

UNIVERSIDAD NACIONAL SAN CRISTÓBAL DE HUAMANGA
FACULTAD DE INGENIERÍA DE MINAS, GEOLOGIA Y CIVIL
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



“Aplicación Machine Learning con Visión Artificial al juego de Piedra, Papel o Tijera”

DOCENTE : Ing. HUARANCCA ÑAUPARI, Yonny

ASIGNATURA : Sistemas Expertos

SIGLA : IS-442

ESTUDIANTES :

- LOPEZ YUPANQUI, Yulisa Daniela
- YANCCE CAHUANA, Jessica

AYACUCHO – PERÚ

2023

ÍNDICE

INTRODUCCIÓN	3
CONTEXTO DEL PROBLEMA	3
JUSTIFICACIÓN E IMPORTANCIA.....	4
OBJETIVOS.....	5
OBJETIVO GENERAL.....	5
OBJETIVOS ESPECÍFICOS	5
MARCO TEÓRICO	5
INTELIGENCIA ARTIFICIAL	5
MACHINE LEARNING	6
REDES NEURONALES	7
REDES NEURONALES CONVOLUCIONALES	7
IMÁGENES DIGITALES	11
DESARROLLO DEL PROYECTO.....	12
DIAGRAMAS	12
DIAGRAMA DE DESCOMPOSICIÓN DEL TRABAJO.....	12
VISTA FÍSICA DEL COMPUTADOR.....	13
CASOS DE USO	14
HERRAMIENTAS UTILIZADAS	15
MEDIA PIPE HANDS	15
MEDIAPIPE	15
OPEN CV2	16
NumPy	17
MATH.....	18
RANDOM.....	19
DESARROLLO DE INTERFAZ	19
CONCLUSIÓN	25
BIBLIOGRAFÍA.....	26

INTRODUCCIÓN

El juego “piedra papel o tijera” es un juego clásico, histórico que se disfrutó por largo tiempo y se percibió como sencilla y divertida al tomar decisiones al azar. A pesar de esta aparente simplicidad, el juego encierra un desafío para la inteligencia artificial y el aprendizaje automático, en el presente proyecto nos adentraremos en el mundo de la inteligencia artificial y el aprendizaje automático “machine learning” para desarrollar un sistema capaz de jugar el juego “piedra papel o tijera”.

Adoptaremos un enfoque del uso de técnicas de aprendizaje automático, que permitirá entrenar al algoritmo.

Veremos como un juego simple de “Piedra, papel o tijeras” es una buena plataforma para experimentar con algoritmos de inteligencia artificial.

CONTEXTO DEL PROBLEMA

El juego de piedra, papel o tijera es ampliamente conocido y se juega en diferentes culturas y edades. Tradicionalmente, se juega de forma manual, donde los jugadores forman una de las tres figuras con sus manos y luego muestran su elección al mismo tiempo para determinar el ganador según las reglas del juego: la piedra vence a las tijeras, las tijeras vencen al papel y el papel vence a la piedra.

Sin embargo, con los avances en visión artificial y el uso de algoritmos de machine learning, es posible automatizar el proceso y permitir que un sistema inteligente juegue contra un ser humano. El desafío principal es el reconocimiento preciso y confiable de las formas realizadas por las manos del jugador humano, ya que esto es crucial para tomar decisiones informadas sobre qué jugada realizar.

Para abordar este problema, se necesitaría un modelo de machine learning entrenado con un conjunto de datos etiquetado que contenga imágenes de manos realizando cada una de las tres formas del juego. El modelo debe ser capaz de extraer características relevantes de las imágenes y aprender a diferenciar entre piedra, papel y tijera con alta precisión. Además, el sistema debería ser lo suficientemente rápido para tomar decisiones en tiempo real, lo que requiere un equilibrio entre precisión y eficiencia.

JUSTIFICACIÓN E IMPORTANCIA

La credibilidad e importancia de la implementación de visión artificial y machine learning en el juego de Piedra Papel o Tijera radica en varios aspectos, que van desde el desarrollo tecnológico hasta la diversión y la interacción humana. El juego implica la creación de algoritmos y modelos de aprendizaje automático para reconocer y comprender los gestos humanos. Este es un gran avance en la visión artificial y abre la puerta a aplicaciones más sofisticadas para reconocer y comprender los gestos humanos en una variedad de contextos.

Los juegos con visión artificial pueden ofrecer una experiencia interactiva única en la que los jugadores juegan contra una inteligencia artificial que “entiende” y responde a los gestos de las manos el cual crea una sensación de conexión entre el hombre y la máquina.

La creación de juegos con visión artificial y Machine Learning también pueden tener aplicaciones educativas. Por ejemplo, se puede utilizar como una herramienta para enseñar a los estudiantes de la tecnología los conceptos básicos de la visión artificial y el aprendizaje automático.

El desarrollo del movimiento y los modelos para identificar piedras, papel o tijeras puede ser un punto de partida para un estudio más amplio en la inteligencia artificial y tecnología de

robótica. Este progreso se puede utilizar en los campos como la robótica asistencial, la interacción humano-robot y la automatización de tareas complejas.

OBJETIVOS

OBJETIVO GENERAL

Implementar un sistema o algoritmo de “machine learning” que sea capaz de aprender y mejorar la estrategia del juego “Piedra, papel o tijera”, utilizando técnicas de aprendizaje cada vez que juega contra distintos contrincantes o experimente diferentes situaciones de juego para así adaptarse a patrones de juegos cambiantes.

OBJETIVOS ESPECÍFICOS

- Crear un algoritmo preciso que tenga la capacidad de analizar imágenes de manos humanas e identificar los correspondientes movimientos de Piedra, Papel y Tijera, lograr una alta tasa de aciertos en el procesamiento de imágenes y aprendizaje automático.
- Implementación de un modelo de aprendizaje automático; se entrenará utilizando un conjunto de datos que contiene ejemplos de movimientos de Piedra Papel o Tijera.
- Lograr que el sistema tenga un tiempo de respuesta rápido con una baja latencia.
- Mejorar la precisión con la que el sistema detecta los movimientos para evitar posibles errores.

MARCO TEÓRICO

INTELIGENCIA ARTIFICIAL

La inteligencia artificial es una rama de la informática que tiene su enfoque en el desarrollo de sistemas y algoritmos, estos sistemas buscan reflejar o imitar la inteligencia

humana, el término nace en el año 1956 en el congreso de Dartmouth. La existencia de los ordenadores a comienzos del siglo XX hizo crecer enormemente su popularidad.

Según Marvin Minsky, que es considerado como uno de los padres de la inteligencia artificial, la define como: “Una disciplina que se encarga de crear máquinas programadas que sean capaces de hacer cosas que requieren la misma inteligencia que si fuese hechas por humanos”.

