



Inteligencia Artificial: Para la detección de suelos y paredes en imágenes

Daniel Alejandro Rodriguez Garcia

Trabajo de grado para optar al título de ingeniero de sistemas

Asesor

Joan Andrés Hasper Tabares, Magister en Administración de negocios

Universidad de Antioquia

Facultad de Ingeniería

Ingeniería de Sistemas

Medellín

2022

Cita

Rodriguez Garcia, Daniel Alejandro [1]

Referencia

- [1] Rodriguez, Daniel, "Inteligencia Artificial: Para la detección de suelos y paredes en imágenes", Practica empresaria, Ingenieria de Sistemas, Universidad de Antioquia, Medellín, 2022.

Estilo IEEE (2020)



Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano/Director: Jesús Francisco Vargas Bonilla.

Jefe departamento: Diego José Luis Botía Valderrama.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Agradecimientos

Agradezco a mi familia por ayudarme y darme apoyo para llegar a este punto, a mi esposa por la ayuda que me proporciono para finalizar este trabajo y a Immersivo por darme la oportunidad de desarrollar este proyecto con ellos.

TABLA DE CONTENIDO

| | |
|--------------------------|----|
| RESUMEN | 6 |
| ABSTRACT | 7 |
| I. INTRODUCCIÓN | 8 |
| II. OBJETIVOS | 9 |
| A. Objetivo general | 9 |
| B. Objetivos específicos | 9 |
| III. MARCO TEÓRICO | 10 |
| IV. METODOLOGÍA | 12 |
| V. RESULTADOS Y ANALISIS | 13 |
| VII. CONCLUSIONES | 24 |
| REFERENCIAS | 25 |

LISTA DE FIGURAS

| | |
|---|----|
| Fig. 1 API | 13 |
| Fig. 2 Función de generar máscara..... | 14 |
| Fig. 3 Imagen original (izq.), Máscara generada (der.)..... | 15 |
| Fig. 4 Imagen original (izq.), Máscara generada (der.)..... | 15 |
| Fig. 5 Render Substance 3D (izq.), Máscara colorida (der.)..... | 16 |
| Fig. 6 Máscara colorida (izq.), Máscara generada (der.)..... | 16 |
| Fig. 7 Imagen original (izq.), Máscara generada (der.)..... | 17 |
| Fig. 8 Imagen original (izq.), Máscara generada (der.)..... | 18 |
| Fig. 9 Imagen original (izq.), Máscara generada (der.)..... | 18 |
| Fig. 10 Función de cambio de color | 19 |
| Fig. 11 Imagen original (izq.), Imagen con color cambiado (der.) | 20 |
| Fig. 12 Imagen original (izq.), Imagen con color cambiado (der.) | 20 |
| Fig. 13 Render original (izq.), Imagen con color cambiado usando máscara generada (der.)..... | 21 |
| Fig. 14 Imagen original (izq.), Imagen con color cambiado (der.) | 21 |
| Fig. 15 Imagen original (izq.), Imagen con color cambiado (der.) | 22 |
| Fig. 16 Imagen original (izq.), Imagen con color cambiado (der.) | 22 |
| Fig. 17 Funciones firebase storage | 23 |

RESUMEN

La inteligencia artificial hoy día se ha convertido en un pilar del avance en muchos campos de la industria, principalmente en el manejo y análisis de datos para diferentes funciones como lo son los anuncios que se personalizan para un usuario, también ha tenido su impacto en la industria automotor ya que se ven cada vez más y mejores vehículos con un nivel de asistencia al conducir o la capacidad de conducirse solos.

Pero en donde no se ha visto muy aplicada es en las experiencias del usuario al momento de comprar un producto y con la visión que plantea el Metaverse actualmente se ve en la necesidad de que los clientes tengan la capacidad desde la comodidad de sus hogares ver cómo se verían los productos en sus espacios, como probar combinaciones de colores en pisos y paredes de acuerdo al ambiente o estilo que se le quiere dar a un espacio para así tomar una decisión más informada y porque no, más entretenida sobre los productos que se adquieren para lograr esa visión.

