

**Instituto Tecnológico de Estudios
Superiores de Monterrey**



TC5035 Proyecto Integrador
Modelo para la Detección de Colores Dominantes
de Objetos en Cámaras de Seguridad

Semana 9 - Actividad 1
Avance 7 – Resumen Ejecutivo

Equipo #43

Ronald David Zuniga Sánchez

A01686240

16/Junio/2024

Contenido

Introducción	3
Síntesis del problema	3
Contexto del Proyecto	3
Plataforma Flex AI	4
Formulación del Problema	4
Propuesta de solución.....	5
Hallazgos más importantes del análisis exploratorio de datos	6
Descripción del conjunto de datos:.....	6
Ingeniería de características	7
Resultados de análisis de datos	8
Modelos generados y razones de la elección del modelo final	8
Etapas de preprocesamiento de imágenes	8
Modelos propuestos	9
Análisis del modelo seleccionado	12
Recomendaciones clave para implementar la solución	13
Requerimientos para implementación	13
Recomendaciones para implementación.....	13
Plan de acciones a futuro	13
Análisis Costo – Beneficio	14
Costos asociados al proyecto	14
Beneficios esperados del proyecto	15
Riesgos y Desafíos Potenciales	17
Referencias	18

Ilustraciones

Ilustración 1. Portafolio de soluciones en nube de Hanwha Vision.	4
Ilustración 2. Proceso actual para entrenamiento de objetos en Flex AI.	4
Ilustración 3. Diagrama de Solución Propuesta.....	5
Ilustración 4. Ejemplo de Solución Propuesta.	6
Ilustración 5. Ejemplo de imágenes en dataset CIFAR10.	6
Ilustración 6. Ejemplo de visualización de etiquetas de color generadas por OpenClip.	7
Ilustración 7. Resultados de creación de etiquetas de color en CIFAR10.....	7
Ilustración 8. Ejemplo de imágenes etiquetadas por color dominante en CIFAR10.	8
Ilustración 9. Procesamiento de normalización de valores RGB.	8
Ilustración 10. Arquitectura de Modelo 1 – Base Twin Network Model	9
Ilustración 11. Arquitectura de Modelo 2 - Single Network Model	10
Ilustración 12. Arquitectura de Modelo 3 - Updated Twin Network Model	10
Ilustración 13. Arquitectura de Modelo 4 - VGG-16 Based Single Network.....	11
Ilustración 14. Arquitectura de Modelo 5 - Transfer Learning VGG16.....	11

Introducción

El presente informe presenta la metodología y resultados obtenidos en el desarrollo de la investigación para la creación de un modelo para la Detección de Colores Dominantes de Objetos en Cámaras de Seguridad patrocinado por la empresa Hanwha Vision, proveedor global de soluciones de captura y procesamiento de video.

Se desarrolla en los siguientes capítulos la descripción del contexto del problema y requerimientos funcionales del modelo. Se describen los hallazgos más importantes encontrados en la exploración de datos utilizados para el entrenamiento de los modelos propuestos, así como los resultados relevantes y las razones para la selección del modelo final preliminar.

Finalmente se recomiendan las estrategias a futuro sugeridas para la implementación del modelo y un análisis de costos asociados al desarrollo de la presente investigación. Concluyendo en una reflexión sobre los riesgos y beneficios potenciales de la solución propuesta en el contexto de implementación en casos reales.

Síntesis del problema

Contexto del Proyecto

La aplicación de Inteligencia Artificial en análisis de imágenes y video se denomina visión computacional, y en los últimos años las técnicas de clasificación, detección y segmentación semántica en 2D han mejorado exponencialmente [4]. La adopción de esta tecnología ha impulsado la función de detección de objetos a ser una característica estándar en el mercado de cámaras de seguridad.

Detected Objects	Attribute Category	Supported Attribute Items
 Person	Gender	Female, Male
	Upper garment	(Up to 2 colors simultaneously)
	Lower garment	Black, Gray, White, Red, Orange, Yellow, Green, Blue, Purple
	Bag	Carries a bag, does not carry a bag
 Face	Gender	Female, Male
	Age	Young/Adult/Middle/Senior
	Glasses	Wearing glasses, not wearing glasses
	Face mask	Wearing mask, not wearing mask
 Vehicle	Type	Car (sedan/SUV/van/etc.), bus, truck, motorcycle, bicycle
	Color	(Up to 2 colors simultaneously) Black, Gray, White, Red, Orange, Yellow, Green, Blue, Purple
 License plate	License plate	Presence of a license plate

Tabla 1. Tipos de objetos y atributos estándar en cámaras de seguridad.

Sin embargo, esta aplicación de tecnología se ha limitado a detectar objetos relacionados con unos pocos casos de uso y verticales como la detección de personas y vehículos y sus atributos (ver Tabla 1) [5]. La empresa (Hanwha Vision) ha identificado la oportunidad en usuarios finales que han expuesto múltiples casos de uso que requieren más especialización. Es decir, la creación de una solución para la detección de objetos específicos a partir del entrenamiento de modelos de visión computacional de acuerdo con los objetos de interés en las escenas del usuario final.

Plataforma Flex AI

El proceso de negocio a impactar es en el nuevo ofrecimiento de servicios en nube de Hanwha Vision que busca abrir nuevas aplicaciones para los sistemas de visión (cámaras, grabadores y software de administración de video). En 2024, Hanwha Vision lanzará esta serie de servicios bajo el nombre “Hanwha Vision Cloud Portal”, que incorporará servicios de administración de cámaras en nube (DM Pro), servicio de grabación de video en nube (On Cloud), plataforma para agregación y visualización de datos para inteligencia de negocios (Sightmind), y un servicio para la creación y entrenamiento de modelos de visión computacional personalizados (Flex AI) [3].

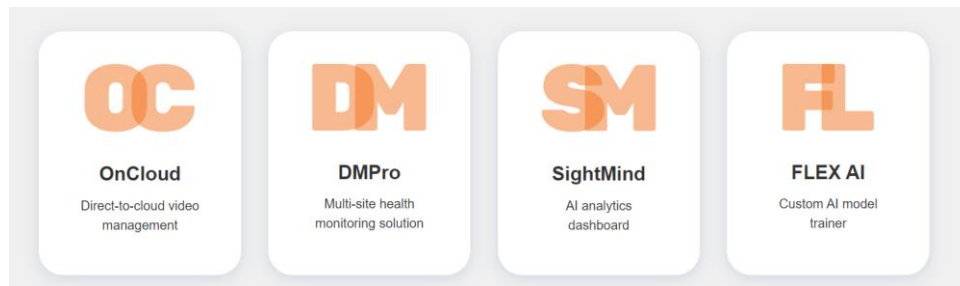


Ilustración 1. Portafolio de soluciones en nube de Hanwha Vision.

