

**COMPARACIÓN DE ALGORITMOS BASADOS EN INTELIGENCIA ARTIFICIAL PARA
LA DETECCIÓN DEL USO DEL TAPABOCAS**

Sebastián Barbudo González

Trabajo Dirigido

Tutores

**Carlos Andrés Cifuentes PhD
Marcela Múnera PhD**



**UNIVERSIDAD DEL ROSARIO
ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO
PROGRAMA DE INGENIERÍA BIOMÉDICA**

BOGOTÁ D.C
TABLA DE CONTENIDO

1. INTRODUCCIÓN	5
2. OBJETIVOS	9
2.1. General	9
2.2. Específicos	9
3. METODOLOGÍA.....	10
3.1. Definición de variable	10
3.2. Tipo de estudio y diseño general.....	10
3.3. Universo de estudio	10
3.4. Equipos e instalaciones	11
3.5. Propuesta.....	11
4. RESULTADOS Y DISCUSIÓN	13
5. RECOMENDACIONES Y TRABAJOS FUTUROS	16
5.1. Propuesta futura	18
6. CONCLUSIONES.....	19
REFERENCIAS	20

LISTA DE TABLAS

<i>Tabla 1. mAP y FPS de los modelos entrenados en el Moxa3K.....</i>	<i>6</i>
<i>Tabla 2. Conteo valores estadísticos por algoritmo</i>	<i>14</i>
<i>Tabla 3. Comparación de algoritmos de detección de tapabocas por medio de sus resultados estadísticos.....</i>	<i>14</i>

LISTA DE FIGURAS

<i>Figura 1. Muestras de imágenes del RMFD.</i>	6
<i>Figura 2. Comparación de los modelos entrenados con algunas muestras.</i>	7
<i>Figura 3. Entrada de la Escuela Colombiana de Ingeniería Julio Garavito.</i>	12
<i>Figura 4. Representación ilustrativa de valores estadísticos.</i>	13
<i>Figura 5. Detección de tapabocas.</i>	15
<i>Figura 6. Diagrama de control con módulos de imagen a color y cámara térmica.</i>	16
<i>Figura 7. Salidas del módulo A y B con su respectiva superposición C.</i>	17

1. INTRODUCCIÓN

Con la llegada del SARS-CoV-2, el uso del tapabocas se ha convertido en un nuevo hábito [1]. Actualmente los diferentes gobiernos han decretado normas en donde el uso de este es indispensable para las personas mientras estén en contacto con otros, es por esto por lo que los establecimientos públicos han empezado a controlar el ingreso de la gente, asegurando mediante diferentes métodos que tengan el tapabocas puesto correctamente y que no presenten síntomas de esta enfermedad.

La inteligencia artificial basada en “Machine learning” y “Deep learning” ha sido una gran herramienta para enfrentar la pandemia, debido a que los investigadores al tener acceso a grandes cantidades de información para prevenir la propagación del virus pueden desarrollar mecanismos de defensa o de alerta para la actual o futuras pandemias. El uso de bases de datos y datasets es fundamental para la aplicación de la IA en estas situaciones, ya que, si se establecen imágenes representativas de personas que porten o no el tapabocas, por medio de redes neuronales los sistemas computarizados pueden aprender a identificar si una persona cumple las reglas o no facilitando el control de masas en interiores públicos [2].

Varios estudios que buscan diseñar modelos o probar sistemas que permitan detectar el uso del tapabocas han aplicado estos conceptos de IA en sus investigaciones y han logrado desarrollar nuevos o mejorar los ya existentes algoritmos que permiten el continuo aprendizaje mediante algunas librerías como TensorFlow, Keras y OpenCV, Pytorch y Deep learning para un aprendizaje automático. Uno de los sistemas de reconocimiento de objetos más usados en este campo ha sido el modelo YOLO (You Only Look Once) el cual es un sistema de código abierto que funciona en tiempo real que se caracteriza por hacer uso de una red neuronal convolucional (CNN) que permite detectar objetos específicamente en imágenes.

A lo largo del documento se podrán apreciar las diferentes versiones existentes del YOLO y como ha contribuido en diversos proyectos. Otras estrategias para la detección de rostros y uso del tapabocas consisten en el SSD (Single Shot MultiBox Detector) implementando esta vez a la red ResNet permitiendo el entrenamiento de las redes neuronales profundas [3].

Los avances actuales con respecto a la detección de objetos, que, en este caso se aplica directamente sobre el uso del tapabocas, tienden a seguir metodologías parecidas. El uso de la inteligencia artificial en estos proyectos parece ser constante, además todos incluyen a las redes neuronales como el motor fundamental que permite que se logre el aprendizaje correcto dentro del sistema y así poder obtener resultados óptimos. Un término que generalmente aparece cuando se utilizan grandes cantidades de datos es “Dataset”, que dentro de las ramas de la IA es demasiado importante por su gran influencia sobre las redes neuronales y su proceso de aprendizaje. Por ejemplo, crear un dataset de imágenes con personas con o sin tapabocas para utilizarlas como referencia y que el programa recuerde los patrones y pueda predecir acciones en el futuro, así como realizaron en la Universidad de California en Estados Unidos en donde los desarrolladores tomaron un dataset de la RMTD (Real- World Masked Face Dataset) y lo implementaron en su algoritmo para instruir al programa y alimentar a la red neuronal [4]. El modelo a desarrollar de ellos consistía