En este documento se puede observar la fase preliminar para la implementación de la inteligencia artificial para permitir, por medio de fotos o renders, la personalización de espacios por medio del cambio de color de las superficies que se detectan en las imágenes proporcionadas.

***Palabras clave* — Inteligencia Artificial, Aprendizaje de Maquina, Metaverse, Segmentación Semántica, Tensorflow.**

ABSTRACT

Artificial intelligence today has become a pillar of progress in many fields of the industry, mainly in the management and analysis of large amounts of data for different functions such as personalized ads for a user, it has also had its impact on the automotive industry since more and more capable vehicles are seen with a level of assistance when driving or the ability to drive themselves.

But where it has not been widely applied is in user experiences when buying a product and the Metaverse introducing a new way to interact with the world, it's in the need for customers to have the ability from the comfort of their homes and to see how the products would look in their spaces, such as testing color combinations on floors and walls according to the environment or style that you want, giving the space and ability to make a more informed decision when you buy, and why not, a more entertaining process about selecting the products that are acquired to achieve that vision.

In this document the preliminary phase for the implementation of artificial intelligence can be observed to allow, through photos or renders, the personalization of spaces through the change of color of the surfaces that are detected in the images provided.

Keywords — **Artificial intelligence, Machine learning, Metaverse, Semantic Segmentation, Tensorflow.**

I. INTRODUCCIÓN

La inteligencia artificial actualmente se ve implementada en muchas áreas de la vida cotidiana, desde análisis de correos para identificación de spam pasando por el análisis de perfiles para la selección de anuncios hasta la detección de rostros en fotos para identificar personas, esta última aplicación no solo se puede realizar con personas, sino también con objetos, y con el uso de la tecnología actual de realidad virtual y realidad aumentada se puede llevar a las salas de venta de nuestros clientes dándoles la capacidad de ver ambientes y modificarlos en tiempo real.

Se busca con este proyecto, iniciar con la capacidad de identificar y modificar elementos, como paredes y pisos, en fotos y escenas renderizadas en 3D, estos últimos siendo diseñados por la empresa a petición del cliente, para en un futuro realizar lo mismo en tiempo real con ambientes y objetos más complejos.

II. OBJETIVOS

A. Objetivo general

Realizar el desarrollo de un aplicativo que permita la detección de paredes, pisos y ángulos en fotos y renderizados 3D.

B. Objetivos específicos

- Entrenar e implementar el algoritmo para la detección de paredes o planos verticales.
- Entrenar e implementar el algoritmo para la detección de pisos o planos horizontales.
- Mejorar el algoritmo con la capacidad de diferenciar entre diferentes ángulos y separaciones entre los planos y separadores verticales.

III. MARCO TEÓRICO

Machine Learning: El aprendizaje de máquina es la aplicación de inteligencia artificial que permite el análisis de datos que le da la capacidad de aprender de la información dada y mejorar el algoritmo entregado a partir de la predicción y detección de patrones dentro del mismo para así mejorar su precisión [1]. Estos tipos de aprendizaje se separan en 3 grandes categorías:

- **Supervisado:** Es alimentado por información previamente clasificada dada para así poder aprender a clasificar nueva información de acuerdo con los parámetros aprendidos [1].
- **Sin supervisión:** Es alimentado por información en la cual hay patrones o agrupaciones de los cuales se pueden identificar lo cual permite identificar nueva información dada [1].
- **Semi supervisada:** Es la combinación de las mejores características de ambas, ya que por medio del uso de una parte de la información entregada de manera previamente clasificada aprende a clasificar o identificar el resto de la información [1].

Procesamiento digital de imágenes: Es una de las aplicaciones de Machine learning la cual permite identificar objetos previamente aprendidos a partir de imágenes o mejorar la calidad de estas [2].

