DruBot: Prototipo robótico para autenticación por comparación de proporciones faciales para el control de asistencia y detectar la suplantación en evaluaciones

DruBot: Robotic prototype for authentication and comparison of facial proportions for assistance control and impersonation detection in evaluations

Neisser Ale[1], Abigail Huisacayna[1], Tereza Yallico[1]*, Marks Calderón[1]

[1] *Ingeniería en Tecnologías de la Información y Sistemas, Universidad ESAN, Jr. Alonso de Molina 1652, Santiago de Surco, Lima, Perú*

Recibido (Recieved): 16/10/2018 Aceptado (Accepted): 06/06/2019

RESUMEN

Este trabajo describe el desarrollo del prototipo robótico llamado DruBot que busca reconocer los rostros de las personas que ingresan a un aula en específico, un área privada o un examen, comparándolos con una base de datos para cada caso (para distinguirlos a partir de las características extraídas de la foto del carnet universitario y los frames obtenidos del video de bienvenida de cada estudiante) y determinar si la persona que se ve a través de la cámara tiene o no acceso al área, emitiendo una señal distinta si se le permite el ingreso o no. Aplicamos técnicas de visión artificial (Haar cascade para la detección de rostros en toda la imagen capturada por la cámara en tiempo real y Face Landmarks para encontrar los puntos clave del rostro humano detectado, calcular sus proporciones a partir de distancias euclidianas y comparar para el reconocimiento de cada persona en específico) y comunicación serial con dispositivos electrónicos a fin de que los presentes noten cuándo hay un intruso o cuándo ha reconocido bien a alguien para tomarle asistencia.

Palabras Clave: Face Landmarks, Haar Cascade, Distancia Euclidiana, Autenticación, Visión Artificial

ABSTRACT

The work 'DruBot: Robotic prototype for authentication and comparison of facial proportions for assistance control and impersonation detection in evaluations' describes the development of the robotic prototype called DruBot that seeks to recognize the faces of the persons who join to a classroom specific, a private area or an examination, comparing them with a database for each case (to distinguish them from the characteristics extracted from the photo of the university identification and the frames obtained of the video of welcome of every student) and to determine if the image of the person which camera is capturing has or hasn't access to the area, issuing a different sign if his or her access is allowed or not. We apply technologies of artificial vision (Haar cascade for the detection of faces in the whole image captured by camera in real time and Face Landmarks to find the key points of human detected face, to calculate his proportions with Euclidean distances and to compare for the recognition of every person in specific) and serial communication with electronic devices so that the presents notice when there is an intruder or when the student has been recognized well and register his or her assistance.

Keywords: Face Landmarks, Haar Cascade, Euclidean Distance, Authentication, Computational Vision

1. INTRODUCCIÓN

Una de las capacidades más desarrolladas de los seres humanos es reconocer personas u objetos a pesar de las circunstancias, fondo e iluminación. Sin embargo, esta tarea que puede parecer tan sencilla para nosotros, es bastante más complicada al querer ejecutarla desde una computadora. [1] Esto debido a que, el algoritmo necesita ser lo suficientemente entrenado para primero poder detectar un rostro humano y en base a este, usando diversas técnicas computacionales, predecir la identidad de una persona. El que un robot pueda reconocer personas como miembros de un grupo predefinido puede servir como una forma rápida de verificación de acceso a lugares donde es necesario

* Corresponding author:
E-mail: 15101410@ue.edu.pe
DOI: https://doi.org/10.21754/tecnia.v29i1.561

clasificar a grandes cantidades de personas en poco tiempo y con un bajo porcentaje de error.

Con lo explicado, el objetivo de esta investigación es la creación de un robot que mantenga seguro el tráfico de acceso a áreas restringidas por personas no autorizadas, y así prevenir el fraude estudiantil (la suplantación de alumnos para evaluaciones) y, de paso, registrar la asistencia a clase de los estudiantes a una clase específica.

El proyecto se llevó a cabo bajo la utilización de diversas técnicas de visión artificial. Se utilizó como método de detección de rostros los clasificadores en cascada (Haar Cascade) de Viola & Jones [11] que recorren toda la imagen para encontrar patrones de manera secuencial en las facciones del rostro como ojos, cejas, nariz, boca, etc. (ver Figura 1).

