en esta fase de modelado II, en la cual pasamos mucho tiempo pensando que era la buena metodología, y entrenando diferentes modelos con diferentes hiper parámetros.

6. Modelado III: Esta tercera iteración, no salió como se esperaba considerando que el modelo híbrido de YOLO v8 demostró un desempeño significativamente superior en precisión, recall y mAP en todas las condiciones de iluminación. Además, el entrenamiento y ajuste de Faster R-CNN presentó limitaciones en tiempo, dificultando su capacidad para cumplir con los objetivos del proyecto. En futuros proyectos, tendríamos que priorizar modelos con mejor desempeño en condiciones mixtas, como YOLO.

Siguientes pasos:

De acuerdo a los resultados evaluados acerca de las iteraciones realizadas a lo largo del proyecto, se ha determinado la decisión de finalizar el proyecto y avanzar hacia la fase de entrega.

Para ser más precisos, la realización de diferentes modelos dentro de las dos arquitecturas evaluadas las cuales son YOLOv8 y Faster R-CNN, nos ayudó a recolectar la información necesaria para reconocer cuál modelo es el más adecuado para cumplir con los objetivos de negocio del CAETEC y los objetivos de minería de datos.

El proceso seguido por parte del equipo de trabajo también nos permitió tomar esta decisión, desde el inicio conseguimos definir de manera correcta y confiable los objetivos de negocio y de minería de datos con los que construimos el proyecto, de manera que cada uno de los pasos tomados fueron tomando en consciencia el resultado que queríamos obtener para el socio formador.

El buen entendimiento de los datos, su preparación y la construcción de cada una de las soluciones propuestas nos dieron datos cuantificables para reconocer como fuimos progresando hasta llegar a este punto, donde nuestro objetivo final es dar una entrega completa de la solución final.

1. El modelo de detección de objetos se puede extender para su uso en rastreo (tracking) de objetos.
 - a. **Razones a favor**
 - i. Generar datos de movimiento de las vacas nos podría permitir analizar más a detalle los patrones de aglomeración que se presentan en la fila.
 - ii. Puede que no se requiera de entrenar de un nuevo modelo de visión puesto a que el tracking utiliza los datos de la detección de objetos entre fotogramas y redes pre entrenadas junto con algoritmos ligeros en comparación con la detección de objetos para determinar el movimiento de las cajas.
 - b. **Razones en contra**
 - i. El dataset actual cuenta con fotografías tomadas cada 5 minutos, esto presenta un problema para evaluar la efectividad del rastreo de objetos ya que estos algoritmos requieren que los objetos detectados entre fotogramas se encuentren en proximidad. Esto implicaría esperar a conseguir fotos tomadas en intervalos más cortos ej. 5 segundos.
 - ii. Los algoritmos de rastreo utilizan redes neuronales pre entrenadas para extraer “mapas de características” que permiten comparar la similitud de los objetos. Y aunque estas pueden funcionar de manera decente, no es garantía que presenten una buena detección en el dominio del ganado sin antes haber realizado un proceso de entrenamiento.

1. Decisión

- a. *Posibles acciones*

- i. Analizar el potencial de entrega de cada resultado: tenemos un proyecto con alto potencial, en lo que es del plan visual, tenemos una interfaz que respeta los requerimientos del cliente, el cual la aprobó, y es relativamente agradable para una utilización futura. En lo que es del modelo y de sus resultados, es bastante robusto, con un 95% de precisión, lo cual es bastante complicado ya que en muchas ocasiones las condiciones de conteo son complicadas: las vacas se amontonan (muchas intersecciones) o la imagen es muy oscura y se hace difícil hasta para un humano poder contarlas.
- ii. Estimar el potencial de mejora del proceso actual: para lo que es del modelo, con un poco más de recursos o/y tiempo, se podría entrenar el modelo con una versión más compleja, ya que actualmente estamos usando la versión de YOLO las más pequeña (nano).
- iii. Revisar los recursos restantes para determinar si se permiten iteraciones del proceso adicionales o si puede disponerse de recursos adicionales: para lo que es del modelo, para un potencial refinamiento, lastimosamente no habrá tiempo, ya que estamos con recursos limitados (hablando de GPUs) y con el tiempo que es limitado (entrenar el modelo híbrido con nuestras más de 7000 imágenes nos toma más de 8 horas), pero por otro lado, el modelo híbrido actual, tiene métricas bastante buenas: 95% de precisión.
- iv. Recomendar continuaciones alternativas: ya que el objetivo principal del cliente es ampliar el proyecto expandiendo a otros ranchos, estaría bien entrenar el modelo con imágenes de otras fuentes. La versión de YOLO v8n que vamos a entregar si tiene una precisión de 95%, y si la va a tener cuando esté desplegada, pero en cuanto cambien la cámara de sitio, poniéndola por ejemplo en otro rancho, las métricas de predicciones serán las mismas.
- v. Refinar el plan del proceso: Obtener más imágenes nocturnas para poder entrenar mejor el modelo.

Decisión

Se recomienda avanzar con el modelo híbrido para implementación piloto, acompañado de las siguientes estrategias:

1. Captura de imágenes con menor intervalo de tiempo para explorar la capacidad de rastreo.
2. Refinar el modelo nocturno mediante aumento de datos y preprocesamiento específico.
3. Planear una segunda fase del proyecto que incluya entrenar con imágenes de otros ranchos, mejorando así la capacidad de generalización.

Este enfoque garantiza resultados más confiables y sostenibles en el largo plazo, además de alinearse con los objetivos del cliente.