básicamente en dos partes, la primera involucraba Deep learning con el ResNet50 (red residual), y por medio de lo que se conoce como “Extracción de Características” (Feature extraction) transformaban la información de las imágenes del dataset (Figura 1) en características para poder enviarlas a la red e identificarlas posteriormente, mientras que la segunda parte consistía en el “Machine learning” tradicional cuyos clasificadores son, árboles de decisión y máquinas de vectores de soporte que una vez entrenadas se hacían las pruebas de precisión sobre cada clasificador mostrando que todos lograban alcanzar más del 96% de precisión en la detección del tapabocas [5].



Figura 1. Muestras de imágenes del RMFD [5].

El método anterior de detección a través de redes residuales no es el único existente para la detección de objetos, de hecho, desarrolladores de India han propuesto el monitoreo del uso del tapabocas mediante el uso de otros algoritmos como el YOLOv3, YOLOv3Tiny, SSD y Faster R-CNN y entrenarlos con en el dataset “Moxa3K” usando diferentes tamaños de imagen. Estos métodos en general tuvieron resultados positivos con un mAP (mean of average precision) cabe mencionar que estos valores cambiaban con respecto al tamaño de las imágenes (Tabla 1) y que cada modelo permitía captar información en tiempo real con sus respectivos FPS [6].

Tabla 1

mAP Y FPS DE LOS MODELOS ENTRENADOS EN EL MOXA3K

Modelo	mAP@50	FPS
YOLOv3 414x414	63.99	21.2
YOLOv3 608x608	66.84	10.9
YOLOv3 832x832	61.73	6.9
YOLOv3Tiny 414x414	56.27	138
YOLOv3Tiny 608x608	55.08	72
YOLOv3Tiny 832x832	56.57	46.5
SSD 300 MobliNetv2	46.52	67.1
F-RCNN 300 Inceptionv2	60.5	14.8

Finalmente, para probar el modelo completo hicieron pruebas sobre imágenes descargadas de internet en donde pudieron diferenciar todos los métodos. En la Figura 2, se puede ver el resultado final de utilizar estos algoritmos y ponerlos en práctica. Adicionalmente, se pueden ver que las imágenes captadas por los diferentes modelos generan un recuadro verde en el rostro de las personas cuando se detecta el tapabocas y un recuadro de otro color ya sea rosado o azul si no se porta el tapabocas.

