

Mask-net: Identificación del uso correcto de mascarilla mediante visión por computador

Alexander Kalen-Targa, Alberto Landi-Cortiñas,
Nicolas Araque-Volk, Alejandro Marcano Van-Grieken

Universidad Metropolitana,
Facultad de Ingeniería,
Venezuela

{alexanderkalen, albertolandi}@correo.unimet.edu.ve,
{naraque, amarcano}@unimet.edu.ve

Resumen. Este trabajo se enfoca en crear un sistema de reconocimiento de uso correcto de una mascarilla mediante técnicas de visión por computador. Se realizó una investigación con el objetivo de establecer los lineamientos para el uso correcto e incorrecto de la mascarilla, los cuales permitieron determinar los criterios para la creación de los conjuntos de los datos “KLD” e “IMKLD”. Dichos conjuntos sirvieron para entrenar, validar y probar los modelos “Mask-net” e “I-Mask-net” respectivamente. Estos son modelos de arquitectura de Aprendizaje Profundo, donde se aplicó la técnica de “Aprendizaje por Transferencia” al utilizar “MobileNet” como una base para la extracción de atributos. Los resultados de los entrenamientos arrojaron que el ajuste de hiperparámetros realizado fue el adecuado en ambos casos, mientras que las pruebas hechas demostraron que los modelos tienen un buen porcentaje de exactitud.

Palabras clave: Mascarilla, visión por computador, inteligencia artificial, aprendizaje profundo, redes neuronales.

Mask-net: Detection of Correct Use of Masks through Computer Vision

Abstract. This paper focuses on creating a system for recognizing the correct use of a mask through computer vision techniques. Research was carried out with aims of establishing guidelines for the correct and incorrect use of a mask, which allowed for determining the criteria for the creation of the “KLD” and “IMKLD” datasets. These datasets were used to train, validate and test the “Mask-net” and “I-Mask-net” models respectively. The results given by training showed that the fine tuning carried out was adequate in both cases, while the tests carried out showed that the models have an acceptable level of accuracy.

Keywords: Masks, computer vision, artificial intelligence, deep learning, neural networks.

1. Introducción

El 11 de marzo del 2020, la Organización Mundial de la Salud (OMS), declaró el COVID-19, también conocido como coronavirus, como una pandemia a nivel mundial. Actualmente, miles de millones de personas han sido afectadas directa o indirectamente por el coronavirus. A nivel mundial, se han notificado más de 100 millones de casos confirmados de COVID-19, incluidas más de 2 millones de muertes, notificadas a la OMS [1]. Las medidas de emergencia obligatorias para controlar y evitar una mayor propagación del virus han causado cambios importantes en el estilo de vida de los seres humanos.

Uno de los problemas respecto al cumplimiento de medidas de bioseguridad es el uso incorrecto de la mascarilla; estudios preliminares realizados indican que [2] “la enfermedad se propaga principalmente de persona a persona a través de las gotículas que salen despedidas de la nariz o la boca de una persona infectada al toser, estornudar o hablar”. Por lo que el uso incorrecto de la misma pudiese generar más contagios, saturando los sistemas de salud y causando una mayor cantidad de fallecimientos.

Actualmente, el uso de la mascarilla es una de las medidas de prevención más efectivas a nivel mundial, y los primeros estudios realizados respecto a su uso arrojan resultados prometedores, como el realizado por Mitze et al. [3], donde afirman en un estudio hecho en Jena, Alemania, que el uso de la mascarilla ayudó a reducir en un 60 % la tasa de crecimiento diario de infecciones.

Ante estos beneficios, es indispensable extender la normativa de uso obligatorio de la mascarilla, fundamentalmente en espacios cerrados y/o de gran concentración de personas, y controlar que sean utilizadas correctamente. Los modelos de visión computacional son utilizados para resolver problemáticas complejas relevantes a imágenes o videos y pudiesen ser diseñados y entrenados para predecir el uso correcto de una mascarilla en una persona y así ayudar a solucionar este problema.