Realidad Virtual (VR) y Realidad Aumentada (AR): Siendo ambas tecnologías que de una manera u otra permiten la interacción con un mundo virtual, VR enfocándose en simular un mundo virtual generado por computadora, en la cual un individuo puede interactuar con este ambiente por medio del uso de otros dispositivos físicos, por otro lado, AR permite integrar componentes de un mundo virtual al mundo real por medio de estímulos sensoriales como elementos visuales y auditivos [3].

Deepfake: El ejemplo más cercano a lo que se busca lograr con este desarrollo son los Deepfakes, los cuales son una implementación de la Inteligencia artificial a fotos de personas o personajes, con la cual se pueden reemplazar los rostros de los personajes o personas en la foto por otros personajes o personas, dado que la implementación que se busca con el desarrollo es realizar lo mismo, pero con superficies [4].

Segmentación semántica: Es un algoritmo de Deep learning en el que se asocia una etiqueta o categoría a cada píxel en una imagen, principalmente usado para la detección de elementos dentro de una imagen al reconocer grupos de píxeles con la misma etiqueta o categoría [5].

Tensorflow: Es una plataforma de código abierto para el machine learning. Cuenta con un ecosistema de herramientas, bibliotecas y recursos comunitarios los cuales están a disposición de los desarrolladores y entusiastas para impulsar el machine learning en las diferentes áreas de aplicación [6].

Metaverse: Es la nueva evolución que Facebook plantea a las redes sociales, un universo dentro de un espacio computarizado que permite la conexión entre personas, empresas y desarrollos tecnológicos por medio de AR y VR, minimizando la brecha entre nuestra realidad y la virtual [7].

IV. METODOLOGÍA

Por medio del uso de Machine learning, específicamente el procesamiento digital de imágenes, el cual permite la detección de objetos y la capacidad de identificarlos de acuerdo con la información que se le provee al algoritmo, ya sean obtenidos por medio de fotos o renders 3D, permitiendo la detección de los diferentes componentes que comprenden una escena o habitación.

Con la información tomada a partir del análisis y clasificación de imágenes y modelos 3D, permitir la modificación de estos en tiempo real por medio del uso de pantallas, realidad aumentada (AR) y realidad virtual (VR) las cuales permiten de una manera u otra traer elementos del mundo virtual al real y la capacidad de interactuar con ellos hasta cierto nivel.

Para así permitir una experiencia más inmersiva y facilitar el proceso al cliente al momento de tomar una decisión con respecto a los productos ofrecidos por el uso del desarrollo dándole un valor agregado a nuestros desarrollos.

Dado que se necesita un dispositivo con ciertas especificaciones técnicas para el uso de AR y VR se explorarán los requisitos mínimos para que el proyecto se desempeñe de manera fluida en aquellos que soporten la tecnología.

V. RESULTADOS Y ANALISIS

El modelo usado para la detección de los diferentes objetos y planos en una imagen se entrenó usando herramientas de Google y Python, como lo es Tensorflow, con el cual se obtuvo el modelo entrenado para la detección, y se implementó dentro de una API REST (Fig. 1) con la que realizan peticiones con imágenes y/o renders para generar una máscara, que cuando se envían con un valor en hexadecimal de un color nos permite ver las superficies previamente determinadas del color enviado, como pisos, paredes y en algunos casos tapetes.