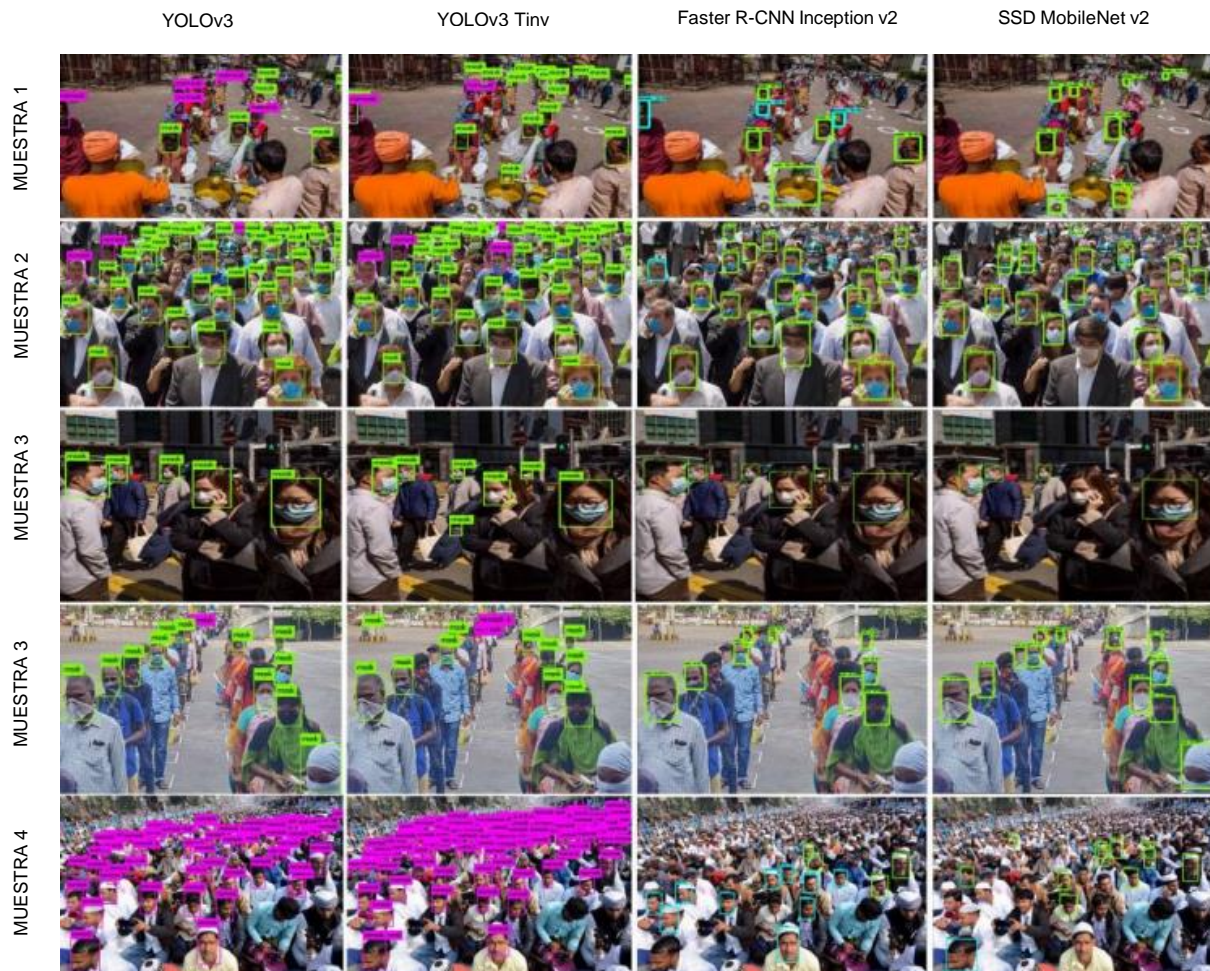


Figura 2. Comparación de los modelos entrenados con algunas muestras [7].

Desde la aparición del Covid-19 se ha incrementado el uso de estas mascarillas debido a que es uno de los métodos más efectivos y comunes para evitar el contagio, por lo que se ha ido desarrollando con mayor fuerza el diseño de técnicas que permitan detectar su uso. Los trabajos relacionados que han realizado los investigadores es evaluar el rendimiento de los algoritmos que desarrollan y compararlos con los modelos preexistentes mediante indicadores estadísticos de eficacia como precisión, exactitud y memoria. Para ello han realizado pruebas en diferentes espacios públicos como en supermercados, calles, aeropuertos o en transporte público.

En [7] los autores desarrollaron un nuevo método de identificación de la condición de uso de máscara facial, fusionándolo con la red SR y con la clasificación de imágenes faciales de la red de clasificación (SRCNet). Dividieron el algoritmo en cuatro pasos: imagen y preprocesamiento, detección y recorte de rostros, SR e identificación de las condiciones de uso de la mascarilla. El SRCN logró una precisión del 98,70% al detectar tres categorías que son: uso correcto de mascarilla facial, uso incorrecto de mascarilla facial y no uso de mascarilla facial.

Esta investigación busca evaluar algunos de los modelos de detección de objetos orientados al uso de máscaras mediante inteligencia artificial con el fin de conocer el desempeño de los algoritmos existentes y conocer las posibles diferencias entre ellos que conduzcan a un mejor porcentaje de eficiencia. y así poder determinar qué parámetros son los más requeridos para una detección de mascarillas más exitosa y garantizar un seguimiento social seguro y más efectivo.

Se pretende replicar estas pruebas en las instalaciones de la Escuela Colombiana de Ingeniería Julio Garavito, evaluar el desempeño de cuatro algoritmos cuyos modelos neuronales son ResNet y MobileNet y determinar qué diferencias existen y qué hacen a unos más precisos que a otros.

2. OBJETIVOS

2.1. General

Evaluar y comparar los diferentes sistemas computarizados, modelos y algoritmos que se han utilizado en la detección de tapabocas para el monitoreo social en espacios públicos.

2.2. Específicos

- a) Estudiar el desempeño y funcionamiento de los diferentes métodos de detección de tapabocas al implementarlos.
- b) Analizar y evaluar cuantitativamente los criterios estadísticos de precisión, exactitud y sensibilidad de las causas asociadas a las variaciones de precisión y desempeño general entre algoritmos.

3. METODOLOGÍA

Durante este estudio, se utilizaron las instalaciones de la Escuela de Ingeniería de Colombia Julio Garavito para probar los diferentes métodos de detección de mascarillas seleccionados para realizar sus evaluaciones y comparaciones. En el ingreso a la universidad se hicieron diferentes videos capturando a personas que estaban por entrar y salir de la institución, luego las grabaciones pasaron por ciertos pasos. En primer lugar, se procesaron los cuatro algoritmos elegidos. Luego, cada video se dividió en cuadros y se analizó cuadro por cuadro. Finalmente, los criterios estadísticos de efectividad como precisión, exactitud y sensibilidad se determinaron contando los falsos positivos, falsos negativos, verdaderos positivos y verdaderos negativos para determinar el desempeño de los algoritmos.

3.1. Definición de variable

Para evaluar los resultados de este estudio, se medirán las siguientes variables:

- **Exactitud:** La Exactitud permite medir el porcentaje de casos que el modelo ha acertado.
- **Precisión:** La Precisión nos permite medir la calidad del modelo.
- **Exhaustividad o Sensibilidad:** La Exhaustividad o Sensibilidad permite encontrar la cantidad de datos que el modelo permite identificar.
- **Valor F1:** El Valor F1 como una medida general, permite combinar los resultados de Precisión y Exhaustividad en un solo valor para determinar el rendimiento en general del modelo.
- **Verdadero positivo:** Número de casos donde se detecta tapabocas y realmente se dispone del tapabocas.
- **Verdadero negativo:** Número de casos donde no se detecta tapabocas y realmente no se dispone del tapabocas.
- **Falso positivo:** Número de casos donde se detecta tapabocas y realmente no se dispone del tapabocas.
- **Falso negativo:** Número de casos donde no se detecta tapabocas y realmente se dispone del tapabocas.