Figura 1. Detección del 'rostro de Lena' usando los clasificadores en cascada de Viola & Jones

Posteriormente, usamos una técnica muy conocida para la extracción de 68 puntos faciales (Face Landmarks) [2] mediante Dlib detector [12][14], con ello, se halló las proporciones importantes del rostro. Finalmente, utilizando diversas proporciones halladas con los la librería Face_Recognition [10] en base a los puntos clave del rostro de entrada se realizará la comparación final con el método de Distancia Euclidiana (1), que se define como la lejanía entre dos puntos hallada con la fórmula basada en el teorema de Pitágoras, siendo el resultado la hipotenusa en caso de situar los puntos en solo dos ejes [13].

$$\mathrm{dist}(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (1)$$

Sin embargo, para nuestro cálculo se tomará en cuenta cada proporción facial del individuo comparada con la de otros sujetos en la base de datos en varias dimensiones. El que tenga menor distancia total resultará ser el individuo más parecido al evaluado, y de acuerdo a un factor de tolerancia (0.5) se le permitirá, o no, el acceso al área específica.

El presente trabajo de investigación está dividido de la siguiente forma. En la Sección 2, mostraremos un recuento de la literatura, explicando proyectos pasados y técnicas muy importantes para llevar a cabo esta investigación. Posteriormente, en la Sección 3, detallaremos la metodología a utilizar. En seguida, mostraremos los resultados y experimentación de nuestra investigación. Finalmente, en la última sección daremos nuestras conclusiones.

2. ESTADO DEL ARTE

Un objetivo clave de los investigadores de visión artificial es crear sistemas de reconocimientos faciales automatizados que puedan igualar, y eventualmente superar, el desempeño humano [3]. Para este fin, es importante conocer los hallazgos clave de los estudios experimentales de reconocimiento de rostros humanos, tales como: reconocimiento de expresiones faciales, reconocimiento de la posición del rostro, características importantes de los rostros, etc. [4][5]

Los sistemas robustos de reconocimiento de rostros actualmente se usan en proyectos relacionados con la seguridad, interacción humano-computadora, etc. En el diseño de los sistemas de reconocimiento facial, se deben tener en cuenta al menos tres tareas:

- *Verificación.* Un sistema de reconocimiento determina si la persona fotografiada en una imagen de rostro coincide con una identidad reclamada.
- *Identificación.* Un sistema de reconocimiento determina la identidad de una persona con una imagen de la cara.
- *Lista de observación.* Un sistema de reconocimiento determina si la persona en una imagen de rostro aparece en una lista de observación y, de ser así, identifica a esa persona. [6]

Muchos sistemas de seguridad trabajan bajo estos 3 focos mencionados, que proveen formas distintas de seguridad. Sin embargo, todos cumplen el mismo objetivo de hacer comparaciones entre rostros y detectar un criterio de identidad entre una imagen de entrada y una imagen almacenada.

Una técnica muy popular y de fácil implementación para la detección de rostro es el uso de Haar Cascade [7][11]. También existen diferentes métodos relacionados verificación e identificación de rostros. El esquema de reconocimiento facial basado en modelos tiene como objetivo construir un modelo del rostro humano, que sea capaz de capturar las variaciones faciales. [5]

El conocimiento previo del rostro humano se u mucho para diseñar el modelo. Por ejemplo, la coincidencia basada en características deriva características de distancia y posición relativa de la colocación de elementos faciales internos (ojos, nariz, cejas, etc). En visión artificial, para el reconocimiento de rostros basados en imágenes, existen métodos como: PCA, LDA, KPCA, Elastic Bunch Graph, etc. [8].

Por otro lado, es necesario en muchos casos obtener características importantes en los rostros de las personas, para hacer una posible comparación o detección de alguna emoción. Una de las técnicas que trabajan en ello es Face Landmarks o puntos faciales del rostro (ver Figura 2). Los puntos de referencia faciales son un conjunto de puntos sobresalientes, generalmente ubicados en las esquinas, puntas o puntos medios de los componentes faciales como: ojos, cejas, boca, nariz, oreja, entro otros [2].

Los puntos de referencias faciales confiables y sus algoritmos de detección y seguimiento asociados se pueden usar ampliamente para representar las características visuales importantes para el reconocimiento facial, el reconocimiento de expresiones y la discriminación de rostros. La detección automática de puntos de referencia en imágenes fijas es útil en muchas tareas de visión artificial en las que se necesita reconocimiento de objetos o determinación de la pose con alta confiabilidad. Su objetivo es facilitar la localización de correspondencia de puntos entre imágenes y un modelo conocido con las características naturales [9].

Figura 2. Obtención de Face Landmarks en diferentes posiciones del rostro

3. METODOLOGÍA

Desarrollo del algoritmo de reconocimiento facial

Los rasgos faciales de las personas tienden a cambiar en el tiempo debido a factores como la edad o marcados cambios de peso (ver Figura 3 y 4).