Formulación del Problema

El problema por resolver tiene dos puntos de vista, el problema a resolver para los usuarios finales y el problema a resolver para la empresa que desarrolla el producto (Flex AI). Se entiende como usuarios finales a las empresas, de diferentes rubros, que actual o potencialmente pueden utilizar los productos y servicios de la empresa para video vigilancia y procesamiento de video.

- **Problema de usuarios finales**
Personalización de video analíticas para la detección y clasificación de objetos utilizando el criterio de color para la diferenciación de los objetos de acuerdo con la industria o caso de uso.
- **Problema de la empresa/producto.**
Investigación de soluciones aplicables para agregar la variable de color como atributo de personalización en el flujo de entrenamiento de modelo de detección de objetos para el producto en desarrollo (Flex AI).

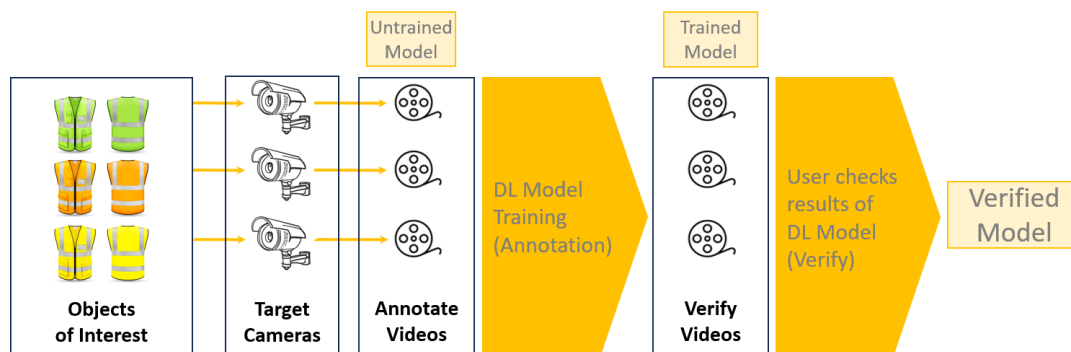


Ilustración 2. Proceso actual para entrenamiento de objetos en Flex AI.

Para la empresa, Hanwha Vision, el problema a resolver es de carácter técnico donde se requiere investigar técnicas relevantes para el uso de color del objeto como atributo para el entrenamiento e inferencia en vivo del modelo en cámaras de seguridad. Los factores para considerar son los siguientes:

1. Solución con un rendimiento/exactitud (accuracy) de por lo menos un 90% como parámetro base, y del 99% como parámetro deseado.
2. Modelo optimizado para ejecutar inferencias en borde (en procesador de cámaras de seguridad) por lo que requiere optimización de recursos de memoria y tiempo de ejecución.
 - a. Tiempo de ejecución ≤ 10 ms.
 - b. Total número de parámetros de modelo final ≤ 1 M.
3. Diseño para funcionar con la menor cantidad de imágenes de entrenamiento posibles.

Propuesta de solución

Considerando que el resultado actual del modelo Flex AI es la detección de un objeto de interés que incluye:

- Un cuadro delimitador (bounding box)
- Categoría de tipo de objeto
- Imagen recortada del objeto detectado

Se propone utilizar directamente la imagen recortada del objeto detectado como la entrada directa para el modelo propuesto para detectar el Color Dominante del objeto.

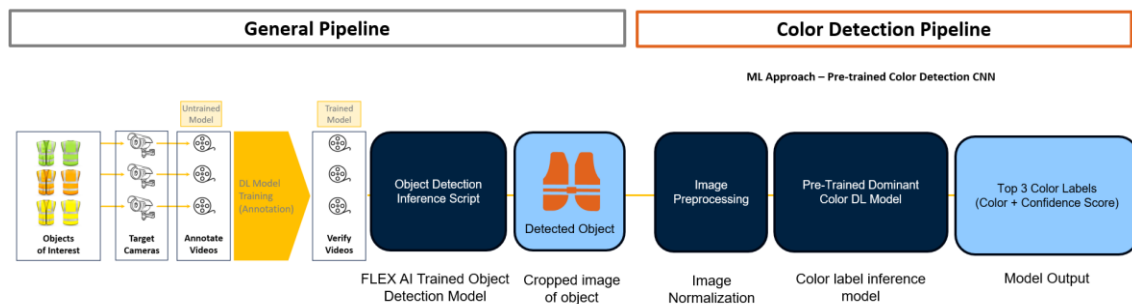


Ilustración 3. Diagrama de Solución Propuesta.

Consideraciones:

- El color dominante se define como el color más prevalente de la imagen.
- Un problema identificado para este enfoque son los objetos que tienen una distribución de color cercana al 50/50.
- El modelo propuesto será diseñado para inferir el color dominante y proporcionar un top 3 de los colores identificados.

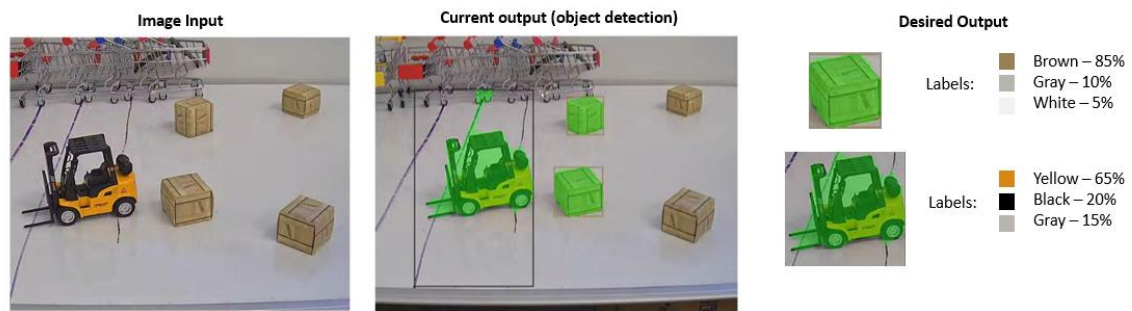


Ilustración 4. Ejemplo de Solución Propuesta.