3.2. Tipo de estudio y diseño general

El estudio presentado en este documento es de carácter transversal, descriptivo y prospectivo.

3.3. Universo de estudio

El estudio está orientado para realizarse en espacios públicos en tiempos de pandemia por COVID-19 para garantizar el cumplimiento de las normas del uso del tapabocas. Principalmente, se desarrollará en las instalaciones de la Escuela Colombiana de Ingeniería Julio Garavito, sin embargo, el universo de estudio podría ampliarse al implementar las pruebas en otras entidades.

3.4. Equipos e instalaciones

- Cámara de CCTV.
- Computador portátil con los algoritmos seleccionados.
- Lenguaje de programación Python.
- Grabación objetivo.
- Entrada e instalaciones de la universidad.

3.5. Propuesta

Se instaló una cámara de circuito cerrado de televisión en la entrada de la institución con el fin de realizar una grabación durante todo el día de las personas que ingresan a la universidad y así garantizar el cumplimiento de las normas de uso de máscaras. El objetivo fue crear un conjunto de datos fragmentando los videos para poder recrear diferentes escenas y evaluar el rendimiento de los algoritmos de detección de máscaras. La Figura 3 muestra la entrada de la universidad donde se realizó la prueba, la cámara fue incorporada apuntando hacia la carretera y colocada cerca de los torniquetes, de esa manera tenía suficiente alcance para detectar miembros de la institución que están llegando y pueden incluir más de una persona dentro del rango de visión de la cámara. Algunos de los videos fueron tomados en otros espacios interiores dentro de la institución como algunos laboratorios.

En cuanto al proceso de evaluación del algoritmo, una vez que el video ha sido fragmentado en sus respectivas escenas, los cuatro algoritmos fueron analizados escena por escena, es decir que cada modelo tuvo que reproducirse. las mismas escenas para la detección de la máscara facial, una vez ejecutados los códigos, se contabilizaron los verdaderos negativos (VN), verdaderos positivos (VP), falsos positivos (FP) y falsos negativos (FN) para determinar estadísticamente la evaluación de desempeño criterios (exactitud, precisión, sensibilidad y Valor F1) que se describen mediante las siguientes ecuaciones:

$$Exactitud = \frac{VP+VN}{VP+VN+FP+FN} \cdot 100 \% \quad (1)$$

$$Precisión = \frac{VP}{VP+FP} \cdot 100 \% \quad (2)$$

$$Sensibilidad = \frac{VP}{VP+FN} \cdot 100 \% \quad (3)$$

$$Valor - F1 = 2 \cdot \frac{Sensibilidad \cdot Precisión}{Sensibilidad + Precisión} \quad (4)$$



Figura 3. Entrada de la Escuela Colombiana de Ingeniería Julio Garavito.

4. RESULTADOS Y DISCUSIÓN

Todas las grabaciones se almacenaron en un archivo que contiene los videos originales, los videos procesados y los fotogramas de cada video divididos en carpetas de los diferentes algoritmos que facilitan la organización de los datos y el análisis de la información.

Cada algoritmo tiene un tipo diferente de red neuronal. Los algoritmos 1 y 2 utilizan el modelo MobileNetV2, mientras que los Algoritmos 3 y 4 son construidos por la red neuronal residual ResNet50V2. Sin embargo, todos muestran una estructura de código similar: (1) importar bibliotecas y definir variables y modelos, (2) bucle de grabación de cámara en vivo, (3) condiciones de detección de máscara facial.

Vale la pena mencionar que todas las interfaces muestran un rectángulo que cambia su color cuando el código se está ejecutando. Cuando el algoritmo detecta una máscara facial el rectángulo se vuelve verde, en caso contrario se vuelve rojo.

En la Figura 4 se puede ver de forma ilustrativa cómo se muestran los valores estadísticos. Por ejemplo, la figura 4a muestra un caso verdadero positivo, 4b muestra un verdadero negativo, 4c muestra un falso positivo y finalmente 4d muestra un caso falso negativo. Adicionalmente, se puede observar el color que toma el rectángulo en cada caso.

