



**RECONOCIMIENTO AUTOMÁTICO DE EMOCIONES FACIALES
PLANTEADAS POR EKMAN PARA APOYAR EL DIAGNÓSTICO
O TRATAMIENTO DE LA DEPRESIÓN EN ADULTOS JÓVENES**

TRABAJO DE GRADO

**PAULA ANDREA ACOSTA SÁNCHEZ
PAULA XIMENA BALLESTEROS BALLESTEROS**

**PROGRAMA DE INGENIERÍA DE SISTEMAS
FACULTAD DE INGENIERÍA
UNIVERSIDAD CATÓLICA DE COLOMBIA BOGOTÁ, MAYO
2022**

RECONOCIMIENTO AUTOMÁTICO DE EMOCIONES FACIALES
PLANTEADAS POR EKMAN PARA APOYAR EL DIAGNÓSTICO O
TRATAMIENTO DE LA DEPRESIÓN EN ADULTOS JÓVENES

PAULA ANDREA ACOSTA SÁNCHEZ PAULA
XIMENA BALLESTEROS BALLESTEROS

TRABAJO DE GRADO PARA OPTAR AL TÍTULO DE:
INGENIERO DE SISTEMAS Y COMPUTACIÓN

TUTOR: JUAN CARLOS BARRERO CALIXTO
jcbarrero@ucatolica.edu.co

CO-TUTORA: VERÓNICA CASTRO MUNAR
vcastro@ucatolica.edu.co

ALTERNATIVA: TRABAJO DE INVESTIGACIÓN TECNOLÓGICA

PROGRAMA DE INGENIERÍA DE SISTEMAS
FACULTAD DE INGENIERÍA
UNIVERSIDAD CATÓLICA DE COLOMBIA
BOGOTÁ, MAYO
2022



Atribución-NoComercial 4.0 Internacional (CC BY-NC 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Advertencia.](#)

Usted es libre de:

Compartir — copiar y redistribuir el material en cualquier medio o formato

Adaptar — remezclar, transformar y construir a partir del material

La licenciente no puede revocar estas libertades en tanto usted siga los términos de la licencia

Bajo los siguientes términos:



Atribución — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciente.



NoComercial — Usted no puede hacer uso del material con [propósitos comerciales](#).

No hay restricciones adicionales — No puede aplicar términos legales ni [medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia](#).

<https://creativecommons.org/licenses/by-nc/4.0/deed.es>

NOTA DE ACEPTACIÓN

Jurado 1

Jurado 2

Juan Carlos Barrero Calixto
Tutor

Verónica Castro Munar
Co-Tutora

Bogotá, mayo 27 del 2022

AGRADECIMIENTOS

Agradecemos a los ingenieros Juan Barrero y Verónica Munar por la ayuda, acompañamiento y conocimientos brindados para la ejecución y finalización de este proyecto.

A la Universidad de Pittsburgh y a cada una de las personas que sumaron su apoyo en la construcción del conjunto de imágenes.

A la Universidad Católica de Colombia, a nuestros profesores y compañeros que fueron parte fundamental para la construcción de nuestra vida profesional.

DEDICATORIA

El presente trabajo de investigación está dedicado:

A nuestros padres Maritza Ballesteros, Roberto Ballesteros, Martha Sánchez y Arquímedes Acosta por el acompañamiento brindado a lo largo de nuestra vida, la guía, los conocimientos, consejos, el apoyo tanto emocional como económico y la paciencia, siendo unas de las motivaciones más importantes durante el proceso formativo.

A nuestras hermanas Laura Ballesteros y Carolina Acosta, a nuestras familias y a la memoria de aquellas personas que ya no están con nosotros, por brindarnos bases para nuestra formación íntegra, la motivación para completar este proceso, comprensión y escucha.

A Daniel Amaya, por su amor, paciencia, apoyo incondicional y ayuda durante todo este proceso, siendo una motivación muy importante.

A Andrés Solano, por su acompañamiento, comprensión y voz de aliento en cada una de las etapas y retos de este proceso.

TABLA DE CONTENIDO

RESUMEN	13
ABSTRACT	14
INTRODUCCIÓN	15
1. PROBLEMA DE INVESTIGACIÓN.....	17
1.1. PLANTEAMIENTO DEL PROBLEMA	17
1.2 FORMULACIÓN DE LA PREGUNTA PROBLEMA	19
1.3 SISTEMATIZACIÓN DEL PROBLEMA	19
2. JUSTIFICACIÓN.....	20
2.1 ALCANCE Y LIMITACIONES	21
3. OBJETIVOS.....	22
3.1 OBJETIVO GENERAL	22
3.2 OBJETIVOS ESPECÍFICOS	22
4. MARCO DE REFERENCIA	23
4.1 MARCO TEÓRICO	23
4.1.1 Análisis de Componentes Principales (PCA)	23
4.1.2 Análisis Discriminante Lineal (LDA)	24
4.1.2.1 Discriminante lineal de Fisher	25
4.1.2.2 Python.....	25
4.1.2.3 Keras de TensorFlow.....	27
4.1.2.4 Convolutional neural networks (CNN)	29
4.1.2.5 Tensor.....	30
4.1.2.6 Matriz de confusión.....	30
4.1.2.7 Épocas (Epoch)	33
4.2 MARCO CONCEPTUAL	33
4.2.1 Inteligencia Artificial (IA)	33
4.2.2 Machine Learning:	34
4.2.3 Deep learning:	34
4.2.4 Aprendizaje no supervisado.....	34
4.2.5 Procesamiento lenguaje natural	34
4.2.6 Dataset	35
4.2.7 EigenFace.....	35
4.2.8 Loadings	35
4.2.9Salud mental.....	35
4.2.10 Depresión	36
4.2.11 Emociones	36
4.2.12 Diagnóstico	36
4.3 ESTADO DEL ARTE	37

