



Universidad
Nacional
de Loja

UNIVERSIDAD NACIONAL DE LOJA

FACULTAD DE LA ENERGÍA, LAS INDUSTRIAS Y

LOS RECURSOS NATURALES NO RENOVABLES

INGENIERA EN COMPUTACIÓN

ITINERARIO: SISTEMAS INTELIGENTES

Human Computer Vision

Sistema de Detección Automática de Dorsales en Carreras y Maratones

AUTORES:

Diego Fernando Lojan Tenesaca

Angel Jahir Román Sánchez

Cecilia Fernanda Trueba Reyes

DOCENTE:

Ing. Luis Chamba-Eras Mg. Sc.

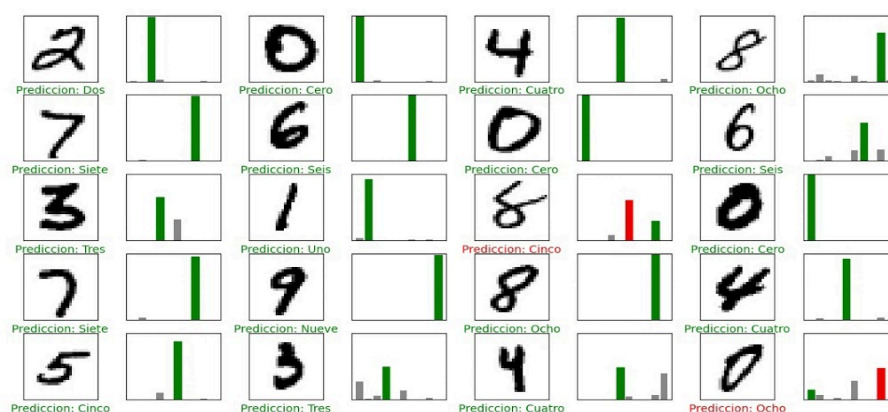
CICLO Y PARALELO:

Octavo “A”

LOJA – ECUADOR

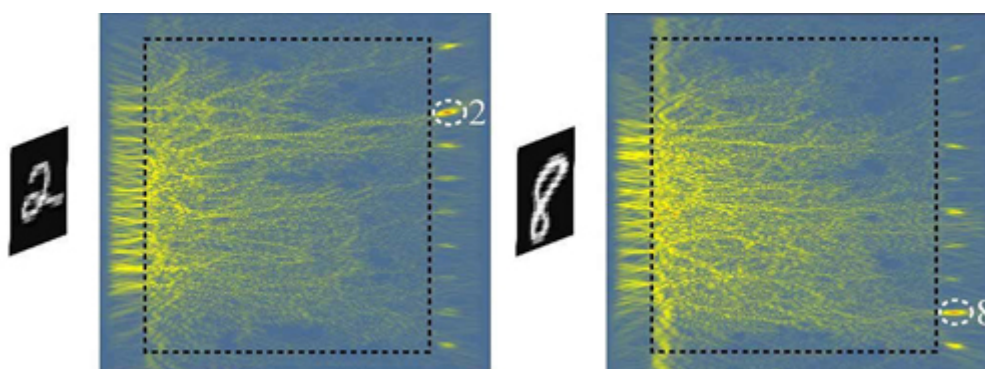
2023 - 2024

1. Reconocimiento de números escritos a mano.



El reconocimiento de números escritos a mano en modelos de visión por computador generalmente involucra procesos de aprendizaje automático o aprendizaje profundo donde un sistema se entrena para identificar y clasificar imágenes de dígitos. Estos modelos son entrenados usando grandes conjuntos de datos de imágenes etiquetadas que contienen ejemplos variados de dígitos manuscritos, permitiendo que el sistema aprenda a reconocer patrones y características específicas de cada número.

Un proceso típico comienza con la preprocesamiento de la imagen, que puede incluir la normalización de tamaño, la conversión a escala de grises, y a veces la binarización, que transforma la imagen en blanco y negro para simplificar el análisis. Este preprocesamiento ayuda a reducir la variabilidad entre imágenes que podrían deberse a diferencias en tamaño, orientación o color. Después, la imagen procesada se introduce en un modelo de red neuronal, que es donde ocurre el reconocimiento.