MACHINE LEARNING

El Machine Learning o aprendizaje automático es una rama de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos capaces de aprender patrones y realizar tareas sin ser programadas específicamente para ello. En lugar de que el algoritmo siga reglas predefinidas, este tipo de sistemas utilizan datos para mejorar el desempeño y así se adapte a nuevas situaciones.

Para que un sistema que tiene como base pueda funcionar se necesita de experiencia, la cual debe ser entregada por el programador, ya sea por imágenes, sonidos, colores o simplemente de un base de conocimiento, esto con el fin de comprender su entorno.

El aprendizaje de este algoritmo se puede llevar a cabo de distintas maneras:

Aprendizaje supervisado: la siguiente técnica realiza un etiquetamiento previo de los datos, para que luego el algoritmo tenga claro qué entradas deben coincidir con qué salidas.

Aprendizaje no supervisado: en esta técnica los datos ya no se entregan con sus etiquetas correspondientes, por lo que el algoritmo debe encontrar patrones o similitudes entre la información recibida.

Aprendizaje por refuerzo: este algoritmo debe aprender por su propia experiencia, mediante un sistema de recompensas basado en las respuestas correctas e incorrectas.

REDES NEURONALES

Las redes neuronales son una clase de algoritmos de aprendizaje automatizado inspirados en el funcionamiento del cerebro humano. Están diseñados para procesar datos a partir de ejemplos proporcionados en el entrenamiento.

Las redes neuronales se organizan mediante una arquitectura multicapa, en el cual existe una capa de entrada, donde las neuronas presentes reciben señales del entorno, propagándose a las siguientes capas. Luego viene la capa oculta, que puede ser una o múltiples capas. La última capa es la de salida, que es donde la predicción realizada por la red es entregada.

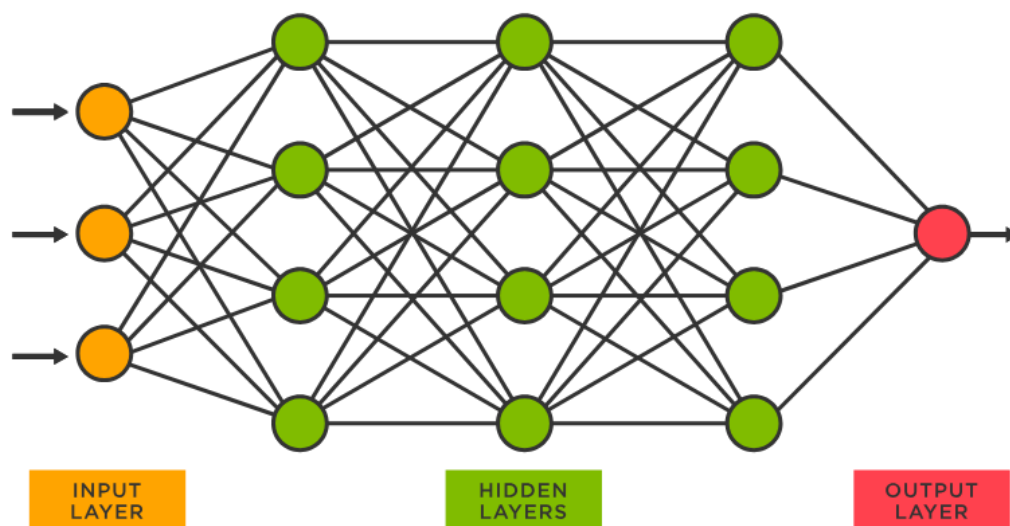


Figura 01: Capas de las redes neuronales

REDES NEURONALES CONVOLUCIONALES

Las redes neuronales convolucionales (CNN), son un tipo de red creada específicamente para el reconocimiento de patrones presentes en imágenes. Estas redes emulan el sistema de visión humano para poder identificar objetos previamente entrenados.

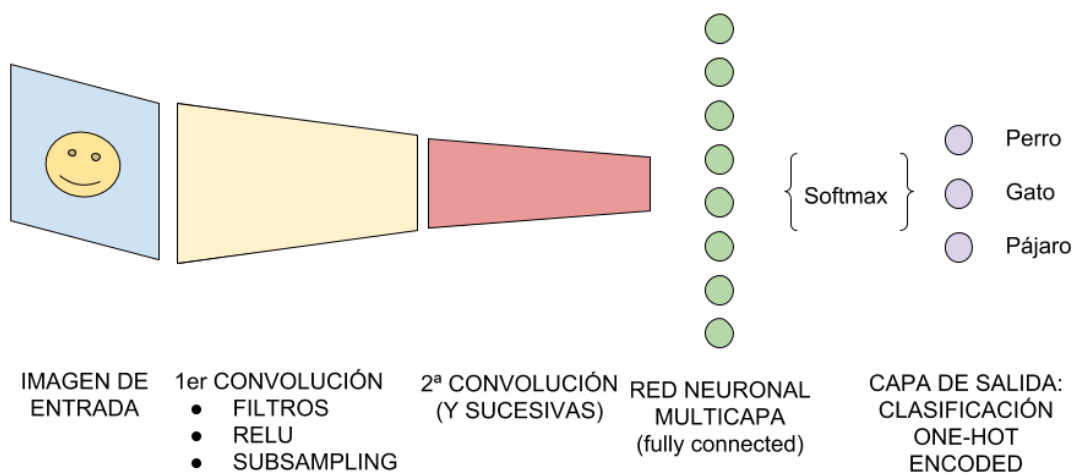


Figura 02: Etapas de las redes neuronales convolucionales.

Para el siguiente proceso, estas redes cuentan con múltiples capas ocultas, con las que, de manera jerárquica, van reconociendo desde filas y curvas hasta objetos completos. Para lograr ello es necesario un entrenamiento previo, en el cual se le debe contar grandes cantidades de imágenes.

También se debe tener en cuenta que estas redes perciben imágenes de una manera distinta a como lo haría el ojo humano, y es mediante el uso de matrices que contienen la información de cada píxel presente en la imagen.