4.4 HERRAMIENTAS UTILIZADAS	39
4.4.1 Software.....	40
4.4.2 Hardware	42
5. METODOLOGÍA.....	43
5.1 FASE 1: RECOLECCIÓN DE DATOS	44
5.2 FASE 2: IDENTIFICAR MÉTODO.....	44
5.3 FASE 3: DESARROLLO	44
5.4 FASE 4: VALORACIÓN DEL MÉTODO.....	45
5.5 FASE 5: OBSERVACIÓN DE RESULTADOS	45
6. CONSTRUCCIÓN DEL DATASET	45
7. SELECCIÓN DE MÉTODO AUTOMÁTICO.....	50
8. DISEÑO DE LOS MÉTODOS.....	53
8.1 MÉTODO SECUENCIAL CON KERAS	53
8.1.1 Extracción de características	54
8.1.2 Clasificación.....	55
8.1.3 Implementación.....	55
8.1.4 Porcentual.....	55
8.1.5 Diagrama de barras	56
8.2 MÉTODOS CON OPENCV	56
8.2.1 Entrada	57
8.2.2 Procesamiento	57
8.2.3 Entrenamiento	57
8.3 Ajuste de los parámetros de los métodos seleccionados	59
8.3.1 Método secuencial de Keras.....	59
8.3.2 Ajuste de métodos con OpenCV	60
9. EVALUACIÓN DE MEDIDAS DE DESEMPEÑO.....	61
9.1 MÉTODO SECUENCIAL DE KERAS.....	61
9.1.1 Matriz de confusión de las pruebas	61
9.1.2 Matriz de confusión Dataset7030.....	62
9.1.3 Matriz de confusión Dataset7525.....	63
9.1.4 Matriz de confusión Dataset8020.....	63
9.2 MÉTODOS DE OPENCV	65
9.2.1 Observaciones de estudio de parámetros para clasificador Haarcascade.....	65
9.2.2 Matriz de confusión de las pruebas	67
10. RESULTADOS Y ANÁLISIS DE RESULTADOS.....	72
CONCLUSIONES	76
TRABAJO FUTURO.....	77
BIBLIOGRAFÍA.....	78

TABLA DE FIGURAS

Figura 1: Tasas de años de vida con discapacidad (por 100.000) por país para 17	17
los trastornos depresivos en 2010. 17	17
Figura 2: Fórmula BatchNormalization 28	28
Figura 3: Arquitectura de una red neuronal convolucional 30	30
Figura 4: Matriz de confusión binaria..... 31	31
Figura 5: Matriz de confusión multivariable 32	32
Figura 6: Metodología propuesta 43	43
Figura 7: Arquitectura método secuencial 53	53
Figura 8: Definición de las etapas de la arquitectura del método secuencial 54	54
Figura 9: Ejemplo visualización porcentual..... 55	55
Figura 10: Ejemplo visualización porcentual..... 56	56
Figura 12: Matriz de confusión prueba 3 con el Dataset7030 y 90 épocas 62	62
Figura 13: Matriz de confusión prueba 6 con el Dataset7525 y 90 épocas 63	63
Figura 14: Matriz de confusión prueba 9 con el Dataset8020 y 90 épocas 64	64

ÍNDICE DE TABLAS

Tabla 1. Ejemplos de fotos recibidas para el dataset.....	46
Tabla 2. Consolidado de fotos aceptadas para el dataset propio.	46
Tabla 3. Consolidado de imágenes extraídas de datasets externos	47
Tabla 4. Conformación de los 3 diferentes datasets.....	48
Tabla 5. Fotografías de personas con Gafas incluidas por categoría en el dataset 7030.....	48
Tabla 6. Fotografías de personas con Gafas incluidas por categoría en el dataset 7525.....	49
Tabla 7. Fotografías de personas con Gafas incluidas por categoría en el dataset 8020.....	49
Tabla 8. Métodos comúnmente utilizados	50
Tabla 9. Parámetros de ejecución método secuencial	59
Tabla 10. Parámetros de ejecución modelo haarcascade	60
Tabla 11. Parámetros de ejecución modelos diseñados con OPENCV.....	60
Tabla 12. Consolidado de porcentaje de precisión por cada emoción por prueba con método secuencial con Keras.....	64
Tabla 13. Tabulación de resultados según parámetros definidos para el clasificador (mínimo tamaño: 100 y máximo tamaño:1000) aplicado en la submuestra.....	65
Tabla 14. Tabulación de resultados según parámetros definidos para el clasificador (mínimo tamaño: 30 y máximo tamaño:1300) aplicado en dentro de la submuestra	66
Tabla 15. Tabulación de resultados según parámetros definidos para el clasificador (mínimo tamaño: 30 y máximo tamaño:950) aplicado en la submuestra.....	66
Tabla 16. Consolidado de exactitud por cada emoción y prueba en dataset7030 por FisherFaces.....	67
Tabla 17. Consolidado de exactitud por cada emoción y prueba en dataset7525 por FisherFaces.....	68
Tabla 18. Consolidado de exactitud por cada emoción y prueba en dataset8020 por FisherFaces.....	68
Tabla 19. Consolidado de exactitud por cada emoción y prueba en dataset7030 por EigenFaces.....	69
Tabla 20. Consolidado de exactitud por cada emoción y prueba en dataset7525 por EigenFaces.....	69
Tabla 21. Consolidado de exactitud por cada emoción y prueba en dataset8020 por EigenFaces.....	70
Tabla 22. Consolidado de exactitud por cada emoción y prueba en dataset7030 por LBPH	70

Tabla 23. Consolidado de exactitud por cada emoción y prueba en dataset7525 por LBPH	71
Tabla 24. Consolidado de exactitud por cada emoción y prueba en dataset8020 por LBPH	71
Tabla 25. Comparación de precisión entre las mejores pruebas de los métodos secuenciales de Keras y OpenCV	73
Tabla 26. Resultados obtenidos en la prueba 9 del método secuencial de Keras.....	74
Tabla 27. Tipos de visualización de detección de emociones	74

TÍTULO

Reconocimiento automático de emociones faciales planteadas por Ekman para apoyar el diagnóstico o tratamiento de la depresión en adultos jóvenes.

ALTERNATIVA

Teniendo en cuenta Acuerdo 265/2018: “Por el cual se aprueban los lineamientos y las opciones de grado para los programas académicos de la Facultad de Ingeniería de la Universidad Católica de Colombia”. Se elige la alternativa de trabajo de investigación.

LÍNEA DE INVESTIGACIÓN

Modelo de deep learning con redes neuronales convolucionales, aplicando el método Secuencial de Keras para identificación de emociones propuestas por Ekman en adultos jóvenes.

EJE TEMÁTICO

La temática principal abarca generar un modelo de aprendizaje automático que permita identificar las emociones presentadas por una persona en tiempo real para el apoyo en el diagnóstico o consultas de depresión en adultos jóvenes.

RESUMEN

La depresión en Colombia, según un boletín de la Organización Mundial de la Salud (OMS)¹, se encuentra por encima de las cifras a nivel mundial, ya que indican que en el mundo un 4,4% de la población padece de esta enfermedad, mientras que en Colombia es el 4,7%; adicional, información recopilada del mismo boletín por el periódico el Tiempo² indica que, aunque el trastorno se presenta en niños y adolescentes, tiene un nivel mayor de existencia en los adultos. Por lo que, se hace necesario el tratamiento y diagnóstico oportuno de dicho trastorno, de modo tal que sea confiable y el paciente reciba la atención necesitada de forma rápida; para ello, la Universidad Católica de Colombia, cuenta con una unidad de servicios psicológicos en donde se atienden este tipo de trastornos, sin embargo, la misma, no cuenta con ninguna herramienta tecnológica especializada que apoye a los psicólogos en este proceso.