Seguidamente, se conoce que las redes neuronales convolucionales (CNN) son las redes más comunes en el campo de la visión por computador. Para Goodfellow et al. [4], son un tipo de red neuronal para el procesamiento de la data con una topología matricial conocida en forma de cuadrícula, siendo tremendamente exitosas en aplicaciones prácticas. La red emplea una operación matemática llamada convolución, siendo tremendamente exitosa para esta aplicación.

Utilizar un modelo de inteligencia artificial para detectar el uso correcto de la mascarilla sin intervención humana puede ser de gran utilidad en espacios públicos y privados de alto tránsito de personas, ya que se evitaría la revisión manual de la misma, promoviendo así las medidas de distanciamiento social con una revisión fiable. Según Johns Hopkins Medicine & Maragakis [5], se considera el uso correcto de la mascarilla aquellos casos en los que se tapan la nariz, la boca y la barbilla de una cara en su totalidad, consiguiendo así la mayor eficacia de la mascarilla facial.

Por otro lado, se considera uso incorrecto de la mascarilla cualquier caso en el que la nariz, la boca o la barbilla de un rostro estén expuestos, parcial o totalmente. Partiendo del argumento anterior, el objetivo de la presente investigación es implementar un sistema de reconocimiento de uso correcto de una mascarilla médica mediante técnicas de visión por computador.



Fig. 1. Cambios de color de mascarillas en muestras.

Para ello, se desarrolló el modelo de aprendizaje profundo “Mask-net” utilizando la arquitectura de red neuronal convolucional para reconocer el uso correcto, incorrecto, o no uso de la mascarilla. Adicionalmente, se creó otro modelo de aprendizaje profundo “I-Mask-Net”, que utiliza la salida del modelo anterior cuando se identifica uso incorrecto de la misma, este modelo indica si la mascarilla está en la barbilla, en la boca y la barbilla, o en la nariz y la boca, para poder brindar una sugerencia al usuario sobre cómo portarla correctamente.

2. Metodología

2.1. Conjunto de datos

Puesto que no existe un conjunto de datos preconstruido que se adhiera a los parámetros preestablecidos, la creación y diseño de conjuntos de datos personalizados de alta calidad se convirtió en uno de los enfoques de mayor importancia para el desarrollo de los modelos de aprendizaje automático Mask-net e I-Mask-net. Dicho lo anterior, se construyeron dos conjuntos de datos: KLD e IMKLD.

KLD es un conjunto de datos compuesto por un total de 11.003 imágenes, con el objetivo de entrenar y validar el modelo Mask-net. Las imágenes seleccionadas se originaron de múltiples conjuntos de datos de dominio público disponibles en internet, y algunas de autoría propia, entre las cuales se incluyen:

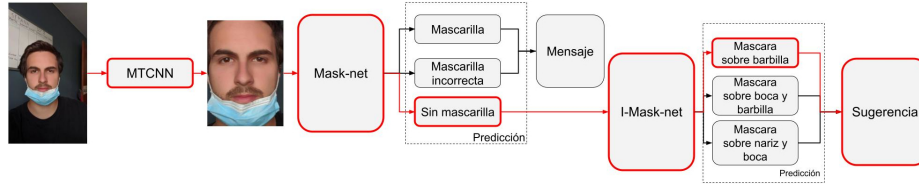


Fig. 2. Implementación combinada de los modelos Mask-net e I-Mask-net.

Tabla 1. Matriz de confusión de Mask-net.

		PREDICCIÓN		
		Mascarilla	Mascarilla incorrecta	Sin mascarilla
CLASE REAL	Mascarilla	55	1	0
	Mascarilla incorrecta	13	42	2
	Sin mascarilla	4	3	50

- Medical Mask Dataset [6], que proporcionó imágenes de personas reales usando mascarilla correctamente.
- Flickr-Faces-HQ [7], que proporcionó imágenes de personas reales, sin mascarilla.
- MaskedFace-net [8], que proporcionó imágenes de personas usando mascarilla médica de manera correcta e incorrecta.
- Autoría propia, un conjunto de imágenes que captaron situaciones del mundo real, tomando en cuenta ángulos de toma, fondos y uso de la mascarilla, así como bordados, colores y tipos utilizadas en el conjunto de pruebas. Dichas imágenes fueron proporcionadas por familiares, amigos y conocidos.