En el núcleo del proceso de reconocimiento se encuentra la red neuronal convolucional (CNN), que es especialmente eficaz para tareas de visión por computador. Una CNN típicamente consta de varias capas que incluyen capas convolucionales, de activación como ReLU, capas de agrupación (pooling) y capas completamente conectadas. Las capas convolucionales actúan como detectores de características que extraen atributos importantes de la imagen, como bordes y formas específicas de los números. Las capas de agrupación reducen la dimensionalidad de los datos conservando las características esenciales, y las capas completamente conectadas se utilizan para clasificar la

imagen en una de las categorías de dígitos, basándose en las características extraídas.

2. El modelo en funcionamiento no reconoce números a mano.

YOLOv4-tiny puede no ser la opción más adecuada para la detección de números escritos a mano debido a las diferencias en la naturaleza de la tarea y las limitaciones de la arquitectura del modelo, la escritura a mano puede variar significativamente entre diferentes personas, lo que hace que la tarea de detección sea aún más desafiante. Los modelos de detección como YOLOv4-tiny pueden tener dificultades para generalizar y reconocer correctamente una amplia variedad de estilos de escritura a mano, además adaptar YOLOv4-tiny para la detección de números escritos a mano requeriría un conjunto de datos grande y diverso que contenga ejemplos de números escritos a mano en una variedad de estilos y contextos.

En su lugar, pueden ser más apropiados enfoques específicos diseñados para el reconocimiento de caracteres escritos a mano, como las redes neuronales convolucionales (CNN) diseñadas para esta tarea específica.

Method	FPS	mAP(%)
YOLOv3	49	52.5
YOLOv4	41	64.9
YOLOv3-tiny	277	30.5
YOLOv4-tiny	270	38.1
Proposed method	294	38.0

- YOLOv4: Registra un FPS de 41 y el mAP más alto de la tabla con un 64.9%. Aunque ofrece la mejor precisión, es más lento en velocidad de procesamiento.
- YOLOv4-tiny: Ofrece un balance mejor entre velocidad y precisión, con un FPS de 270 y un mAP del 38.1%.

Cuando se trata de detectar números de dorsales, especialmente en escenarios dinámicos o en condiciones de imagen variadas, tener una red que pueda ajustarse finamente a las características de los objetos pequeños es crucial. YOLOv4 con su capacidad para utilizar características detalladas y su robustez general sería más confiable, mientras que YOLOv4-tiny ofrece ventajas en entornos donde la velocidad de procesamiento y la eficiencia operativa son más críticas que la precisión absoluta [1].



YOLOv4-tiny, diseñado primordialmente para la detección de objetos de tamaño y forma relativamente consistentes en imágenes y videos, podría enfrentar dificultades al detectar números escritos a mano debido a la variabilidad significativa en el estilo, tamaño y forma de estos. Dado que los números manuscritos pueden variar enormemente de una persona a otra y suelen presentarse en tamaños pequeños, YOLOv4-tiny, con su estructura simplificada y entrenamiento típicamente enfocado en objetos más grandes y uniformes, no está optimizado para captar los detalles finos y las sutiles diferencias requeridas para un reconocimiento preciso de dígitos manuscritos.

3. **Análisis de errores en el reconocimiento de números.**

Los modelos diseñados para el proyecto tienen dos componentes principales: primero identifica los dorsales en las imágenes de corredores y, posteriormente, detecta los números en esos dorsales. Esto es clave para asegurarse de que el modelo solo reconozca números asociados con los dorsales y no otros números que puedan estar presentes en la imagen.

3.1. **¿Por qué el modelo no reconoce esos números?**

El modelo no reconoce números escritos a mano porque fue entrenado específicamente con un conjunto de datos como el Street View House Numbers (SVHN), que incluye principalmente números impresos en formatos estándar, similares a los que se encuentran en las direcciones residenciales. Este entrenamiento lo hace ideal para identificar números impresos, como los que comúnmente aparecen en los dorsales de los corredores. Sin embargo, los números escritos a mano presentan variaciones significativas en estilo y forma, las cuales no coinciden con los patrones de los datos de entrenamiento. Esta diferencia en el estilo puede llevar a que el modelo tenga una baja precisión o incluso falle completamente al intentar reconocer estos números escritos a mano.

