

Reentrenamiento de Modelo Clasificador por Imágenes para Museografía Interactiva

Image Classifier Model Retraining for Interactive Museography

Presentación: 23 y 24 de octubre de 2024

Francisco Levrino

Universidad Tecnológica Nacional, Facultad Regional San Francisco, Museo Interactivo de Ciencias (MuIC), Argentina
flevrino@facultad.sanfrancisco.utn.edu.ar

Sofía B. Bovo

Universidad Tecnológica Nacional, Facultad Regional San Francisco, Museo Interactivo de Ciencias (MuIC), Argentina
sbovo@facultad.sanfrancisco.utn.edu.ar

Hugo A. Pipino

Universidad Tecnológica Nacional, Facultad Regional San Francisco, Museo Interactivo de Ciencias (MuIC), Argentina
hpipino@sanfrancisco.utn.edu.ar

Resumen

El Museo Interactivo de Ciencias (MuIC) de la Universidad Tecnológica Nacional, Facultad Regional San Francisco, combina recreación y aprendizaje a través de su muestra interactiva "ConCiencia". Para mejorar la accesibilidad y la independencia de los visitantes mediante guías audiovisuales, el MuIC desarrolla herramientas innovadoras a través del uso de dispositivos móviles. En este sentido, se implementa un sistema de reconocimiento por imágenes utilizando la red neuronal YOLO V8, con la misma se reentrena el modelo clasificador, mejorando las capacidades del modelo implementado con anterioridad. Esto aumenta la precisión en la detección y amplía la capacidad del modelo para reconocer las experiencias en diversos entornos. Además, se integran imágenes de nuevas experiencias para enriquecer la experiencia de los visitantes al interactuar con los módulos. Los resultados son prometedores y sugieren un potencial significativo para futuras mejoras, enriqueciendo la interacción y el aprendizaje en el MuIC.

Palabras clave: Modelo Clasificador, Experiencias Interactivas, YOLO, Museografía

Abstract

The Interactive Science Museum (MuIC) at National Technological University, San Francisco Regional Faculty, combines recreation and learning through its interactive exhibit "ConCiencia". To enhance accessibility and visitor independence through audiovisual guides, MuIC develops innovative tools using mobile devices. In this context, an image recognition system utilizing the YOLO V8 neural network is implemented, which retrains the existing classification model, improving its capabilities. This increases detection accuracy and expands the model ability to recognize experiences in various environments. Additionally, images of new experiences are integrated to enrich the visitor experience when interacting with the modules. The results are promising and suggest significant potential for future improvements, enhancing interaction and learning at MuIC.

Keywords: Classifier Model, Interactive Experience, YOLO, Museography

Introducción

El Museo Interactivo de Ciencias (MuIC) de la Universidad Tecnológica Nacional (UTN), Facultad Regional San Francisco, es un grupo de investigación conformado por docentes, estudiantes y graduados de varias carreras de ingeniería. Su muestra interactiva, "ConCiencia", presenta una variedad de experiencias relacionadas con diferentes áreas de la ciencia, con el objetivo de crear un espacio que combine recreación, estímulo intelectual y aprendizaje. Esta muestra está dirigida tanto a estudiantes de diversos niveles educativos como al público en general, de San Francisco y sus alrededores.

Dado que los objetivos del MuIC incluyen la adaptación constante a las nuevas tecnologías y sus aplicaciones, es fundamental destacar el desarrollo de herramientas innovadoras para acceder a la museografía interactiva mediante dispositivos móviles. Estas herramientas están diseñadas para ofrecer una experiencia más dinámica y accesible, permitiendo a los visitantes interactuar con guías audiovisuales específicas para cada experiencia (Gilli et al., 2023).

En el contexto de visitas con grandes grupos, esta herramienta es beneficiosa, ya que reduce la necesidad de guías en formato papel y permite a los visitantes moverse con mayor independencia dentro de la muestra, como acceder simultáneamente a la guía de un módulo.

