

Sistema de reconocimiento facial para localización de personas desaparecidas utilizando cámaras hiperespectrales.

Alejandro Hernández Villanueva.

Dr. Doricela Gutierrez Cruz

*Universidad Autónoma del Estado de México, Nezahualcóyotl,
Estado de México, México.*

alexhdzv00@gmail.com, dgutierrezcr@uaemex.mx

Resumen.

El presente trabajo detalla la creación y aplicación de un sistema de reconocimiento facial utilizando cámaras de espectro infrarrojo y bibliotecas de Python. Su objetivo es localizar personas desaparecidas o no encontradas, identificando características faciales específicas como el iris, tamaño de ojos, contorno de labios y tamaño de cejas. Se emplea una malla poligonal para tridimensionalizar el rostro, permitiendo mediciones precisas de dimensiones, distancias entre ojos y profundidades. Este proceso es crucial para entrenar la red CNN y lograr así una identificación precisa del objetivo, obteniendo un reconocimiento efectivo con una tasa de certeza que oscila entre el 90% y el 95% en entornos con poca iluminación y una aplicación correcta de la malla poligonal.

Palabras clave: Reconocimiento facial, Python, cámaras espectro infrarrojo, malla poligonal, CNN.

Abstract.

The present work details the creation and application of a facial recognition system using infrared spectrum cameras and Python libraries. Its objective is to locate missing or missing people, identifying specific facial characteristics such as the iris, eye size, lip contour and eyebrow size. A polygonal mesh is used to three-dimensionalize the face, allowing precise measurements of dimensions, distances between eyes and depths. This process is crucial to train the CNN network and thus achieve accurate target identification, obtaining effective recognition with a certainty rate ranging between 90% and 95% in low-light environments and correct application of the polygonal mesh. ..

Keywords: Facial recognition, Python, infrared spectrum cameras, polygonal mesh, CNN.

1. INTRODUCCIÓN

Los secuestros en México han crecido un 3.2% al cierre de 2023 con respecto al año anterior, contabilizando 1630 nacionales y 772 personas identificadas como migrantes. Siendo un problema con respecto a la administración y plan de ejecución dispuesto por el gobierno federal por el gobierno federal. Siendo la zona norte de la república mexicana con mayor índice de secuestros, encabezado por Sonora siguiendo por San Luis Potosí y Tamaulipas. Teniendo un promedio de 3.3 secuestros por día. [1] Los planes de acción cuando una persona es privada de su libertad constan en: iniciar una búsqueda en el último lugar conocido de su ubicación, realizar entrevistas a familiares, conocidos y posibles testigos de la desaparición o supervisar rutas habituales, entre otras [2]. En todo caso, este protocolo toma por último la revisión de cámaras de seguridad públicas o privadas. En la posibilidad de encontrar una cámara cercana, el reconocimiento facial sería una herramienta favorable ya que puede ser capturado el rostro, no sólo de la víctima, sino también de los victimarios.

Es importante señalar que, debido a la velocidad con la que pueden moverse los secuestradores y las víctimas, es complicado encontrarlos en las primeras 72 horas. Por ello, se deben aprovechar las herramientas tecnológicas disponibles para incrementar las posibilidades de encontrarlos. Entre ellas se encuentran las cámaras hiperespectrales de ondas infrarrojas que permiten percibir lo que el ojo humano no puede ver a simple vista. Además, se pueden utilizar distintos tipos de entrenamiento para analizar patrones de secuestro y mejorar el tiempo de respuesta. Todo esto contribuiría a reducir los índices de personas desaparecidas y aumentar las posibilidades de encontrar a quienes han sido privados de su libertad.