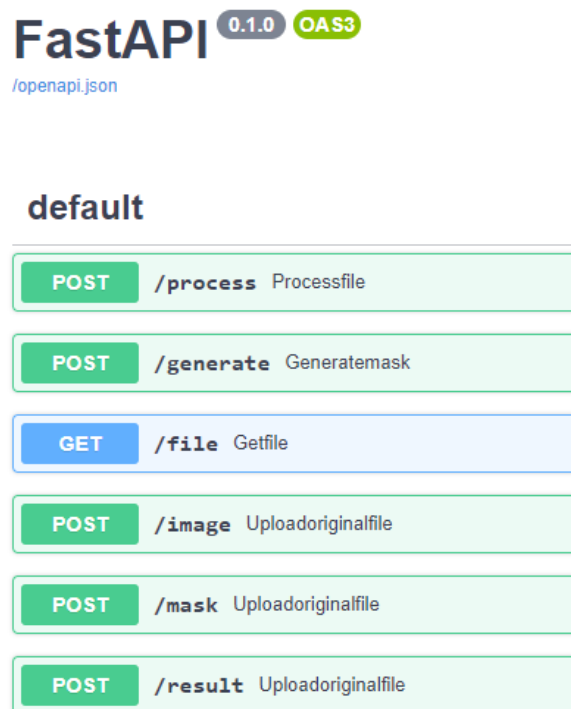


Fig. 1 API

La API tiene las siguientes funciones que se explicaran de manera detallada a continuación:

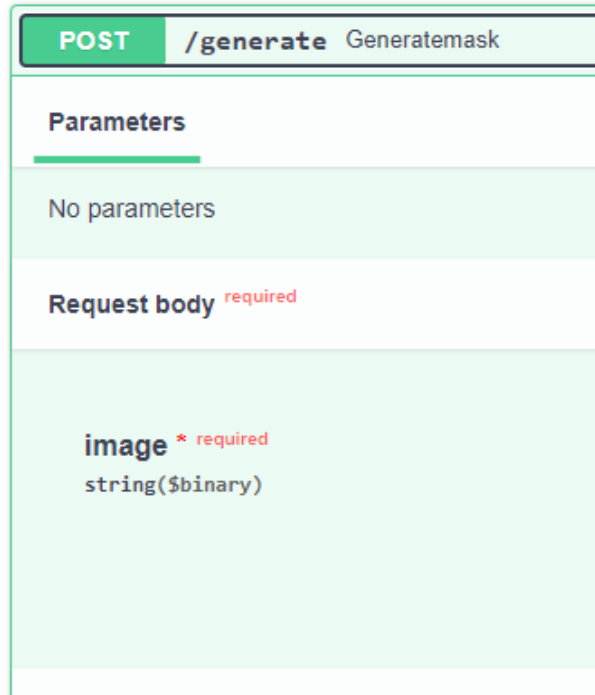


Fig. 2 Función de generar máscara

Generar máscara: Esta función recibe 1 archivo, idealmente una foto que se desea procesar aunque también permite generar la máscara de un render, se recomienda usar la máscara generada por la herramienta que inicialmente creó el render, en nuestro caso esa herramienta es Substance 3D, ya que la máscara que se genera son arreglos de números representados en una imagen y cada número representa un objeto, en el caso de los renders estos números fueron asignados al momento de generarlo por lo que el proceso de cambio de color es mucho más rápido y eficiente, finalmente esta función retornará la máscara en forma de imagen.

Se envía una imagen, con lo que se genera una máscara, a continuación, se presentan varios ejemplos con el lado izquierdo siendo la imagen original y el derecho la máscara generada por el modelo.