Hallazgos más importantes del análisis exploratorio de datos

Considerando la necesidad de un conjunto de datos con imágenes de tamaño uniforme que contengan un solo objeto por imagen en diferentes entornos, se ha seleccionado el conjunto de datos CIFAR 10.

Descripción del conjunto de datos:

El conjunto de datos CIFAR-10 consiste en 60,000 imágenes en color de 32x32 en 10 clases, con 6000 imágenes por clase. Hay 50000 imágenes de entrenamiento y 10000 imágenes de prueba.

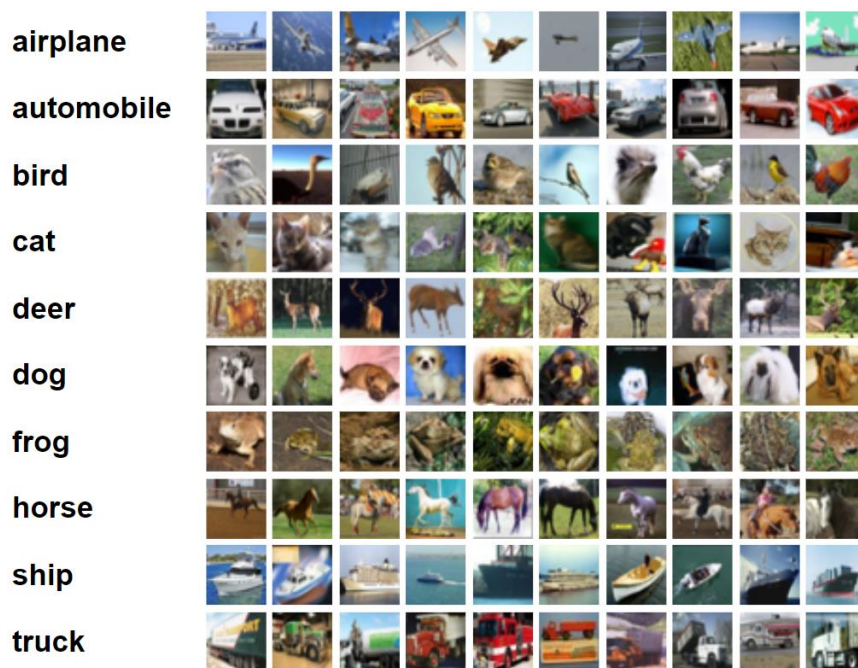


Ilustración 5. Ejemplo de imágenes en dataset CIFAR10.

El conjunto de datos está dividido en cinco lotes de entrenamiento y un lote de prueba, cada uno con 10000 imágenes. El lote de prueba contiene exactamente 1000 imágenes seleccionadas aleatoriamente de cada clase. Considerando la necesidad de un conjunto de datos equilibrado, solo se considera el lote de imágenes de entrenamiento.

Ingeniería de características

Considerando que el objetivo del modelo a entrenar es el de identificar el color dominante del objeto de interés, se consideraron dos enfoques para crear las etiquetas de color para el conjunto de datos CIFAR 10:

1. Agrupamiento por K-Means para la detección del color dominante basado en valores RGB.
2. Etiquetado del color dominante mediante el uso de un modelo semántico multimodal (OpenCLIP) para la clasificación de imágenes según etiquetas predefinidas.

Aunque ambos enfoques produjeron resultados precisos, se utilizó el modelo multimodal ya que proporcionó una manera más rápida de crear iteraciones de las inferencias de etiquetas de color.

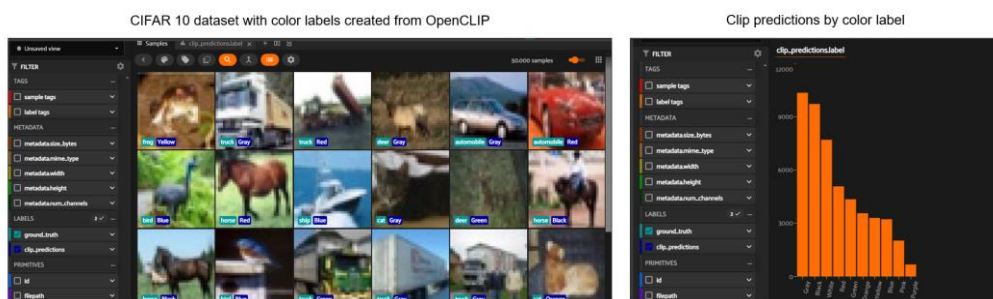


Ilustración 6. Ejemplo de visualización de etiquetas de color generadas por OpenClip.

Se exploraron cuatro (4) versiones del conjunto de datos CIFAR10 con etiquetas de color creadas con OpenCLIP. El objetivo de esta exploración fue encontrar las etiquetas de color que produjeran el conjunto de datos más equilibrado en términos de muestras por clase.

Las combinaciones de etiquetas utilizadas son:

1. 10 Etiquetas – Colores primarios y secundarios + Gris y Rosa.
2. 10 Etiquetas – Colores primarios y secundarios + Gris y Marrón.
3. 8 Etiquetas – Colores primarios y secundarios.
4. 8 Etiquetas – Colores primarios y secundarios (–) Púrpura + Marrón.

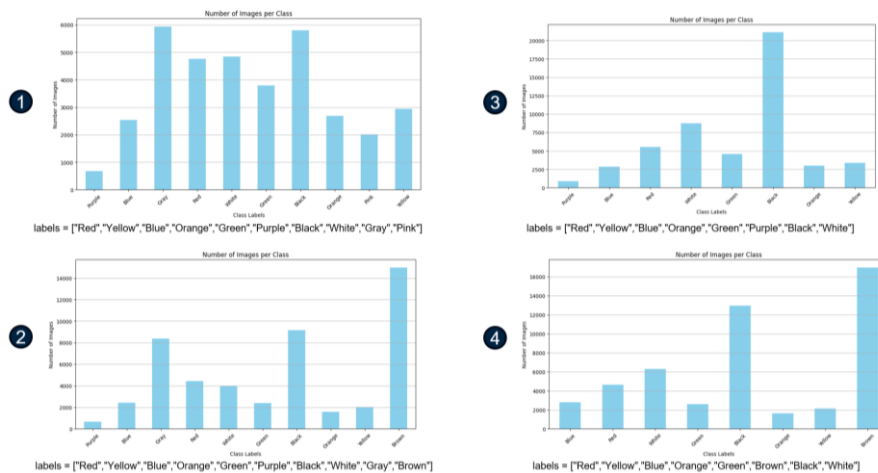


Ilustración 7. Resultados de creación de etiquetas de color en CIFAR10.