Con el fin de dar solución a la problemática se propone el diseño de un método automático, que permite la identificación en tiempo real de las emociones planteadas por el psicólogo Paul Ekman: ira, alegría, asco, tristeza, sorpresa, miedo; con el fin de realizar un apoyo en el diagnóstico y/o tratamiento para la depresión en adultos jóvenes.

Esta opción de diagnóstico se estructuró como una alternativa no presencial, a partir del aislamiento que exigió la pandemia del Covid-19. Y se convierte en una alternativa viable, para atender los casos de depresión en menor cantidad de tiempo (Teniendo en cuenta variables como desplazamiento, disponibilidad de tiempo, etc) propiciando así un ambiente más cercano entre el psicólogo y el paciente; con el fin de evitar que la falta de tratamiento conduzca a implicaciones más graves en la salud del mismo.

Para la implementación del método, se realizó la construcción de un dataset propio, el cual se unió con datasets públicos que cumplieran con las características necesarias con el fin de aumentar la cantidad de imágenes del mismo, teniendo un total de 7920 imágenes y 1320 por cada emoción. Posteriormente, se realizó la indagación de los diferentes algoritmos de detección facial existentes, eligiendo los que más se acoplaron a los resultados esperados, con el fin de ejecutarlos con el dataset generado. Teniendo los resultados, se procedió a realizar el análisis de métricas de cada uno de los algoritmos identificando que, para esta investigación, el algoritmo secuencial de Keras, es el más óptimo en cuanto a resultados y funcionamiento, obteniendo de este modo un 62% de exactitud en el reconocimiento de emociones. Finalmente, se entrega un método que puede reconocer las emociones faciales y dar visualización de dos formas diferentes, por medio de porcentaje de acercamiento a la emoción o mediante un diagrama de barras en donde se ve presentado el índice de coincidencia con determinada emoción.

Palabras Claves:

Depresión, detección, emociones, inteligencia artificial, redes neuronales, método secuencial de Keras, matriz de confusión.

¹ ORGANIZACIÓN MUNDIAL DE LA SALUD, OMS. Depresión. [Consultado el 18 de agosto de 2021]. Disponible en: <http://www.who.int/topics/depression/es/>

² PERIÓDICO EL TIEMPO. La depresión, un mal que crece en el mundo. [Consultado el 18 de mayo de 2022]. Disponible en: <https://www.eltiempo.com/archivo/documento/CMS-16826965>

ABSTRACT

Depression in Colombia, according to a bulletin of the World Health Organization (WHO), is above the figures worldwide, indicating that in the world 4.4% of the population suffers from this disease, while in Colombia it is 4.7%; additionally, information compiled from the same bulletin by the newspaper El Tiempo indicates that although the disorder occurs in children and adolescents, it has a higher level of existence in adults. Therefore, it is necessary the timely treatment and diagnosis of this disorder, so that it is reliable and the patient receives the needed attention quickly; for this, the Catholic University of Colombia, has a unit of psychological services where this type of disorders are treated, however, it does not have any specialized technological tool to support psychologists in this process.

In order to solve the problem, the following is proposed the design of an automatic method, which allows the identification in real time of the emotions proposed by the psychologist Paul Ekman: anger, joy, disgust, sadness, surprise, fear; in order to support the diagnosis and/or treatment of depression in young adults in psychological services unit of the Catholic University of Colombia doesn't currently have any technological tool to support psychologists in these processes.

This diagnostic option was structured as a non face-to-face alternative, from the isolation required by the Covid-19 pandemic. And it becomes a viable alternative, to attend cases of depression in less time (taking into account variables such as travel, time availability, etc.) thus promoting a closer environment between the psychologist and the patient; in order to avoid that the lack of treatment leads to more serious implications in the patient's health.

For the implementation of the method, a own dataset was built, which was joined with public datasets that met the necessary characteristics in order to increase the number of images of the same, having a total of 7920 images and 1320 for each emotion. Subsequently, an investigation of the different existing face detection algorithms was carried out, choosing the ones that best matched the expected results, in order to run them with the generated dataset. Having the results, we proceeded to perform the analysis of metrics of each of the algorithms, identifying that for this research, the sequential algorithm of Keras, is the most optimal in terms of results and performance, thus obtaining a 62% accuracy in the recognition of emotions. Finally, a method is provided that can recognize facial emotions and give visualization in two different ways, by means of a percentage of approach to the emotion or by means of a bar chart where the index of coincidence with a certain emotion is presented.

Keywords: depression, detection, emotion, artificial intelligence, neural networks, sequential method, confusion matrix.

INTRODUCCIÓN

En la actualidad, existen diferentes alternativas tecnológicas que buscan ayudar a mejorar la calidad de vida de las personas mediante la optimización de tiempos y recursos en cierta tarea específica. Estas tecnologías son desarrolladas cada vez más en el contexto de la revolución digital 4.0 de Colombia, donde se desarrollan diferentes técnicas innovadoras no solo para la optimización de procesos tecnológicos sino también en entornos relacionados con la salud. Esto, considerando las iniciativas de nuevas tecnologías para combatir distintas enfermedades durante su tratamiento o la mejora de procesos de diagnóstico que conllevan brechas de precisión de gran magnitud y por lo tanto dudosa fiabilidad de parte de los pacientes.

El caso principal de estudio en esta investigación se centra en el área de la salud mental relacionada con la depresión; la cual, es una enfermedad que alarma a nivel mundial dado el incremento de casos. De acuerdo con el NIH (National Institute of Mental Health) ³Es posible deducir que es una enfermedad que no distingue entre raza, edad, ingresos, nivel educativo o cultura. Además, establece a la depresión como aquella enfermedad con posibilidad de derivar síntomas graves que afectan cómo se siente, piensa y coordina sus actividades diarias como dormir, comer o trabajar. De la anterior definición, es posible interpretar que las implicaciones de esta enfermedad pueden llegar a afectar todos los aspectos de la vida de las personas.

Por otro lado, dentro de las alternativas que han surgido debido a las diferentes restricciones impuestas por la pandemia de COVID 19 y el estudio de las características que presenta la depresión, se encuentran las psicoterapias como alternativa de tratamiento que en este contexto permitieron ser una prioridad en combinación con la telemedicina. De esta manera, el énfasis de la investigación busca encontrar las manifestaciones físicas (como las emociones) que puedan ser medidas mediante diferentes algoritmos; teniendo en cuenta, apreciaciones realizadas por la pesquisa de la Universidad de Buenos Aires⁴ las emociones resultan de respuestas químicas neuronales que surgen de forma natural y son esenciales para el desarrollo de la vida cotidiana de los seres humanos. Se componen de expresión motora, de elementos cognitivos, neurofisiológicos, entre otros, que a su vez dependen de diferentes subsistemas internos del organismo.