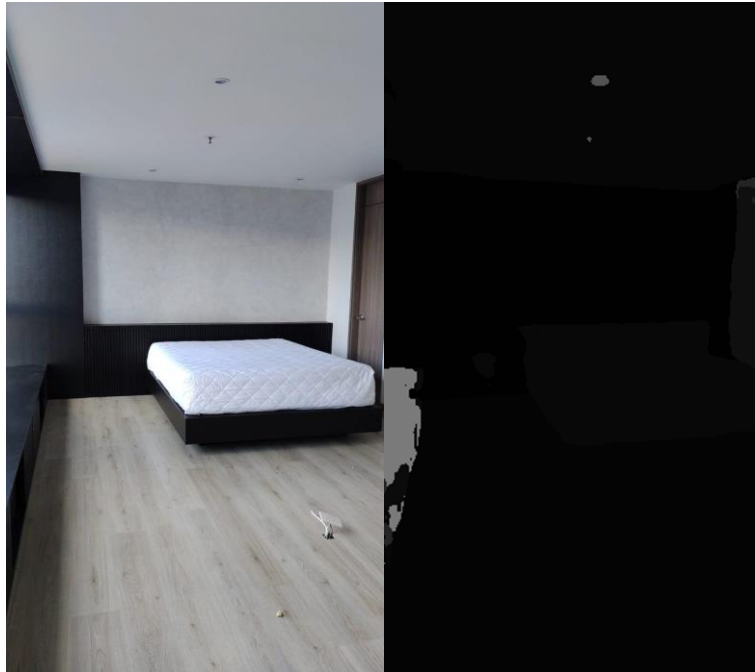


Fig. 3 Imagen original (izq.), Máscara generada (der.)

Como se puede notar en la imagen, algunos elementos son identificables en la máscara, como las luces del techo y parte de la iluminación del lado izquierdo de la habitación.



Fig. 4 Imagen original (izq.), Máscara generada (der.)

Y para el caso de los renders, la imagen creada y la máscara que genera al momento de crearse.



Fig. 5 Render Substance 3D (izq.), Máscara colorida (der.)

Como se puede observar en la Fig. 5, la máscara que se genera con los renders ya viene con algunos colores, esto se hizo intencionalmente para facilidad de identificación de los objetos, pero la máscara real no tendrá estos colores y se verá en escala de grises similar a las generadas en la versión final.

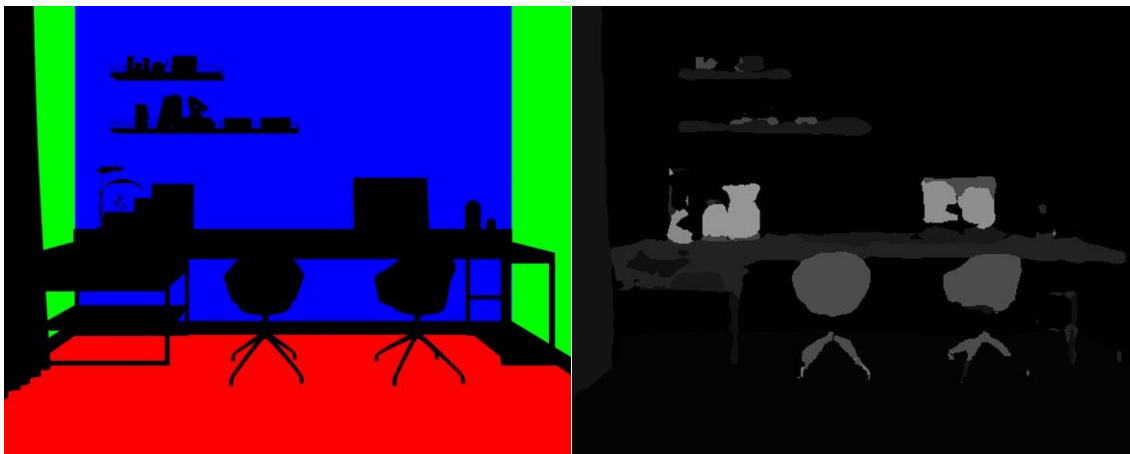


Fig. 6 Máscara colorida (izq.), Máscara generada (der.)

Claramente se pueden notar diferencias en la Fig. 6 entre ambas máscaras, principalmente en la nitidez con la que se separan los diferentes objetos los unos de los otros e incluso algunas de las formas de estos.

Hasta el momento las pruebas realizadas con espacios exteriores no han dado buenos resultados como se puede observar en la Fig. 7 y el resultado del cambio de color en la Fig. 14.



Fig. 7 Imagen original (izq.), Máscara generada (der.)