Figura 3. Cambio facial por edad. Tejido óseo facial perdurable

Figura 4. Cambio facial por peso. Tejidos blandos no perdurables

Sin embargo, según la ciencia médica, estos cambios suelen ser más pronunciados en el tejido blando del cuerpo humano (reservas lípidas faciales, tersura o flacidez de la piel, tono muscular, contracción de tendones) que en el óseo o estructural, que casi no cambia en el tiempo por tratarse de las proporciones entre los rasgos del cráneo. En el desarrollo del algoritmo tuvimos en cuenta las proporciones perdurables en el tiempo.

Para el desarrollo del proyecto utilizamos los face landmarks (puntos clave faciales) facilitados por la librería Dlib (ver Figura 5), y calculamos las distancias euclidianas entre las proporciones faciales capturadas de ambos tipos de rasgos (duraderos y no duraderos) pero con un umbral de tolerancia que asegura que a pesar de que la persona haya cambiado, el clasificador apoyándose en los duraderos y habiendo sido entrenado con gestos de la misma persona sea capaz de predecir quién es.

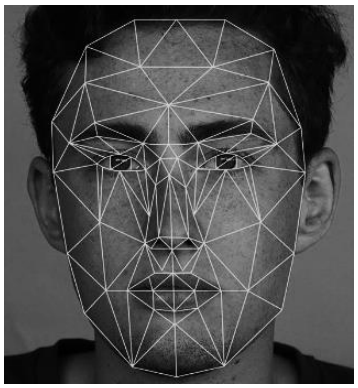

Figura 5. Proporciones de referencia para la comparación facial

El algoritmo de reconocimiento facial primero capta la imagen de prueba a través de la cámara web instalada en el DruBot, luego extrae los Face Landmarks de esa imagen y calcula las distancias euclidianas entre las proporciones faciales para darnos un vector

característico (de 128 dimensiones) que será comparado con los centroides (uno por persona en la data) de los vectores característicos de todas las imágenes de una misma persona, que fueron obtenidos en la fase de entrenamiento (ver Figura 6), a través de varios frames obtenidos del video de cada sujeto, donde se les capturó hablando y gesticulando, colaborando con que sea un algoritmo resistente a estas situaciones) de la data.

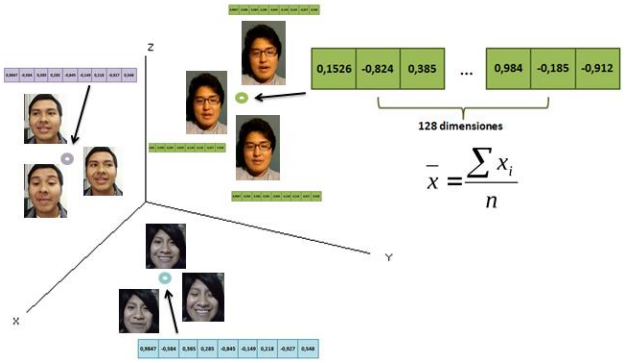

Figura 6. Centroides de los vectores característicos por persona

La comparación se hace a través de distancias euclidianas y la que obtenga menor resultado será la persona más parecida, si de todas las distancias ninguna es inferior al umbral (0.5) se tomará como persona desconocida (ver Figura 7).

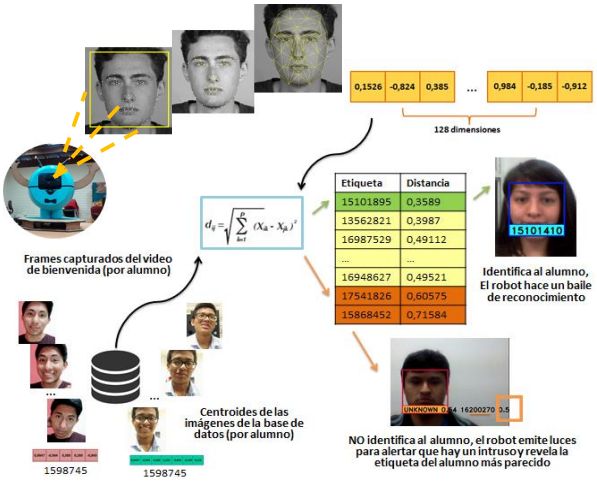

Figura 7. Proceso del algoritmo de reconocimiento facial

Diseño y construcción del prototipo: Las piezas fueron diseñadas en Autodesk Inventor y luego impresas en 3D usando PLA (Ácido Poli-láctico, filamento de impresión 3d que no emite gases nocivos en su proceso y es relativamente ecológico al estar hecho a base de plantas como el maíz) para que sea resistente y le dé la estabilidad necesaria al DruBot (ver Figura 8).