IMKLD es un conjunto de datos compuesto por un total de 1.841 imágenes, con el objetivo de entrenar y validar el modelo I-Mask-net. Las imágenes seleccionadas se originaron en el conjunto de datos MaskedFace-net, el cual proporcionó las tres clases de imágenes de uso incorrecto de la mascarilla. Además, algunas imágenes de autoría propia fueron incorporadas a dicho conjunto para incrementar la cantidad de casos reales.

Las imágenes están clasificadas en 3 clases según las partes del rostro cubiertas por la mascarilla: barbilla, boca y barbilla, y nariz y boca. Los conjuntos de datos KLD e IMKLD fueron almacenados localmente, donde los archivos fueron separados en conjuntos de entrenamiento, validación y prueba, cada uno subdividido en subdirectorios según las clases de dicho conjunto de datos.

KLD se distribuyó en 8.804 imágenes para entrenamiento, 2.199 imágenes para validación y 170 imágenes para prueba. Mientras que el conjunto de imágenes IMKLD se distribuyó en 1.610 imágenes de entrenamiento, 231 imágenes de validación y 57 imágenes de prueba.

Tabla 2. Métricas de precisión (presicion), exhaustividad (recall) de Mask-net.

Clase	Presición	Exhaustividad
Mascarilla	0,7638	0,9821
Mascarilla incorrecta	0,9130	0,7368
Sin mascarilla	0,9615	0,8771

2.2. Preprocesamiento de imágenes

Con el objetivo de mantener un orden en los conjuntos de datos, es necesario establecer una nomenclatura común. Para ello, se creó un algoritmo con el que se enumeraron y renombraron cada una de las miles de imágenes pertenecientes a ambos conjuntos de datos. También se consideró fundamental reducir la información innecesaria (tales como fondos) que pudiese generar sesgo en los datos, así como una reducción en la velocidad de entrenamiento del modelo.

Esta problemática se resolvió al establecer una Región de Interés (RoI) que envolviese los rostros de las muestras suministradas. Utilizando un algoritmo en Python, se procesaron las imágenes de ambos conjuntos de datos. Alimentando iterativamente cada una de las imágenes a MTCNN [9], se marcaron los bordes detectados de la Región de Interés, procediendo a recortar y posteriormente guardar la imagen de salida, teniendo como resultado solo los rostros de los sujetos para cada una de las muestras.

Por último, un problema de sesgo encontrado en ambos conjuntos corresponde al color azul de la mascarilla quirúrgica, ya que estas imágenes forman un porcentaje importante de la totalidad disponible. Para mitigar esto, fue necesario suministrar la mayor variación en colores y patrones de mascarillas de manera que el modelo a entrenar generalice lo mejor posible.

Una solución a esta problemática involucró la construcción de un algoritmo, que detecta zonas con la presencia de un rango de color seleccionado (en el espacio de color HSV), en este caso, tonalidades de color azul claro presente en las mascarillas quirúrgicas. Luego de detectar dichas zonas de coincidencia en las imágenes suministradas, se comenzó a plasmar sobre tales áreas otras coloraciones y tonalidades proporcionadas por valores aleatorios del espacio de color BGR, de tal manera que la variedad de mascarillas encontradas fuese más cercanas a la realidad.

2.3. Arquitectura de la red neuronal

Mask-net e I-Mask-net fueron diseñados con arquitecturas similares haciendo uso del Aprendizaje por Transferencia, lo que implicó el uso de una arquitectura predeterminada, siendo MobileNet el modelo preferido para construir la arquitectura base.