Otros ejemplos de la generación de máscara se presentarán a continuación:

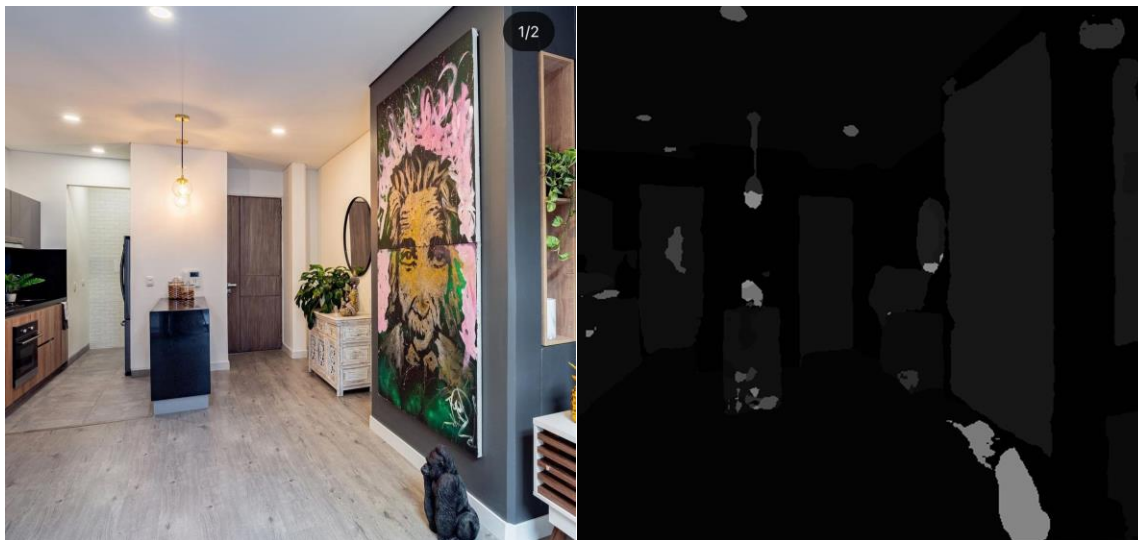
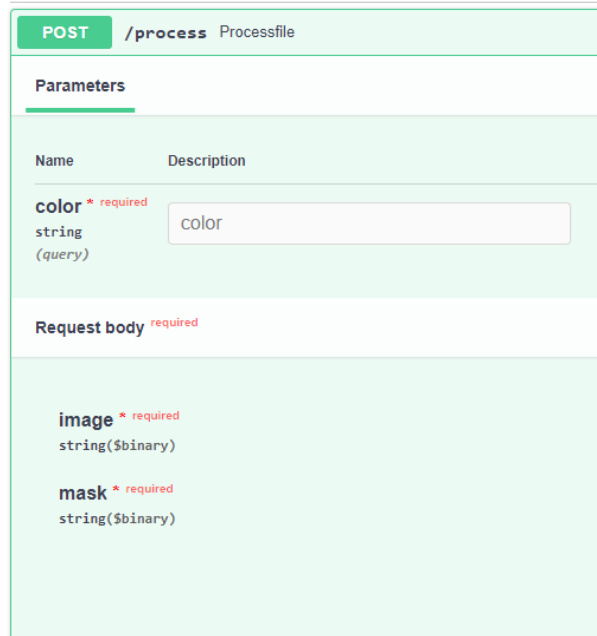


Fig. 8 Imagen original (izq.), Máscara generada (der.)



Fig. 9 Imagen original (izq.), Máscara generada (der.)

Una vez generada la máscara se pasa a la siguiente función de la API, en la cual se tomarán la imagen original y la máscara generada para cambiar el color de las superficies.