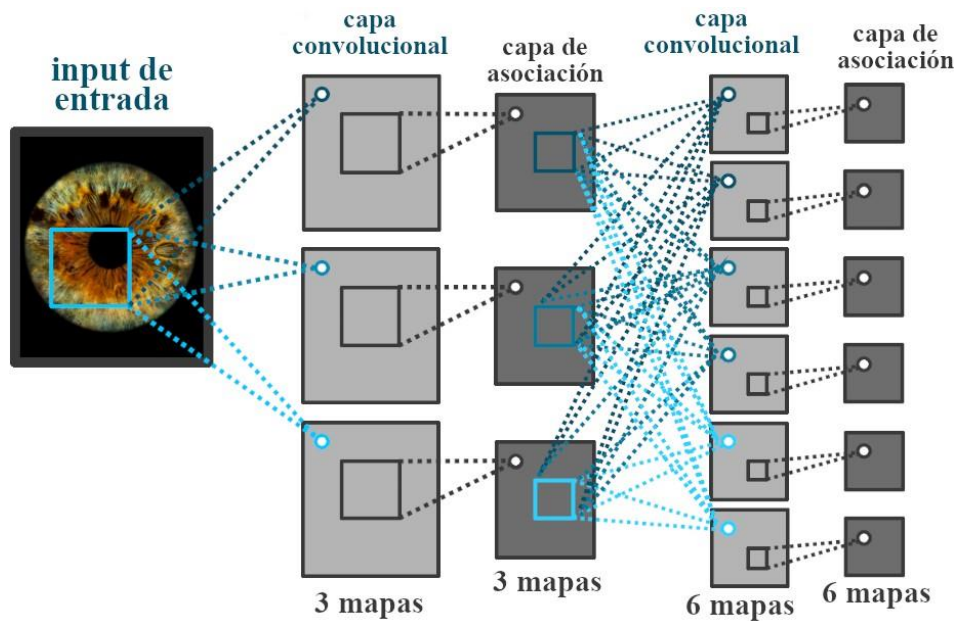


Figura 03: Capas de las redes neuronales convolucionales

La conclusión se produce mediante el uso de Kempes, que son datos que, al inicio, se encuentran dispuestos de manera aleatoria, pero luego serán ajustados mediante lo que se conoce como “black propagation”. Esta técnica se realiza mediante la programación de los resultados finales hacia las primeras neuronas.

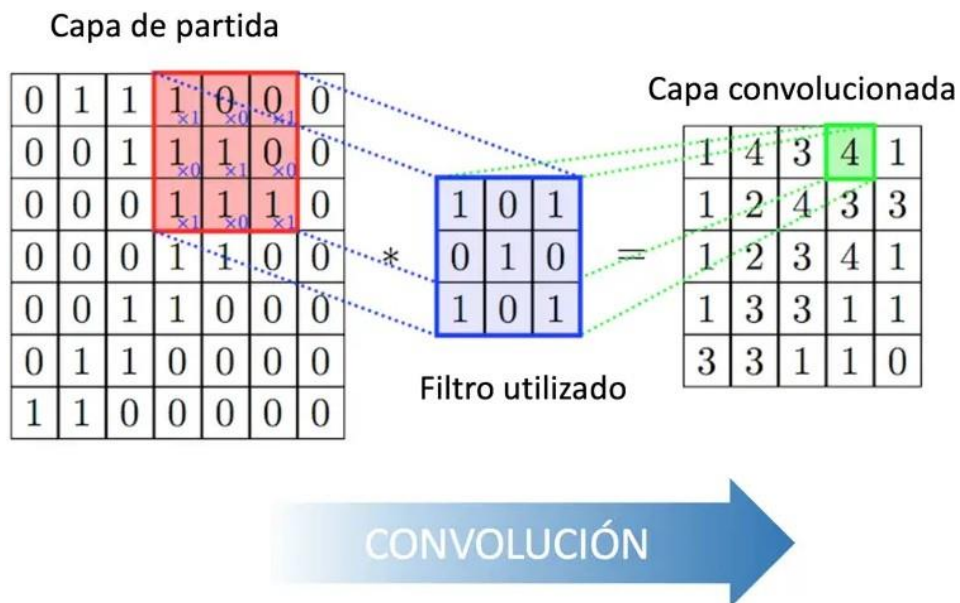


Figura 04: Filtro usado en las redes neuronales convolucionales.

Entonces, las matrices son utilizadas mediante un producto escalar con la imagen original, lo que genera unas nuevas matrices de salida, denominadas feature maps. Estas matrices contienen información de las principales características encontradas.

Durante la convolución, con valores que son tanto negativos como positivos, es por ello por lo que se debe usar una función de activación, la cual suprime los valores negativos y deja los positivos sin aplicar ningún cambio.

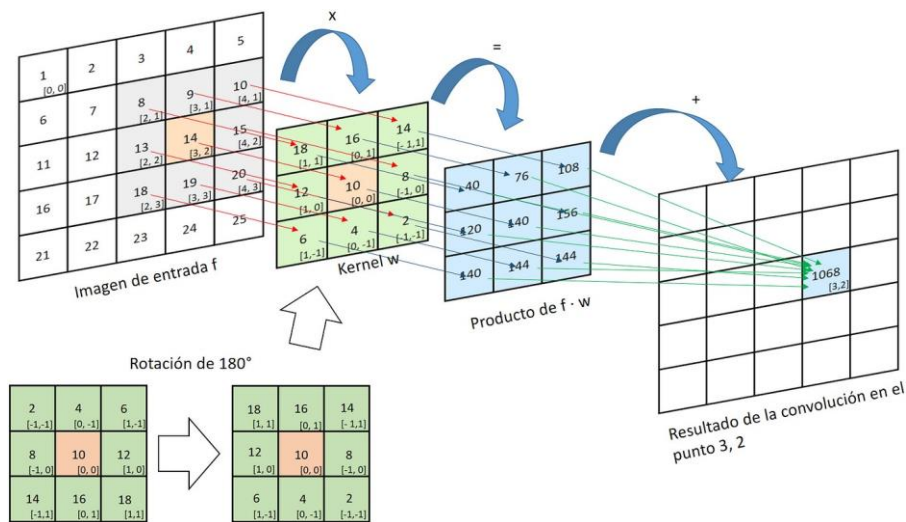


Figura 05: Redes neuronales convolucionales.

Finalizando la convolución, el número de salidas es equivalente a la cantidad de filtros utilizados, es decir, si por ejemplo utilizamos 8, la salida de una imagen serán 8 nuevas matrices, las cuales deben ser procesadas de manera individual y necesitarán la misma capacidad de cómputo que la original para cada una. Por esto se usa una técnica denominada subsampling, con la cual podremos reducir el tamaño de cada una de estas, almacenando las características más importantes.

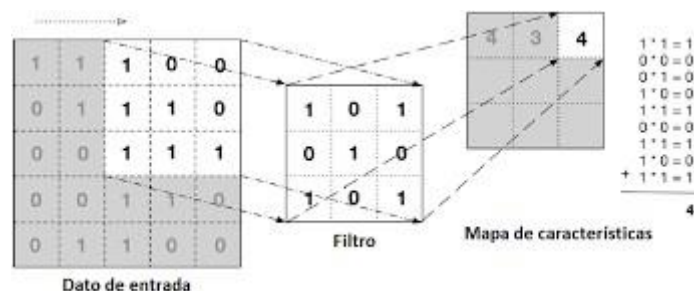


Figura 06: Datos de entrada y filtro en las redes neuronales convolucionales.

Al momento de producirse la última convolución, es decir al pasar a la última capa oculta de la red, se aplica la función llamada, en este caso, softmax que realiza la conexión con la capa de salida, donde el número de neuronas corresponderá al total de clases. Indicando así el porcentaje de precisión a la clase correspondiente según la predicción.