³ NATIONAL INSTITUTE OF MENTAL HEALTH . Depresión. [Sitio Web]. EEUU. [Consultado el 18 de junio de 2021]. Disponible en: <https://www.nimh.nih.gov/health/publications/espanol/depression-sp>

En función de lo planteado, se desarrolló un trabajo de investigación interdisciplinario entre la facultad de psicología y de ingeniería de la Universidad Católica de Colombia, en pro de contribuir a la sociedad, especialmente a las personas que utilizan la unidad de servicios psicológicos, implementando un desarrollo tecnológico basado en machine learning que a partir de la detección de emociones con base en expresiones faciales durante una consulta virtual, se obtiene como resultado, un informe con las emociones presentadas por los pacientes, coadyuvando así a la detección de la depresión temprana en adultos jóvenes que realiza el psicólogo durante la sesión anteriormente mencionada.

El componente tecnológico novedoso del presente proyecto de investigación se basa en descifrar los patrones de comportamiento del paciente, demuestra un impacto tecnológico significativo, que ofrece por otro lado, la posibilidad de apoyar al psicólogo en el diagnóstico de la depresión en adultos jóvenes, favoreciendo los tiempos de respuesta y la efectividad frente a la entrega de un diagnóstico psicológico, lo cual denota un importante impacto social hacia los usuarios que requieran este servicio de psicología al interior de la universidad.

El capítulo 1 define el problema de investigación, el capítulo 2 justifica la importancia del proyecto mediante la determinación de las ventajas que conlleva la correcta implementación, el capítulo 3 precisa los objetivos de la investigación con los cuales se demarca el proceso de desarrollo del proyecto, el capítulo 4 describe el desarrollo de los marcos de referencia que permiten dar el contexto técnico del desarrollo de la investigación, el capítulo 5 detalla la metodología empleada para dar solución a las metas de la presente investigación, el capítulo 6 describe las características que se tuvieron en cuenta para la construcción del dataset empleado por los métodos propuestos, capítulo 7 especifica las principales propiedades de los métodos aplicados, capítulo 8 contiene la información relacionada con las arquitecturas de los métodos, capítulo 9 detalla las métricas de desempeño por cada algoritmo, capítulo 10 desarrolla los análisis de los resultados y capítulo 11 finaliza con las conclusiones de la investigación.

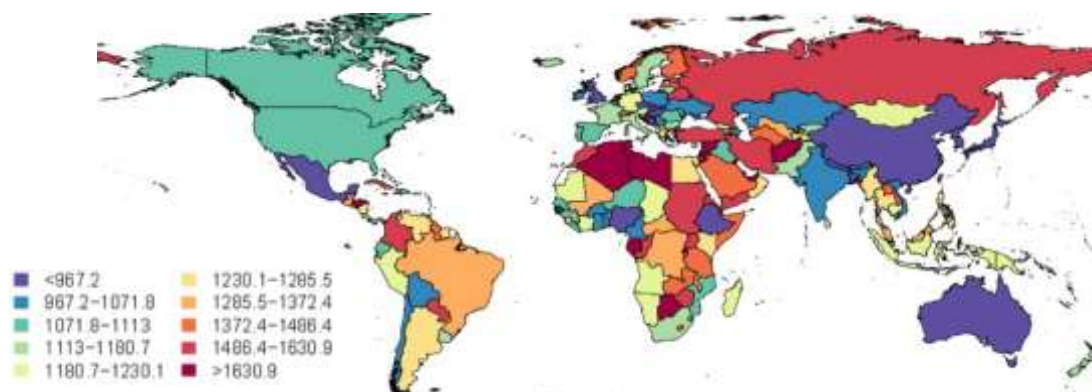
⁴ COSSIN, Florencia. RUBINSTEIN, Wanda. POLITIS, Daniel. Universidad de Buenos Aires. ¿Cuántas son las emociones básicas? Estudio preliminar en una muestra de adultos mayores sanos. En el Anuario de Investigaciones, vol. XXIV, pp. 253-257, 2017.

1. PROBLEMA DE INVESTIGACIÓN

1.1. PLANTEAMIENTO DEL PROBLEMA

La depresión es un trastorno mental que en la actualidad ha aumentado y se constituye como un padecimiento preocupante a nivel mundial, tal como se muestra en el artículo de la revista Semana⁵ titulado: El mundo triste de la depresión, en el cual se afirma que esta enfermedad es la segunda causal de generación de incapacidades en el mundo. Así mismo, este fenómeno se demuestra mediante el siguiente mapa como se ve en la figura 1 Tasas de años de vida con discapacidad (por 100.000) por país para los trastornos depresivos en 2010, tomado de un estudio publicado por la revista PLOS Medicine:

Figura 1: Tasas de años de vida con discapacidad (por 100.000) por país para los trastornos depresivos en 2010.



Fuente: Plos Medicine, 2013

En la figura 1 Tasas de años de vida con discapacidad (por 100.000) por país para los trastornos depresivos en 2010, puede evidenciarse la prevalencia de la depresión clínica diagnosticada mundialmente según cada país, también, se observa que los colores con los cuales se demarcan los países van organizados de menor AVD, en color azul, a mayor AVD, en color rojo como es el caso de Colombia, información proveniente de cada país analizada hasta 2010. De manera similar, según la OMS⁶, se estimó que para el año 2019 en el mundo existían 350 millones que padecían depresión, adicional que, como desenlace de ello, 800 mil personas entre los 18 y 29 deciden quitarse la vida. Esta situación no es ajena a Colombia, pues como se evidencia en la anterior imagen, en años posteriores al análisis de 2010, Medicina Legal⁷ había confirmado un incremento del 10% en el índice de muertes autoinfligidas en el país, además, en el 2019 la causa de muerte con mayor incremento de casos fue el suicidio, especialmente, en el mismo grupo etario que había confirmado la OMS.

⁵ SEMANA S: A.: Salud Mental. El mundo triste de la depresión. [Sitio Web]. [Consultado el 05 de agosto de 2021]. Disponible en: <https://www.semana.com/vida-moderna/articulo/depresion-en-paises-violentos/364663-3/>

Dado al incremento en casos de lesiones autoinfligidas, a causa de sintomatología relacionada con la depresión mencionado anteriormente y a las presiones psicológicas impulsadas por la pandemia en el diario vivir de toda la sociedad, el centro de salud Mente a Mente⁸ indica que han incrementado alrededor del 30% las consultas en servicios de salud mental. De esta manera es posible introducir los tratamientos virtuales del trastorno depresivo, en sus dos formas principales de tratamiento: la terapia psicológica o psicoterapia y la medicación antidepresiva. Donde es posible resaltar que dado a las restricciones debidas a la pandemia y además, con el surgimiento de diferentes alternativas alineadas con la situación mundial; el desarrollo de psicoterapias en la modalidad de telemedicina ha marcado un nuevo rumbo en la sociedad.

