



Tipo de artículo: Artículos originales  
Temática: Inteligencia artificial  
Recibido: 20/08/2022 | Aceptado: 24/09/2021 | Publicado: 30/09/2022

Identificadores persistentes:  
ARK: [ark:/42411/s9/a74](https://nbn-resolving.org/urn:nbn:org:ark:42411/s9/a74)  
PURL: [42411/s9/a74](https://nbn-resolving.org/urn:nbn:org:ark:42411/s9/a74)

# Sistema de identificación de emociones a través de reconocimiento facial utilizando inteligencia artificial

## *Emotion identification system through facial recognition using artificial intelligence*

Alexandra Paricela Canazas<sup>1</sup>, Johnnathan Jimmy Ramos Blaz<sup>2</sup>, Patricio Dante Torres Martínez<sup>3</sup>, Xiomara Jaquehua Mamani<sup>4</sup>

<sup>1</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [aparicelac@unsa.edu.pe](mailto:aparicelac@unsa.edu.pe)

<sup>2</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [jramosb@unsa.edu.pe](mailto:jramosb@unsa.edu.pe)

<sup>3</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [ptorresmar@unsa.edu.pe](mailto:ptorresmar@unsa.edu.pe)

<sup>4</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [xjaquehua@unsa.edu.pe](mailto:xjaquehua@unsa.edu.pe)

\* Autor para correspondencia: [aparicelac@unsa.edu.pe](mailto:aparicelac@unsa.edu.pe)

---

### Resumen

El presente artículo tiene como principal objetivo el desarrollo de un sistema que permita identificar las emociones de una persona mediante el reconocimiento de rostros utilizando inteligencia artificial. Para el desarrollo del sistema se tuvo como base el algoritmo básico de Eigenfaces o Análisis de Componente Principal, el cual es uno de los modelos de reconocimiento de rostros más utilizados. Así mismo fue utilizado el lenguaje Python y algunas de sus librerías disponibles como Numpy, OpenCV y Sklearn para la implementación.

**Palabras clave:** Expresiones faciales; Emociones; Visión por computadora; Aprendizaje automático; Inteligencia artificial.

### Abstract

*The main objective of this paper is to develop a system to identify the emotions in a person using face recognition using artificial intelligence. The system development was based on the basic algorithm of Eigenfaces or Principal Component Analysis, one of the most widely used face recognition models. In addition, Python language and some of its available libraries such as Numpy, OpenCV and Sklearn were used for the implementation.*

**Keywords:** Facial expressions; Emotions; Computer Vision; Machine Learning; Artificial Intelligence.

---

## Introducción

La influencia de las emociones es crucial en el proceso de aprendizaje, y es además, inherente al lenguaje y las formas empleadas para la transmisión de la información; por consiguiente, es importante comprender cómo es que las emociones modifican el proceso de aprendizaje y cómo a partir de esto se diseñan modelos que buscan optimizar los logros de aprendizaje. Para esto se consideran cinco factores o estímulos: fisiológicos (auditivo, visual, táctil y cenestésico), sociológicos, ambientales, psicológicos y emocionales que afectan el aprendizaje de un individuo [1]. Así, podemos destacar aspectos como la personalidad, el estado de ánimo y las emociones que juegan un papel importante en los entornos de aprendizaje.

Por otra parte, también es claro que los niños recurren a sus emociones con el fin de dar sentido a su pertenencia al participar en el proceso de aprendizaje; y es este sentido de pertenencia el que define la forma en la que encuentran significativa su participación [2]; algo muy asociado a la idea de libre elección y expresión emocional.

A pesar de esto, dado el caso, puede darse que no exista relación entre las emociones y la ganancia proporcional de aprendizaje; sin embargo, una significativa correlación entre las emociones y la precisión de algunos procesos cognitivos y metacognitivos [3].

Enfocado en la psicología: En los últimos años y con el avance de la tecnología se han dado en mayor parte las consultas a psicólogos o tutores privados por internet u online sin embargo el desempeño psicológico de un estudiante puede llegar a verse afectado por el sobrecargo de tareas o la escuela para notar estos rasgos es necesario un análisis completo de su comunicación no verbal como lo que expresa los gestos que muestra [4]. Para ayudar en este ámbito tanto a los tutores como psicólogos u otro oficio relacionado se plantea una IA que permite entrenar un sistema en el que a través de cámaras se registre los gestos fáciles y a través de esta se identifiquen las emociones de los estudiantes en instituciones educativas [5].