Para optimizar aún más la experiencia del visitante, se emplea un sistema de reconocimiento por imágenes basado en técnicas de visión artificial. Este sistema es capaz de identificar cada experiencia de manera precisa, permitiendo al visitante acceder a información detallada y relevante sobre cada una, facilitando así una interacción más enriquecedora y personalizada. De esta manera, durante la muestra, los visitantes se enfrentan a las experiencias, crean sus propias hipótesis y las validan a través de la interacción con ellas. Para una mayor comprensión, los visitantes acceden a las guías audiovisuales a través de sus dispositivos móviles, lo que les permite conocer el correcto uso del módulo, plantearles interrogantes y repetir la experiencia.

Para lograr lo mencionado, se emplea la red neuronal YOLO (Ultralytics, 2024) en su versión V8 utilizando como *backbone* CSPDarknet53 (Wang et al., 2020), que emplea conexiones parciales entre etapas para mejorar el flujo de información entre capas y aumentar la precisión. Esta red permite de manera sencilla y rápida el entrenamiento de un modelo basado en un *dataset* personalizado. El *dataset* está compuesto por imágenes que fueron tomadas de cada una de las experiencias adoptadas como casos de estudio, y clasificadas mediante la aplicación “labelImg” (Tzutalin, 2015). Además, cada imagen está acompañada por un archivo de texto que identifica el objeto y su posición en la imagen.

En este trabajo, se extienden los resultados obtenidos por Levrino y col. (Levrino et al., 2023), mediante la mejora del modelo y un reentrenamiento del modelo clasificador, integrando nuevas experiencias de las expuestas en las muestras del MuIC. Este proceso ha permitido no solo una mayor precisión en la detección de las experiencias contempladas en el modelo previo, sino también una mejor capacidad para reconocer una gama más amplia de entornos gracias a la incorporación de imágenes adicionales y la actualización de la red neuronal utilizada.

Desarrollo

Experiencias seleccionadas

Las experiencias seleccionadas como casos de estudio incluyen “Jugando con el Submarino”, “Cosa de Parejas” y “Tangram”, que ya han sido clasificadas en el modelo de entrenamiento inicial descrito por Levrino y col. (Levrino et al., 2023). En esta fase, se incorporan nuevas imágenes de estos módulos y se añaden experiencias adicionales: “Torres de Hanói”, “Viaje por la Manguera”, “Lupa Rodante”, “Atrapa el Chanchito” y “Espectroscopio”. Es importante destacar que todas estas experiencias forman parte de la muestra permanente del MuIC. En la Figura 1 se presentan las experiencias contempladas.



a. Experiencia “Jugando con el Submarino”.



b. Experiencia “Cosa de Parejas”.



c. Experiencia "Tangram".



d. Experiencia "Torres de Hanoi".



e. Experiencia "Viaje por la manguera".



f. Experiencia "Lupa rodante".



g. Experiencia "Atrapa el chanchito".



h. Experiencia "Espectroscopio".

Figura 1. Experiencias seleccionadas para el modelo clasificador.

Modelo clasificador

Para el entrenamiento del modelo clasificador se utilizan imágenes de las experiencias detalladas anteriormente, especificando el área donde se encuentra el módulo y el nombre por el cual se identifica la experiencia. Estas imágenes se agrupan en un conjunto de datos (*dataset*), el cual se divide en dos grupos: el conjunto de datos de entrenamiento (*train*) y el conjunto de datos de evaluación (*test*). La Tabla 1 resume el número de etiquetas generadas para cada experiencia, tanto para el entrenamiento como para la evaluación del modelo clasificador.

El entrenamiento se desarrolla en iteraciones, donde se suministran los datos del grupo *train* al modelo y se evalúa su respuesta para que pueda aprender de las etiquetas adjuntas a cada imagen. Una vez terminada esta etapa de aprendizaje, el modelo es alimentado con datos del grupo *test* y se comparan los resultados que se generan, con las etiquetas realizadas por las personas. Esto permite evaluar la precisión del modelo en la detección.

Tabla 1. Cantidad de etiquetas agregadas al *dataset*.

