



Evaluación del estado de somnolencia a través de la ecuación de relación de aspecto del ojo (EAR) mediante un algoritmo adaptativo implementado en teléfonos inteligentes

Evaluation of sleepiness state through the eye aspect ratio equation (EAR) using an adaptive algorithm implemented on smartphones

Avaliação do estado de sonolência através da equação da proporção dos olhos (EAR) utilizando um algoritmo adaptativo implementado em smartphones

Mateo Granja^I

jmgranja1@espe.edu.ec

<https://orcid.org/0009-0004-3536-3159>

Daniel Yáñez^{II}

dayanez8@espe.edu.ec

<https://orcid.org/0000-0002-5425-8918>

Javier Montaluisa^{III}

fjmontaluisa@espe.edu.ec

<https://orcid.org/0000-0002-9816-1990>

Patricio Navas^{IV}

mpnavas@espe.edu.ec

<https://orcid.org/0000-0001-5862-9475>

Correspondencia: jmgranja1@espe.edu.ec

Ciencias Técnicas y Aplicadas
Artículo de Investigación

* **Recibido:** 08 de diciembre de 2023 * **Aceptado:** 11 de enero de 2024 * **Publicado:** 29 de febrero de 2024

- I. Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.
- II. Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.
- III. Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.
- IV. Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.

Resumen

Cuando se pretende evaluar la somnolencia mediante el uso de la cámara de un teléfono inteligente, uno de los principales inconvenientes que se presentan es la rapidez del parpadeo esto hace que la detección automática sea difícil. Un estudio reciente propone una técnica para identificar somnolencia a través de un algoritmo de visión artificial que puede identificar la somnolencia en tiempo real mediante la cámara frontal de un teléfono inteligente. La técnica consiste en determinar la posición de los puntos de referencia faciales en cada fotograma, extraer la distancia vertical entre los párpados y calcular un único valor escalar conocido como relación de aspecto del ojo (EAR). El objetivo de esta investigación es determinar qué valor del umbral EAR es un indicador confiable y si un algoritmo adaptativo implementado en teléfonos inteligentes es una herramienta útil para evaluar la somnolencia. Los resultados de las pruebas experimentales indican que un umbral EAR más alto disminuye la precisión y que 0.25 es el umbral óptimo en esta investigación.

Palabras clave: ojo; parpadeo; procesamiento de imágenes; somnolencia; teléfono inteligente.

Abstract

When attempting to evaluate drowsiness using a smartphone camera, one of the main drawbacks that arise is the rapidity of blinking, which makes automatic detection difficult. A recent study proposes a technique to identify sleepiness through a computer vision algorithm that can identify sleepiness in real time using the front camera of a smartphone. The technique involves determining the position of facial landmarks in each frame, extracting the vertical distance between the eyelids, and calculating a single scalar value known as the eye aspect ratio (EAR). The goal of this research is to determine which EAR threshold value is a reliable indicator and whether an adaptive algorithm implemented on smartphones is a useful tool for assessing sleepiness. The experimental test results indicate that a higher EAR threshold decreases the accuracy and that 0.25 is the optimal threshold in this research.

Keywords: eye; flicker; image processing; drowsiness; smartphone.

Resumo

Ao tentar avaliar a sonolência por meio da câmera de um smartphone, uma das principais desvantagens que surge é a rapidez do piscar, o que dificulta a detecção automática. Um estudo recente propõe uma técnica para identificar a sonolência por meio de um algoritmo de visão

computacional que consigue identificar a sonolência em tempo real usando a câmara frontal de um smartphone. A técnica envolve determinar a posição dos pontos de referência faciais em cada quadro, extrair a distância vertical entre as pálpebras e calcular um único valor escalar conhecido como relação de aspecto do olho (EAR). O objetivo desta pesquisa é determinar qual valor do limiar EAR é um indicador confiável e se um algoritmo adaptativo implementado em smartphones é uma ferramenta útil para avaliar a sonolência. Os resultados dos testes experimentais indicam que um limite EAR mais alto diminui a precisão e que 0,25 é o limite ideal nesta pesquisa.

Palavras-chave: olho; cintilação; processamento de imagem; sonolência; telefone inteligente.

Introducción

Según Rosales Mayor y Rey De Castro Mujica (2010), en el artículo "Somnolencia: Qué es, qué la causa y cómo se mide", la evaluación del estado de somnolencia es de vital importancia en numerosos ámbitos, como la seguridad vial, la medicina del sueño y la salud en general. Detectar y prevenir la somnolencia excesiva puede tener un impacto significativo en la prevención de accidentes y la mejora de la calidad de vida de las personas. En este sentido, el avance de la tecnología ha permitido el desarrollo de métodos no invasivos y de bajo costo para evaluar el estado de somnolencia. Uno de estos enfoques prometedores es la utilización de la ecuación de relación de aspecto del ojo (EAR, por sus siglas en inglés) a través de un algoritmo adaptativo implementado en teléfonos inteligentes. La ecuación de relación de aspecto del ojo se basa en el análisis de las características del ojo, como el tamaño y la forma de la pupila, para determinar el grado de somnolencia de una persona (Granda, 2021). La pupila es una parte clave del sistema visual que se modifica en función del estado de alerta y la somnolencia. En estudios anteriores, se ha demostrado que la relación de aspecto de la pupila puede ser un indicador confiable de la somnolencia. La implementación de esta ecuación en teléfonos inteligentes ofrece una serie de ventajas significativas. En primer lugar, los teléfonos inteligentes son dispositivos ampliamente disponibles y utilizados en la actualidad, lo que facilita su adopción y aplicación generalizada. Además, hay dispositivos que cuentan con cámaras de alta resolución y potentes capacidades de procesamiento, lo que permite la captura precisa de imágenes del ojo y el cálculo eficiente de la relación de aspecto. En este artículo, presentamos un algoritmo adaptativo para la evaluación del estado de somnolencia a través de la ecuación de relación de aspecto del ojo implementándolo en teléfonos inteligentes. El algoritmo utiliza técnicas de aprendizaje automático y procesamiento de imágenes para detectar

y analizar las características oculares relevantes. Posteriormente, se lleva a cabo una validación experimental del algoritmo, así como una evaluación de la precisión de la ecuación EAR, utilizando un grupo de participantes. Los resultados obtenidos fueron comparados y se estableció el umbral óptimo para la evaluación del estado de somnolencia.