POST /process Processfile

Parameters

| Name | Description |
|---------------------------------------|-------------|
| color * required string (query) | color |

Request body required

image * required
string(\$binary)

mask * required
string(\$binary)

Fig. 10 Función de cambio de color

Cambio de color: Está función recibe dos imágenes, la imagen original y la máscara generada por la función anterior o la máscara generada por Substance 3D, además recibe un color, determinado por su código en hexadecimal, como parámetros para la ejecución, y retorna una imagen en la cual se pueden ver las superficies con el color seleccionado.

Los resultados de ingresar los ejemplos previamente mostrados son los siguientes:
Cambiando tanto paredes como pisos en la Fig. 11.



Fig. 11 Imagen original (izq.), Imagen con color cambiado (der.)

Cambiando solo paredes en la Fig. 12.



Fig. 12 Imagen original (izq.), Imagen con color cambiado (der.)

Cambiando solo paredes en la Fig. 13.



Fig. 13 Render original (izq.), Imagen con color cambiado usando máscara generada (der.)

Cambiando solo paredes en la Fig. 14.



Fig. 14 Imagen original (izq.), Imagen con color cambiado (der.)

Cambiando solo pisos en la Fig. 15.

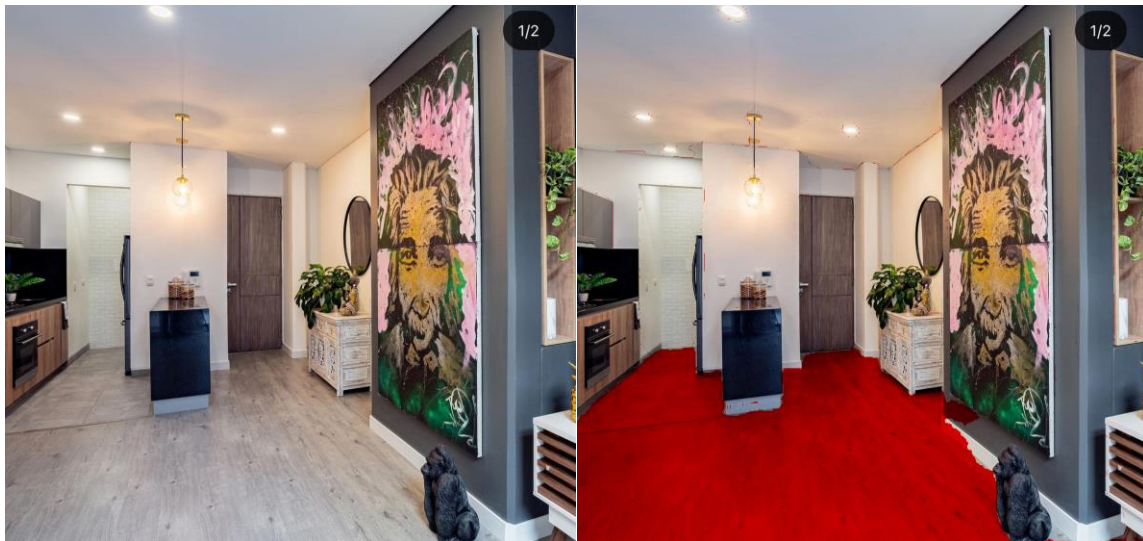


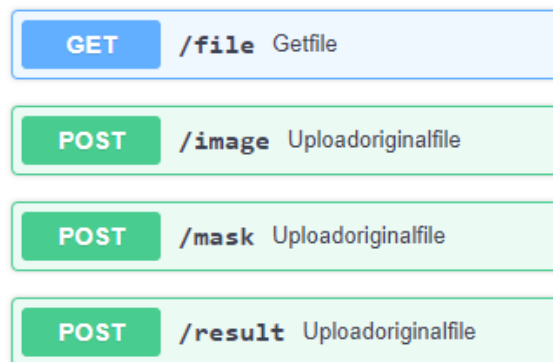
Fig. 15 Imagen original (izq.), Imagen con color cambiado (der.)