	Cantidad Entrenamiento	Cantidad Prueba
Jugando con el submarino	418	122
Cosa de parejas	360	117
Tangram	429	123
Espectroscopio	694	422
Torres de Hanói	560	312
Viaje por la manguera	625	422
Lupa rodante	295	246
Atrapa al Chanchito	641	408

A medida que se realiza el entrenamiento se obtienen los datos de cada época, en las siguientes figuras se observa la progresión de mejora en la precisión del modelo (Fig. 2) y la disminución en la función de pérdidas (Fig. 3). El entrenamiento fue realizado con un máximo de 150 épocas con un *batch-size* de 16, dando un total de 2400 iteraciones. Aproximadamente en la época 75 se linealizan las curvas de mAP% (*mean Average Precision*) y de la función de pérdidas. A partir de estos valores no se producen grandes variaciones con el incremento del tiempo y los valores se estabilizan próximos al 90% para detecciones con una confianza entre 0 y 0.5, cercana al 70% para un umbral entre 0.5 y 0.95, y la función de pérdidas se encuentra entre 0.5 y 1.25.

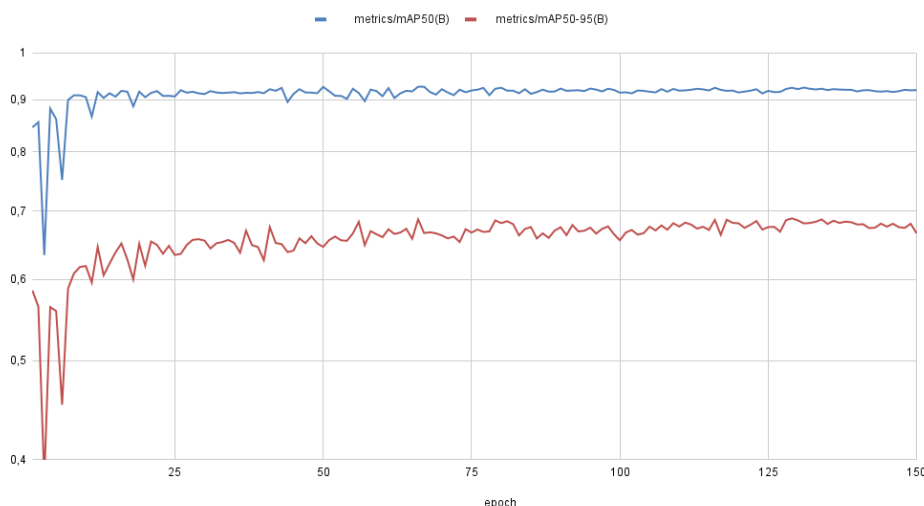


Figura 2. Media de los promedios de precisión.

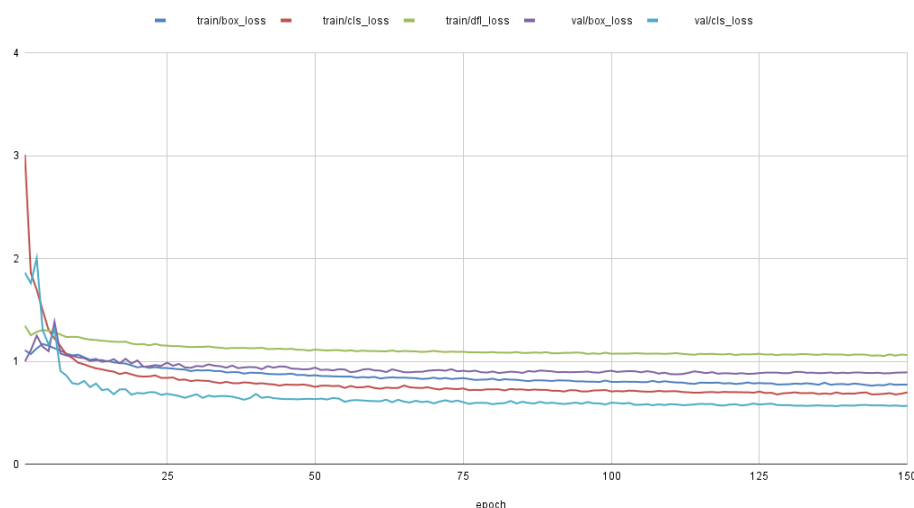


Figura 3. Funciones de pérdida.