Por otro lado, de acuerdo con el banco BBVA⁹, la tecnología ha permitido mejorar las condiciones de vida y generar mejores trabajos, dado a las diversas oportunidades que ofrece como la optimización de recursos, tiempos, distancias, almacenamiento, precisión de resultados, entre otros. Esto hace que varios sectores se apropien de estas herramientas y vean en la tecnología un mecanismo de solución a diferentes problemáticas en pro del bienestar de la sociedad.

Partiendo de un análisis con docentes de la facultad de psicología de la Universidad Católica de Colombia se encontró que los psicólogos de la unidad de servicios psicológicos, han empleado alternativas virtuales como una nueva forma de seguimiento a sus pacientes. Sin embargo, dado a las magnitudes e implicaciones que ha tenido la prevalencia de la depresión y la necesidad de implementación de tecnología en pro de mejorar diferentes aspectos de la vida diaria, se propone la implementación de una herramienta tecnológica de apoyo en consultas virtuales que permita al psicólogo conocer con base a la expresión facial de su paciente la emoción que presenta, generando de este modo un algoritmo de valor para el diagnóstico, principalmente en la modalidad de telemedicina.

⁶ MINSALUD. Depresión: hablemos. [Sitio Web] [Consultado el 26 de julio de 2021]. Disponible en: <https://www.minsalud.gov.co/Paginas/Depresi%C3%B3n-hablemos.aspx>

⁷ PINZÓN, Jorge. Bienestar Colsanitas. Suicidio en Colombia: alerta amarilla. [Sitio Web]. [Consultado el 05 de agosto de 2021]. Disponible en: <https://www.bienestarcolsanitas.com/articulo/suicidio-en-colombia-alerta-amarilla.html>

⁸ LÓPEZ, David. SANZ, Natalia. Mente a Mente. Terapia para la depresión. [Sitio Web]. [Consultado el 06 de agosto de 2021]. Disponible en: <https://www.menteamente.com/terapia-depresion>

⁹ SAUNDERS, Adam. BBVA. El impacto de la tecnología en el crecimiento y el empleo. [Consultado el 25 de noviembre de 2021]. Disponible en: <https://www.bbvaopenmind.com/articulos/el-impacto-de-la-tecnologia-en-el-crecimiento-y-el-empleo/>

1.2 FORMULACIÓN DE LA PREGUNTA PROBLEMA

Para el desarrollo del presente trabajo se plantea la siguiente pregunta de investigación:

¿Qué elementos deben considerarse para la implementación de un método automático que identifique las emociones faciales planteadas por Ekman en adultos jóvenes (18 a 29 años) para apoyar en el diagnóstico o tratamiento de la depresión utilizando técnicas de Machine Learning?

1.3 SISTEMATIZACIÓN DEL PROBLEMA

Durante el planteamiento del problema y el análisis de la pregunta de investigación surgieron preguntas cómo:

- a. ¿Qué criterios deben tener las imágenes que conformarán el dataset?
- b. ¿Cómo se deben recolectar las fotografías por emoción para la creación del dataset?
- c. ¿Es posible realizar la detección de las 6 diferentes emociones planteadas por Ekman: ira, alegría, asco, tristeza, sorpresa, miedo?
- d. ¿Cómo es posible garantizar la validez del algoritmo?
- e. ¿Es posible que el algoritmo detecte más de un rostro y sea útil en terapias grupales?

Entre otros aspectos a tener en cuenta, se considera la necesidad de participación de forma voluntaria de más de 100 personas entre los 18 y 29 años que a partir de la firma del consentimiento informado, realicen el envío de 6 fotos de su rostro, representando cada emoción propuesta por Ekman; teniendo en cuenta, que es vital para la elaboración del dataset que enmarca las características principales de la población objetivo de la investigación.

2. JUSTIFICACIÓN

El presente proyecto de investigación tiene como finalidad el desarrollo de una herramienta de apoyo a los psicólogos y terapeutas de la Unidad de Servicios Psicológicos de la Universidad Católica de Colombia, en la detección de la depresión en pacientes que asisten a este servicio. Colombia, según estudio de la revista *Pesquisa de la Universidad Pontificia Javeriana*¹⁰, evidenció que durante la pandemia las mujeres, adultos jóvenes entre 18 y 29 años y personas de bajos recursos se vieron afectadas emocionalmente, siendo el 35% de la población quienes presentaron síntomas de depresión.

Por otro lado, con la llegada de la pandemia por COVID 19 en el año 2020, se hizo necesario evolucionar la forma en que se realizan los diagnósticos psicológicos y sus tratamientos, cambiando la forma de atención al paciente de presencial a virtual. Donde, mediante el apoyo de la tecnología, fue posible llegar a todas aquellas personas que cuentan con dispositivos móviles, acceso a internet y red de cobertura en general, recibiendo atención desde cualquier lugar y posiblemente a cualquier hora, también eliminando o disminuyendo los tiempos que son necesarios para el traslado y movilización de los pacientes hacia los centros asistenciales, adicional a los recursos con los que debe contar un paciente a la hora de asistir al servicio. De esta manera, más personas podrán verse beneficiadas y acceder a estos servicios fácilmente.

A partir de lo anterior, conviene agregar, que este proyecto ayuda a mitigar el impacto ambiental, ya que se propone un uso completamente virtual en cuanto a lo que encierran las terapias y diagnósticos, evitando de esta manera el uso de papel.

La implementación de este proyecto beneficiará a todas las personas entre 18 y 29 años que usan la unidad de servicios psicológicos de la Universidad Católica de Colombia dado que optimizará el manejo de tiempos empleados en el desarrollo de la consulta como se abordó anteriormente; así como, a los profesionales que ofrecen el servicio de consulta, por tener una herramienta que les permite identificar de mejor manera la sintomatología que expresa el paciente mediante la detección de ciertos rasgos en su expresión facial durante la consulta. Lo anterior, partiendo de que el entregable final de este proyecto, servirá a su vez como apoyo al trabajo de investigación titulado “Diseño y validación de un software de apoyo para la detección de sintomatología

¹⁰GUZMAN, Juan, TAMAYO, Alejandro. Ansiedad y depresión, ‘verdugos’ de los jóvenes en la pandemia. [Consultado el 30 de agosto de 2021]. [Sitio Web]. Disponible en: <https://www.javeriana.edu.co/pesquisa/ansiedad-depresion-jovenes-pandemia/>

emocional en narrativas textuales y orales” en donde se pretende articular los entregables de diferentes investigaciones relacionadas con el apoyo al diagnóstico emitido por los profesionales de salud Mental de la Universidad Católica de Colombia.