3.2. **¿Qué condiciones deben tener los dorsales para ser reconocidos?**

Para que los dorsales sean reconocidos efectivamente por el modelo, deben cumplir ciertos criterios de visibilidad y formato:

- **Presencia del dorsal:** La imagen debe contener un dorsal claramente visible para que el modelo pueda operar. Si no hay dorsal en la imagen, el modelo no procederá con el reconocimiento de números.
- **Claridad y visibilidad:** Los dorsales deben estar claramente visibles y no estar obstruidos ni parcialmente cubiertos por ropa, brazos del corredor u otros objetos.
- **Orientación y distorsión:** Idealmente, los dorsales deben estar orientados frontalmente hacia la cámara para minimizar la distorsión. Las inclinaciones o curvaturas extremas pueden dificultar el reconocimiento.
- **Contraste y legibilidad:** Los números en los dorsales deben tener un contraste suficiente respecto al fondo del dorsal para facilitar la detección. Esto incluye tener un tamaño adecuado, un color que resalte y una tipografía clara.
- **Formato estándar:** Los dorsales que sigan un formato estándar en cuanto a la disposición y tamaño de los números tienden a ser más

fácilmente reconocidos por el modelo.

4. Listado de limitaciones de los dos modelos.

Los dos modelos utilizados en el proyecto de detección de dorsales y números en imágenes de corredores presentan varias limitaciones que son importantes considerar para su uso eficaz:

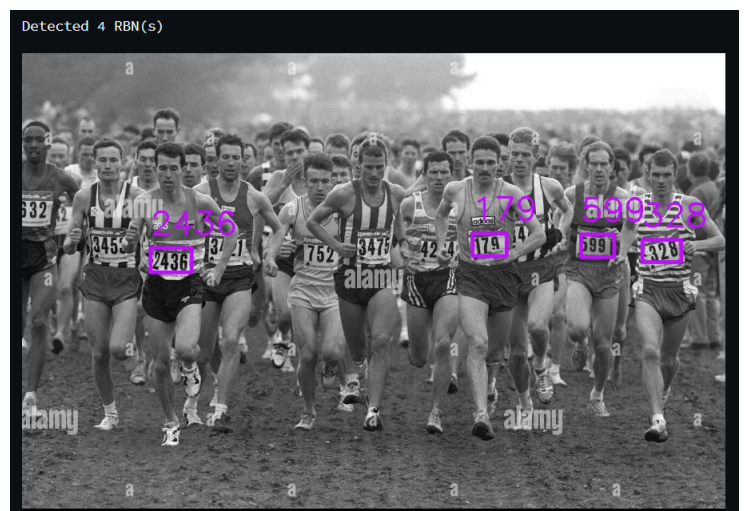
A. Modelo de Detección de Dorsales

- Dependencia de la visibilidad: El modelo requiere que los dorsales estén claramente visibles y no obstruidos por otros objetos, como ropa o brazos del corredor.
- Sensibilidad a la orientación: Los dorsales que no están orientados directamente hacia la cámara o que están distorsionados (por ejemplo, por el movimiento del corredor) pueden no ser detectados correctamente.
- Necesidad de un formato estándar: El modelo está optimizado para detectar dorsales que siguen un formato estándar en cuanto a tamaño y disposición de los números. Dorsales no estándar o con diseños creativos pueden no ser reconocidos.
- Condiciones de iluminación: Variaciones en la iluminación, como sombras fuertes o contraluces, pueden afectar la capacidad del modelo para identificar correctamente los dorsales.

B. Modelo de Reconocimiento de Números.

- Limitado a números impresos: El modelo está entrenado principalmente con el conjunto de datos SVHN, que contiene números impresos. No está optimizado para reconocer números escritos a mano o estilos de números altamente estilizados.
- Sensibilidad al contraste y claridad: Requiere que los números en los dorsales tengan suficiente contraste con respecto al fondo y sean legibles, lo que implica que los números borrosos o de bajo contraste pueden no ser reconocidos.
- Dependencia del primer modelo: Su eficacia depende directamente de la correcta detección de dorsales por el primer modelo. Si el dorsal no es detectado, el modelo de reconocimiento de números no puede operar.
- Errores en la segmentación: Errores en la segmentación de números individuales dentro de un dorsal pueden llevar a errores en la identificación final del número, especialmente si los números están muy juntos o se superponen.
- Ambos modelos, aunque potentes, tienen sus restricciones operativas que deben ser consideradas al implementarlos en aplicaciones reales, como eventos deportivos donde la precisión y la confiabilidad son críticas.

Pruebas



Referencias

- [1] Z. Jiang, L. Zhao, S. Li, y Y. Jia, "Real-time object detection method for embedded devices", Arxiv.org. [En línea]. Disponible en: <https://browse.arxiv.org/pdf/2011.04244v2>.