Según datos del Instituto Nacional de Estadísticas y Geografía (INEGI), la encuesta nacional de seguridad pública urbana en México del cuarto trimestre del año 2022, en el mes de diciembre el 64.2% de la población mayor a 18 años considera inseguro vivir en su ciudad. Para proporcionar una medición periódica, amplia, integral y oportuna de seguridad pública en el país, la ENSU cuenta con una periodicidad trimestral.[3]

Una herramienta utilizada para el fin de localización de personas de interés es el Registro de Personas desaparecidas y no localizadas (RNPNO); es una herramienta que utilizan las entidades federativas, haciendo uso de los registros previos aportado por el Centro Nacional de Planeación, Análisis e

Información para el Combate a la Delincuencia (CENAPI), Registro Nacional de Datos de Personas Extraviadas o Desaparecidas (RNPED) e información de diversas fuentes no gubernamentales con el fin de optimizar la búsqueda, localización e identificación.[4]

“Visión es un proceso que produce a partir de las imágenes del mundo exterior una descripción que es útil para el observador y que no tiene información irrelevante”, Marr [5].

La Visión por Computadora (VC) se puede resumir como la abstracción de un fenómeno del mundo real a lo digital, conforme sea el interés, podemos eliminar información irrelevante que altere la descripción o interpretación del resultado al que se pretende llegar.

Un área muy ligada es el procesamiento de imágenes la cual es indispensable para que la interpretación de la visión por computadora sea más exacta, ya que puede ayudar a remover problemas como: ruido en la imagen y el desenfoque, además, mejorar las propiedades de la imagen como: la exposición de la luz, contraste, brillo, sombras, estructura o agregar colores a una imagen monocromática o en escala de grises.

2. MATERIALES Y MÉTODOS

Se propone implementar algoritmos de reconocimiento facial del Lenguaje Python en puntos de revisión estratégicos utilizando las librerías OpenCV, MediaPipe Face Mesh, y Face_Recognition incorporado a cámaras hiperespectrales de ondas infrarrojas.

Es necesario llevar un de tratamiento de la imagen, el entrenamiento requiere una imagen en las mejores condiciones para que la red neuronal sea estable. Las etapas involucradas en sistema de reconocimiento se pueden resumir en la Fig. 1. Las técnicas de procesamiento de imágenes (PDI) basadas en color permiten representar adecuadamente la información de una imagen por medio de la computadora.[6]

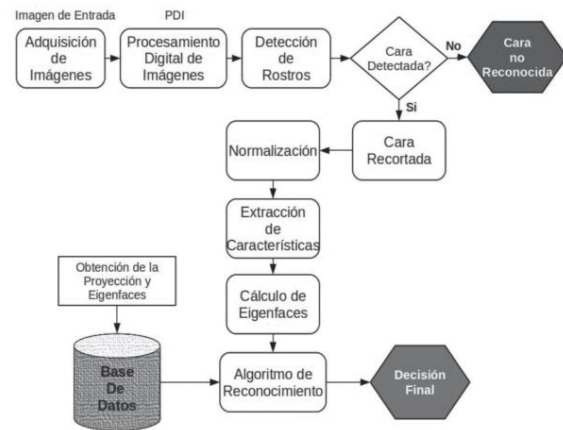


Figura 1. Estructura de sistema de reconocimiento facial.
Fuente: PEDRAZA PICO, B. O., RONDÓN, P., & ARGUELLO, H. (2011). Sistema de reconocimiento facial basado en imágenes con color. Revista UIS Ingenierías, 10(2), 113-122.

2.1. Tratamiento de vídeo con OpenCV

OpenCV (Open source computer vision library) es una biblioteca de código abierto que incluye cientos de algoritmos de visión por computadora. Basado en el lenguaje de programación C++ de la versión 2.5 en adelante. Al poseer una estructura modular, podemos invocar funciones específicas de esta librería en este caso utilizamos las siguientes:

- Procesamiento de imágenes. Encargada en el filtrado de imágenes lineales y no lineales, transformación de imágenes.
- Procesamiento de vídeo. Considera la estimación del video, resta del fondo y algoritmos de seguimiento.
- Marco de características. Detector de características sobresalientes, descriptores y cargadores de descriptores.[7]