Recapitulando, el presente estudio tiene como objetivo desarrollar y evaluar un algoritmo adaptativo basado en la ecuación de relación de aspecto del ojo implementado en teléfonos inteligentes para la evaluación del estado de somnolencia. Se espera que los resultados de esta investigación contribuyan al avance de las tecnologías no invasivas de evaluación de la somnolencia y tengan un impacto significativo en la seguridad y la salud de las personas.

Herramientas y métodos

Herramientas Software

Las herramientas software que han permitido el desarrollar el aplicativo son los siguientes:

Java: Es un lenguaje de programación de alto nivel y orientado a objetos, que permite desarrollar aplicaciones móviles. Es uno de los lenguajes de programación más populares de la industria.

Android Studio: Es una herramienta desarrollada por Google, ampliamente utilizada para el desarrollo de Software en diferentes lenguajes.

Google Vision: Es una API desarrollada por Google que permite integrar funcionalidades de Machine Learning en aplicaciones desarrolladas en Java o Kotlin.

SQLite: Es un motor de base de datos que se integra con Java para la persistencia de los mismos

Planteamiento del sistema

El sistema será implementado en un dispositivo móvil, por lo cual la entrada de imagen será la cámara frontal. Para el desarrollo del algoritmo se ha realizado una división por etapas, las cuales se pueden apreciar en la *Fig. 1*.

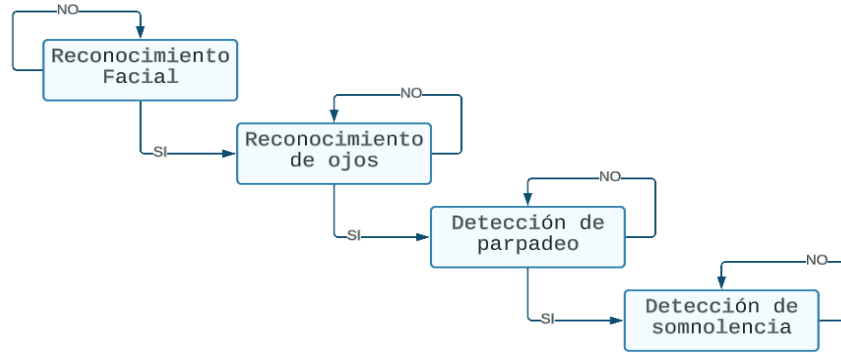


Fig. 1: Etapas del sistema.

Reconocimiento facial: En esta etapa se busca reducir los tiempos de procesamiento y mejorar la eficacia del algoritmo. Esto se logra mediante una detección preliminar del rostro en la imagen capturada, con el propósito de establecer el punto de inicio del algoritmo y, por ende, realizar un análisis más exhaustivo de la imagen. Además, permite extraer un área de interés de la imagen, lo que acorta los tiempos de procesamiento y garantiza la existencia de una región ocular, mejorando así la eficiencia de la aplicación.

Reconocimiento de ojos: En esta etapa se logra ubicar los ojos en el rostro, una vez que se ha delimitado el rostro en la etapa previa (reconocimiento facial), es más sencillo aproximar el reconocimiento de ojos usando el método dentro de una zona delimitada.

Detección de parpadeo: Esta etapa es la más crítica debido a que evaluará si el individuo presenta sus ojos abiertos o cerrados, también lo realiza con ayuda de la etapa previa, el reconocimiento de ojos, en la cual ya está delimitada la zona a evaluar por lo que se busca determinar si los ojos están abiertos o cerrados.

Detección de somnolencia: Esta etapa analiza, mediante la detección de parpadeo, el tiempo que la persona mantiene sus ojos cerrados, en la etapa previa se determina cuándo están abiertos o cerrados, en esta se evalúa por cuánto tiempo los mantiene cerrados para establecer que el individuo tiene somnolencia.

Métodos de detección

A continuación, en la *Fig. 2.* se presentan los métodos de detección facial, de ojos, de parpadeo y de somnolencia más relevantes:

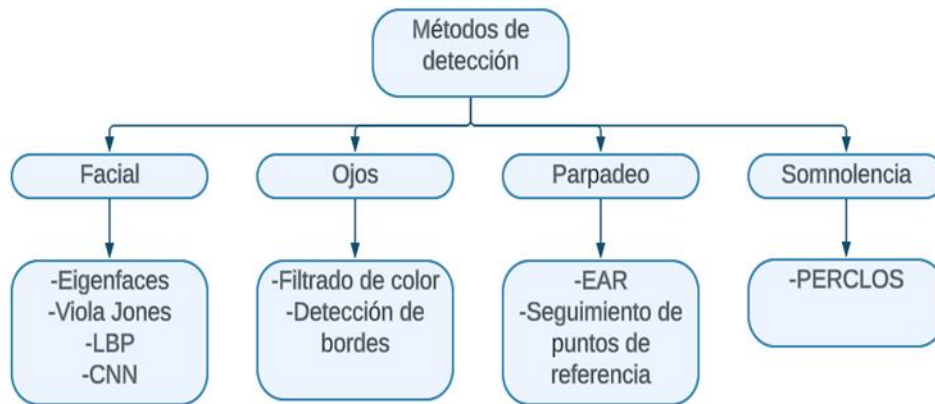


Fig. 2: Métodos de detección.