Comúnmente los estudiantes pueden tener problemas familiares, sociales, personales que lo pueden afectar mental y físicamente; teniendo esto en cuenta los profesores/encargados no notarían siempre, o que obvian esto, sin darse cuenta de las consecuencias psicológicas que pueden ocasionar a lo largo de los años. Por ello se elaboraría un sistema que de precaución con conocimientos similares a los que poseen los humanos. [6] ante las expresiones de los estudiantes que será clasificado con las expresiones faciales prototípicas básicas reconocidas universalmente. Estas son: enojo, disgusto, miedo, felicidad, tristeza y sorpresa. [7].

En este artículo se tratará de abordar un sistema enfocado a base en reconocimiento facial buscando los patrones y características de los rasgos faciales captados dentro de la imagen, para que esta sea recortada solo al contorno de la cara [8], entre ellas se debe hacer una identificación y análisis de cada tipo de emociones con su intensidad para así calcular el porcentaje y la intensidad de cada emoción [9] [10]

## **Materiales y métodos o Metodología computacional**

### Estado del Arte

#### **A. Reconocimiento facial con deep learning y Python**

El presente trabajo [6] describe e implementa los pasos necesarios para crear un programa capaz de identificar a las personas presentes en una imagen, video o webcam haciendo uso de las librerías disponibles en python y aplicando los conceptos de Deep Learning y redes neuronales.

#### **B. Detección de emociones y reconocimiento facial utilizando aprendizaje profundo**

El presente trabajo de titulación [4] tiene como finalidad implementar un sistema de detección de emociones y reconocimiento facial a través del aprendizaje profundo, específicamente utilizando la técnica de redes neuronales convolucionales y la librería face recognition.

De este modo se crea una arquitectura de aprendizaje profundo con dataset de Kaggle denominado “Learn facial expressions from an image” con el algoritmo Frontal Face elaborada en Django.

**C. Uso de patrones de reconocimiento de las emociones para apoyar la didáctica de enseñanza aprendizaje**

El artículo [5] presenta el desarrollo de una investigación aplicada en el ámbito de la teoría educativa apoyado por procesos en comunicación, con el fin de permitir el mejoramiento de las estrategias de enseñanza en el ámbito superior técnico profesional o nivel terciario en educación, considerando el uso de patrones de reconocimiento para capturar las emociones según el registro de reacciones faciales haciendo uso de los algoritmos “Eigenfaces” y “Fisherfaces” y a través de ellos aplicaciones de filtros para regular las distorsiones de la digitalización.

**D. Diseño de un Sistema de Reconocimiento de rostros aplicando inteligencia y visión artificial**

El presente artículo [9] se centra en el proceso de diseño, desarrollo e implementación de un sistema de reconocimiento de rostros mediante la hibridación de técnicas de reconocimiento de patrones, visión artificial e inteligencia artificial con el uso de redes neuronales. Así mismo es recopilado el producto de la unión de las técnicas de visión artificial y las técnicas de inteligencia artificial y sus implicaciones en múltiples aplicaciones tales como el control de robots de interacción social.<sup>1</sup>

**E. Sistema de reconocimiento de gestos faciales captados a través de cámaras para analizar el nivel de satisfacción de clientes en restaurantes**

La presente investigación [8] tiene como principal objetivo el desarrollo de un sistema que reconozca la satisfacción o insatisfacción de un cliente en un restaurante con base en los gestos que este mismo realiza al momento de recibir el servicio brindado por el establecimiento. En la etapa de preprocesamiento de los datos se hizo uso de una máquina de soporte vectorial como clasificador, a su vez el histograma de gradientes fue usado en la detección de rostro dentro de la imagen.