El modelo Mask-net cuenta con 85 capas profundas, de las cuales 82 son capas de extracción de atributos de MobileNet y 3 son capas de ajuste propio (fine tuning), que corresponden a una capa oculta de 512 nodos, seguida de una capa de Dropout con un índice de 0,2, culminando con una capa de salida para las tres clases determinadas, con función de activación Softmax.



Fig. 3. Predicciones correctas del Modelo Mask-net por clases.

Por otra parte, I-Mask-net cuenta con 84 capas profundas y 2 de ajuste propio, donde se conectó una capa de Dropout con un índice de 0,6, culminando con una capa de salida de función de activación Softmax, para las tres clases determinadas. Posteriormente, se congelaron ambos modelos con pesos correspondientes al preentrenamiento realizado por MobileNet con el conjunto de datos ImageNet, siendo entrenables las últimas 23 capas de Mask-net, y las últimas 34 capas en el caso de I-Mask-net.

Al congelar las capas, se permitió al modelo retener conocimientos de atributos generales extraídos de ImageNet, conjunto de datos que presenta clases similares a KLD (mask, respirator, gas mask, gas helmet, ski mask, oxygen mask), y diferentes en el caso de IMKLD, debido a que no existen clases específicas.

2.4. Entrenamiento de los modelos

Para el modelo Mask-Net, el entrenamiento final fue realizado localmente en Jupyter Notebooks, haciendo uso de aceleramiento de hardware por GPU “Nvidia GTX 1050ti with Max-Q Design”.



Clase verdadera: Mascarilla correcta									
									
Mascarilla incorrecta									
Clase verdadera: Mascarilla incorrecta									
									
Mascarilla	Mascarilla	Mascarilla	Sin mascarilla	Mascarilla	Mascarilla	Sin Mascarilla	Mascarilla	Mascarilla	Mascarilla
Clase Verdadera: Sin mascarilla									
									
Mascarilla	Mascarilla	Mascarilla	Mascarilla incorrecta	Mascarilla incorrecta	Mascarilla	Mascarilla incorrecta			

Fig. 4. Predicciones incorrectas del Modelo Mask-net por clases.

Tabla 3. Matriz de confusión de I-Mask-net.

		PREDICCIÓN		
		Barbilla	Boca y barbilla	Nariz y boca
CLASE REAL	Barbilla	16	2	1
	Boca y barbilla	2	14	3
	Nariz y boca	0	0	19

Se usó la inicialización de parámetros “random uniform”, con media 0 y desviación estándar 1, implementado sobre la capa lineal de 512 nodos; además, se regularizaron los pesos de dicha capa con regularizador L2 igualado a 0,01. El modelo se compiló utilizando el optimizador “Adam” con tasa de aprendizaje igualada en 0,0001 y función de pérdida “Categorical Cross Entropy”.

Se hizo uso de técnicas de aumento de datos implementando volteado horizontal y la rotación de imágenes respecto a su eje vertical hasta 20°. Además, el uso de la capa de Dropout que prevendría el sobreajuste durante el entrenamiento de Mask-net, el cual se llevó a cabo durante 10 iteraciones para evitar el sobreaprendizaje del modelo, tomando así 18 minutos para completarse.

Matriz de Confusión del modelo Mask-net, en I-Mask-net, se compiló utilizando el optimizador “Adam” con tasa de aprendizaje igualada a 0,00001 para frenar la velocidad de entrenamiento y con función de pérdida “Categorical Cross Entropy”.



Fig. 5. Predicciones correctas del Modelo I-Mask-net por clases.

Se compiló utilizando el optimizador “Adam” con tasa de aprendizaje igualada a 0,00001 para frenar la velocidad de entrenamiento y con función de pérdida “Categorical Cross Entropy”.