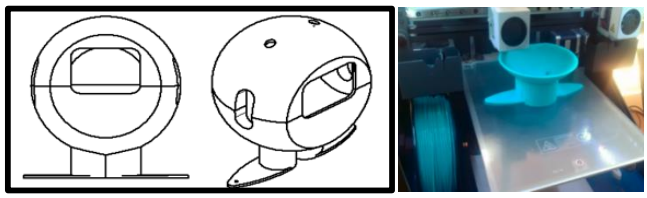

Figura 8. Plano del robot ensamblado y su impresión 3D

Otras piezas fueron hechas en MDF con una cortadora láser, algunas de estas han sido utilizadas en el soporte de los servomotores y otras en la plataforma que sostiene el robot para darle el equilibrio adecuado.

Integración electrónica: Al unir las piezas tanto cortadas con láser como impresas y tener lista la carcasa, se unió con los dispositivos electrónicos (ver Figura 9) controlados a través de una placa de desarrollo Tegra Jetson TK1, ampliamente usada para prototipos con visión artificial por ser un CPU, GPU e ISP en un mismo procesador.

Para capturar las imágenes a probar se utilizó una cámara web Logitech c525 con resolución máxima de 720p/30, conectada por USB al Jetson TK1.

Figura 9. Hardware y electrónica del proyecto

Según el funcionamiento del robot al detectarse si el alumno está en la BD, los servomotores se moverán de un lado a otro, sujetando los "brazos" del robot, simulando un pequeño "baile de reconocimiento" (agita sus brazos rítmicamente con el movimiento de los motores) que indica que tiene el acceso permitido y su asistencia será registrada, de no tenerlo se encenderán los leds de forma titilante para indicar que esa persona no puede ingresar al aula por no tener permiso de acceso.

4. EXPERIMENTOS Y RESULTADOS

Diseño CAD: En el diseño del robot se optó hacerlo de una forma circular para ser agradable a la vista. Se hizo la carcasa del robot con material PLA por medio de la impresión 3D para que sea más resistente (ver Figura 10) y parte de los soportes se fabricaron en una cortadora láser (ver Figura 11).

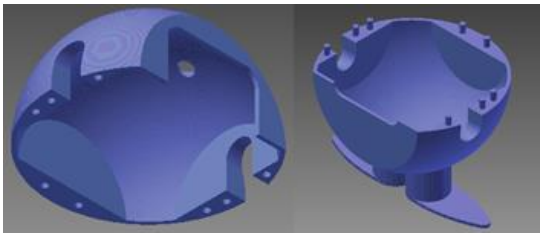

Figura 10. CAD de la tapa de la carcasa del robot

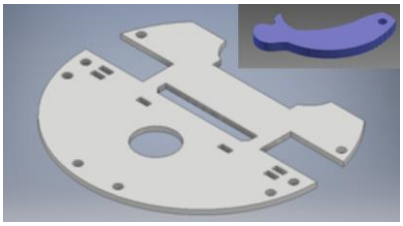

Figura 11. CAD de las piezas cortadas con láser

Pruebas: En esta etapa, se probó con frames capturados de la webcam (captados en tiempo real) para calcular las proporciones faciales y sus vectores característicos a fin de encontrar la mayor cercanía que exista entre este y alguno de los centroides de la base de datos (ya obtenidos en la fase de entrenamiento). A fin de saber si la persona captada tiene acceso o no al área restringida y notificar a través de los actuadores Se observaron tiempos de respuesta rápidos (casi inmediatos a pesar de que el alumno hacía gestos), en las pruebas se imprimieron en pantalla el código del alumno más parecido al que la cámara capta en el momento (ver Figura 12) y notificando si es que no es reconocido por el clasificador al no encontrarse en la base de datos de alumnos con acceso al área (ver Figura 13)

Figura 12 - Pruebas Finales con el algoritmo de reconocimiento facial

Figura 13. Funcionamiento del prototipo robótico

A continuación se muestran los resultados obtenidos en la fase de prueba (ver Tabla 1), llevados a cabo con videos de alumnos hechos por ellos mismos (videos diferentes a los de entrenamiento) o capturas de la Web-cam en tiempo real.