## **Fundamentación Teórica**

### **Reconocimiento Facial**

Para el desarrollo del proyecto, centrado en el reconocimiento de emociones, era indispensable la consideración de un sistema de reconocimiento facial, y más específicamente, el algoritmo seleccionado para esto fue el de Eigenfaces, uno de los modelos de reconocimiento de rostros más utilizado, el cual está basado en las propiedades matemáticas de la imagen digitalizada, que captura características invariantes de los rostros. No fue hasta la década de 1960 que se desarrolló el primer sistema semiautomático de reconocimiento facial basado en un método mediante el cual el observador localizaba rasgos faciales en fotografías, a partir de esto, se buscó el cálculo de distancias y proporciones con el fin de realizar comparaciones. Ya en la década de 1970, se propuso el uso de características específicas para el reconocimiento facial y posteriormente en 1988, Sirvoich y Kirby propusieron el método de análisis de componentes principales (PCA), dando inicio a la mejora en este ámbito. Es por esto que actualmente se cuenta con diversas técnicas para el procesamiento de imágenes y el reconocimiento de patrones, convirtiéndose en un área de investigación activa [11].

### **Reducción de Dimensionalidad**

La reducción de dimensionalidad se refiere a técnicas o algoritmos que buscan una proyección de la representación original de muchas dimensiones de nuestros datos a una de menor dimensión, reduciendo el número de variables de entrada en los datos de entrenamiento. Es decir, buscan una superficie de baja dimensión, incrustada en el espacio de datos de alta dimensión, proyectando los datos a un subespacio dimensional más bajo que capturan lo más importante de los datos, de modo que al tratar menos dimensiones también se trabaje con menos parámetros, simplificando el modelo de aprendizaje automático [12]. Es por esto que la reducción de la dimensionalidad no implica pérdida de información; sin embargo, pueden emplearse representaciones con pérdida y a pesar de esto obtener un buen control, lo que requerirá un equilibrio cuidadoso entre preservar aspectos importantes de las distribuciones y usar la menor cantidad de dimensiones posible.

## **Análisis de Componentes Principales**

Una de las técnicas de reducción de dimensionalidad más comunes es el Análisis de Componentes Principales, en el que, dado un conjunto de datos, se busca la representación lineal de menor dimensión de los datos conservando la varianza de los datos reconstruidos. De esta forma, esta técnica hace uso del hiperplano de baja dimensión tal que, al proyectar los datos en este hiperplano, los cambios en la varianza de los datos sean mínima [13]. Así, aplicando PCA, el vector de primera base,  $u_1$ , apunta en la dirección en la que los datos tienen la mayor varianza. En otras palabras, las proyecciones de los puntos de datos sobre  $u_1$ , dadas por  $u_1^T x_1, \dots, u_1^T x_N$ , tienen la varianza muestral más grande posible, para cualquier elección del vector unitario  $u_1$ . El segundo vector base  $u_2$ , apunta en la dirección de máxima varianza sujeto a la restricción de que  $u_2$  sea ortogonal a  $u_1$  (es decir,  $u_2^T u_1 = 0$ ). En general, el  $i$ -ésimo vector base, también llamado  $i$ -ésimo componente principal, es la dirección de máxima varianza que es ortogonal a los  $i - 1$  componentes principales anteriores.

## **Eigenfaces**

En 1991, Turk y Pentland propusieron el enfoque Eigenfaces para el reconocimiento de rostros que utiliza conceptos de álgebra lineal y reducción de dimensionalidad para reconocer rostros, tras los fundamentos planteados por Sirovich y Kirby en 1987 en su artículo Procedimiento de baja dimensión para la caracterización de rostros humanos, convirtiéndose en un enfoque fundamental en la historia de la visión computacional.

En este enfoque, al ingresar un conjunto de datos de  $Z$  imágenes de rostros es necesario que cada rostro se represente en escala de grises como un mapa de bits  $N \times N$  de los píxeles. Es necesario formar un solo vector a partir de la imagen, y esto se consigue simplificando cada una en un vector de características de intensidades de píxeles sin procesar; posteriormente se concatenan todas las filas juntas, formando una única y larga matriz de intensidades de píxeles en escala de grises de  $Z$  filas y  $N^2$  columnas. A partir de esa matriz se calcula la media de cada columna en la matriz, obteniendo el valor de intensidad de píxel promedio para cada

coordinada en el conjunto de datos de la imagen. Es necesario centrar los datos en la media restando está a cada columna para obtener la matriz de covarianza y calcular los valores propios y los vectores propios de la matriz de covarianza.