Reconocimiento Facial

Eigenfaces: Es un algoritmo utilizado en el reconocimiento facial que utiliza un conjunto de imágenes de entrenamiento para generar una representación compacta de las caras. Estas representaciones, llamadas eigenfaces, capturan las características más distintivas de las caras y se utilizan para comparar y reconocer caras en nuevas imágenes. (Calli, 2015)

Viola Jones: Es un algoritmo popular utilizado para la detección rápida y precisa de objetos en imágenes. Es conocido especialmente por su eficacia en la detección de rostros en tiempo real, utilizando características llamadas "cascadas de características" y un clasificador basado en aprendizaje automático. (Hernández, Cabrera & Sanchez, 2012)

LBP (Local Binary Patterns): Es un descriptor de características utilizado en el análisis de imágenes para representar texturas y patrones locales. Se basa en la comparación de píxeles vecinos en una imagen y su relación binaria para codificar información sobre texturas y formas locales. (Bernal, 2018)

CNN (Convolutional Neural Network): Es una arquitectura de red neuronal profunda diseñada específicamente para el procesamiento de imágenes. Las CNN utilizan capas de convolución para extraer características relevantes de una imagen y aprenden automáticamente patrones y características a diferentes niveles de abstracción. (Wu, Hassner, Kim, Medioni & Natarajan, 2017).

Reconocimiento de ojos

Filtrado de color: El filtrado de color es una técnica empleada para seleccionar o resaltar objetos o regiones en una imagen, utilizando como base su información cromática. Esta técnica es comúnmente utilizada en aplicaciones de visión por computadora para segmentar y extraer información específica de una imagen en función de sus componentes de color.

Detección de bordes: La detección de bordes es un proceso utilizado para identificar los límites o transiciones abruptas entre diferentes regiones en una imagen. Los bordes contienen información crucial sobre la estructura y los contornos de los objetos presentes en la imagen, por lo que su detección resulta fundamental en numerosas aplicaciones de procesamiento de imágenes.

Reconocimiento de parpadeos

EAR: Es una métrica utilizada en el análisis de ojos para cuantificar y detectar cambios en la apertura de los ojos. Se basa en la relación geométrica entre los puntos clave de los ojos y puede ser utilizado en aplicaciones como el reconocimiento de parpadeo o la detección de somnolencia.

Seguimiento de puntos de referencia: El seguimiento y localización de puntos es una técnica empleada para rastrear y ubicar puntos específicos en una secuencia de imágenes. En el contexto del reconocimiento facial, estos puntos de referencia pueden ser características como los ojos, la nariz, entre otros.

Reconocimiento de somnolencia

PERCLOS (Percentage of Eye Closure): Es una métrica utilizada en la detección de somnolencia para medir el porcentaje de cierre de los ojos en una secuencia de imágenes. El PERCLOS se utiliza como indicador de somnolencia y puede ser utilizado para alertar o prevenir situaciones de riesgo, como la conducción somnolienta.

Una vez que se han descrito las etapas y sus potenciales métodos de implementación es hora de determinar cuál de las siguientes se conviene ser desarrollada, por lo que a priori se realizó la combinación **Fig. 3.** de cada elemento perteneciente a cada grupo de detección, los resultados son los siguientes:

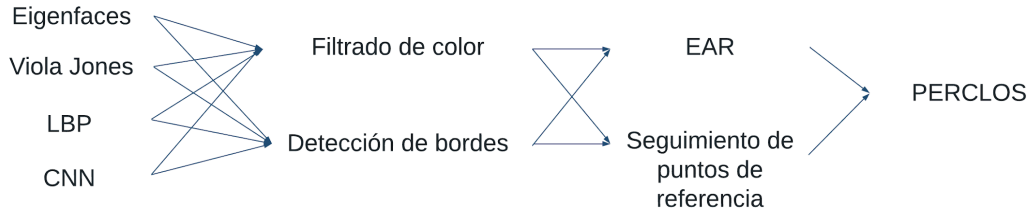


Fig. 3: Combinación de cada metodología de evaluación.

- Tomando en cuenta las tecnologías existentes para el desarrollo en dispositivos móviles es importante seleccionar una que permita implementar alguna de las combinaciones encontradas. Una de ellas es Google Vision API la cual en su documentación detalla que permite integrar el reconocimiento facial y de ojos mediante una red neuronal Convolutiva (Fig. 4.) y el uso de detección de bordes en la parte de reconocimiento de ojos, por lo que quedarían dos etapas a implementar en el desarrollo, el reconocimiento de la apertura de los ojos y determinar si el individuo está o no con somnolencia.

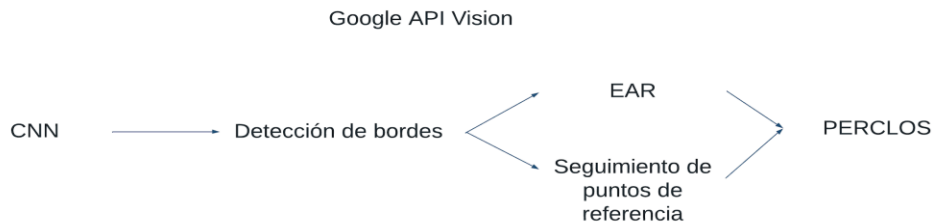


Fig. 4: Características de Google Vision API.

Esta selección nos limita a determinar si utilizar EAR o el seguimiento de puntos de referencia para determinar el parpadeo. Esta selección se explica dentro del diseño del apartado 2.4.1 *diseño del algoritmo*

Implementación

A continuación, se muestra un esquema general de implementación Fig. 5. que marca el resultado a obtener en cada etapa del algoritmo.

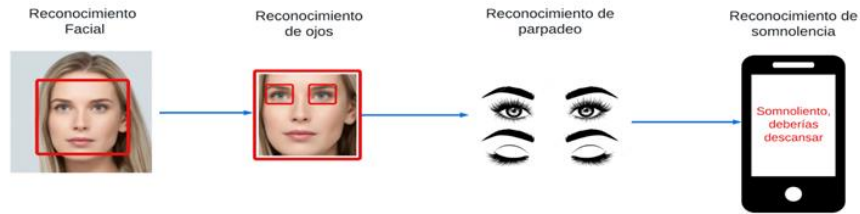


Fig. 5: Etapas de implementación

Diseño del algoritmo

De acuerdo a la revisión de aplicativos similares implementados en ordenadores, es recomendable utilizar el EAR debido a su rápida implementación y validación (Patel,2018), por lo que la combinación a desarrollar será mostrada en la **Fig. 6**.