Resultados de análisis de datos

El conjunto de datos #4 fue seleccionado considerando que proporcionó la mayor cantidad de muestras en el color con menos etiquetas (naranja) con 1600 muestras.

Utilizando un método de imputación aleatoria, se creó una versión equilibrada del conjunto de datos con 1600 muestras por color y un total de 12,800 muestras para entrenamiento y validación.

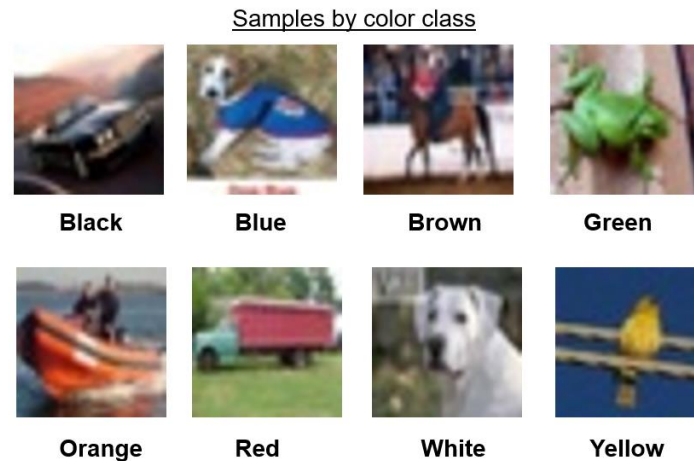


Ilustración 8. Ejemplo de imágenes etiquetadas por color dominante en CIFAR10.

Modelos generados y razones de la elección del modelo final

Etapa de preprocesamiento de imágenes

Se propone un preprocesamiento de normalización de canales calculando los valores medios y la desviación estándar de los valores RGB en el conjunto de datos para ser utilizados como valores fijos al procesar nuevas imágenes con los siguientes objetivos:

- Equilibrar la dominancia de características – reduciendo la escala en los canales dominantes para prevenir pesos sesgados.
- Mejorar la generalización – reduciendo el sobreajuste en rangos específicos de valores de píxeles.
- Ajuste de contraste – para hacer que los colores parezcan más distintos entre sí.

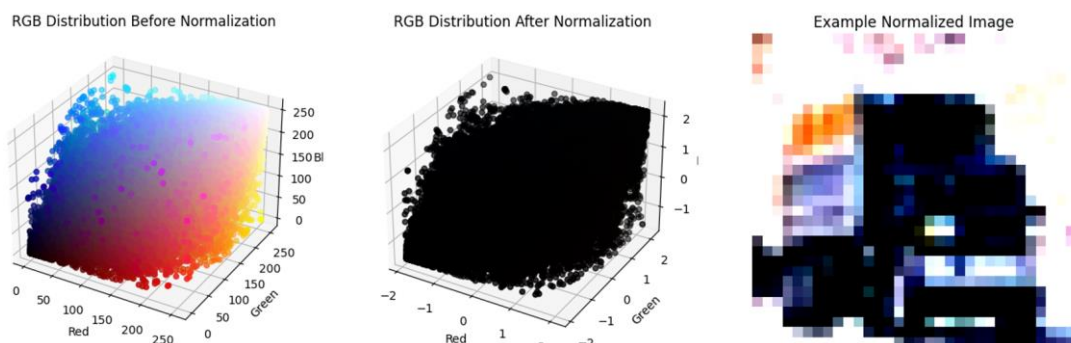


Ilustración 9. Procesamiento de normalización de valores RGB.

Los valores de promedio (mean) y desviación estándar RGB se calculan solo en el subconjunto de entrenamiento del conjunto de datos para evitar la fuga de características desde los subconjuntos de validación y prueba que harían que el modelo se sobre-entrene y disminuya su capacidad de generalización.

Los valores calculados se determinan como parte del proceso en el paso de preprocesamiento de imágenes aplicando lo siguiente:

- $\text{image} = \text{Normalize}(\text{mean}=\text{mean_rgb}/255.0, \text{std}=\text{std_dev_rgb}/255.0)(\text{image})$
- $\text{mean_rgb} = [124.75003924, 123.10938383, 114.89393648]$
- $\text{std_dev_rgb} = [63.66785235, 62.47050596, 67.81292227]$

Modelos propuestos

Para la creación de los 5 modelos propuestos para entrenar desde cero, se utilizan los siguientes artículos como referencia:

1. Reconocimiento de Color de Vehículos utilizando CNNs – propone un modelo de redes gemelas.
2. Reconocimiento de Color en Entornos de Iluminación Desafiantes con CNNs – propone un modelo de red base única.
3. Transfer learning with VGG16 - modificado a través del aprendizaje por transferencia para reentrenar la capa final.