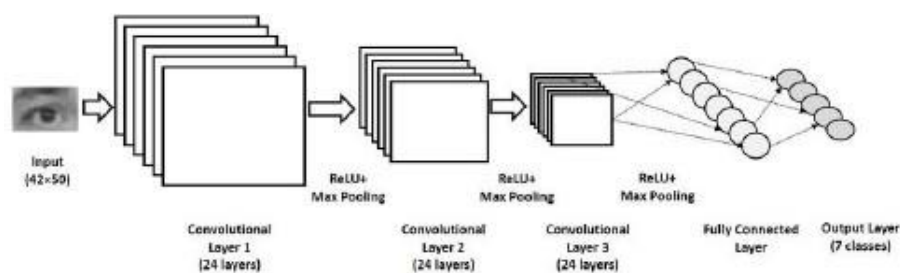


Figura 08: Procesos de las redes neuronales convolucionales.

IMÁGENES DIGITALES

Las imágenes digitales son representaciones visuales de objetos, escenas o personas que se capturan, almacenan y visualizan en formato digital utilizando dispositivos electrónicos, como cámaras digitales, escáneres o sensores de imagen. Estas imágenes se componen de una cuadrícula de píxeles que contienen información sobre el color y la intensidad de la luz en cada punto de la imagen.

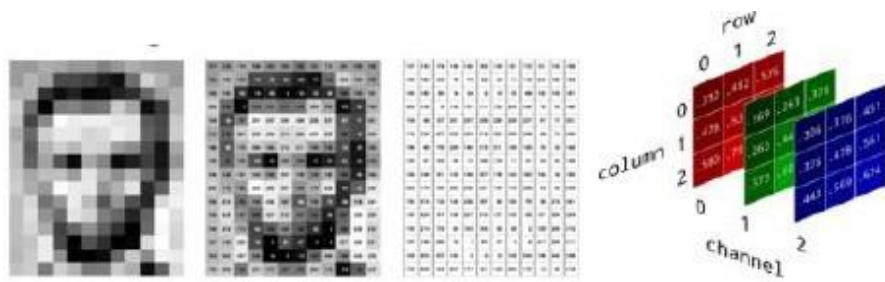


Figura 09: Imágenes digitales en cuadrícula de píxeles.

DESARROLLO DEL PROYECTO

DIAGRAMAS

DIAGRAMA DE DESCOMPOSICIÓN DEL TRABAJO

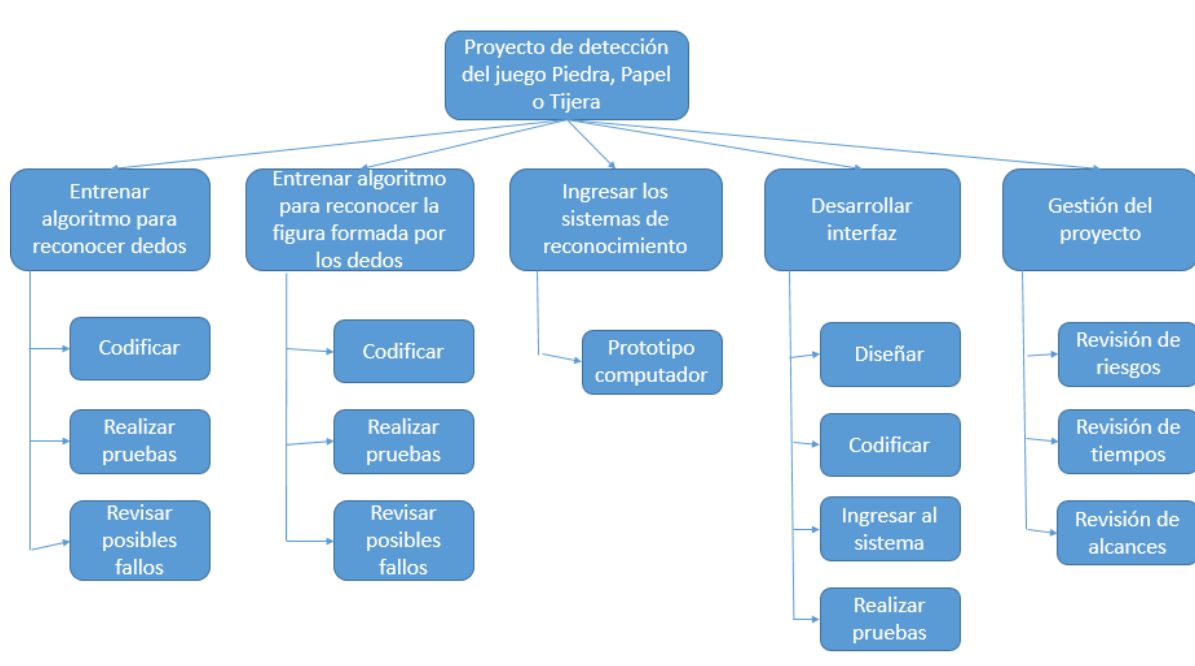


Figura 10: Diagrama de descomposición del trabajo

Fuente: Elaboración propia

VISTA FÍSICA DEL COMPUTADOR

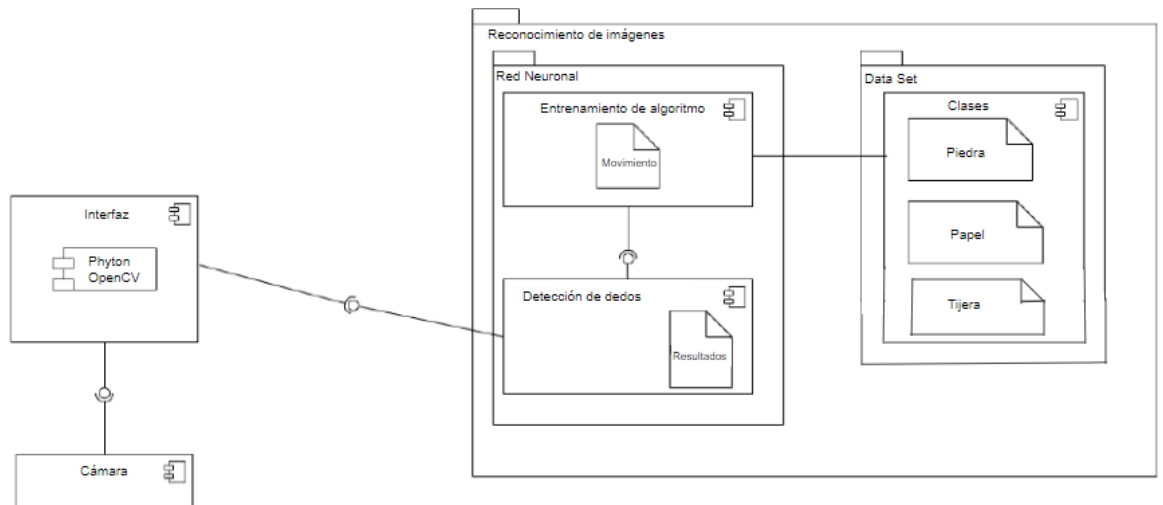


Figura 11: Gráfica de Vista física del computador

Fuente: Elaboración propia

Con respecto a la vista de desarrollo, obtenemos una visión general de la solución y ahondando en lo que es el sistema de reconocimiento de imágenes, en la cual se muestra la red junto con el dataset que contiene las clases propuestas.