Fig.6: Diseño del Algoritmo

Red Neuronal Convolutacional

Las redes neuronales convolucionales (CNN) han demostrado ser muy efectivas para detectar rostros en imágenes (Li, Z., et al. 2021). Para implementarlas es necesario seguir un conjunto de pasos recomendados:

Conjunto de datos: Primero, se necesita un conjunto de datos con imágenes de rostros y sus correspondientes ubicaciones o cajas delimitadoras (bounding boxes) que indiquen la posición y tamaño de los rostros en cada imagen (Detección de la cara, 2021). Este conjunto de datos se utilizará para entrenar la red neuronal.

Arquitectura de la red: La arquitectura de la CNN para detección de rostros puede variar, pero usualmente suelen componerse de capas convolucionales para extraer características y capas de detección para identificar y localizar los rostros (Saquicela et al., 2022).

Extracción de características: Las primeras capas convolucionales de la red se encargan de extraer características relevantes de la imagen, como bordes, texturas y patrones faciales. Estas

características se obtienen mediante la aplicación de filtros convolucionales, los cuales se van refinando a medida que la red se profundiza.

Capas de detección: Después de la etapa de extracción de características, se utilizan capas especializadas para detectar y localizar los rostros en la imagen. Una de las técnicas comunes para la detección de objetos es la utilización de capas de convolución en cascada.

Aprendizaje supervisado: Durante el entrenamiento, la red neuronal ajusta los pesos de sus conexiones para minimizar la diferencia entre las ubicaciones de los rostros detectados y las ubicaciones reales de los rostros proporcionadas en el conjunto de datos etiquetado. Esto se realiza mediante el uso de algoritmos de optimización, como el descenso de gradiente estocástico, y se basa en la retro propagación del error a lo largo de la red.

Evaluación y predicción: Una vez entrenada, la red neuronal se puede utilizar para detectar rostros en nuevas imágenes. La red examina la imagen de entrada utilizando ventanas deslizantes de diferentes tamaños y realiza predicciones sobre la presencia y ubicación de rostros en cada ventana. Posteriormente, se aplican técnicas de post-procesamiento para filtrar y refinar las detecciones, como la eliminación de detecciones superpuestas o la utilización de algoritmos de ajuste preciso.

Es importante destacar que en esta investigación no se abordó el proceso de implementación de una Red Neuronal Convolutiva (CNN), ya que nuestro objetivo se centra en la evaluación del estado de somnolencia mediante el uso de la ecuación EAR. No obstante, resulta relevante mencionar que el proceso descrito anteriormente es utilizado en la detección de rostros, lo cual puede resultar útil para investigaciones futuras.

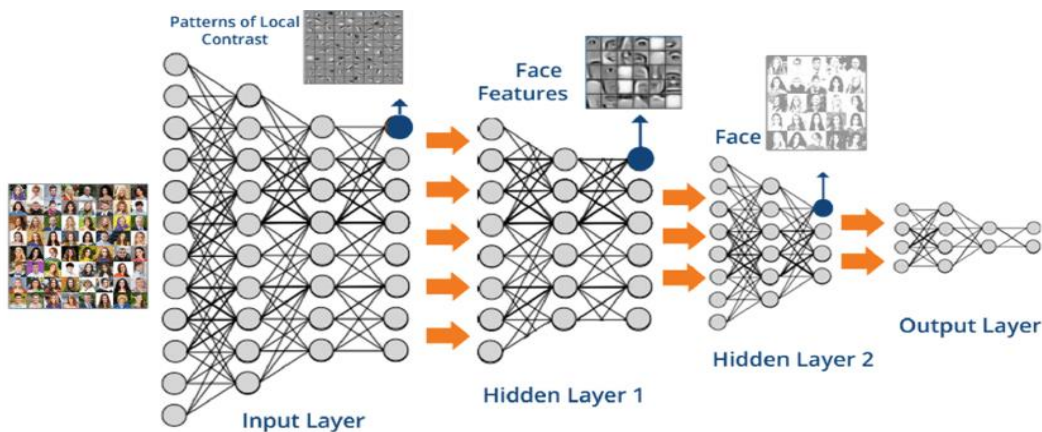


Fig. 7: Red Neuronal Convolutiva con varias capas. (Buigas, 2017)

Detección de Bordes

Cuando la capa de reconocimiento facial se encuentra completada el modelo será capaz de evaluar el rostro de la persona y con ayuda de un lienzo (canvas), dibujar los puntos de interés (Wu, Y., Hassner. et al., 2017) del individuo **Fig. 8** en tiempo real. En esta etapa es importante reconocer los ojos y delimitarlos con los puntos de interés, estos puntos se usarán posteriormente en el análisis de la apertura.

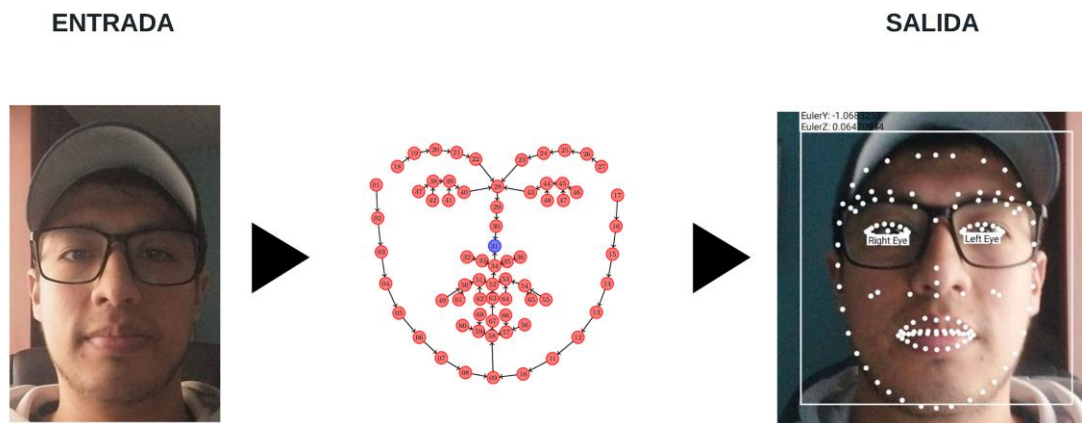


Fig. 8: Detección de borde de rasgos faciales

La API de Google Vision proporciona la numeración de los puntos basándose en el modelo Haarcascade, en su documentación podemos encontrar el detalle de cada rasgo facial **Tabla 1**.