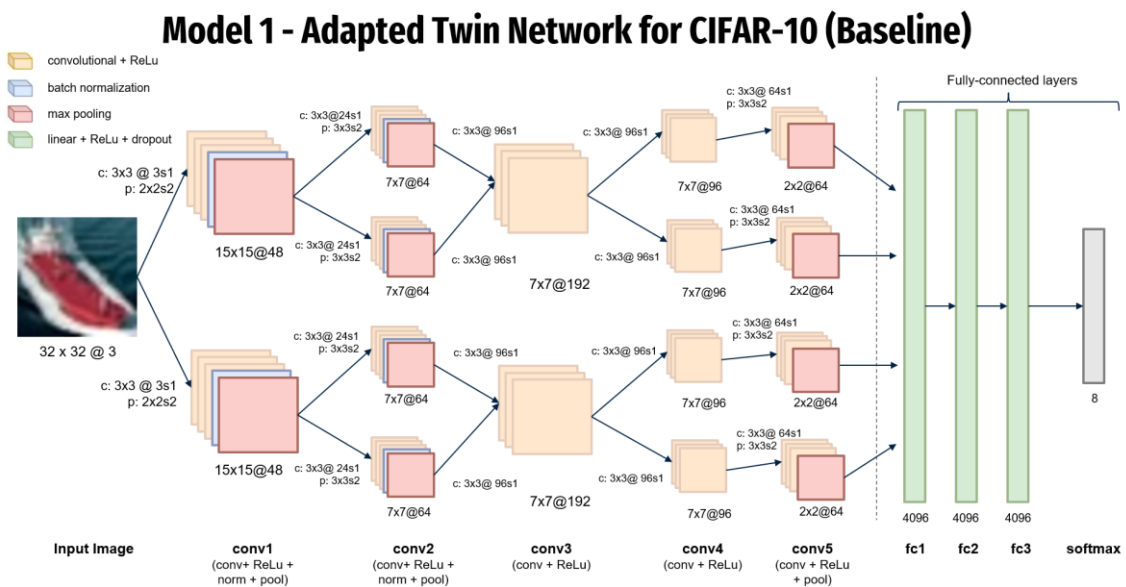


Ilustración 10. Arquitectura de Modelo 1 – Base Twin Network Model

Model 2 - Single Base Network

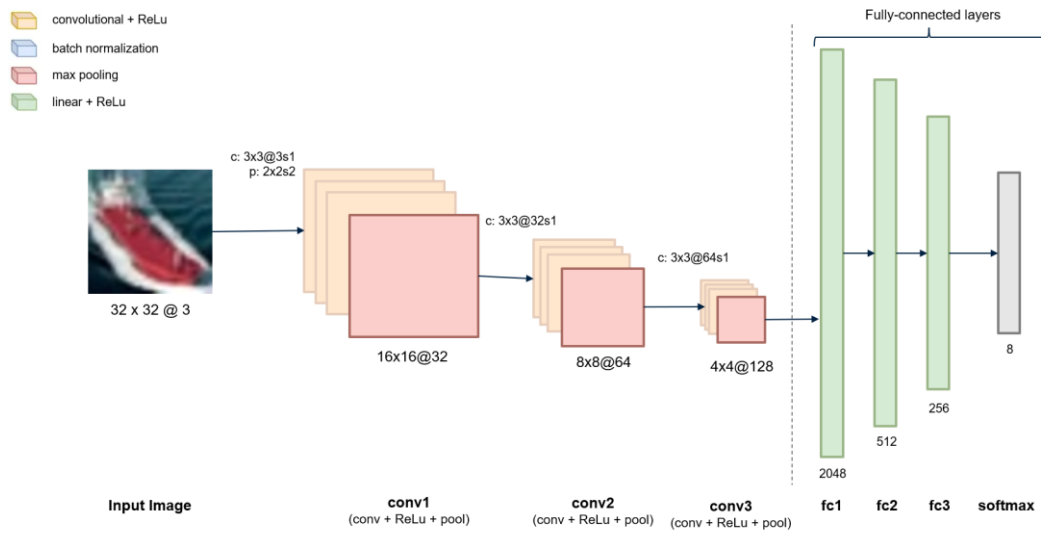


Ilustración 11. Arquitectura de Modelo 2 - Single Network Model

Model 3 - Simplified Twin Network

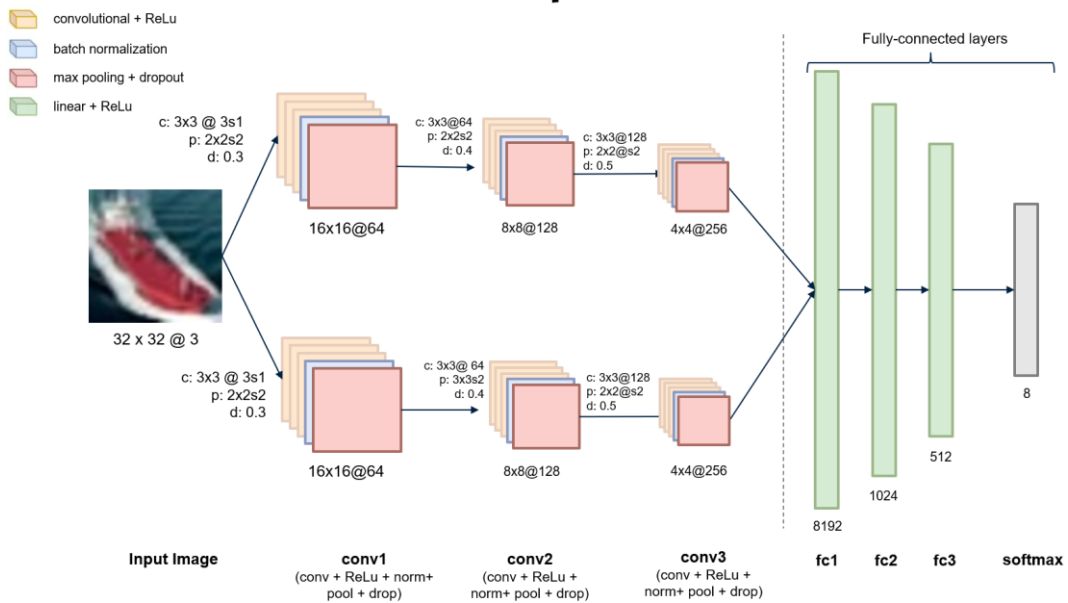
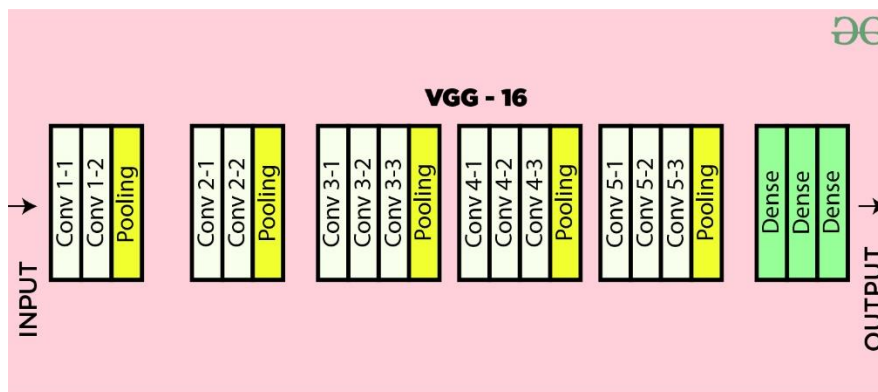
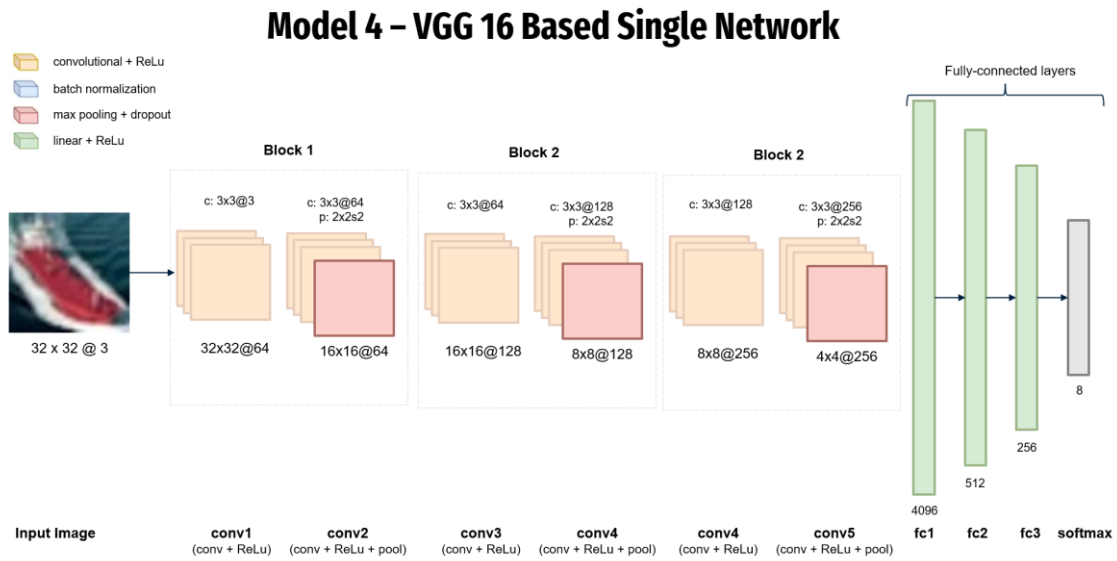