CASOS DE USO

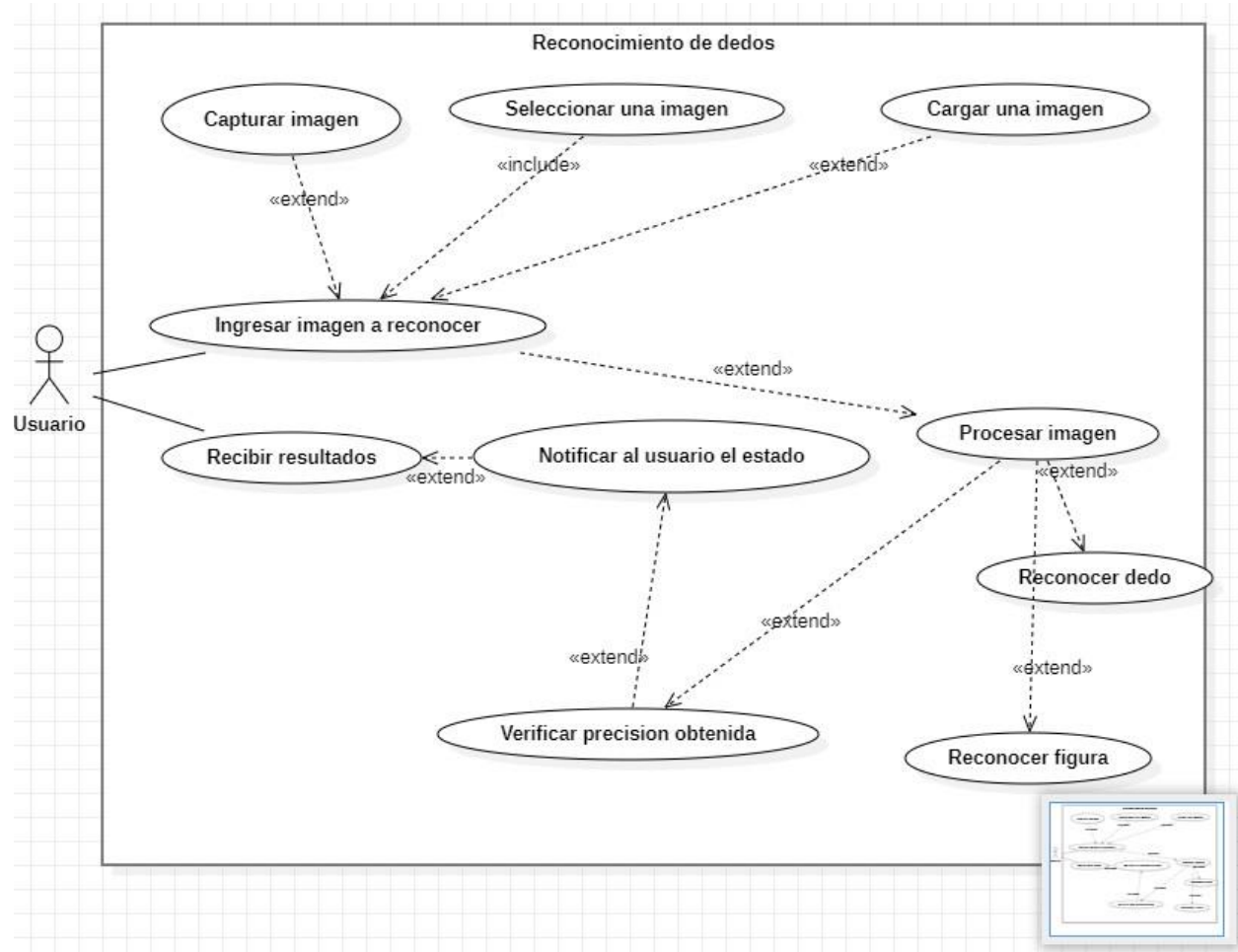


Figura 12: Diagrama de Casos de uso

Fuente: Elaboración propia

En el diagrama del caso de uso se muestra al usuario que interactúa con 2 partes de la solución, en el ingreso de la imagen que se produce en la selección de la imagen capturada y posteriormente es entregada al algoritmo para así detectar los dedos y la figura que este realiza. Por otro lado, se encuentran los resultados, que una vez ejecutado el análisis se le entrega mediante una notificación el cual indica la figura formada por los dedos en base al análisis de la imagen.

HERRAMIENTAS UTILIZADAS

MEDIA PIPE HANDS

MediaPipe Hands utiliza una canalización de aprendizaje automático que consta de varios modelos que trabajan juntos: un modelo de detección de palma que opera en la imagen completa y devuelve un cuadro delimitador de mano orientado. Un modelo de punto de referencia de la mano que opera en la región de la imagen recortada definida por el detector de la palma y devuelve puntos clave de la mano en 3D de alta fidelidad.

Después de la detección de la palma en toda la imagen, nuestro modelo de punto de referencia de la mano posterior realiza una localización precisa de puntos clave de 21 coordenadas 3D de los nudillos de la mano dentro de las regiones de la mano detectadas mediante regresión, es decir, predicción directa de coordenadas. El modelo aprende una representación consistente de la postura interna de la mano y es robusto incluso ante manos parcialmente visibles y autooclusiones.



Figura 13: Puntos de referencia de manos

MEDIAPIPE

Mediapipe es una biblioteca de visión por computadora y procesamiento de medios desarrollada por Google, que ofrece una variedad de soluciones y herramientas para el análisis de imágenes y videos mediante el uso de técnicas de aprendizaje automático.

Mediapipe ofrece una amplia gama de módulos predefinidos que se pueden combinar y personalizar según las necesidades del proyecto a realizar. Algunos de los módulos más populares incluyen:

- Detección de caras y seguimiento facial.
- Detección de manos y seguimiento de gestos de las manos.
- Estimación de poses humanas (sistemas de puntos clave del cuerpo).
- Detección de objetos y reconocimientos de acciones.
- Seguimiento de objetos y rastreo de movimiento.

Además, Mediapipe proporciona herramientas para la corrección de lentes, el seguimiento de objetos en 3D, la superposición de gráficos y muchas otras funcionalidades relacionadas con el procesamiento de medios. Dado que Mediapipe es una biblioteca de código abierto, se puede acceder a su código fuente en el repositorio público de GitHub y es utilizado tanto por desarrolladores individuales como por equipos de investigación y proyectos comerciales para construir aplicaciones y sistemas que requieren análisis y procesamiento de medios avanzado. La biblioteca está diseñada para ser fácil de usar, eficiente y flexible, y se ha utilizado en una variedad de aplicaciones.