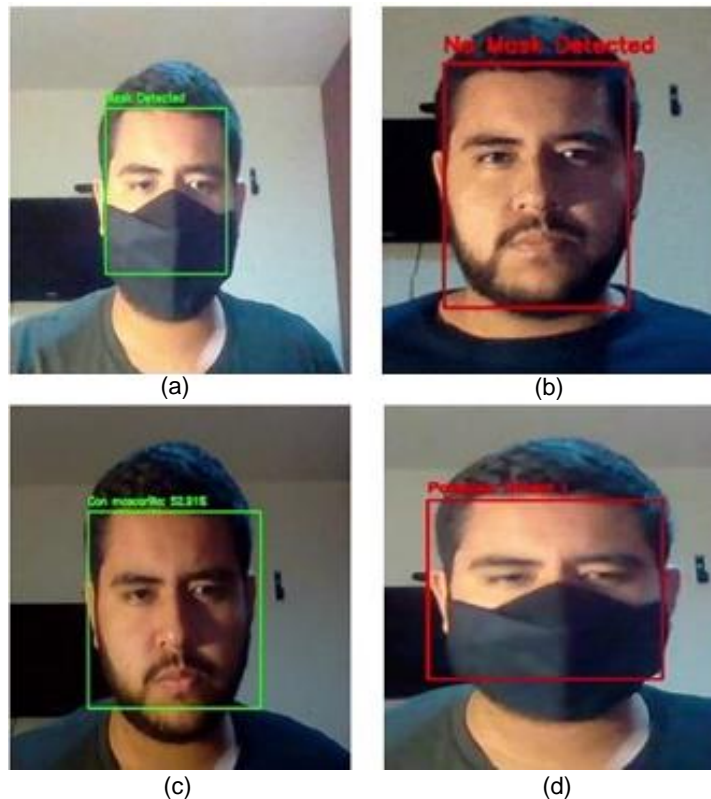


Figura 4. Representación ilustrativa de valores estadísticos. a) Verdadero positivo b) Verdadero negativo c) Falso positivo d) Falso negativo.

Se analizaron 28142 fotogramas. En la Tabla 1 se muestra el recuento de los valores estadísticos y los resultados generales de los algoritmos con respecto a los valores (TP, TN, FP, FN). Se puede observar que en general los verdaderos positivos son la mayoría con 18263 fotogramas en total, lo que nos lleva a buenos y aceptables sistemas de detección.

Tabla II
CONTEO DE VALORES ESTADÍSTICOS POR ALGORITMO

Método	VP	VN	FP	FN
Algoritmo 1	4270	584	11	3840
Algoritmo 2	1321	307	2	775
Algoritmo 3	7857	1093	8	1359
Algoritmo 4	4815	838	7	1055
General	18263	2822	28	7029

El algoritmo 3 presentó 7857 verdaderos positivos en total siendo el que obtuvo el mayor número de detecciones. Es importante señalar que debido a que los videos fueron tomados en espacios públicos, la probabilidad de encontrar a alguien sin máscara es muy baja, por lo tanto, la suma de los verdaderos negativos es baja y esto afecta los resultados.

Tabla III
COMPARACIÓN DE ALGORITMOS DE DETECCIÓN DE TAPABOCAS POR MEDIO DE SUS RESULTADOS ESTADÍSTICOS

Método	Exactitud	Precisión	Sensibilidad	Valor F1
Algoritmo 1	55,76105686	99,74305069	52,65104809	68,92099104
Algoritmo 2	67,69230769	99,84882842	63,02480916	77,27405674
Algoritmo 3	86,75002423	99,89828353	85,25390625	91,99695568
Algoritmo 4	84,18466121	99,85483202	82,02725724	90,06734007

Con respecto al desempeño general, la Tabla 2 muestra que los algoritmos 3 y 4 fueron los más destacados en el estudio. Ambos mostraron porcentajes de exactitud del 86,8% y 84,2% respectivamente. Cabe recordar que estos dos algoritmos utilizaron la red residual ResNet50V2. Mientras que los otros dos, aunque tenían buenos porcentajes de precisión del 55,8% y 67,7% respectivamente, podrían ser mejores.

Para los porcentajes de sensibilidad y Valor F1, los algoritmos 3 y 4 volvieron a ser los ganadores, alcanzando por encima del 82% para la sensibilidad y el 90,1% para el valor F1, mientras que los algoritmos 1 y 2 alcanzaron el 52,7%, el 63% para la sensibilidad y el 68,9%. 77,3% para el valor F1 respectivamente.

Es necesario reconocer que las redes neuronales convolucionales como MobileNetV2 son inferiores a las redes neuronales residuales como ResNet50V2. Algo que es importante decir es que la precisión de todos los algoritmos fue del 99%, significa que estos algoritmos lograron el objetivo principal de la investigación.

Los 4 códigos completos permitieron realizar una correcta detección de objetos, que en este caso fueron rostros y máscaras que mostraron una alta calidad a la hora de trabajar lo que explica la razón por la cual el número de falsos positivos es demasiado bajo. Todos estos porcentajes se calcularon a partir de las ecuaciones (1) a (4).

La Figura 5 muestra los cuatro algoritmos trabajando en un fotograma tomado del mismo video. Aquí se puede ver que los cuatro pudieron hacer una buena detección de verdaderos positivos en 2 personas al mismo tiempo.