Tabla 1: Detalle de rasgos faciales por rango. Adaptado de (Conceptos de detección de rostro | ML Kit | Google for Developers, s. f.).

	RANGO	CANTIDAD	CAPAS	TOTAL PUNTOS
ROSTRO	0-35	1	1	36
CEJAS	0-4	2	2	20
NARIZ	0-1	1	1	2
BOCA	0-8	1	4	36
OJOS	0-15	2	1	32
				126

Eye Aspect Ratio

El "Eye Aspect Ratio" (EAR) es una métrica utilizada en el campo de la visión por computadora y el reconocimiento facial para medir la apertura de los ojos en una imagen o video.

El EAR se basa en la posición de los puntos de referencia de los ojos, como las esquinas interna y externa de los ojos, y la parte superior e inferior del párpado. **Fig 9.**

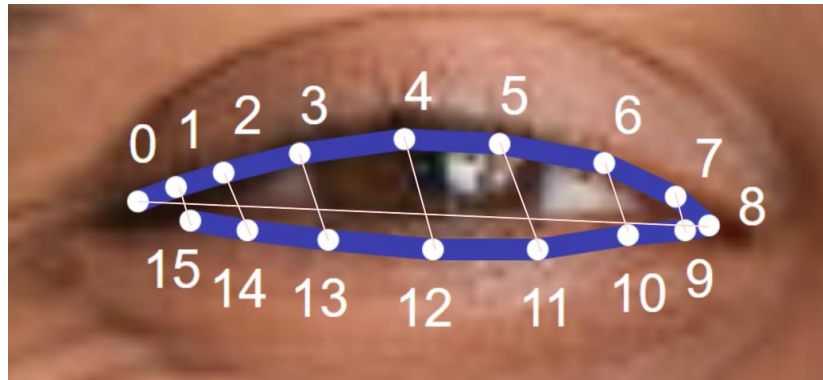


Fig. 9: Distancia entre los puntos de interés que delimitan el ojo

Al calcular la relación entre las distancias verticales y horizontales entre estos puntos de referencia, se puede determinar si los ojos están abiertos, cerrados o parcialmente cerrados. Su implementación es sencilla debido a que la Red Neuronal arroja en su detección un vector de reconocimiento con los puntos de interés de cada zona, véase **Tabla 1**, posterior a esto se debe realizar una operación matemática **Ecuación 1** durante cada frame, los valores encontrados se deben almacenar en otro vector para ser analizados en la siguiente etapa. Estos valores pueden oscilar entre 0 y 1, lo cual significa que los ojos están cerrados o abiertos respectivamente.

$$EAR = \frac{\sum_{n=17}^{i=2, i < \frac{n-1}{2} + 1} |P_i - P_{n-i+1}|}{2 |P_{i-1} - P_{\frac{n}{2}+1}|}$$

Ecuación 1: Relación de Aspecto del Ojo (EAR)

Perclos

El PERCLOS (Percent of Time Eyelids are Closed) es un índice utilizado para medir el estado de somnolencia basado en el porcentaje de tiempo en el que los ojos están cerrados durante un intervalo determinado. Se utiliza principalmente para evaluar la somnolencia de conductores,

operadores de maquinaria u otras personas en situaciones en las que la somnolencia puede representar un riesgo. (García, 2011)

Para calcular el índice PERCLOS, se utiliza la siguiente fórmula:

$$PERCLOS = \frac{N_{\text{Fotogramas con los ojos cerrados}}}{T_{\text{Número total de fotogramas}}} \times 100$$

Ecuación 2: Índice Perclos

En este cálculo, se capturan "n" fotogramas durante un período de tiempo específico utilizando la cámara frontal del dispositivo móvil. Se cuenta el número de fotogramas en los que los ojos están cerrados (EAR) y se divide por el número total de fotogramas. Luego, se multiplica por 100 para obtener el porcentaje de tiempo en el que los ojos están cerrados, lo cual representa el índice PERCLOS.

La **Fig. 10** muestra la implementación final del algoritmo en el lenguaje de programación Java.

```
//calculate EAR(Eye Aspect Ratio)
double EAR = (face.getIsLeftEyeOpenProbability() + face.getIsRightEyeOpenProbability())/2;
EAR = Math.round(EAR*100.0)/100.0;
if (EAR<0.25){
    Log.d("Ojos cerrados (EAR)",Double.toString(EAR));
    counter++;
    if (counter>=EYE_AR_CONSEC_FRAMES){

        canvas.drawText("SOMNOLIENTO ",
            400,
            1175, mIdPaint);
        canvas.drawText("Deberias descansar " ,
            375,
            1275, mIdPaint);

        try {
            Uri notification = RingtoneManager.getDefaultUri(RingtoneManager.TYPE_NOTIFICATION);
            Ringtone r = RingtoneManager.getRingtone(co, notification);
            r.play();
        } catch (Exception e) {
            e.printStackTrace();
        }
    }
}
}else{
    Log.d("Ojos abiertos (EAR)",Double.toString(EAR));
    counter=0;
}
```

Fig. 10: Muestra de implementación del algoritmo en JAVA

Pruebas

Para las pruebas de rigor se ha decidido testear tanto el reconocimiento de parpadeo en primera instancia como paso previo al reconocimiento de la somnolencia y la evaluación del parpadeo en

un rango de tiempo, esto debido a que el dato de entrada para evaluar la somnolencia es el número de parpadeos en un lapso de tiempo determinado.

Parpadeo

En la evaluación del parpadeo interviene el umbral EAR que determina si la persona está con el ojo abierto o cerrado. En estas pruebas se iniciará con un valor EAR de 0.3 el cual se ajustará de acuerdo a los resultados obtenidos. La prueba consiste en colocarse frente al dispositivo móvil y parpadear, contar el número total de parpadeos y comparar con el número de reconocimientos que ha realizado el algoritmo. Para mayor precisión se realizarán 3 iteraciones de 5 evaluaciones cada una. Posterior se obtendrá el porcentaje de aciertos.