OPEN CV2

La librería “cv2” se refiere a OpenCV, que significa “Open Source Computer Vision Library”. OpenCV es una biblioteca de código abierto de visión por computadora y procesamiento de imágenes que proporciona una amplia gama de herramientas y algoritmos para trabajar con imágenes y videos en tiempo real. Esta librería fue desarrollada por Intel en 1999 y desde entonces fue ampliamente utilizada en aplicaciones de visión por computadora en diversos campos, como la robótica, la realidad aumentada, la vigilancia, la detección de objetos, el reconocimiento facial, entre otros.

Las principales características y funcionalidades que ofrece OpenCV incluyen:

1. Manipulación de imágenes. Permite cargar, mostrar, guardar y procesar imágenes en diferentes formatos.
2. Procesamiento de imágenes: ofrece una amplia gama de funciones para aplicar filtros, transformaciones y operaciones matemáticas en imágenes.
3. Detección y seguimiento de objetos: proporciona herramientas para la detección y el seguimiento de objetos en imágenes y videos.
4. Estimación de poses humanas: permite la detección y el seguimiento de puntos clave en el cuerpo humano para la estimación de poses.
5. Aprendizaje automático: ofrece integración con bibliotecas de aprendizaje automático como TensorFlow y PyTorch.
6. Procesamiento de videos: permite leer y escribir videos, así como procesar fotogramas individuales.

NumPy

NumPy es una popular librería de código abierto para Python que se utiliza ampliamente en el ámbito de la ciencia de datos, el análisis numérico y el cómputo científico. El nombre NumPy es una abreviatura de Numerical Python. Esta librería proporciona una potente estructura de datos, arreglo multidimensional (llamado ndarray), que permite realizar operaciones matemáticas y numéricas de manera eficiente. NumPy es ampliamente utilizado por científicos, ingenieros, investigadores y desarrolladores que trabajan con datos numéricos y tareas de cómputo científico. Su eficiencia y versatilidad lo convierten en una herramienta esencial para el procesamiento y análisis de datos en Python.

Principales características:

- Arreglos multidimensionales: NumPy ofrece el objeto “ndarray”, que es un arreglo multidimensional que permite representar datos en forma de matrices y vectores de N dimensiones.
- Operaciones matemáticas: NumPy proporciona una amplia gama de funciones matemáticas y operaciones vectorizadas que se pueden aplicar directamente a los arreglos lo que permite realizar cálculos numéricos de manera eficiente.
- Indexación y filtrado: es posible acceder y manipular los elementos de los arreglos mediante indexación y filtrado, lo que facilita el procesamiento de datos.
- Broadcasting: NumPy implementa broadcasting, que permite realizar operaciones entre arreglos de diferentes formas y dimensiones, lo que evita tener que escribir bucles explícitos.
- Álgebra lineal: NumPy incluye funciones para realizar operaciones de álgebra lineal, como multiplicación de matrices, resolución de sistemas de ecuaciones lineales y descomposición de valores singulares.
- Generación de números aleatorios: NumPy proporciona herramientas para generar números aleatorios de diversas distribuciones, lo que es útil para simulaciones y experimentos.
- Integración con otras bibliotecas: NumPy se integra bien con otras bibliotecas de Python utilizadas en ciencia de datos y cómputo científico, como SciPy, Matplotlib y Pandas.

MATH

La librería Math es un módulo incorporado que proporciona funciones matemáticas y constantes para realizar operaciones numéricas y cálculos matemáticos. No es necesario instalar nada adicional para utilizar esta librería ya que viene incluida en la instalación estándar de Python.

Esta librería es muy útil cuando se necesita realizar cálculos matemáticos más complejos o cuando trabajas con funciones trigonométricas, exponenciales y logarítmicas. Es una herramienta poderosa que forma parte del conjunto estándar de módulos de Python y es ampliamente utilizada en aplicaciones científicas, ingeniería y análisis numérico.

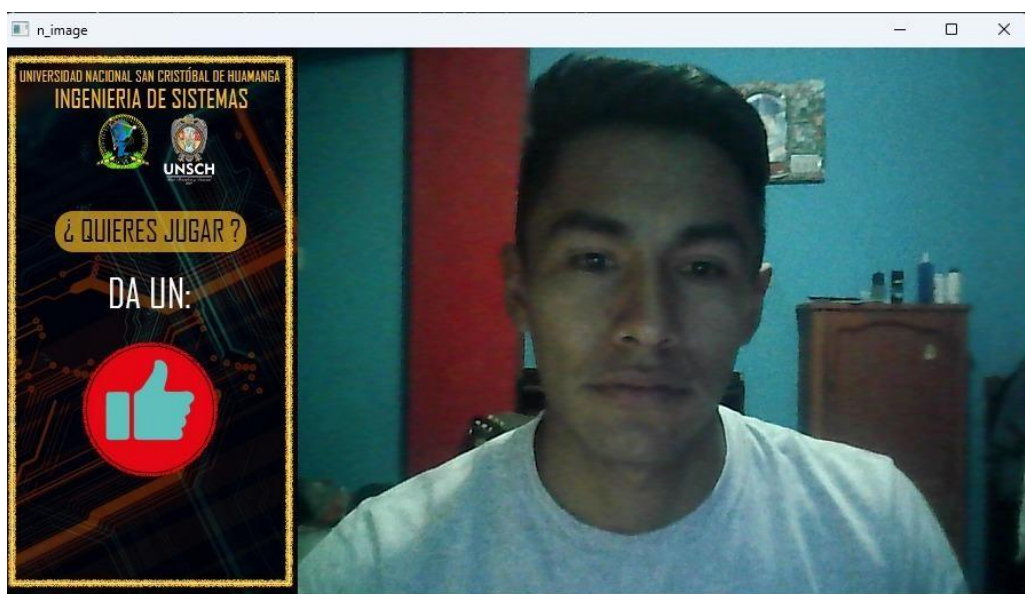
RANDOM

La librería random en Python es otro módulo integrado que proporciona funciones para generar número aleatorios y realizar operaciones relacionadas con la aleatoriedad. Al igual que la librería math, random también es parte de la biblioteca estándar de Python y no requiere instalación adicional. Permite realizar operaciones relacionadas con la generación de valores aleatorios, como obtener números enteros aleatorios, números de punto flotante aleatorios y seleccionar elementos aleatorios de una lista.