Es evidente que dentro del proyecto de investigación se realiza la aplicación de varios conceptos y conocimientos tecnológicos recibidos en el proceso de formación integral como ingenieros de sistemas dentro de la universidad y el compromiso de la responsabilidad social con las personas con depresión, que son tratadas por la unidad de servicios psicológicos de la Universidad Católica de Colombia, de esta forma se articula el trabajo en equipo con las facultades de psicología e ingeniería, que contribuye a la mejora de la calidad de vida de las personas para generar un impacto positivo en sus vidas y en su salud tanto física como mental.

2.1 ALCANCE Y LIMITACIONES

Por otro lado, el presente proyecto de investigación presenta las siguientes limitaciones:

- En el desarrollo de este proyecto solo se emplean las emociones definidas por el psicólogo Paul Ekman.
- La calidad de las imágenes está sujeta a la resolución de la cámara empleada por los participantes.
- Se trabaja con el GPU virtual ya que los modelos seleccionados son los que se puedan procesar con esta.

3. OBJETIVOS

3.1 OBJETIVO GENERAL

Diseñar un método automático para el apoyo en la identificación de emociones en tratamientos psicológicos de depresión en adultos jóvenes a partir de técnicas de aprendizaje automático.

3.2 OBJETIVOS ESPECÍFICOS

- 3.2.1 Construir el dataset, a partir de fotografías tomadas directamente y datasets públicos de las emociones faciales postuladas por Ekman en jóvenes adultos para entrenar los algoritmos de machine learning.
- 3.2.2 Identificar el método automático de clasificación, con base a la selección de técnicas de extracción de características y de machine learning, con el fin de detectar las emociones faciales en adultos jóvenes propuestas por Ekman.
- 3.2.3 Desarrollar el método automático, a partir del entrenamiento de los algoritmos seleccionados con el dataset y el ajuste de parámetros para validar el diseño propuesto.
- 3.2.4 Evaluar el desempeño del método automático desarrollado, a partir de métricas de modelos de clasificación para validar el método y el ajuste de sus parámetros.

4. MARCO DE REFERENCIA

Teniendo en cuenta que el presente trabajo se centra en la implementación de un método automático a través de visión artificial, cabe dar explicación a algunos conceptos básicos relacionados con el desarrollo del presente proyecto.

4.1 MARCO TEÓRICO

A continuación, se exponen términos específicos para la realización del presente trabajo de investigación.

4.1.1 Análisis de Componentes Principales (PCA)

Según Joaquín Amat¹¹, hace parte de las técnicas de aprendizaje no supervisado, en el cual se “condensa” la información en pocos componentes. Es principalmente aplicado en técnicas estadísticas de regresión, clustering entre otros. En el cual, se parte de la existencia de una muestra con “n” individuos cada uno con “p” variables, se aplica la técnica y resultan “z” variables que permiten caracterizar los n individuos con menos variables. Dentro de esta técnica se aplican los eigenvectores y los eigenvalores, teniendo en cuenta que en el método PCA, cada una de las componentes se corresponde con un eigenvector, y el orden de componente se establece por orden decreciente de eigenvalue. Así pues, el primer componente es el eigenvector con el eigenvalor asociado más alto. Además, dado un conjunto de datos X con n observaciones y p variables, el procedimiento para calcular el primer componente principal es el siguiente:

- Centralización de la variable: restar a cada valor la media de la variable a la que pertenece. De este modo se garantiza que la media de cada variable sea cero.
- Se resuelve un problema de optimización para encontrar el valor de las cargas en el que se maximiza la varianza. Una forma de resolver el problema de optimización es calcular el valor integrado de los vectores propios de la matriz de covarianza.

4.1.1.1 EigenVectores

Los eigenvectores de una matriz son todos aquellos vectores que, al multiplicarlos por dicha matriz, resultan en el mismo vector o en un múltiplo entero del mismo. Los eigenvectores tienen una serie de propiedades matemáticas específicas:

¹¹ AMAT, J. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado 11 de octubre de 2021]. Disponible en: https://www.cienciadedatos.net/documentos/35_principal_component_analysis

- Los eigenvectores solo existen para matrices cuadradas y no para todas. En el caso de que una matriz $n \times n$ tenga eigenvectores, el número de ellos es n .
- Si se escala un eigenvector antes de multiplicarlo por la matriz, se obtiene un múltiplo del mismo eigenvector. Esto se debe a que, si se escala un vector multiplicándose por cierta cantidad, lo único que se consigue es cambiar su longitud, pero la dirección es la misma.
- Todos los eigenvectores de una matriz son perpendiculares (ortogonales) entre ellos, independientemente de las dimensiones que tengan¹².

4.1.1.2 Eigenvalores

Cuando se multiplica una matriz por alguno de sus eigenvectores se obtiene un múltiplo del vector original, es decir, el resultado es ese mismo vector multiplicado por un número. Al valor por el que se multiplica el eigenvector resultante se le conoce como eigenvalue. A todo eigenvector le corresponde un eigenvalue y viceversa.¹³

4.1.1.3 Cálculo de las componentes principales

La primera componente principal de un grupo de variables (X_1, X_2, \dots, X_p) es la combinación lineal normalizada de dichas variables que tiene mayor varianza:

$$Z_1 = \phi_{11} * X_1 + \phi_{21} * X_2 + \dots + \phi_{p1} X_p$$

Los términos $\phi_{11}, \dots, \phi_{p1}$ reciben en el nombre de loadings y son los que definen a la componente. ϕ_{11} es el loading de la variable X_1 de la primera componente principal.

4.1.2 Análisis Discriminante Lineal (LDA)

Técnica de Análisis Multivariante que permite asignar o clasificar nuevos individuos dentro de grupos previamente reconocidos o definidos, donde el objetivo es obtener una función capaz de clasificar a un nuevo individuo a partir del conocimiento de los valores de ciertas variables discriminadoras¹⁴.

¹² AMAT, J. Eigenvectores. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado el 11 de octubre de 2021]. Disponible en https://www.cienciadedatos.net/documentos/35_principal_component_analysis

¹³ AMAT, J. Eigenvalores. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado el 11 de octubre de 2021]. Disponible en https://www.cienciadedatos.net/documentos/35_principal_component_analysis

Adicionalmente, LDA utiliza el discriminante lineal de Fisher para reducir la dimensionalidad de los datos mientras maximiza la separación entre clases. Lo hace maximizando la distancia entre las medias y minimizando la variación dentro de la clase.