2.2 Malla Poligonal con MediaPipe Face Mesh

MediaPipe Face Mesh es una solución de detección de 468 puntos faciales de referencia en un plano 3D en tiempo real, emplea aprendizaje automático para inferir la superficie facial tridimensional sin la necesidad de un sensor de profundidad dedicado. Al establecer un espacio 3D métrico y utilizar la posición de la pantalla como punto de referencia para estimar la transformación del rostro dentro del mismo espacio

dando como resultado una matriz de transformación del rostro y una malla facial triangular, además de la utilización de un método estadístico llamado Análisis de Procrustes que conduce la lógica robusta de manera eficiente optimizando el uso de la CPU y memoria RAM.

La tubería de ML implementada consta de dos modelos de redes neuronales profundas en tiempo real que funcionan juntos: Un detector que opera en la imagen completa y calcula las ubicaciones de la cara y un modelo de punto de referencia de la cara en 3D que opera en esas ubicaciones y predice la superficie 3D aproximada a través de la regresión. [8]

2.3 Reconocimiento facial con algoritmo Face Recognition

Face_recognition es una de las líneas de comandos con la biblioteca de reconocimiento facial más común del mundo. Construido a partir de la librería Dlib con aprendizaje profundo. Este modelo tiene una precisión de 99.38% en los rostros que se etiquetan por puntos de referencia estratégicos, la forma de trabajar de este algoritmo consiste en transformar la imagen en una matriz tipo NumPy.[9] [10]

Se realiza una codificación de los rasgos faciales para realizar una comparativa entre sujetos y determinar si existen coincidencias. Para tener un resultado favorable es indispensable contar con una base de datos de las imágenes a comparar y sea fácil de consultar.

2.4 Conexión de algoritmos a sensor Kinect.

Para establecer la comunicación entre el sensor Kinect y la PC, es necesario instalar el SDK v1.8.0, Kinect Studio v1.8.0 y el software Developer ToolKit Browser v1.8.0[11], para que, sea posible visualizar la cámara infrarroja que trabaja en una frecuencia de onda de los 780 nm a 2500 nm, ya que mide 830nm y obtener una imagen que detalle de manera más precisa las facciones del rostro.[12]

2.5 Redes neuronales

CNN: Red neuronal convolucional es un tipo de red que utiliza como entrada imagen y audio, son usadas principalmente para tareas de clasificación y visión artificial, aprovechando los principal de álgebra lineal (matrices), identificando los patrones de una imagen. consta de 3 capas

- Capa convolucional
- Capa de agrupamiento
- Capa totalmente conectada

las capas convolucionales pueden seguir de otra capa convolucional o agrupadas donde la completamente conectada es la capa final, conforme aumenta la capa la red se vuelve más compleja, donde las primeras capas se enfocan en en las características simples como el color y bordes , continua con el reconocimiento de elementos más grandes hasta que llega al objeto deseado. [13]

3. DESARROLLO

Se realizó la conexión de una cámara de espectro visible para verificar el correcto funcionamiento de nuestro sistema, ajustamos la malla poligonal, corrección de colores, entrenamiento y reconocimiento a cargo de la red neuronal.

Para la creación y desarrollo de este sistema se utilizó programación orientada a objetos en el lenguaje Python con ayuda de las librerías ya antes mencionadas y explicadas de manera general, además de la conexión y controladores necesarios para obtener una imagen hiperespectral y poder realizar las pruebas necesarias. Se tuvo que obtener un adaptador de CA-USB además de la instalar el controlador SDK developer que hace posible la conexión de la cámara de ondas infrarrojas y la visualización de imagen en el espectro infrarrojo.

```
with mp_face_mesh.FaceMesh(
    max_num_faces=15,
    refine_landmarks=True,
    min_detection_confidence=0.9,
    min_tracking_confidence=0.05) as face_mesh:
```