Tras completar la etapa de entrenamiento, se obtiene un modelo que demuestra la capacidad para detectar las experiencias con una buena precisión, considerando la cantidad de imágenes y las configuraciones empleadas. En la Figura 4 se muestra cómo el modelo logra detectar cada módulo, levantando el modelo con un *script* de Python utilizando la librería de OpenCV. Es importante destacar que las imágenes utilizadas no forman parte del *dataset* de entrenamiento, sino que representa una situación real dentro del ámbito del museo.

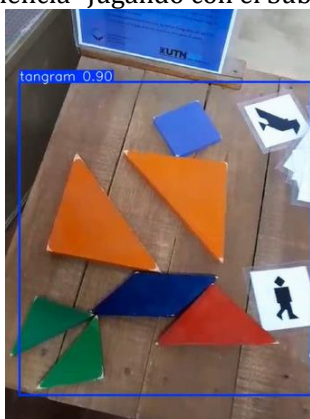
Por otro lado, en la Figura 4 se puede observar que al colocar las guías en formato papel, presentes en el museo, detrás de las experiencias “Cosa de Parejas” y “Jugando con el Submarino”, se mejora la detección de cada una respecto a los resultados presentados por Levrino y col. (Levrino et al., 2023).



a. Experiencia “Jugando con el Submarino”.



b. Experiencia “Cosa de Parejas”.



c. Experiencia “Tangram”.



d. Experiencia “Torres de Hanoi”.



e. Experiencia “Viaje por la manguera”.



f. Experiencia “Lupa rodante”.



g. Experiencia "Atrapa el chanchito".



h. Experiencia "Espectroscopio".

Figura 4. Detección de cada experiencia seleccionada.

Los resultados obtenidos muestran una confianza aceptable para la aplicación en la detección de experiencias en museografía interactiva. Sin embargo, para mejorar la precisión y robustez del modelo, se considera incorporar una mayor cantidad de imágenes de las experiencias interactivas en diversos entornos durante futuros entrenamientos. La inclusión de un conjunto de datos más variado permite al clasificador adaptarse mejor a las diferentes condiciones de iluminación, ángulos y configuraciones, aumentando la confianza en la detección.

Conclusiones

La implementación de la nueva versión del modelo mejoró, de manera significativa, la detección de las experiencias presentadas en la muestra del MuIC, considerando también la integración de sus respectivas guías. La incorporación de nuevas experiencias amplía la capacidad del modelo clasificador, lo que enriquece la experiencia de los visitantes al permitirles acceder, de manera rápida y eficiente, a las guías interactivas de una mayor cantidad de experiencias. Se destaca que las guías audiovisuales permiten a los visitantes visualizar de manera clara y concisa cómo interactuar con las distintas experiencias e incorporar preguntas estratégicas con el objetivo de fomentar la comprensión y el pensamiento crítico por parte de los visitantes.

Los resultados obtenidos hasta el momento son altamente prometedores y sugieren un notable potencial para futuras mejoras. La continua integración de avances en el modelo abre nuevas oportunidades para perfeccionarlo y ampliar su impacto, mejorando así la interacción y el aprendizaje de los visitantes.

Agradecimientos

A la Universidad Tecnológica Nacional (UTN) por financiar el Proyecto de Investigación y Desarrollo (PID) SIPPSF0009830: "Análisis e implementación de técnicas de reconocimiento por imágenes en la experiencia de usuario de ConCiencia - Museo Interactivo de Ciencias", en cuyo contexto se enmarca el presente trabajo.

Referencias

Gilli L., Cordoba R., & Pipino H. A. (2023). *Desarrollo de Museografía Interactiva Animada*. Jornadas de Ciencias y Tecnología 2023. San Francisco (Córdoba).

Levrino F., Castellano L., Fantin M., Mulassano M., & Pipino H. (2023). *Museografía Interactiva con Acceso mediante Código QR*. Jornadas de Ciencias y Tecnología 2023. San Francisco (Córdoba).

Tzotalin (2015). *Labellmg*. Git code. Disponible en: <https://github.com/tzotalin/labellmg>.

Ultralytics (2024). Ultralytics YOLOv8. Disponible en: <https://github.com/ultralytics/ultralytics>.

Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 390-391).