Iteraciones

Tabla 2: Iteraciones de parpadeo evaluado por el aplicativo

EAR = 0.3 i1	EAR = 0.35 i2	EAR = 0.25 i3	
Evaluación 1		Evaluación 1	
Total parpadeos	10	Total parpadeos	10
Reconocidos	7	Reconocidos	5
Sin Reconocer	1	Sin Reconocer	5
Evaluación 2		Evaluación 2	
Total parpadeos	10	Total parpadeos	10
Reconocidos	8	Reconocidos	4
Sin Reconocer	2	Sin Reconocer	6
Evaluación 3		Evaluación 3	
Total parpadeos	10	Total parpadeos	10
Reconocidos	9	Reconocidos	10
Sin Reconocer	1	Sin Reconocer	0

Total parpadeos	10
Reconocidos	6
Sin Reconocer	4
Evaluación 4	
Total parpadeos	10
Evaluación 4	
Reconocidos	7
Sin Reconocer	3
Evaluación 5	
Total parpadeos	10
Reconocidos	8
Sin Reconocer	2

Total parpadeos	10
Reconocidos	5
Sin Reconocer	5
Evaluación 4	
Total parpadeos	10
Evaluación 4	
Reconocidos	6
Sin Reconocer	4
Evaluación 5	
Total parpadeos	10
Reconocidos	7
Sin Reconocer	3

Total parpadeos	10
Reconocidos	8
Sin Reconocer	2
Evaluación 4	
Total parpadeos	10
Evaluación 4	
Reconocidos	9
Sin Reconocer	1
Evaluación 5	
Total parpadeos	10
Reconocidos	10
Sin Reconocer	0

Véase el análisis de resultados en la sección *RESULTADOS*

Evaluación de somnolencia

Una vez que se ha obtenido un valor EAR confiable se continúa al diseño y evaluación de la prueba de evaluación de somnolencia. Teniendo en cuenta que el índice PERCLOS se obtiene mediante una operación matemática, **Ecuación 2**, es necesario entonces evaluar en cada fotograma el estado del individuo a lo largo de un tiempo determinado. Como la evaluación de la somnolencia debe ser rápida y certera, se ha decidido empezar la evaluación con un tiempo de 1 segundo, los fotogramas a analizar dependen del equipo que proporciona la imagen de entrada.

Se realizaron 2 evaluaciones con cada dispositivo, la prueba consiste en mantener los ojos cerrados por el periodo de tiempo indicado y obtener los valores EAR promedio de cada fps.

Iteraciones

Para las siguientes pruebas se utilizarán 2 teléfonos inteligentes por lo que a continuación se detallan las especificaciones técnicas relevantes para esta investigación.

Tabla 3: Especificaciones técnicas

Características relevantes para la investigación	Equipo A	Equipo B
Modelo	Samsung Galaxy A23	Samsung Galaxy A31
Cámara frontal	8 MP, f/2.2, (wide) - 1080p 30fps	20 MP, f/2.2, (wide) - 1080p 60fps
Procesador	Qualcomm SM6375 Snapdragon 695 5G (6 nm)	Mediatek MT6768 Helio P65 (12nm)
CPU	Octa-core (2x2.2 GHz Kryo 660 Gold & 6x1.7 GHz Kryo 660 Silver)	Octa-core (2x2.0 GHz Cortex-A75 & 6x1.7 GHz Cortex-A55)
GPU	Adreno 619	Mali-G52 MC2
S.O.	Android 12	Android 10

Las siguientes pruebas consisten en que el individuo debe mirar fijamente a la cámara frontal del dispositivo para luego cerrar los ojos por un periodo de tiempo determinado. **Fig. 11.**

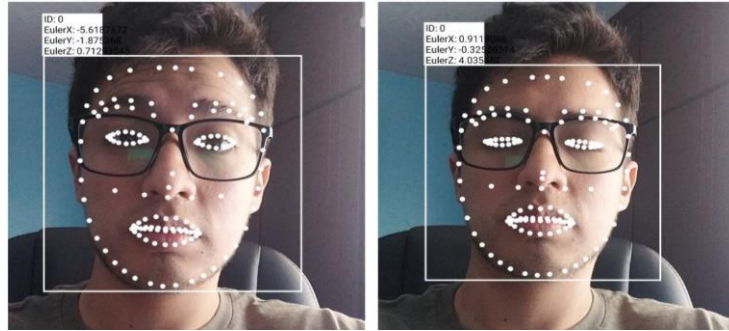


Fig. 11: Cerrar los ojos durante un intervalo de tiempo

Tabla 4: Pruebas de evaluación de somnolencia

ITERACIÓN	1
EQUIPO	A
TIEMPO (T)	1s
FPS	30
MUESTRAS	30
UMBRAL EAR	0,25
PERCLOS	91.41
RESULTADO	SOMNOLIENTO

ITERACIÓN	1
EQUIPO	B
TIEMPO (T)	1s
FPS	60
MUESTRAS	60
EMBRAL EAR	0,25
PERCLOS	96.67
RESULTADO	SOMNOLIENTO

ITERACIÓN	2
EQUIPO	A
TIEMPO (T)	3s
FPS	30
MUESTRAS	90
UMBRAL EAR	0.25
PERCLOS	88.89
RESULTADO	SOMNOLIENTO

ITERACIÓN	2
EQUIPO	A
TIEMPO (T)	3s
FPS	60
MUESTRAS	180
UMBRAL EAR	0,25
PERCLOS	93.89
RESULTADO	SOMNOLIENTO

Resultados

Se presenta la demostración de la ejecución del proceso que se lleva a cabo gracias a la operación matemática EAR, propuesta a través del análisis empírico (véase sección 2.5.1) cuando el valor obtenido por EAR es menor a 0.25 podemos afirmar que existe somnolencia) y mayor a 0.25 el estado del individuo es de vigilia, el análisis se lo puede ver en la **Tabla 5**.