3. Resultados

Utilizando un grupo de 170 imágenes correspondientes al conjunto de prueba de KLD, se corrieron predicciones sobre el modelo Mask-net. De la matriz de confusión visible en la Tabla 1 y de los datos de exhaustividad de la Tabla 2, se infiere que Mask-net predijo acertadamente en el 98,21 % de las veces que se le presenta un rostro portando una mascarilla correctamente. Cuando se le presentaron rostros portando la mascarilla incorrectamente, Mask-net fue capaz de predecir ciertamente dicho caso en el 73,68 % de las veces.

Cuando se le presentó uno sin mascarilla, Mask-net predijo correctamente dicho caso en el 87,71 % de las veces, lo que da un porcentaje de exactitud del 86,47 % sobre todo el conjunto de prueba (overall accuracy). Como se observa en la Figura 3, Mask-net es capaz de predecir correctamente imágenes diversas, con personas de diferentes etnicidades y edades, portando variedad de mascarillas en patrones y colores, accesorios y objetos obstruyendo su rostro.

Incluso, se notó que Mask-net puede acertar en algunas ocasiones aunque usuarios intenten engañarlo. Como se observa en la Figura 4, Mask-net presentó algunas fallas particularmente en las clases “Mascarilla incorrecta” y “Sin mascarilla”. Las fallas en mascarilla incorrecta se dieron sobre todo en imágenes donde la barbilla estaba descubierta e imágenes cuyos sujetos no muestran su rostro de manera completamente frontal.

Tabla 4. Métricas de precisión (presicion), exhaustividad (recall) de I-Mask-net.

Clase	Presición	Exhaustividad
Barbilla	0,8888	0,8421
Boca y barbilla	0,8750	0,7368
Nariz y boca	0,8260	1,0000

Las fallas en rostros sin mascarilla son pertinentes a individuos intentando engañar al modelo utilizando algún objeto o inclusive su mano, situaciones de carácter particular. Por otro lado, Utilizando un grupo de 57 imágenes correspondientes al conjunto de prueba de IMKLD, se corrieron las siguientes predicciones sobre el modelo Mask-net. Ver Tabla 2. De la matriz de confusión de la Tabla 3 y de los datos de exhaustividad de la Tabla 4, se entiende que I-Mask-net predijo acertadamente en el 84,21 % de las veces que se le presentó un rostro portando una mascarilla sobre la barbilla.

Cuando se le presentó al modelo situaciones donde se cubría la boca y barbilla y se dejaba expuesta la nariz, I-Mask-net predijo correctamente dicho caso en el 73,68 % de las veces. Cuando se le presentaron rostros portando la mascarilla sobre la nariz y boca y se dejaba la barbilla expuesta, I-Mask-net fue capaz de predecir ciertamente dicho caso en el 100 % de las veces. Lo que da un porcentaje de exactitud del 85,96 % sobre todo el conjunto de prueba (overall accuracy).

Como se observa en la Figura 5, I-Mask-net es capaz de predecir correctamente imágenes de las diferentes clases en la vasta mayoría de los casos estudiados. Como se observa en la Figura 6, I-Mask-net cometió errores en la clase “barbilla”, cuando el sujeto tiene la mascarilla cercana a su boca.

Por otra parte, la clase “boca y barbilla” sufrió mayor cantidad de errores en aquellas imágenes donde la mascarilla estaba superpuesta sobre la boca del individuo. Tomando en cuenta lo expuesto anteriormente, vale la pena acotar que el resto de los errores del modelo I-Mask-net se debe a una falta de generalización en algunas de las imágenes de prueba. Esto probablemente ocurre debido a que el conjunto de entrenamiento y validación de IMKLD no es lo suficientemente extenso para generalizar correctamente.

4. Estado del arte

A continuación se comparan los resultados experimentales de Mask-net con resultados de trabajos publicados destacados en la identificación de mascarillas. Es importante acotar que estos trabajos no resuelven la misma problemática que resuelve Mask-net, ya que la misma se enfoca no solamente en detectar si una persona tiene la mascarilla puesta, sino también si la tiene puesta de manera correcta.