La línea max_num_face especifica el numero máximo de rostros que el modelo puede identificar en una imagen.

refire_landmarks=True habilita la mejora de puntos faciales clave

min_detection_confidence=0.9 establece un nivel de confianza mínimo (valor entre 0 y 1)

min_tracking_confidence es el valor de confianza mínimo para el seguimiento del rostro, El uso del contexto with ... as face_mesh: garantiza que los recursos utilizados por FaceMesh sean correctamente liberados después de que el bloque de código dentro de este contexto se haya ejecutado

4. RESULTADOS

Describiendo la figura 2, se visualiza una imagen en el espectro visible añadiendo una malla poligonal de 468 vértices sobre el rostro, identificando los rasgos faciales como tamaño de las cejas, ojos, iris, contorno de labios y concavidades del rostro que pueden ser identificadas donde los puntos son más cercanos entre ellos.

Posteriormente se duplica la imagen que permite examinar de manera detallada el rostro, verificando las áreas donde los puntos son más cercanos entre sí y etiquetando el rostro con el nombre del sujeto, la prueba fue realizada con obstáculos que dificultan la comparativa de rostros y una segunda imagen sin malla que permite visualizar la apariencia, color de piel o si lleva algún distintivo como perforación o tatuajes

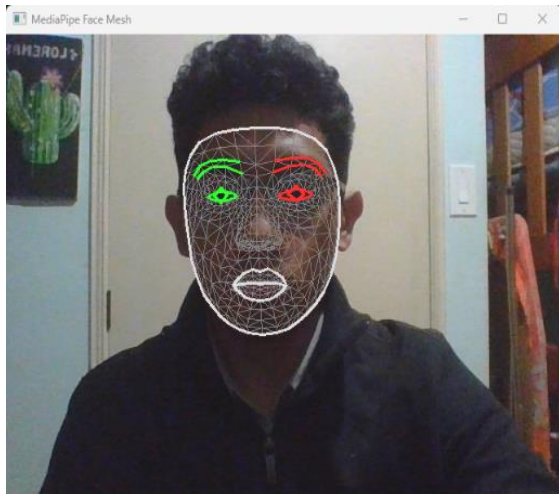


figura 3.

Figura 2. Malla poligonal sobre un rostro en la frecuencia de espectro visible.

Fuente: Autoría propia (2024)

Una vez confirmada la correcta ejecución y salida deseada del proceso, se realiza la transferencia de cámara de ondas del espectro visible a una de ondas de espectro infrarrojo, repitiendo el procedimiento anterior. Se recomienda realizar anotaciones de los

detalles del rostro que deseemos ver en la siguiente prueba, para modificar la red neuronal si es necesario. Describiendo los resultados de la figura 4.

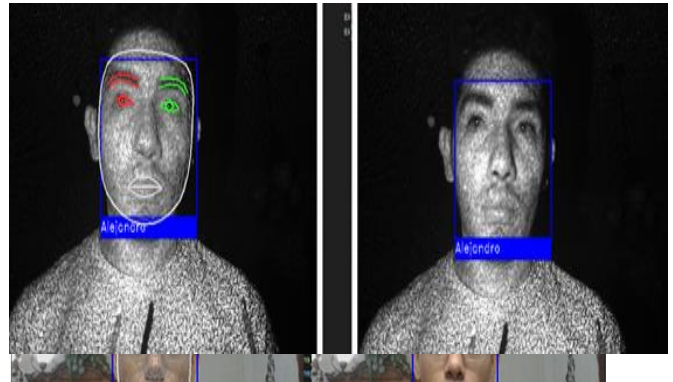


Figura 4. Detección y etiquetado de rostro conocido en una cámara de ondas infrarrojas.

Fuente: Autoría propia.(2024).

Figura 3. Detección y etiquetado de un rostro conocido en la frecuencia de espectro visible.