TABLA 1. Experimentos y resultados de las pruebas

Descripción de las condiciones en los experimentos	
Cantidad de sujetos en el Dataset de entrenamiento	135 sujetos
Cantidad de imágenes por sujeto en la fase de entrenamiento	20 a 30 imágenes capturadas de los videos
Resultados de las pruebas	
Tiempo de respuesta	0.5 segundos
Distancia entre la persona evaluada y la cámara del prototipo robótico	40 centímetros aproximadamente
Accuracy obtenido o porcentaje de acierto del detector de rostros unido al clasificador de identidad	0.96 equivalente a 96% de acierto

CONCLUSIONES Y TRABAJOS A FUTURO

El reconocimiento visual suele verse afectado por factores como la iluminación de las imágenes de prueba usualmente distinta a las de referencia, la resolución de la cámara con la que se capta las imágenes, la postura de la persona y las expresiones faciales. Las primeras pruebas que hicimos se vieron distorsionadas por estos factores hasta que los tuvimos más en cuenta al capturar las imágenes de prueba con la cámara web.

Se recomienda, al probar algoritmos de visión computacional tener en cuenta que la luz puede aclarar o difuminar las líneas que tomará el algoritmo como contornos y que la comparación será más certera cuando se tenga un ángulo parecido de iluminación, tampoco debe olvidarse que las emociones o gestos hacen que el rostro se vea distinto aunque sea la misma

persona, se recomienda probar con una expresión serena y mirar directo a la cámara con postura erguida.

El algoritmo de visión artificial se basa en proporciones faciales, tuvimos que hacer varias pruebas para hallar las más significativas y hacer comparaciones certeras, mientras más distancias euclidianas de los puntos perdurables en el tiempo se incluían, más preciso se volvía el algoritmo.

A pesar de nuestros avances, actualmente las redes neuronales artificiales están teniendo una gran participación en la visión computacional, a futuro sería bueno incluirlas en la autenticación facial e incluso podríamos incluir modelos de Deep Learning tales como Openface [14] y Deepface [15] los cuales están basados en arquitectura de redes convolucionales y son invariantes a la posición del rostro. Esto con el fin de conseguir mayor precisión en el reconocimiento de las personas.

Actualmente el robot capta la imagen con la cámara web y la compara con la base de datos de personas con acceso al área restringida a fin de ver si la persona se encuentra o no en la base de fotos de gente con acceso. Actualmente captamos en tiempo real a todas las personas que se vean en la captura y le hace el procesamiento a cada una. Queremos implementar que guarde la foto de las personas que han sido capturadas por la toma pero que no han sido reconocidas como personas con acceso al lugar, algo así como un "repositorio de intrusos" junto con la hora de ingreso y su estado como admitido o no.

AGRADECIMIENTOS

A nuestras madres por su eterno apoyo sin importar la distancia, a los docentes y a nuestra alma mater que nos motiva a crecer.

REFERENCIAS

[1]Learned, R 2011 Introduction to Computer Vision. Department of Computer Science University of Massachusetts, Amherst, MA 01003
[2]Rathod D, Shylaja V & Natarajan S 2014 Facial Landmark Localization - A Literature Survey INPRESSCO Karnataka, India pp 1901-1907
[3]Sucar, R & Gómez, G 2011 Vision Computacional. México
[4]Sinha P, Balas B, Ostrovsky Y & Russell R 2006 Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About IEEE Computer Society pp 1948-1962
[5]Szeliski R 2010 Computer Vision: Algorithms and Applications Springer
[6]Chellappa R, Sinha P & Phillips J 2010 Face recognition by computers and humans *IEEE Computer Society* pp 46-55
[7]Wilson, P & Fernandez, J 2006 Facial Feature Detection Using Haar Classifiers JCSC
[8]Xiaoguang Lu 2004 Image Analysis for Face Recognition *IEEE Computer Society China*
[9]Tie Y & Guan L 2013 Automatic landmark point detection and tracking for human facial expressions. *EURASIP Journal*
[10]Ageitgey A 2017 Face_recognition package. Recuperado de: "http://cort.as/-JHLX"
[11]Viola P & Jones M 2001 Rapid object detection using a boosted cascade of simple features. CVPR (1) pp 511-518
[12]King D 2009 Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research* pp 1755-1758
[13]Danielsson P 1980 Euclidean distance mapping. *Computer Graphics and image processing* pp 227-248
[14]Amos B, Ludwiczuk B & Satyanarayanan M 2016 Openface: A general-purpose face recognition library with mobile aplication. CMU School of Computer Science
[15]Taigman Y, Yang, M, Ranzato, M & Wolf, L 2014 Deepface: Closing the gap to human-level performance in face verfication. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 1701-1708