Tabla 5: Resultados de iteraciones

ITERACIÓN 1			ITERACIÓN 2			ITERACIÓN 3		
Total	50		Total	50		Total	50	
Aciertos	38	0,76	Aciertos	27	0,54	Aciertos	46	0,92
Errores	12	0,24	Errores	23	0,46	Errores	4	0,08
UMBRAL EAR	0,30		UMBRAL EAR	0,35		UMBRAL EAR	0,25	

En la figura 12 (a) se muestra la ejecución de identificación de ojos abiertos y en la figura 12 (b) se muestra la ejecución de identificación de ojos cerrados. En las mismas se puede observar la evaluación de la somnolencia de una persona mediante la cámara frontal de un teléfono móvil, la figura 12 (a) muestra una persona normal con los ojos abiertos (sin signos de somnolencia), y la figura 12 (b) muestra un estado de somnolencia identificado después que el sujeto permaneciera con los ojos cerrados (3 segundos), activando un mensaje de alerta.



Fig. 12 (a): Identificación mediante EAR (ojos abiertos).



Fig. 12 (b): Identificación mediante EAR (ojos cerrados).

El sistema recopila datos en tiempo real basado en los puntos de referencia identificados en los ojos. Luego determina la somnolencia en función de los valores EAR calculados a partir de las imágenes capturadas del usuario. La **tabla 6** muestra el valor medio calculado de EAR, obtenido de las 6 pruebas realizadas. El valor promedio es de 0.60 y 0.10 cuando los ojos se encuentran abiertos y cerrados respectivamente. A partir de este experimento, puede observarse que no hay un valor medio de cero cuando los ojos se encuentran cerrados. No obstante, se puede deducir que si el valor disminuye repentinamente y permanece menor a 0.25 de manera consecutiva, el sujeto se encuentra en estado de somnolencia.

Tabla 6: Valor promedio del EAR mientras los ojos están abiertos y cerrados.

Prueba	EAR (ojos abiertos)	EAR (ojos cerrados)
1	0,70	0,11
2	0,72	0,09
3	0,53	0,12
4	0,39	0,05
5	0,69	0,10
6	0,54	0,11
Promedio	0,60	0,10

Según Patel et al. (2018), cuando el valor del EAR es menor a 0.25 se puede determinar que la persona se encuentra en estado de somnolencia. Este valor es determinado de acuerdo al movimiento de los ojos que cambia de abierto a cerrado; la **Fig. 13** representa el valor resultante del EAR, calculado de la identificación del ojo humano a lo largo del tiempo, las líneas continuas identifican las tres clasificaciones posibles. El comportamiento ocular esperado incluye ojos abiertos (verde): aquí, el EAR es mayor y presenta una gran variación. Los parpadeos (amarillo): el EAR disminuye rápidamente cuando se cierran los ojos y vuelve aumentar de la misma manera al abrirlos. Cuando los ojos permanecen cerrados (azul): el valor del EAR es menor pero el flujo es constante y no presenta una gran variación.

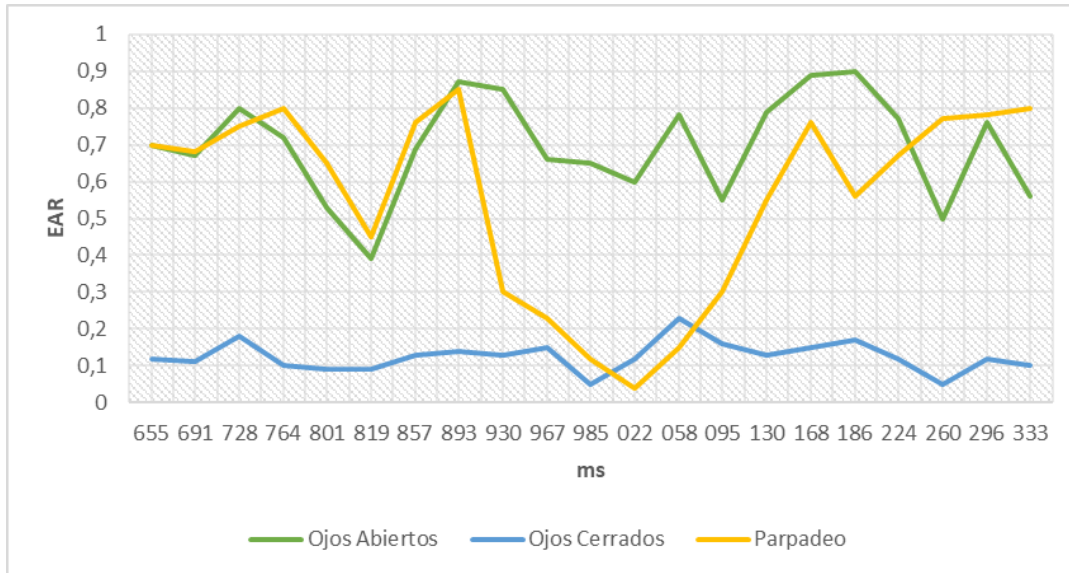


Fig. 13: Comportamiento del valor EAR en tiempo real para las 3 clasificaciones posibles

Discusión

En la presente investigación se evaluó el método Eye Aspect Ratio (EAR) para mejorar la precisión en la evaluación de la somnolencia a través de un algoritmo adaptativo implementado en teléfonos inteligentes. Los resultados obtenidos en base a las pruebas de ejecución realizadas demuestran que el procesamiento de las imágenes capturadas en tiempo real de la persona (para realizar el cálculo del valor EAR) funcionan de forma fluida y determinan de manera correcta el estado que se encuentran los ojos de la persona ya sea abiertos o cerrados, esto permite determinar con certeza cuando una persona se encuentra somnolienta, además que el algoritmo adaptativo implementado funciona correctamente proporcionando una mayor precisión al cálculo, este se adapta a los distintos tipos de rostros reajustando su modelo para detectar los puntos referenciales localizados en los ojos. En otros estudios realizados como el de Zhang et al. (2018), se concluyó que el método EAR tiene un alto porcentaje de precisión. El estudio demostró que el método, es uno de los más precisos para detectar la somnolencia.