Ilustración 12. Arquitectura de Modelo 3 - Updated Twin Network Model



Análisis del modelo seleccionado

En la siguiente tabla se resumen los resultados del entrenamiento y evaluación comparativo de los modelos propuestos:

Criteria	Model 1: Base Twin Network Model	Model 2: Single Network Model	Model 3: Updated Twin Network Model	Model 4: VGG-16 Based Single Network	Model 5: Transfer Learning VGG16
Architecture Summary	Complex multi-layer network with dual paths	Simpler, more compact network	Enhanced version of Model 1 with additional dropout and batch normalization layers	Based on VGG-16, deep with multiple convolutions and fully connected layers	Fixed Feature Extractor Model through transfer learning on VGG16 ImageNet
Total Parameters	22,067,784	1,275,720	9,660,040	3,376,456	134,293,320
Validation Loss (final)	2.122	1.822	1.266	1.899	1.9565
Test Loss (avg)	2.327	1.840	1.343	1.958	2.0398
Validation Accuracy (max)	66.25%	61.82%	66.77%	64.95%	65.62%
Test Accuracy (avg)	63.18%	59.27%	64.48%	64.32%	65.16%
Best Model Size (MB)	86.29	4.98	37.75	13.19	524.59
CPU Inference Time (ms)	9.215	2.644	3.881	3.444	75.46
CUDA Inference Time (ms)	8.08	1.3	3.067	2.515	147.96
Max Memory Usage (CUDA Mem)	23.98 Mb	7.00 Mb	20.00 Mb	28.00 Mb	1.6 Gb

Tabla 2. Métricas de rendimiento de modelos propuestos.

Dadas las métricas de rendimiento y las consideraciones de recursos, el **Modelo 3** se selecciona como el mejor modelo considerando que ofrece el mayor equilibrio entre recursos y rendimiento. Proporciona una precisión del 64,48% junto con demandas de recursos razonables, lo que lo hace adecuado para aplicaciones que requieren un buen equilibrio entre rendimiento y eficiencia.

Descripción de resultados Modelo 3

- Arquitectura y Complejidad: Presenta una versión mejorada del Modelo 1, incluye capas adicionales de abandono (dropout) y normalización por lotes (batch normalization), con el objetivo de mejorar la generalización sobre la Red Gemela Base.
- Rendimiento: Muestra la mejor pérdida de validación (1.266) entre todos los modelos, con una precisión de prueba destacable del 64.48%. Este modelo equilibra la complejidad con el rendimiento, haciéndolo efectivo y relativamente eficiente.
- Adecuación: Adecuado para aplicaciones que requieren un equilibrio entre recursos computacionales y alta precisión considerando un tiempo de inferencia mínimo de 3.067 ms y requerimiento de memoria de 20 Mb.

Recomendaciones clave para implementar la solución

Requerimientos para implementación

Se determina que en el estado actual, el modelo no cumple con todos los parámetros clave de rendimiento solicitado por el sponsor/negocio, se observa que aún hay oportunidad de mejora en el modelo para cumplir con los requerimientos establecidos para implementar el modelo en borde (en una cámara de seguridad):

1. Accuracy $\geq 90\%$ (No cumple)
 - a. Modelo 3 cuenta con un Test Accuracy del 64.48%.
2. Inference time ≤ 10 ms (Cumple)
 - a. Modelo 3 cuenta con un tiempo de inferencia promedio de:
 - i. 3.44 ms en CPU
 - ii. 2.515 en GPU
3. Parameter count ≤ 1 M (after pruning and quantization) (No cumple)
 - a. Modelo 3 cuenta con 3M de parámetros previo a etapas de optimización.

Sin embargo, el modelo muestra potencial para ser utilizado como un elemento de UX en el proceso de entrenamiento de modelos para la selección de color dominante de objetos de manera asistida. Ejemplo “User select based on point-prompt” donde se utiliza el modelo para inferir opciones de colores presentes en el objeto a etiquetar y así definir como variable interna, el valor del color del objeto a entrenar.

Recomendaciones para implementación

Considerando que el objetivo es crear un modelo que pueda generalizar sobre diversos casos de uso y posiciones de captura de imagen (por medio de implementación en borde en cámaras de seguridad) se recomienda optimizar al máximo la precisión de predicción de color dominante y colores complementarios.

Adicionalmente, se recomienda aprovechar el caso de uso de inferencias y entrenamiento dentro de plataforma Flex AI, para la creación de un dataset variado y enfocado a los casos de uso de los usuarios, lo cual permitirá futuros ejercicios de re-entrenamiento del modelo para la mejora de su rendimiento.

Plan de acciones a futuro

Se proponen 3 estrategias para mejorar el rendimiento en la continuación del desarrollo de la investigación a futuro:

1. Mejora del Dataset.

- Mejora en categorización: El conjunto de datos utilizado se produjo utilizando una técnica mixta de creación de etiquetas utilizando OpenClip y validando el color dominante con Kmeans clustering. Sin embargo, aún así se observan algunos datos ambiguos o mal categorizados que pueden estar impactando en el rendimiento del modelo y su capacidad de generalizar.
- Data Augmentation: Mejorar el conjunto de datos con más variaciones como diferentes iluminaciones, ángulos y cambios de fondo podría ayudar a mejorar la robustez y la generalización del modelo.