Fuente: Autoría propia.(2024).

Podemos observar que los detalles trazados por la malla son más eficientes, por ejemplo las cejas, está bien definido el inicio y término de cada una, el iris se pueden ver a mayor detalle, la coloración alrededor del párpado inferior, las áreas de salida de vello facial, forma y tamaño de las orejas. El último procedimiento se realizó en condiciones de poca luz, para verificar si se satisface la necesidad de obtención de imagen por el déficit de iluminación de las cámaras utilizadas actualmente. Dando una respuesta satisfactoria.

Se obtiene la matriz resultante de la red convolucional y la red arroja el mensaje de verdadero una vez que el

rostro ha sido localizado e identificado con éxito figura 5.

Se recupera el resultado del entrenamiento en formato XML basado en eigenvectores sobre el rostro (eigenface) repitiendo las iteraciones al número de imágenes obtenidas figura 6.

necesario realizar un entrenamiento apropiado, disminuir la entropía de nuestro reconocimiento a no más 5% con el motivo de estar trabajando en una dimensión 3D y pasarla a una 2D, la pérdida de información es significativa dejando una toma de decisión poco confiable si no es llevado de buena manera tanto el procesamiento de la imagen como el entrenamiento de la red.

La implementación de este sistema a la vida real y sobre todo en la seguridad pública tendrían un gran alcance, ya que los puntos vulnerables en su mayoría son lugares con mala iluminación o poco transitados, incluir cámaras infrarrojas con una longitud de onda de 3,000nm a 14,000nm nos permite usarlas en horarios diurno y nocturno, haciendo de lado el problema de identificación en condiciones poco favorables, al realzar los rasgos distintivos de rostro, la probabilidad de perder el objetivo es cada vez menor, ya que el sistema reconoce los patrones y al ser una red neuronal sigue aprendiendo cada vez que visualiza su objetivo.

```

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS
> TERMINAL
-0.07139595 -0.17977995 -0.32274303 0.09222341 0.42358565 0.11844037
-0.18870413 0.03300603 -0.1719719 -0.11030342 0.02106275 0.11432043
-0.02615711 -0.00167763 -0.03326095 0.06054643 0.18613417 0.05484102
-0.06039943 0.21393599 -0.01755594 0.01690739 0.02373897 0.11551845
-0.13312495 -0.0380229 -0.12514509 -0.0494183 -0.05561668 -0.04698397
0.01459432 0.09805545 -0.16922876 0.18078229 0.02576401 -0.01130891
    
```

Figura 6. Resultado del entrenamiento método EigenFaces
Fuente: Autoría propia.(2024).

Figura 5. Matriz resultante de la red neuronal
Fuente: Autoría propia.(2024).

```

4 <threshold>1.7976931348623157e+308</threshold>
5 <radius>1</radius>
6 <neighbors>8</neighbors>
7 <grid_x>8</grid_x>
8 <grid_y>8</grid_y>
9 <histograms>
10 <_ type_id="opencv-matrix">
11 <rows>1</rows>
12 <cols>16384</cols>
13 <dt>f</dt>
14 <data>
15 0.00000000 3.08641978e-03 0.00000000 0.00000000 0.00000000
16 3.08641978e-03 6.17283955e-03 3.08641978e-03 0.00000000 0.00000000
17 0.00000000 7.40740746e-02 9.87654328e-02 2.16049384e-02
18 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
19 0.00000000 3.39506194e-02 3.08641978e-03 0.00000000 2.59259254e-01
20 1.26543209e-01 1.85185187e-02 0.00000000 3.08641978e-03 0.00000000
21 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
22 0.00000000 3.08641978e-03 0.00000000 0.00000000 0.00000000
23 0.00000000 1.23456791e-02 0.00000000 0.00000000 4.93827164e-02
24 0.00000000 1.94444448e-01 1.23456791e-02 1.23456791e-02 0.00000000
25 0.00000000 3.08641978e-03 0.00000000 0.00000000 0.00000000
26 0.00000000 0.00000000 0.00000000 6.17283955e-03 0.00000000
27 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
    
```