4.1.2.1 Discriminante lineal de Fisher

Según Diego Unzueta¹⁵, es usado como un clasificador de aprendizaje supervisado, encuentra el conjunto de pesos para determinar un límite de decisión con la clasificación de los datos. Adicionalmente encuentra la máxima distancia entre cada clase de datos, maximiza la distancia de las medias proyectadas y minimiza la varianza dentro de cada una de las clases proyectadas.

4.1.2.2 Python

Es un lenguaje de programación interpretado, multiparadigma y multiplataforma usado, principalmente, en Big Data, AI (Inteligencia Artificial), Data Science, frameworks de pruebas y desarrollo web. Esto lo convierte en un lenguaje de propósito general de gran nivel debido a su extensa biblioteca, cuya colección ofrece una amplia gama de instalaciones. Dentro de sus funcionalidades más destacadas se encuentran: Web development, Machine Learning, DevOps, System Administration, Writing Automation Scripts, Programming of web parsers, scrapers, crawlers, Software Testing, Writing Automated Test, entre otras¹⁶. A continuación, se especifican algunas de sus librerías, módulos y paquetes utilizados en esta investigación.

4.1.2.2.1 Python OS

Este módulo provee una manera versátil de usar funcionalidades dependientes del sistema operativo.¹⁷

¹⁴ UNIVERSIDAD DE GRANADA. Capítulo 8 Análisis Discriminante. [Sitio Web]. [Consultado 21 de octubre de 2021]. Disponible en: https://www.ugr.es/~bioestad/_private/cpfund8.pdf

¹⁵ UNZUETA, D. Towards Data Science. Fundamentos del aprendizaje automático. Discriminante lineal de Fisher: explicación intuitiva. [Sitio Web]. [Consultado 22 de noviembre de 2021].

Disponible en: <https://towardsdatascience.com/fishers-linear-discriminant-intuitively-explained-52a1ba79e1bb>¹⁶ MACHUCA Fernando. ¿Qué es Python? El lenguaje de programación más popular para aprender en 2021. [Sitio Web]. [Consultado el 19 de octubre de 2021]. Disponible en: <https://www.crehana.com/cl/blog/desarrollo-web/que-es-python/>

¹⁷ PYTHON. Documentation. Biblioteca Estándar. Interfaces misceláneas del sistema operativo. [Sitio Web]. [Consultado el 28 de enero del 2022]. Disponible en: <https://docs.python.org/es/3.10/library/os.html>

4.1.2.2.2 OPENCV (Open Source Computer Vision)

Es una biblioteca de visión artificial y código abierto que incluye varios algoritmos especializados en el reconocimiento facial como procesamiento, análisis, calibración de cámara entre otros¹⁸.

Cuenta con más de 500 funciones de computer vision, una biblioteca de aprendizaje de máquinas que a partir de la historia que rescatan sobre la misma Alexis Ospina y Mateo Pulido¹⁹, se establece que surge desde una iniciativa de Intel con el fin de avanzar y optimizar las aplicaciones que hacen uso intensivo de la CPU. Llegando a la determinación de ofrecer una herramienta open source con la cual los códigos que se generen sean portátiles, con rendimiento optimizado y sin costo alguno.

4.1.2.2.3 Numpy

Es un paquete utilizado en la computación científica en Python, proporcionando objetos de matriz multidimensionales, varios objetos derivados como matrices y surtido en rutinas para realizar operaciones matemáticas, lógicas, de manipulación de formas, de ordenación, de selección, de E/S, de transformadas discretas de Fourier, de álgebra lineal básica, de operaciones estadísticas básicas, de simulación aleatoria y mucho más²⁰.

4.1.2.2.4 Pandas

Es una biblioteca de software escrita como extensión de Numpy para manipulación y análisis de datos para el lenguaje de programación Python.²¹

4.1.2.2.5 Matplotlib

Es una completa biblioteca para crear visualizaciones estáticas, animadas e interactivas en Python²².

¹⁸ OPENCV. Introducción. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en : <https://docs.opencv.org/4.5.5/d1/dfb/intro.html>

¹⁹ OSPINA, Alexis. PULIDO, Mateo. Comparar tres métodos convencionales de reconocimiento facial y confrontar el de mejor resultado frente a computación cognitiva. [Sitio Web]. Disponible en: https://repositorio.itm.edu.co/bitstream/handle/20.500.12622/1706/Rep_Itm_pre_Ospina.pdf?sequence=1&isAllowed=y

²⁰ NUMPY. NumPy documentation. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: <https://numpy.org/doc/stable/>

²¹ UNIVERSIDAD DE ALCALÁ. Pandas: Herramienta básica para el data science en Python. [Sitio Web]. [Consultado el 07 de febrero de 2022]. Disponible en: <https://www.master-data-scientist.com/pandas-herramienta-data-science/>

²² Matplotlib. Matplotlib: Visualization with Python. [Sitio Web] [Consultado el 05 de febrero de 2022]. Disponible en: <https://matplotlib.org/>

4.1.2.2.6 Seaborn

Es una librería para Python que permite generar fácilmente elegantes gráficos. Seaborn está basada en matplotlib y proporciona una interfaz de alto nivel que es realmente sencilla de aprender.²³

4.1.2.2.7 Imutils

Según Anaconda²⁴ Es una serie de funciones para realizar el procesamiento básico de imágenes como traslaciones, redimensionamiento, desplegar imágenes de matplotlib, entre otras.

4.1.2.2.8 Scikit Learn

Es una herramienta simple y eficiente para el análisis predictivo de datos, basado en Numpy, SciPy y Matplotlib de código abierto utilizable comercialmente.²⁵

4.1.2.3 Keras de TensorFlow

TensorFlow es una librería de código abierto especializada en el aprendizaje automático, cuenta con diversas librerías, herramientas y recursos que permiten la innovación implementando aplicaciones de alto nivel fácilmente. Una de sus API más reconocidas es Keras, utilizada para entrenar y construir modelos de aprendizaje profundo, se utiliza para la creación de rápida de prototipos, ya que cuenta con diferentes, layers y funciones²⁶, las más relevantes y usadas en la ejecución de este proyecto son:

- **Conv2D:** En esta capa se crea un núcleo de convolución que tiene como fin con la entrada obtener un tensor de salida, en este caso, al ser la primera capa de la arquitectura, se le ingresa el $input_shape = (width_shape, height_shape, 1)$, donde corresponden al ancho, largo y color de la imagen respectivamente, esto especifica las condiciones de la imagen a convolucionar²⁷.

²³ ANALYTICSLANE. Visualización de datos en Python con Seaborn. [Sitio Web]. Consultado el 07 de febrero de 2022]. Disponible en: <https://www.analyticslane.com/2018/07/20/visualizacion-de-datos-con-seaborn/#:~:text=Seaborn%20es%20una%20librer%C3%ADa%20para,defecto%20en%20la%20distribuci%C3%B3n%20Anaconda>.