- Datasets alternativos: El dataset utilizado es el de CIFAR10 con imágenes de 32x32 píxeles. El tamaño pequeño de las imágenes puede estar influyendo en el rendimiento del modelo considerando la información limitada para la generación y abstracción de características en las capas convolucionales. La exploración de modelos alternativos como la reevaluación del uso del dataset COCO 2017 pueden apoyar a mejorar la calidad de información para entrenamiento.

2. Mejora del Modelo.

- Optimización de Hiperparámetros: Ajustar la tasa de aprendizaje, el tamaño del lote, el número de épocas y otros hiperparámetros mediante métodos como la búsqueda en cuadrícula o la optimización bayesiana podría llevar a mejores resultados.
- Modificaciones Arquitectónicas Avanzadas: Incorporar mecanismos de atención o conexiones residuales podría ayudar al modelo a enfocarse mejor en características relevantes y mejorar la eficiencia del aprendizaje.

3. Arquitecturas Alternativas

- Técnicas de Ensemble: Combinar las predicciones de múltiples modelos (incluidos aquellos que no fueron elegidos como el mejor modelo) a menudo puede mejorar el rendimiento predictivo sobre cualquier modelo individual.
- Modelos de Segmentación en lugar de arquitecturas de clasificación como YOLO para la exploración de comportamiento en el aprendizaje y discernimiento de características de color.

Análisis Costo – Beneficio

Costos asociados al proyecto

Nombre de Fase	Subfase	Descripción	Costo Estimado en USD
Comprensión del Negocio	Reuniones y Consultoría	Tiempo del personal y consultores para definir objetivos	Depende de costo operativo del personal.
	Análisis de Factibilidad	Estudios para evaluar viabilidad técnica y económica	Depende de costo operativo del personal.
Comprensión de los Datos	Adquisición de Datos	Costos de adquirir datos si aplicable (CIFAR 10 es gratuito)	\$0
	Herramientas de Análisis	Software para análisis exploratorio de datos	Utilización de herramientas Open Source.
Preparación de los Datos	Manipulación y Limpieza	Herramientas y tiempo de personal para limpiar y preparar datos	Depende de costo de herramientas.
	Costos de Almacenamiento	Almacenamiento en la nube para datos limpios	\$100 - \$200 por mes
Modelado	Uso de Servidores/Cloud	Costos de Google Colab Pro	\$50 - \$100 por modelo
	Licencias de Software de ML	Herramientas y plataformas de ML	Depende de costo de herramientas.
Evaluación	Herramientas de Evaluación	Software para pruebas de validación del modelo	Depende de costo de herramientas.
	Servicios de Cloud para Tests	Costos de prueba en la nube	Depende de costo de herramientas (AWS).
Despliegue	Implementación de Infraestructura	Costos de servidores y plataformas de despliegue	Depende de costo de herramientas (AWS).
	Integración de Sistemas	Desarrollo para integrar el modelo con sistemas existentes	Depende de costo operativo del personal.

Tabla 3. Análisis de costos por etapa CRISPML.

Costos Esperados para Operación y Mantenimiento del Modelo

- Monitoreo Continuo: Costos de herramientas y tiempo de personal para monitorear el rendimiento del modelo en producción.
- Actualizaciones y Retraining: Costos de reentrenamiento periódico del modelo para adaptarse a cambios en los datos o en el entorno.
- Soporte y Resolución de Problemas: Costos de soporte técnico para manejar cualquier problema operacional o de rendimiento.

Beneficios esperados del proyecto

Considerando los casos de uso identificados durante la etapa de comprensión del negocio (descritos en Tabla 4), se resumen los siguientes beneficios esperados para el negocio y para los usuarios finales:

- Mejoras Operativas y Eficiencia: La automatización del reconocimiento de color puede mejorar la velocidad y precisión en procesos donde el color es un factor crítico, reduciendo errores humanos y tiempos de operación.
- Toma de Decisiones Más Informada: Con datos precisos sobre los colores predominantes en imágenes, las decisiones en áreas como marketing, diseño de productos y control de calidad pueden ser más informadas.
- Mejoras en la Experiencia del Cliente: La capacidad para identificar y categorizar rápidamente colores puede mejorar la experiencia del cliente en entornos digitales, como e-commerce, donde la visualización y selección de productos basada en color es esencial.
- Generación de Nuevas Oportunidades de Negocio: Los insights derivados del análisis de color pueden revelar nuevas oportunidades de mercado o mejoras en la segmentación de productos.




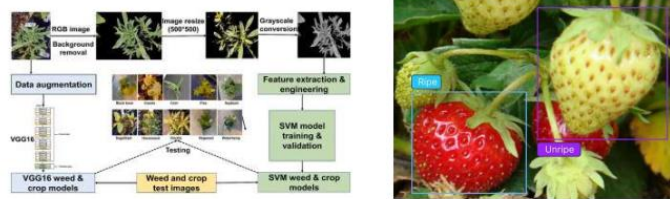
Item	Application	Industry	Applicable Clients	Scenario	Object to be Detected	Behavior to be Monitored	Reference Images
1	Vest/Helmet People Tracking	Manufacturing / Logistics / Construction	Distribution Centers Manufacturing Plants Real Estate Developers	Security alerts regarding unauthorized entry of individuals wearing clothing (vests or helmets) in colors that indicate they are not permitted to enter or move around in the area of interest.	Security Vests/Helmets	Alert of the presence of a color not allowed for a specific period of time.	 <p>SAFETY HELMET COLOR CODE ON CONSTRUCTION SITE</p> <p>WHITE: Supervisors, Engineers, Managers and Foremen</p> <p>YELLOW: Labourers, Workers and Earth Movers</p> <p>BLUE: Electricians, Carpenters and Technicians</p> <p>BROWN: Writers and Workers with High Heat</p> <p>RED: Fire Fighters</p> <p>GREEN: Safety Officers</p> <p>GRAY: Site Visitors</p>
2	FIFO Compliance Monitoring	Logistics / Retail	Distribution Centers Retail Stores	"FIFO" inventory management is applied where the oldest inventory acquired is sold before inventory acquired more recently. Color coded labeling is applied to clarify which items belong to which lots.	Color coded labels that represent FIFO Schedule.	Alerts and reports on color-coded labels detected dependent on the approved color for a specific timeframe.	
3	Shelf Share Calculation	Retail	Retail Stores	Each store must adhere to contractual agreements that mandate a specific percentage of certain brands to be displayed. To address this challenge, certain companies focus on manually counting products to determine the share of different brands on supermarket shelves.	Shelf products differentiated by color	Product counting by type/color for post processing share calculation.	 <p>Share B1 = $3B1 / (3B1 + 5B2) = 37.5\%$</p> <p>Share B2 = $5B2 / (3B1 + 5B2) = 62.5\%$</p> <p>Example of product's share on a shelf</p>
4	Controlled Environment Yield Estimation	Agriculture	Indoor Vertical Farming Hydroponic farms Cannabis farms	Indoor agriculture facilities have already security cameras in place to monitor access and overall operations. Yield and crop status estimation can leverage existing vision sensors to provide alerts based on batches of crops reaching a specific color-based status estimation.	Leaves of the crops exhibiting color shades that correspond to nutrient density and growth stages (e.g., light green for young leaves, dark green for mature and nutrient-rich leaves).	Estimating and categorizing yield quality and growth stages of leafy greens, aiding in harvest planning and nutrient management.	