Cambiando solo paredes en la Fig. 16.



Fig. 16 Imagen original (izq.), Imagen con color cambiado (der.)

Como se puede notar en las imágenes con un cambio de color, en la mayoría de las situaciones el color cubre objetos que no debería, aunque en algunos casos la cantidad cubierta de objetos no objetivo es muy poco como en las Fig. 11, 15 y 16, y en el caso de la Fig. 14 el cambio se aplica a un área mucho mayor a la requerida por lo que se identifica un objetivo a mejorar y reentrenar el modelo para la detección de fotos exteriores.



El diagrama muestra cuatro funciones de Firebase Storage, cada una en un recuadro con un fondo de color y un botón de método HTTP. La primera función es GET /file Getfile con fondo azul claro. Las otras tres son POST /image Uploadoriginalfile, POST /mask Uploadoriginalfile y POST /result Uploadoriginalfile, todas con fondo verde claro.

| | | |
|------|---------|--------------------|
| GET | /file | Getfile |
| POST | /image | Uploadoriginalfile |
| POST | /mask | Uploadoriginalfile |
| POST | /result | Uploadoriginalfile |

Fig. 17 Funciones firebase storage

Funciones firebase storage: La principal razón de estas funciones es el almacenamiento de las imágenes, máscaras y resultados de las dos funciones anteriores para referencia de entrenamiento del modelo y para que el usuario pueda regresar a ellas en caso de requerir usarlas.

VII. CONCLUSIONES

Aunque el entrenamiento que se aplicó al modelo fue suficiente para identificar las superficies objetivo y realizar el cambio de color de manera selectiva, este aun combina las superficies del mismo tipo y en algunas ocasiones como se nota en la imagen 16, puede confundir espacios entre superficies, combinando dos superficies en una más grande.

También se puede notar la necesidad de entrenamiento con ambientes exteriores y al identificar objetos sobrepuestos a las superficies objetivo, como espejos, pantallas y cuadros que van montados en las paredes, además está al ser una versión inicial carece la robustez necesaria para un producto viable dada la cantidad de peticiones que se le pueden realizar una vez que salga al mercado, pero que sienta las bases firmemente para el siguiente paso.

Siendo este siguiente paso el entrenar más el modelo para la detección más precisa de las superficies y la diferenciación entre varias superficies del mismo tipo, por ejemplo, el detectar varias paredes o varios niveles de piso y permitir el cambio de color de manera independiente, además estudiar la posibilidad de implementar un algoritmo extra para aplicar texturas sobre las superficies y dar así la posibilidad de cambiar el material además del color de las superficies de manera individual.

REFERENCIAS

-
- [1] IBM, “¿What is Machine Learning? ,” 2020. <https://www.ibm.com/cloud/learn/machine-learning> (accessed Apr. 30, 2021).
- [2] A. D., “Procesamiento digital de imágenes,” *Perfiles Educativos*, 1996, Accessed: Apr. 30, 2021. [Online]. Available: <https://www.redalyc.org/articulo.oa?id=13207206>
- [3] A. Hayes, “Augmented Reality,” Dec. 02, 2020. <https://www.investopedia.com/terms/a/augmented-reality.asp> (accessed Apr. 30, 2021).
- [4] D. Johnson, “What is a deepfake? Everything you need to know about the AI-powered fake media,” Jan. 22, 2021. <https://www.businessinsider.com/what-is-deepfake> (accessed May 31, 2021).
- [5] MathWorks, “Segmentación semántica Tres cosas que es necesario saber,” *MathWorks*. <https://la.mathworks.com/solutions/image-video-processing/semantic-segmentation.html> (accessed Nov. 30, 2021).
- [6] Tensorflow, “Tensorflow.” <https://www.tensorflow.org/> (accessed Nov. 30, 2021).
- [7] Facebook, “Meta.” <https://about.facebook.com/meta/> (accessed Dec. 31, 2021).