DESARROLLO DE INTERFAZ

Para el desarrollo de la interfaz de la librería Open CV2, en el lenguaje de programación de Python. El interfaz implementado se muestra de la siguiente manera:

Figura 14: Interfaz



Fuente: Elaboración propia

1. Área de detección:

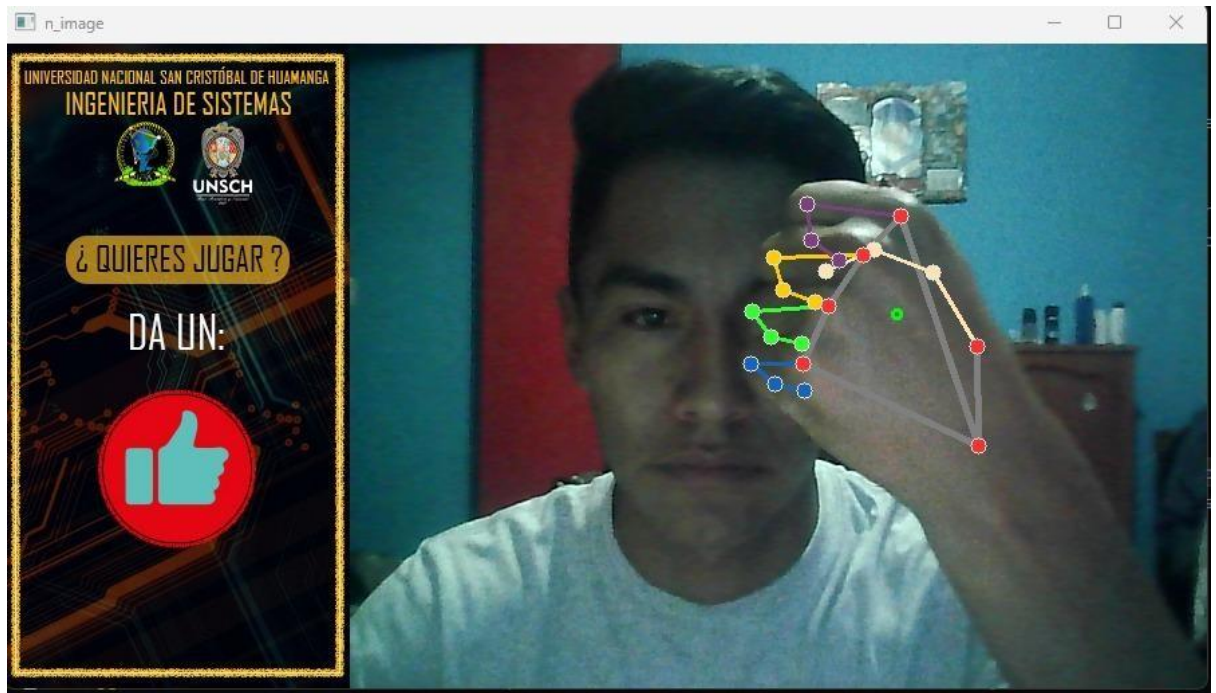
- Muestra en tiempo real la vista de la cámara donde el jugador puede colocar sus manos.

Figura 15: Interfaz



Fuente: Elaboración propia

Figura 16: Interfaz

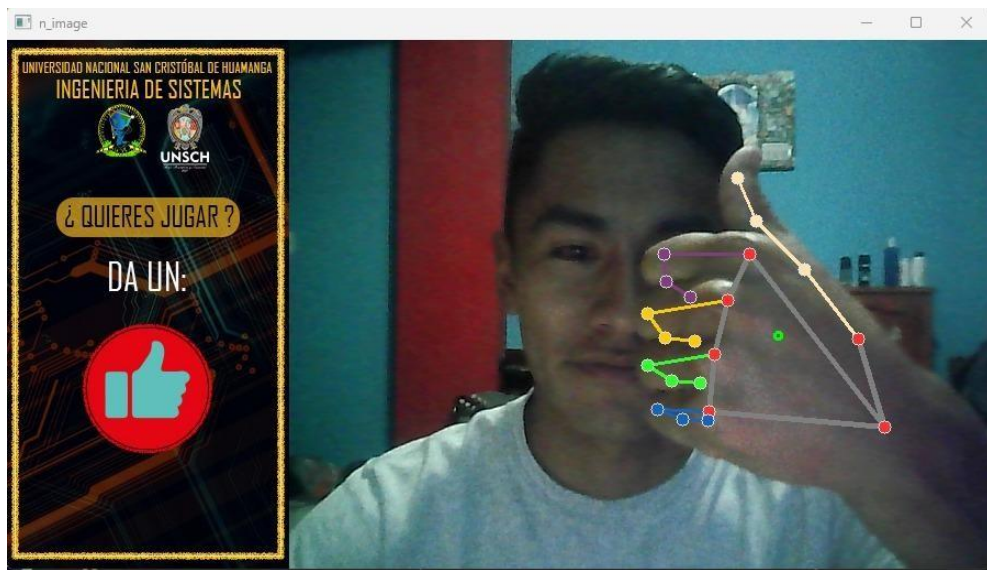


Fuente: Elaboración propia

2. Visualización de elecciones:

- Después de que el sistema detecte un gesto, muestra la imagen del gesto realizado por el jugador (Piedra, Papel o Tijera).
- La interfaz también muestra la elección de la IA para que el jugador pueda ver la competencia.