5. CONCLUSIONES

En conclusión, el presente trabajo demuestra que el uso de cámaras que pueden transitar entre rangos de espectro además del visible, fortalecen el reconocimiento facial, ya que no dependen de condiciones ideales para cumplir su función, mostrando los rasgos distintivos de cada individuo. Al usar algoritmos basados en redes neuronales, es

6. REFERENCIAS

1. “Secuestros en México aumentan un 3.2%; los migrantes son un tercio de las víctimas”. Forbes México. Accedido el 19 de enero de 2024. [En línea]. Disponible: <https://www.forbes.com.mx/secuestros-en-mexico-aumentan-un-3-2-los-migrantes-son-un-tercio-de-las-victimas/>
2. PGR, Protocolo homologado para la búsqueda de personas desaparecidas y la investigación del delito de desaparición forzada. (2014)
3. INEGI, Encuesta Nacional de Seguridad Pública Urbana Cuarto Semestre de 2022, pp. 1, (2023).
4. RNPDO – Comisión Nacional de Búsqueda. (s.f.). Comisión Nacional de Búsqueda – Sitio Oficial. <https://comisionacionaldebusqueda.gob.mx/rnpdno/#:~:text=A1%2015%20de%20junio%20de%202023%20existen%20reportadas%20oficialmente%20como,de%2011%20mil%2042%20personas>
5. Marr D., Visión, (1982).

6. Arguello H., Sistema de reconocimiento basado en la imagen facial, Revista Avances en Sistemas e Informática, vol. 8, núm. 3, pp. 7-15, (2011)
7. Google, https://google.github.io/mediapipe/solutions/face_mesh, (2023).
8. NumPy Developers, <https://numpy.org/doc/stable/user/whatisnumpy.html>, (2008)
9. Geitgey A., Face Recognition, <https://face-recognition.readthedocs.io/en/latest/readme.html>, (2017)
10. OpenCV, <https://opencv.org/>, (2022)
11. Microsoft, <https://www.microsoft.com/en-us/download/details.aspx?id=40278>, (2021)
12. Kinect para Windows - Windows apps. (s.f.). Microsoft Learn: Build skills that open doors in your career. <https://learn.microsoft.com/es-es/windows/apps/design/devices/kinect-for-windows>
13. “¿Qué son las redes neuronales convolucionales? | IBM”. IBM in Deutschland, Österreich und der Schweiz. Accedido el 17 de marzo de 2024. [En línea]. Disponible: <https://www.ibm.com/mx-es/topics/convolutional-neural-networks>
14. Ortega J., Eigenvalores y Eigenvectores, <http://gmc.geofisica.unam.mx/papime2020/index.php/articulos/38-eigenvalores-y-eigenvectores>, (2021)
15. H. Moon and P. J. Phillips, Computational and performance aspects of pca-based face-recognition algorithms. 2001, vol. 30, No. 3, pp. 303-321
16. K. Kim, Face recognition using principal component analysis.
17. N. Morizet, F. Amiel, I. Hamed, and T. Ea, “A comparative implementation of pca face recognition algorithm”. in Electronics, Circuits and Systems, ICECS 2007. 14th IEEE International Conference, pp. 865-868.
18. Pedraza B., Rondón P., Arguello H., Sistema de reconocimiento facial basado en imágenes con color Revista UIS Ingenierías, vol. 10, núm. 2, pp. 113-122, (2011)
19. INEGI, Censo Nacional de Seguridad Pública Estatal 2022, pp 41,60 (2022)
20. *Reconocimiento facial: definición y explicación.* (s.f.). latam.kaspersky.com. <https://latam.kaspersky.com/resource-center/definitions/what-is-facial-recognition> (2024).