Sin embargo, sirven como referencia para medir el rendimiento de los modelos debido a su similaridad al utilizar Aprendizaje por Transferencia y librerías como MTCNN para la detección de los rostros. Las investigaciones disponibles a continuación son comparadas en la Tabla 5:

- RetinaFaceMask: A Face Mask detector [10]. RetinaFaceMask utiliza múltiples mapas de funciones y luego utiliza funciones red piramidal (FPN) para fusionar

Clase Verdadera: Mascarilla en barbilla				
				
nariz y boca	boca y barbilla	boca y barbilla		
Clase verdadera: Mascarilla en boca y barbilla				
				
barbilla	nariz y boca	nariz y boca	nariz y boca	barbilla

Fig. 6. Predicciones incorrectas del Modelo I-Mask-net por clases.

la información semántica de alto nivel. La investigación hace uso del algoritmo del estado del arte para la detección de rostros MTCNN [9], de manera similar a Mask-net, con el cual se presenta mejor precisión al detectar rostros con mascarilla con ángulos de toma diferentes en comparación al usar Haard Cascade [10].

- Multi-Stage CNN Architecture for Face Mask Detection [11]. Es un sistema que consta en una arquitectura CNN de dos etapas, la cual es capaz de detectar rostros con y sin mascarilla. En esta investigación se hace uso de el Aprendizaje por Transferencia con arquitecturas distintas, preentrenando dichas redes con el conjunto de datos ImageNet, de manera similar a Mask-net.

Comparando a RetinaFaceMask, se halla que Mask-net es similar en precisión a RetinaFaceMask en el caso de predecir sobre la clase mascarilla, alcanzando hasta un 98,21 % de exhaustividad. Cuando se presentan rostros descubiertos (sin mascarilla), Mask-net supera a RetinaFaceMask alcanzando un 96,15 % de precisión, manteniendo una exhaustividad similar a este último.

Por otra parte, los resultados de Multi-Stage CNN fueron superiores utilizando un camino parecido al de MaskNet, utilizando el mismo detector de rostros MTCNN y al utilizar una arquitectura CNN con Aprendizaje Por transferencia.

5. Conclusiones

Procesar las muestras utilizando la detección de rostros por MTCNN aumentó la exactitud de Mask-net e I-Mask-net ante casos reales. Igualmente, efectuar cambios de color a un grupo extenso de mascarillas introdujo variedad en los conjuntos de datos KLD e IMKLD, lo cual disminuyó el sesgo de predicción y optimizó la generalización de los modelos entrenados. Esto permitió a los mismos clasificar correctamente sin importar color o bordado de la mascarilla.

Tabla 5. Tabla comparativa de Precisión (precision) y Exhaustividad (recall) de Mask-net con RetinaFaceMask y Multi-Stage CNN para clases equivalentes.

Investigación	Base	Rostro		Mascarilla	
		Precision	Recall	Precision	Recall
RetinaFaceMask	Mobilenet	79,00 %	92,80 %	78,90 %	89,10 %
	Resnet	91,50 %	95,60 %	93,30 %	94,40 %
Multi-Stage CNN	MobilenetV2	N/D	N/D	99,12 %	99,20 %
	Densenet121	N/D	N/D	99,70 %	99,12 %
Mask-net	Mobilenet	96,15 %	87,71 %	76,38 %	98,21 %

Se desarrolló y entrenó el modelo de visión por computador Mask-net para predecir el uso correcto de la mascarilla, presentando una exactitud en predicción de 86,47 % sobre el conjunto de prueba KLD. Además, se desarrolló y entrenó un modelo adicional, I-Mask-net, para complementar a este.

Dicho modelo manifestó una exactitud en predicción de 85,96 % sobre el conjunto de prueba IMKLD. Uno de los principales aportes de Mask-net a diferencia de otros modelos desarrollados es no solo poder identificar mascarillas médicas, sino también identificar diversidad de mascarillas de tela de variedad de colores y bordados, y respiradores como el tipo N95.