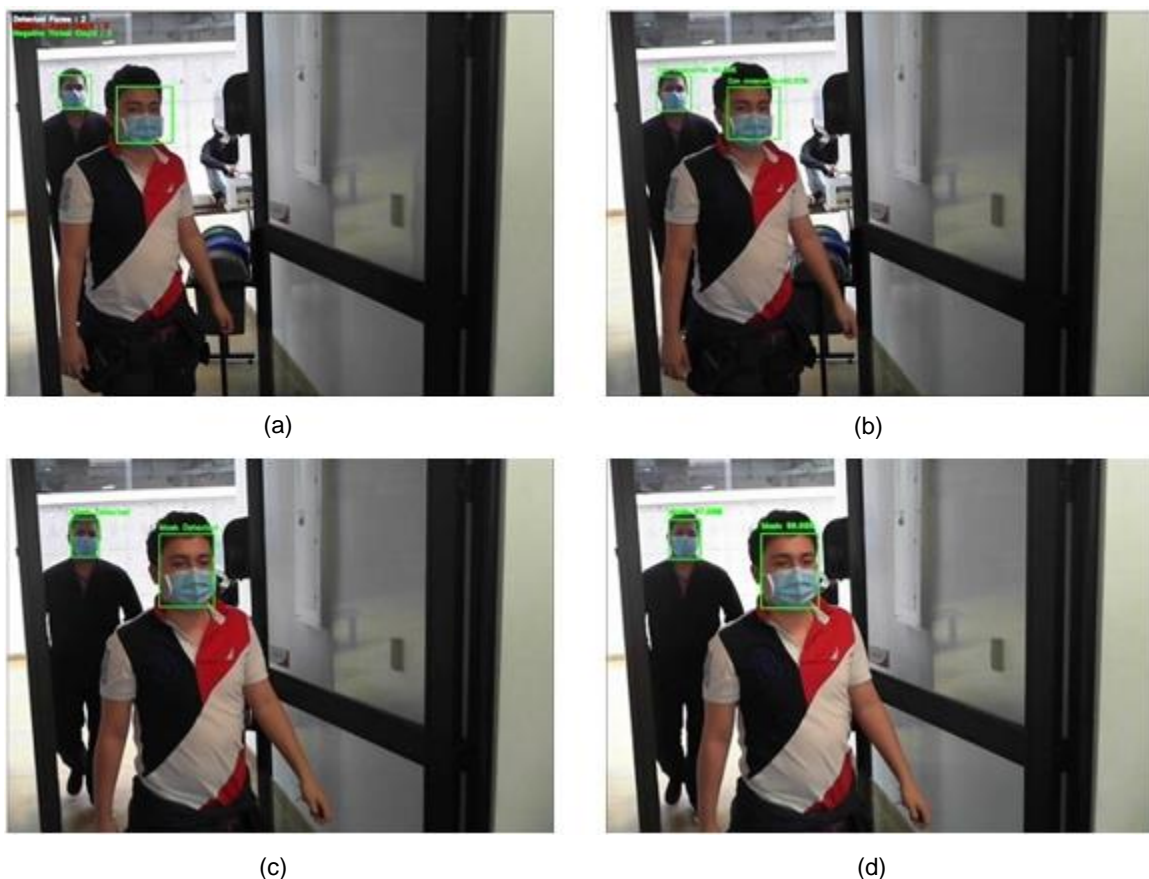


Figura 5. Detección de tapabocas. a) Detección Algoritmo 1 b) Detección Algoritmo 2 c) Detección Algoritmo 3 d) Detección Algoritmo 4.

5. RECOMENDACIONES Y TRABAJOS FUTUROS

Los modelos vistos son solo algunas estrategias que se han implementado y que han tomado fuerza específicamente durante el año pasado y lo que llevamos de este por la influencia que ha tenido el Covid-19 en el desarrollo tecnológico para combatirlo. Es por eso por lo que las innovaciones no se han detenido y se han propuesto nuevas formas para hacerle frente a este virus.

Anteriormente se había comentado acerca de las mejoras de los sistemas actuales de detección, pues en efecto estas están poco a poco implementándose. Aunque no exista mucha literatura sobre el tema, en Indonesia se diseñó un modelo en tiempo real usando Deep learning para la detección de tapabocas y además la temperatura facial de la persona. El detector de objetos usado fue el RetinaNet.

Se construyeron dos módulos diferentes para el modelo del RetinaNet. El primero consiste en la detección de 3 formas diferentes de tener el tapabocas cubriendo la zona facial, y el otro módulo es la detección de la temperatura. Para ello se implementaron cámaras RGB junto con cámaras térmicas que faciliten capturar las temperaturas respectivas [17].

La idea principal de los creadores del modelo es plantear un sistema de control en donde se realicen los diferentes módulos por aparte y luego se junten para resaltar las propiedades que cada uno tiene. En la figura 6 se puede ver el sistema completo en donde las entradas x_1 y x_2 representan los diferentes enfoques de detección, m_1 y m_2 las salidas correspondientes a cada módulo, y finalmente Z como la superposición de ambas salidas como un sistema completo que resalte las funciones de ambos.