A partir de los valores propios más grandes, se calculan los vectores propios correspondientes y se obtiene un lapso dimensional. Haciendo uso de las caras de entrenamiento normalizadas se representa cada vector de rostro en la combinación lineal de los mejores vectores propios [14].

### **Herramientas y Elementos**

**Google Colab** Es una herramienta para escribir y ejecutar código Python en la nube de Google. También es posible incluir texto enriquecido, links e imágenes. En caso de necesitar altas prestaciones de cómputo, el entorno permite configurar algunas propiedades del equipo sobre el que se ejecuta el código. El uso de Google Colab permite disponer de un entorno para llevar a cabo tareas que serían difíciles de realizar en un equipo personal. [15].

**Google Drive** Es un servicio de almacenamiento de datos que son guardados en la nube. Permite copiar archivos desde el ordenador para que sean guardados en la nube [16]. Para el presente trabajo se hará uso del mismo para almacenar las carpetas de nuestro dataset para el entrenamiento de la IA, así mismo las carpetas para la validación y resultados al poner a prueba el reconocedor.

**Python** Es un lenguaje de programación potente y fácil de aprender. Tiene estructuras de datos de alto niveles eficientes y un simple pero efectivo sistema de programación orientado a objetos [17]. Por lo tanto, es ampliamente utilizado en las aplicaciones web, el desarrollo de software, la ciencia de datos y el machine learning (ML). Las librerías utilizadas de este lenguaje en el presente proyecto fueron las siguientes:

**Numpy** NumPy es una librería de Python especializada en el cálculo numérico y el análisis de datos, especialmente para un gran volumen de datos. Es utilizada para trabajar con matrices

permitiendo representar colecciones de datos de un mismo tipo en varias dimensiones, y funciones muy eficientes para su manipulación. NumPy significa Python numérico. [18]

**Os** El módulo OS en Python proporciona funciones para interactuar con el sistema operativo. OS viene bajo los módulos de utilidad estándar de Python. Este módulo proporciona una forma portátil de usar la funcionalidad dependiente del sistema operativo. [19]

**OpenCV-Python** OpenCV-Python es una biblioteca de enlaces de Python diseñada para resolver problemas de visión por computadora. Además, es compatible con las plataformas más utilizadas, Windows, Mac OS y Linux [20].

**Sklearn** Scikit-learn (Sklearn) es la biblioteca más útil y sólida para el aprendizaje automático en Python. Proporciona una selección de herramientas eficientes para el aprendizaje automático y el modelado estadístico, incluida la clasificación, la regresión, el agrupamiento y la reducción de la dimensionalidad a través de una interfaz de consistencia en Python. [21]

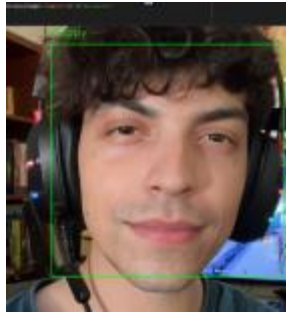
## Información del Dataset

El dataset que se consideró utilizar para el entrenamiento de la IA se encuentra en kaggle con una cantidad de 28 821 imágenes cada una por defecto separada en carpetas con su emoción distinta, estas imágenes ya cuentan con la escala de grises y la medida sugerida para la detección de rostros. Para el caso del entrenamiento se hizo el cambio de unir todas las imágenes separadas en carpetas en una sola, y con esta por sí sola sea el entrenamiento para el reconocedor, y que este identifique por sí mismo los rasgos considerados comunes y con ello distinguiera la diferencia y similitud entre estos para el reconocimiento de próximas imágenes.

```
Precisión 0.2777777777777778  
0.42857142857142855  
0.5  
  
Recall 0.2777777777777778  
0.42857142857142855  
0.42857142857142855  
  
F1 Score 0.25925925925925924  
0.42857142857142855  
0.42857142857142855  
  
Accuracy 0.42857142857142855
```

Figura 1. Resultado





*Figura 2. Resultado- 2*

## **Resultados y discusión**

La calidad del cálculo de la precisión contando el total de verdaderos positivos, falsos negativos y falsos positivos es 0.7142857142857143. Por lo que se interpretaría que un 43 % de los casos en el que se equivocaría el reconocedor.