Ambos modelos mostraron un comportamiento altamente satisfactorio, con algunas excepciones. Mask-net no pudo predecir correctamente imágenes de rostros donde la mascarilla era portada incorrectamente con nariz y boca cubiertas y barbilla expuesta, siendo erróneamente predichas como mascarilla.

Por otra parte, Mask-net e I-Mask-net presentaron ciertas diferencias en precisión debido principalmente a la diferencia en extensión en los conjuntos de datos KLD e IMKLD, así como sus susodichos conjuntos de prueba utilizados.

Considerando las conclusiones realizadas anteriormente, se evidencia que es posible detectar el uso correcto e incorrecto de la mascarilla a través de un sistema de reconocimiento facial utilizando visión por computador. Dicho esto, se efectúan las siguientes recomendaciones:

- Se propone incrementar la cantidad de muestras reales en los conjuntos de datos KLD e IMKLD, para abarcar la mayor cantidad posible de situaciones que se puedan presentar, como más ángulos de toma, y casos particulares, como, por ejemplo: casos de individuos utilizando una mayor variedad de accesorios, estilos de cabello distintos y más opciones de vello facial.
- Se sugiere modificar el conjunto de datos KLD, desglosando la clase “mascarilla incorrecta” en las subsiguientes clases: “barbilla”, “boca y barbilla” y “nariz y boca”, además de las existentes “mascarilla” y “sin mascarilla”, totalizando en cinco clases; con el objetivo de entrenar un modelo multiclase, como Mask-net, y así comparar si esta alternativa es más eficiente que utilizar dos modelos independientemente entrenados.

- Se recomienda probar con diferentes modelos de Aprendizaje por Transferencia para la extracción de características, con el fin de evaluar si el desempeño del modelo mejora, caso similar al de RetinaFaceMask, ya que se vió beneficiado en un incremento de sus niveles de precisión y exhaustividad al implementar el modelo de Aprendizaje por Transferencia Resnet.

Referencias

1. Cabani, A., Hammoudi, K., Benhabiles, H., Mahmoud, M.: Maskedface-net – A dataset of correctly/incorrectly masked face images in the context of COVID-19, vol. 2019, no. 2021, pp. 2352–6483 (2020) doi: 10.1016/j.smhl.2020.100144
2. Chavda, A., Dsouza, J., Badgujar, S., Damani, A.: Multi-Stage CNN architecture for face mask detection. In: Proceedings of the 6th International Conference for Convergence in Technology (2021) doi: 10.1109/i2ct51068.2021.9418207
3. Fan, X., Jiang, M.: Retinafacemask: A single stage face mask detector for assisting control of the COVID-19 pandemic (2020) doi: 10.48550/arXiv.2005.03950
4. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. MIT Press (2016) www.deeplearningbook.org
5. Hopkin, J., Maragakis, L. L.: How to properly Wear a face mask (2020) www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/proper-mask-wearing-coronavirus-prevention-infographic
6. Humans in the loop.: Medical Mask Dataset (2020) humansintheloop.org/resources/datasets/medical-mask-dataset/
7. Jebril, N.: Declared a pandemic public health Mmnce: A systematic review of the coronavirus disease (COVID19). World Health Organization, pp. 12 (2020) www.psychosocial.com/article/PR290311/25748/
8. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. Conference on Computer Vision and Pattern Recognition, pp. 4401–4410 (2019) doi: 10.48550/arXiv.1812.04948
9. Mitze, T., Kosfeld, R., Rode, J., Wälde, K.: Face masks considerably reduce COVID-19 cases in Germany: A synthetic control method approach, vol. 177, no. 51, pp. 32293–32301 (2020) doi: 10.1073/pnas.2015954117
10. World Health Organization: Coronavirus (COVID19) dashboard with vaccination data (2020) covid19.who.int/?gclid
11. Zhang, K., Zhang, Z., Li, Z., Qiao, Y.: Joint face detection and alignment using multi-task cascaded convolutional networks. IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503 (2016) doi: 10.1109/LSP.2016.2603342