Figura 17: Interfaz



Fuente: Elaboración propia

Figura 18: Interfaz

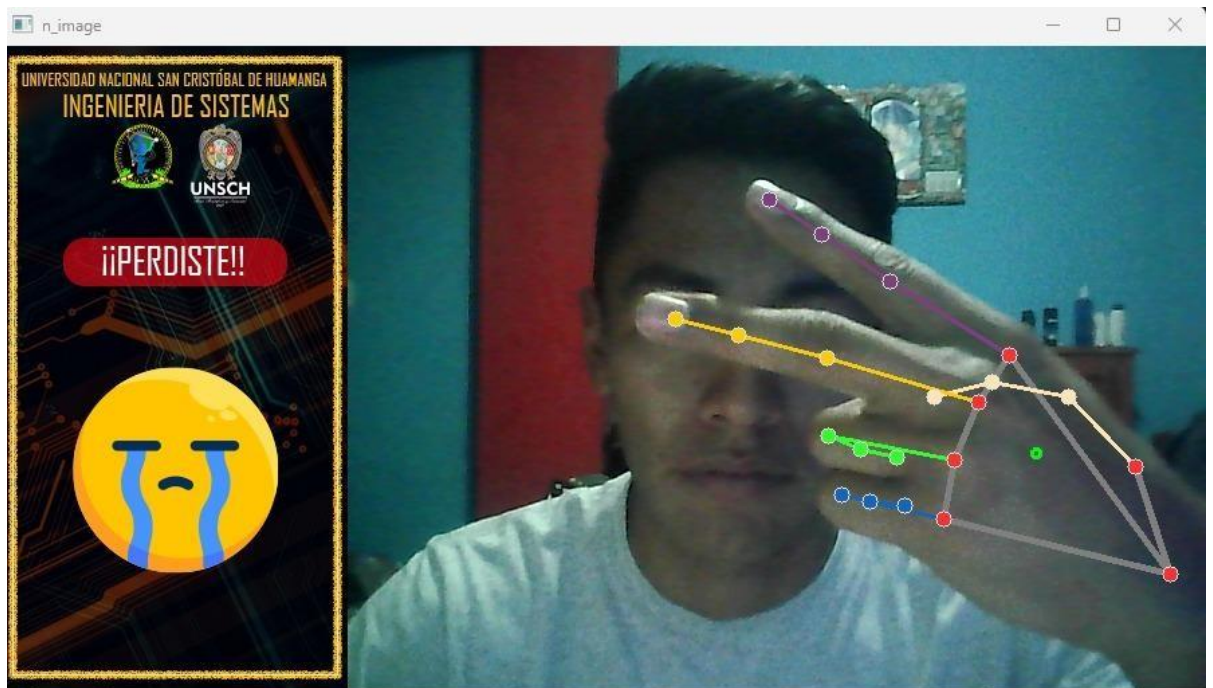


Fuente: Elaboración propia

3. Resultado del juego:

- Después de que ambos jugadores (jugador humano y la Inteligencia Artificial) hayan realizado sus elecciones, muestra el resultado del juego (quién ganó, quién perdió o si hubo un empate).
- Muestra los siguientes mensajes como “Ganaste”, “Perdiste” y “Empate”

Figura 19: Interfaz



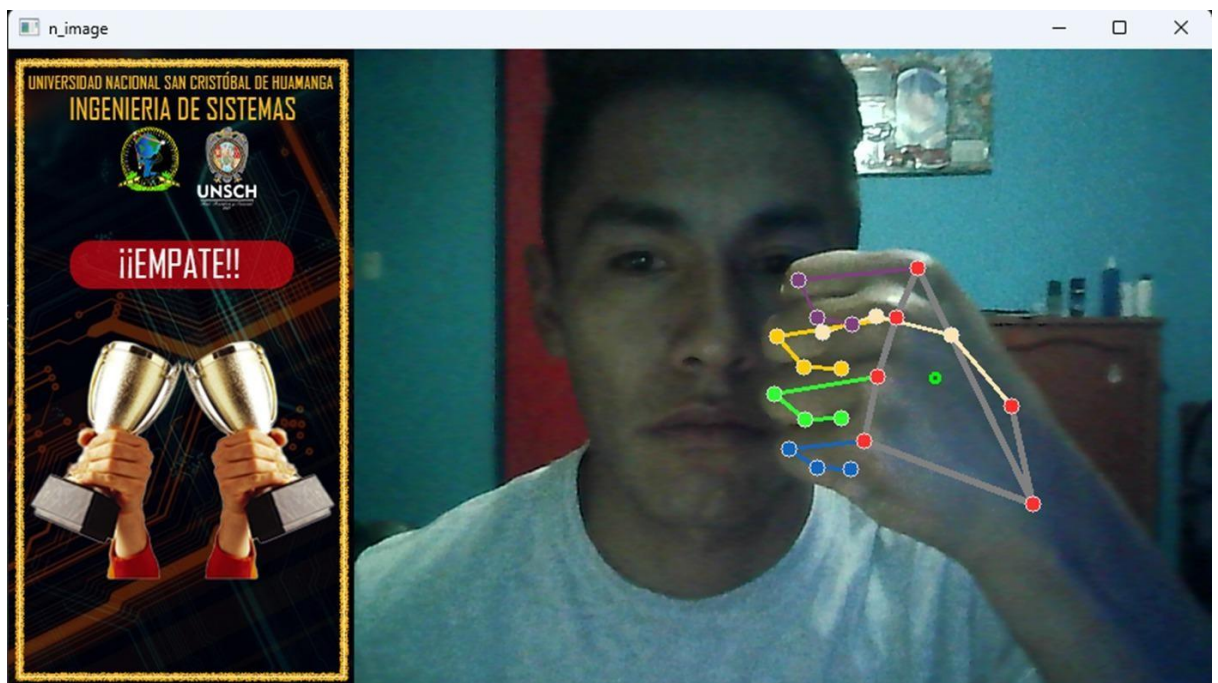
Fuente: Elaboración propia

Figura 20: Interfaz



Fuente: Elaboración propia

Figura 21: Interfaz



Fuente: Elaboración propia

CONCLUSIÓN

Luego de realizar las pruebas sobre la aplicación de visión artificial y machine learning al juego de piedra, papel o tijera, podemos llegar a las diferentes conclusiones:

- La integración de Visión Artificial y Machine Learning en el juego de Piedra, Papel o Tijera ha demostrado la capacidad de enriquecer la experiencia del juego. Al permitir que un usuario se enfrente contra un sistema inteligente capaz de aprender y adaptarse a las estrategias, se crea un desafío más dinámico y emocionante.
- La combinación de estas tecnologías no solo se limita al juego en sí, sino también tiene aplicaciones en otros campos. Desde la detección de gestos y patrones hasta la automatización que fueron adaptadas de manera dinámica y creativa en el juego.
- La combinación de Visión Artificial y Machine Learning promueve una mayor interacción entre los usuarios y máquinas. Esto influye en cómo los jugadores interactúan con la tecnología y cómo confían en la capacidad de la máquina para comprender sus acciones y gestos.

BIBLIOGRAFÍA

- Alonso, R. (09 de Febrero de 2023). HZ. HARD ZONE. Obtenido de IA, Machine Learning, Deep Learning: <https://hardzone.es/tutoriales/rendimiento/diferencias-ia-deep-machine-learning/>
- ELBS. (15 de Octubre de 2020). Escuela de Liderazgo. Obtenido de Conceptos sobre imagen digital: <https://escuelaelbs.com/conceptos-imagen-digital/>
- Thoughtworks. (19 de Mayo de 2020). Radar tecnológico. Obtenido de mediaPipe: [https://www.thoughtworks.com/es-es/radar/languages-and-frameworks/mediapipe#:~:text=MediaPipe%20es%20un%20marco%20de,Web%2C%20y%20dispositivos%20perimetrales\).](https://www.thoughtworks.com/es-es/radar/languages-and-frameworks/mediapipe#:~:text=MediaPipe%20es%20un%20marco%20de,Web%2C%20y%20dispositivos%20perimetrales).)
- Unir, U. e. (03 de Agosto de 2021). Ingeniería y Tecnología. Obtenido de Qué son las redes neuronales: <https://www.unir.net/ingenieria/revista/redes-neuronales-artificiales/>
- XI, A. (21 de Mayo de 2020). COPE, ABSIDE MEDIA. Obtenido de Piedra, papel o tijera: https://www.cope.es/actualidad/vivir/juegos-tradicionales/noticias/juegos-para-aislamiento-piedra-papel-tijera-20200521_728128