En resumen, los resultados indican que el método propuesto puede detectar eficazmente la somnolencia en una persona utilizando la cámara frontal de un teléfono móvil. Además, los valores de EAR proporcionan información útil para determinar el estado de somnolencia en tiempo real, esto es útil para prevenir accidentes y situaciones de riesgo. Sin embargo, algunos de estos indicadores pueden variar debido a la iluminación del lugar o el tipo de resolución de la cámara del

celular, bajo esta premisa es necesario en futuras investigaciones determinar la precisión y la confiabilidad del método en diferentes situaciones y en una población más amplia.

Conclusiones

Al usar el algoritmo propuesto, se pudo rastrear puntos específicos en el rostro humano permitiendo identificar los puntos de referencia en los ojos de la persona, el algoritmo es capaz de adaptarse a los distintos tipos de rostros y a las formas de ojos que existen; a partir de la identificación y localización de estos puntos oculares se obtuvo el valor escalar del EAR que fue mediante la evaluación del uso de su ecuación respectiva en un periodo de tiempo, esto permitió determinar el porcentaje de abertura del ojo humano a lo largo del intervalo (PERCLOS), también se obtuvo que 0.25 es el umbral óptimo para determinar que un ojo se encuentra cerrado.

En base a los resultados obtenidos mediante el uso de la cámara de los dispositivos móviles de prueba, se observó también que el flujo de datos en tiempo real del cálculo del EAR para determinar que una persona se encuentra en estado de somnolencia funciona de forma fluida, sin presentar intermitencias o pérdidas de datos; cuando los ojos se encuentran abiertos el flujo cambia constantemente esto se debe a la iluminación que presente el lugar o al tipo de resolución de la cámara del celular, en cambio cuando el valor es menor a 0.25 y este estado se mantiene por alrededor de 3 segundos; el flujo es constante sin presentar mayor variación debido a la gran similitud que existe entre los rasgos correspondientes al ojo humano cuando se encuentra cerrado lo que facilita la identificación del ojo y la obtención del valor EAR. Cuando existe un parpadeo existe una mayor variación acercándose el valor a 0, demostrando que el tratamiento de la imagen que proporciona la cámara es verificado durante el tiempo que permanezca en uso la aplicación. El rendimiento del algoritmo es óptimo en los distintos celulares independientemente del tamaño, versión del sistema operativo o procesador de los mismos.

Agradecimientos

Después de haber trabajado tanto en este proyecto es necesario expresemos nuestra gratitud al universo por permitirnos perseguir nuestros sueños, a nuestra familia por apoyarnos incondicionalmente, a nuestros amigos por ser parte de este camino y a nuestros profes que han sido grandes guías durante el aprendizaje. Una especial mención a nuestros padres ya que todo lo

que somos y seremos se lo debemos a su incansable coraje, trabajo y enseñanza, son uno de los más grandes ejemplos que tenemos.

Referencias

1. Buigas, J. (2017). Guía rápida de inteligencia artificial. Así funciona Deep Learning. Puentes Digitales. Recuperado en 30 de junio de 2023, de <https://puentesdigitales.com/2017/11/15/guia-rapida-de-inteligencia-artificial-asi-funciona-el-deep-learning/>
2. Bernal Leyva, A. Y. (2018). ANÁLISIS DE MÉTODOS DE RECONOCIMIENTO FACIAL BAJO EL SISTEMA OPERATIVO ANDROID.
3. Calli Olvea, J. (2015). Reconocimiento facial basado en el Algoritmo Eigenface.
4. Conceptos de detección de rostro | ML Kit | Google for Developers. (s. f.). Recuperado 30 de junio de 2023, de <https://developers.google.com/ml-kit/vision/face-detection/face-detection-concepts?hl=es-419>
5. García Daza, I (2011). Detección de fatiga en conductores mediante fusión de sistemas ADAS [Tesis doctoral, Universidad de Alcalá Escuela Politécnica Superior]. Ebuha Biblioteca Digital Universidad de Alcalá. <https://ebuah.uah.es/dspace/handle/10017/16621>
6. Detección de la cara: ¿Qué es y cómo funciona esto Tech. (2021). RecFaces. <https://recfaces.com/es/articles/deteccion-de-la-cara-que-es-y-como-funciona-esto-tech>
7. Granda, F. A. (2021). SISTEMA DE DETECCIÓN DE SOMNOLENCIA A TRAVÉS DE VISIÓN ASISTIDA POR COMPUTADORA EN EL PERIODO ABRIL-SEPTIEMBRE 2021.
8. Hernández, Ernesto del T, Cabrera Sarmiento, Alejandro, & Sánchez Solano, Santiago. (2012). Implementación híbrida hardware software del algoritmo de detección de rostros de Viola-Jones sobre FPGA. Universidad, Ciencia y Tecnología, 16(63), 114-124. Recuperado en 30 de junio de 2023, de http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1316-48212012000200005&lng=es&tlng=es
9. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems.

10. Patel, R., Patel, M., & Patel, J. (2018). Real Time Somnolence Detection System In OpenCV Environment for Drivers. 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 407–410. <https://doi.org/10.1109/ICICCT.2018.8473234>
11. Rosales Mayor, E., & Rey De Castro Mujica, J. (2010). Somnolencia: Qué es, qué causa y cómo se mide. *Acta médica peruana*, 27(2), 137-143.
12. Saquicela, A., Zambrano-Asanza, S., Cajamarca, B., Chitacapa, A., Sanango, J., & Franco, J. (2022). Estimation of Rural Populations without Access to Electricity Through Satellite Images and Deep Learning.
13. Wu, Y., Hassner, T., Kim, K., Medioni, G., & Natarajan, P. (2017). Facial landmark detection with tweaked convolutional neural networks. *IEEE transactions on pattern analysis and machine intelligence*, 40(12), 3067-3074.
14. Zhang, G., Ren, H., Jiang, B., & Zhang, X. (2018). Driver drowsiness detection based on eye state analysis using infrared thermal imaging. *Infrared Physics & Technology*, 94, 269-276.

© 2024 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0) (<https://creativecommons.org/licenses/by-nc-sa/4.0/>).