Tabla 4. Casos de uso para la detección de objetos personalizados basado en colores.

Riesgos y Desafíos Potenciales

El uso de aprendizaje automático para identificar colores dominantes en imágenes implica varios riesgos y desafíos potenciales:

1. Calidad y Disponibilidad de Datos: La precisión del modelo depende en gran medida de la calidad y diversidad de los datos de entrenamiento. Una mala calidad de datos, incluyendo datos incompletos o sesgados, puede limitar la capacidad del modelo para aprender eficazmente y producir predicciones de color precisas.
2. Sobreajuste del Modelo: Existe el riesgo de que el modelo se ajuste demasiado a los datos de entrenamiento, funcionando bien en conjuntos de datos conocidos pero mal en nuevos datos no vistos, impactando de igual manera en las predicciones de color dominante.
3. Sesgo Algorítmico: Si los datos de entrenamiento no representan la diversidad de escenarios del mundo real, el modelo podría desarrollar sesgos que podrían llevar a resultados sesgados o injustos, como favorecer ciertos tonos de color sobre otros.
4. Adaptabilidad a Nuevos Datos: El entorno en el que se toman las imágenes puede variar mucho, afectando el rendimiento del modelo debido a desviaciones de concepto. Este riesgo es en particular mayor considerando la alta variabilidad de casos de uso para los cuales la herramienta Flex AI ha sido diseñada.
5. Complejidad del Sistema e Integración: Integrar el modelo en sistemas existentes podría ser complejo y podría llevar a problemas operativos si no se maneja adecuadamente. Especialmente considerando la implementación del modelo en dispositivos en borde (en cámaras de seguridad) donde puede existir una variabilidad en las arquitecturas de procesadores para los cuales el modelo tenga que ser compilado.

Referencias

1. Project GitHub Repository: <https://github.com/rzunick/MNA-Proyecto-Integrador-Equipo-43>
2. *CIFAR 10- CNN using PyTorch*. (n.d.). Retrieved May 29, 2024, from <https://kaggle.com/code/shadabhussain/cifar-10-cnn-using-pytorch>
3. *Color and Product Type Inspection with Computer Vision*. (2023, October 6). Roboflow Blog. <https://blog.roboflow.com/how-to-analyze-product-color/>
4. De, K., & Pedersen, M. (2021). Impact of Colour on Robustness of Deep Neural Networks. *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 21–30. <https://doi.org/10.1109/ICCVW54120.2021.00009>
5. Flachot, A., & Gegenfurtner, K. R. (2021). Color for object recognition: Hue and chroma sensitivity in the deep features of convolutional neural networks. *Vision Research*, 182, 89–100. <https://doi.org/10.1016/j.visres.2020.09.010>
6. Gowda, S., & Yuan, C. (2019). *ColorNet: Investigating the Importance of Color Spaces for Image Classification* (pp. 581–596). https://doi.org/10.1007/978-3-030-20870-7_36
7. Khan, F., Anwer, R., Weijer, J., Bagdanov, A., Vanrell, M., & López, A. (2012). Color Attributes for Object Detection. *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2012.6248068>
8. Maitlo, N., Noonari, N., Ghanghro, S. A., Duraisamy, S., & Ahmed, F. (2024b). *Color Recognition in Challenging Lighting Environments: CNN Approach* (arXiv:2402.04762; Version 1). arXiv. <https://doi.org/10.48550/arXiv.2402.04762>
9. Maurício, J., Domingues, I., & Bernardino, J. (2023). Comparing Vision Transformers and Convolutional Neural Networks for Image Classification: A Literature Review. *Applied Sciences*, 13(9), Article 9. <https://doi.org/10.3390/app13095521>

10. Rachmadi, R. F., & Purnama, I. K. E. (2018). *Vehicle Color Recognition using Convolutional Neural Network* (arXiv:1510.07391). arXiv. <http://arxiv.org/abs/1510.07391>
11. Smits, T., & Wevers, M. (2021). The agency of computer vision models as optical instruments. *Visual Communication*, 21, 147035722199209. <https://doi.org/10.1177/1470357221992097>
12. Top 10 Datasets for Computer Vision · Syntrix AI. (2022). Retrieved from <https://www.syntrix.ai/post/unlocking-computer-vision-the-10-best-datasets>
13. De, K., & Pedersen, M. (2021). Impact of Colour on Robustness of Deep Neural Networks. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 21–30. <https://doi.org/10.1109/ICCVW54120.2021.00009>
14. FiftyOne by Voxel 51. (n.d.). Retrieved from <https://docs.voxel51.com/>
15. Clustering Images with Embeddings. (n.d.). Retrieved from <https://docs.voxel51.com/tutorials/clustering.html>
16. Gallagher, J. (2024). Color and Product Type Inspection with Computer Vision. Retrieved from <https://blog.roboflow.com/how-to-analyze-product-color/>
17. skimage.color#. (n.d.). Retrieved from <https://scikit-image.org/docs/dev/api/skimage.color.html>
18. Colors Tutorial. W3 Schools. (N.d.). Retrieved from https://www.w3schools.com/colors/colors_wheels.asp#:~:text=The%20%20primary%20colors%20are,Orange%2C%20Green%2C%20and%20Purple.&text=Blue%20and%20Red%20%3D%20Purple
19. Rogers, J. (2024). Artificial Intelligence Risk & Governance. Retrieved from <https://ai.wharton.upenn.edu/white-paper/artificial-intelligence-risk-governance/>