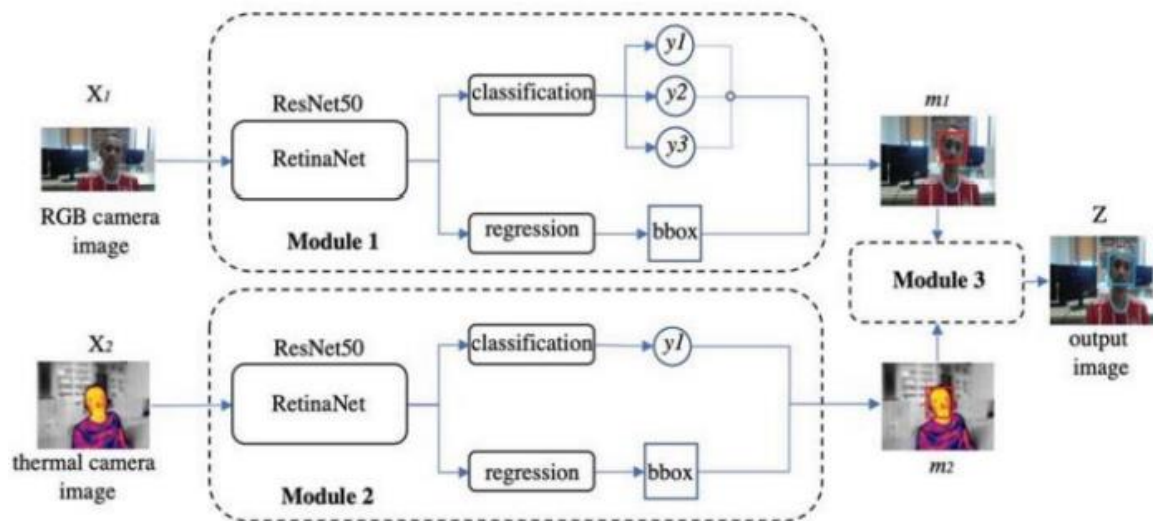


Figura 6. Diagrama de control con módulos de imagen a color y cámara térmica [17].

Este sistema finalmente permite conseguir imágenes como se ven en la figura 7, en donde la figura A representa la salida del primer módulo, la salida B representa la salida del segundo módulo por medio de la cámara térmica y C el resultado de combinar ambas salidas.



Figura 7. Salidas del módulo A y B con su respectiva superposición C [17].

El resultado final fue positivo y tuvo puntuación de confianza mayor al 80%. Logrando exitosamente sus propósitos.

5.1. Propuesta futura

Debido a que no se encuentra mucha información acerca de estos mecanismos mejorados para detectar más de un parámetro, así como la temperatura y el uso del tapabocas, se pretende implementar un sistema que permita la detección de ambos objetos de una manera más eficiente y que avise en caso de que no se cumplan con los requisitos. Se implementará a un algoritmo ya existente con librerías de aprendizaje TensorFlow, Keras y OpenCV el módulo de temperatura para poder sobreponer ambas salidas sin perder la efectividad de ninguna. Adicionalmente se busca mejorar los porcentajes de precisión para alcanzar un modelo más sólido y contundente.

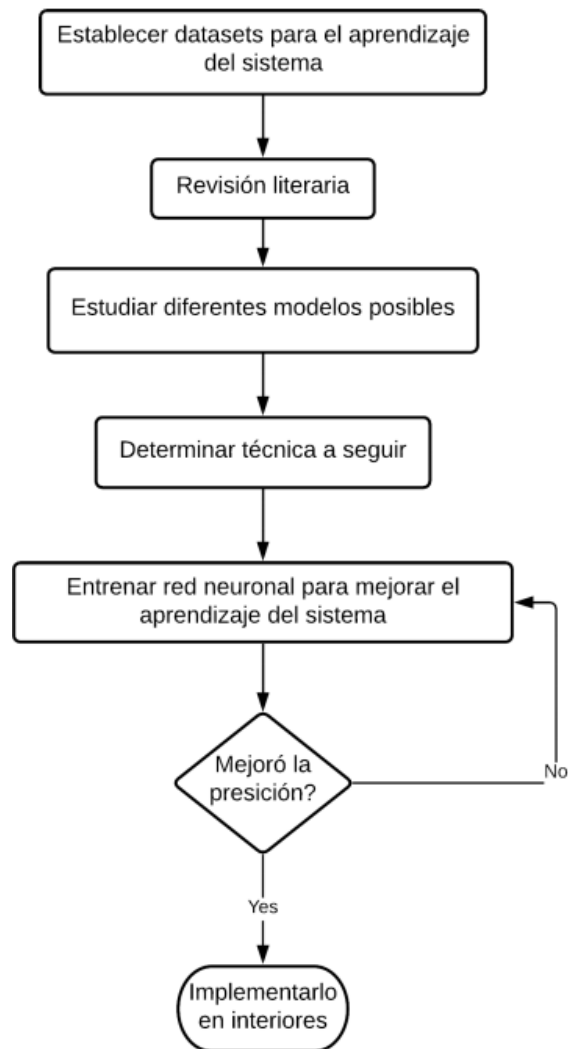


Diagrama del plan a seguir

6. CONCLUSIONES

Se puede evidenciar que la correcta colocación de las cámaras de video permite una mejor detección por algoritmos. Es mejor disparar desde el frente en lugar de disparar desde un ángulo o altura específicos, sin embargo, esto en entornos reales es un poco complicado porque la población experimental no se puede controlar tan fácilmente. Además, el modelo y el tipo de red neuronal son importantes para el rendimiento de los algoritmos.