La cantidad que el reconocedor es capaz de identificar se encuentra el cálculo de forma global  $Recall = 0.5714285714285714$  En el rendimiento promediado de la precisión y la exhaustividad entre varias soluciones del F1 score comparar el rendimiento combinado de la precisión y la exhaustividad entre varias soluciones 0.619047619047619, indicando que el rendimiento del reconocedor.

Como ultimo resultado de la Exactitud (Accuracy) mide el porcentaje de casos que el reconocedor ha acertado, en este caso el resultado es 0.5714285714285714.

## **Conclusiones**

El reconocedor al tener un entrenamiento tiene una mayor identificación de rostros de diferentes medidas y calidades. El lenguaje de programación Python da bastante acceso con las librerías a la elaboración del entrenamiento y del reconocedor de detección de rostros. Las características faciales tienen mucha similitud al momento de la identificación de emociones, por lo que se no se pueden basar con una sola característica para la identificación correcta de la emoción. El entrenamiento debe ser específico según el objetivo que se

quiera llegar, para cada acción específica que se quiera conseguir, se necesita otro entrenamiento enfocado en ese objetivo. El sistema tuvo en total una exactitud de aciertos ligeramente positiva por el entrenamiento que no fue muy eficiente para el modelo de Eigen Face.

## Referencias

- [1] J. S. P. Doulik and I. Simonova, Learning Styles in the e-Learning Environment: The Approaches and Research on Longitudinal Changes., 2017.
- [2] S. Frankel and M. Mountford, In search of meaningful participation: Making connections between emotions and learning., 2021.
- [3] R. R. E. B. C. G. B. M. Taub, R. Azevedo and M. J. Price, How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system?, 2021.
- [4] J. Romero, Detección de Emociones y Reconocimiento Facial utilizando aprendizaje profundo, 2020.
- [5] J. L. Mancilla Monsalve, Uso de patrones de reconocimiento de las emociones para apoyar la didáctica de enseñanza aprendizaje., 2019.
- [6] S. Roy, Face emotion recognition with EfficientNetB2, 2021.
- [7] E. Jesús, Detección de Emociones del Usuario, 2014.
- [8] E. A. Lara, L. Código, H. Alejandro, Q. Cruz, and E. Lara Lévano, Sistema de reconocimiento de gestos faciales captados a través de cámaras para analizar el nivel de satisfacción de clientes en restaurantes., 2019.
- [9] G. O. and S. O., Diseño de un Sistema de Reconocimiento de rostros aplicando inteligencia y visión artificial., 2014.
- [10] S. Roy., Face emotion recognition with EfficientNetB2., 2021.
- [11] P. Kaur, K. Krishan, S. K. Sharma, and T. Kanchan, Facial-recognition algorithms: A literature review. [Online]. Available: <https://doi.org/10.1177/0025802419893168>

- [12] M. Collins and S. Robert, A Generalization of Principal Component Analysis to the Exponential Family. [Online]. Available: [https://www.researchgate.net/publication/2407485 A Generalization of Principal Component Analysis to the Exponential Family](https://www.researchgate.net/publication/2407485_A_Generalization_of_Principal_Component_Analysis_to_the_Exponential_Family)
- [13] I. T. Jolliffe, Principal Component Analysis. [Online]. Available: <https://doi.org/10.1007/b98835>
- [14] M. Turk and A. Pentland, Eigenfaces for Recognition. [Online]. Available: <https://doi.org/10.1162/jocn.1991.3.1.71>
- [15] G. L. Baume, Breve introducci3n a Google Colab., 2021.
- [16] J. M. Uriarte, Google Drive., 2020.
- [17] P. S. Foundation, El tutorial de Python., 2022.
- [18] La librería Numpy. [Online]. Available: <https://aprendeconalf.es/docencia/python/manual/numpy/>
- [19] GeeksforGeeks, M3dulo OS en Python con ejemplos., 2022.
- [20] A. Mordvintsev and A. R. K., Tutoriales de Introducci3n a OpenCV-Python., 2022.
- [21] Scikit, Tutorial de aprendizaje de Scikit. [Online]. Available: <https://www.tutorialspoint.com/scikit-learn/index.htm>