De acuerdo con los resultados observados, las redes residuales mostraron mejores resultados que las redes neuronales convolucionales donde los resultados generales para los algoritmos 3 y 4 que utilizan ResNet50V2 fueron considerablemente mejores que los algoritmos 1 y 2 que utilizaron MobileNetV2.

Los algoritmos de detección de máscaras faciales presentaron en general buenos resultados que nos permiten conocer la importancia de estos en el seguimiento social para controlar masas tanto en lugares públicos como en entidades privadas.

Se están llevando a cabo investigaciones futuras sobre el uso de estos algoritmos de detección de máscaras para crear una interacción humano-robot donde la persona puede recibir comentarios del robot como una notificación del estado de la máscara.

REFERENCIAS

- [1]. M. A. K. Basha, B. N. Pravallika, and E. Shankar, "An Efficient Face Mask Detector with PyTorch and Deep Learning," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 7, no. 25, p. 167843, 2021
- [2]. Y. Said, "Pynq-YOLO-Net: An Embedded Quantized Convolutional Neural Network for Face Mask Detection in COVID-19 Pandemic Era," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, 2020.
- [3]. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, and J. Hemanth, "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," *Sustainable Cities and Society*, vol. 66, p. 102692, 2021.
- [4]. Z. Wang, et al., Masked face recognition dataset and application, arXiv preprint arXiv:2003.09093, 2020.
- [5]. Loey, G. Manogaran, M. H. Taha, and N. E. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Measurement*, vol. 167, p. 108288, 2021.
- [6]. B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas, and T. Das, "MOXA: A Deep Learning Based Unmanned Approach For Real-Time Monitoring of People Wearing Medical Masks," *Transactions of the Indian National Academy of Engineering*, vol. 5, no. 3, pp. 509–518, 2020.
- [7]. B. QIN and D. Li, Identifying facemask-wearing condition using image superresolution with classification network to prevent COVID-19, May 2020, doi: 10.21203/rs.3.rs-28668/v1.
- [8]. A. S. Joshi, S. S. Joshi, G. Kanahasabai, R. Kapil, and S. Gupta, "Deep Learning Framework to Detect Face Masks from Video Footage," *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*, 2020.
- [9]. C. Gupta and N. Gill, "Coronamask: A Face Mask Detector for Real-Time Data," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 5624–5630, 2020.
- [10]. Loey, G. Manogaran, M. H. Taha, and N. E. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustainable Cities and Society*, vol. 65, p. 102600, 2021.

- [11]. M. Rahman, M. M. Manik, M. M. Islam, S. Mahmud, and J.-H. Kim, "An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network," 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 2020.
- [12]. M. R. Bhuiyan, S. A. Khushbu, and M. S. Islam, "A Deep Learning Based Assistive System to Classify COVID-19 Face Mask for Human Safety with YOLOv3," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020.
- [13]. S. K. Addagarla, "Real Time Multi-Scale Facial Mask Detection and Classification Using Deep Transfer Learning Techniques," International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, no. 4, pp. 4402–4408, 2020.
- [14]. S. Li, X. Ning, L. Yu, L. Zhang, X. Dong, Y. Shi, and W. He, "Multi-angle Head Pose Classification when Wearing the Mask for Face Recognition under the COVID-19 Coronavirus Epidemic," 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), 2020.
- [15]. W. Han, Z. Huang, A. kuerban, M. Yan, and H. Fu, "A Mask Detection Method for Shoppers Under the Threat of COVID-19 Coronavirus," 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020.
- [16]. X. Kong, K. Wang, S. Wang, X. Wang, X. Jiang, Y. Guo, G. Shen, X. Chen, and Q. Ni, "Real-time Mask Identification for COVID-19: An Edge Computing-based Deep Learning Framework," IEEE Internet of Things Journal, pp. 1–1, 2021.
- [17]. Farady, C.-Y. Lin, A. Rojanasarit, K. Prompol, and F. Akhyar, "Mask Classification and Head Temperature Detection Combined with Deep Learning Networks," 2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP), 2020.
- [18]. C. Dzien, W. Halder, H. Winner, and M. Lechleitner, "Covid-19 screening: are forehead temperature measurements during cold outdoor temperatures really helpful?," Wiener klinische Wochenschrift, 2020.
- [19]. Bu, Z. Deng, H. Liu, J. Li, D. Wang, Y. Yang, and S. Zhong, "Current methods and prospects of coronavirus detection," Talanta, vol. 225, p. 121977, 2021.
- [20]. Y. Xhou, P. Ghassemi, M. Chen, D. McBride, J. Casamento, J. Pfefer, and Q. Wang, "Clinical evaluation of fever-screening thermography: impact of consensus guidelines and facial measurement location," Journal of Biomedical Optics, vol. 25, no. 09, 2020.

[21]. H.-Y. Chen, A. Chen, and C. Chen, "Investigation of the Impact of Infrared Sensors on Core Body Temperature Monitoring by Comparing Measurement Sites," *Sensors*, vol. 20, no. 10, p. 2885, 2020.