²⁴ ANACONDA. imutils 0.5.4. [Sitio Web]. [Consultado el 03 de febrero de 2022]. Disponible en: <https://anaconda.org/conda-forge/imutils>

²⁵ SCIKIT-LEARN. Machine learning en Python. [Sitio Web]. [Consultado 04 de marzo de 2022]. Disponible en: <https://scikit-learn.org/stable/>

²⁶ KERAS. Deep learning for humans. [Sitio Web] [Consultado el 05 de febrero de 2022]. Disponible en: <https://keras.io/>

²⁷ TENSORFLOW. tf.keras.layers.Conv2D. [Sitio Web] [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Conv2D²⁸

- **BatchNormalization:** En esta capa se realiza una normalización por lotes aplicando una transformación que mantiene la salida media cercana a 0 y la desviación estándar de salida cercana a 1. Durante la etapa de entrenamiento, la capa normaliza la salida usando la media y desviación estándar de la entrada, por lo que por cada canal que se normaliza la capa devuelve²⁸, como se evidencia en la figura 2:

Figura 2: Fórmula BatchNormalization

$$gamma * \frac{(batch - mean(batch))}{sqrt(var(batch))} + epsilon + beta$$

Autor: TensorFlow, sin fecha.

Donde se define :

Gamma = Factor de escala aprendido, inicializado como 1

Batch = Factor de comprensión de lo aprendido, inicializado como 0

Epsilon = Constante pequeña

- **Activation (Relu):** Es un proceso que transforma la entrada permitiendo que el modelo genere la predicción y aprenda más rápidamente, desarrollando un perceptrón convolucional y multicapa²⁹.
- **MaxPool2D:** Reduce la muestra de la entrada por sus dimensiones espaciales, tanto ancho como alto, tomando un valor máximo predefinido para el canal de entrada, desplazando la ventana a strides o saltos (pueden ser tupla de 2 enteros o ninguno), a lo largo de las 2 dimensiones³⁰.

²⁸TENSORFLOW. tf.keras.layers.BatchNormalization.[Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en:

https://www.tensorflow.org/api_docs/python/tf/keras/layers/BatchNormalization

²⁹ TENSORFLOW.. tf.keras.layers.Activation. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Activation

³⁰ KERAS. MaxPooling2D layer. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://keras.io/api/layers/pooling_layers/max_pooling2d/

- **Dropout:** En esta capa se ponen aleatoriamente unidades de entrada a 0 con una frecuencia de tasa en cada paso durante el tiempo de entrenamiento, lo que ayuda a evitar el sobreajuste. Las entradas que no se ponen a 0 se escalan en $1/(1 - \text{tasa})$ de manera que la suma de todas las entradas no cambia. Solo se aplica cuando el entrenamiento se establece en True, de modo que no se eliminan valores durante la inferencia³¹.
- **Flatten:** Es un layer que elimina una dimensión para imágenes multidimensionales, es un layer de transición con el fin de entregar al siguiente layer o capa un dato legible en una sola dimensión³².
- **Dense:** Implementa la operación

$$\text{output} = \text{activation}(\text{dot}(\text{input}, \text{kernel}))^{33}$$

Donde :

Activation = Función de activación por elementos pasada como argumento de activación

Kernel = Matriz de pesos creada por la capa, y bias es un vector de bias creado por la capa (sólo aplicable si use_bias es True).

4.1.2.4 Convolutional neural networks (CNN)

Es un tipo de Red Neuronal Artificial con aprendizaje supervisado que procesa sus capas imitando al córtex visual del ojo humano para identificar distintas características en las entradas que en definitiva hacen que pueda identificar objetos y “ver”. De esta manera, las redes neuronales convolucionales organizan sus capas internas siguiendo una jerarquía, en la cual, las primeras capas pueden detectar líneas, curvas y se van especializando hasta llegar a capas más profundas que reconocen formas complejas como un rostro o la silueta de un animal³⁴.

³¹ KERAS. Dropout layer. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://keras.io/api/layers/regularization_layers/dropout/

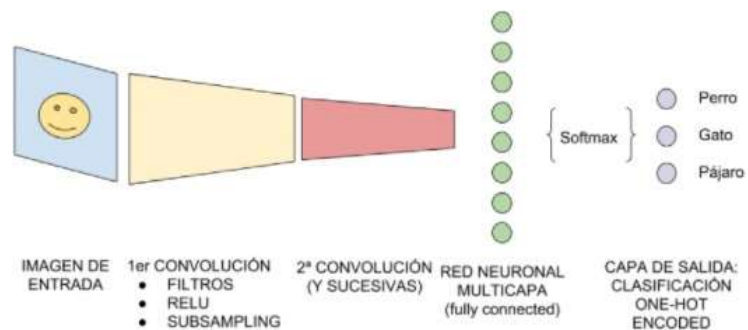
³² KERAS. Flatten layer. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://keras.io/api/layers/reshaping_layers/flatten/

³³ KERAS. Dense layer. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://keras.io/api/layers/core_layers/dense/

³⁴ APRENDER DE MACHINE LEARNING. ¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador. [Consultado el 03 de agosto de 2021]. Disponible en: <https://www.aprendemachinellearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>

A continuación, en la figura 3 Arquitectura de una red neuronal convolucional se muestra la arquitectura general para una red neuronal convolucional (CNN-Convolutional Neural Networks) y los procesos o clasificaciones dentro de las capas de convolución.

Figura 3: Arquitectura de una red neuronal convolucional



Fuente : Aprender machine learning, 2018.

4.1.2.5 Tensor

Los tensores son estructuras de almacenamiento, específicamente son matrices multidimensionales con un tipo uniforme (llamado dtype) que son inmutables; es decir, que no permiten la modificación de su contenido, solo es posible la creación de nuevos tensores. Además, según la guía de Tensor Flow³⁵, los tensores contienen diferentes dimensiones o parámetros como lo son:.

- **Forma:** La longitud (número de elementos) de cada uno de los ejes de un tensor.
- **Rango:** Número de ejes tensoriales. Un escalador tiene rango 0, un vector tiene rango 1, una matriz tiene rango 2.
- **Eje o Dimensión:** Una dimensión particular de un tensor.
- **Tamaño:** el número total de elementos en el tensor, el vector de forma del producto.

4.1.2.6 Matriz de confusión

Es una herramienta utilizada habitualmente para evaluar la calidad de un sistema o modelo de clasificación basado en el aprendizaje

³¹ TENSORFLOW. Introducción a Tensores. [Sitio Web]. [Consultado el 03 de febrero del 2022]. Disponible en : <https://www.tensorflow.org/guide/tensor>

automático. Así, describe cuándo se confunde una categoría con otra, facilitando la comparación de los resultados en función de los distintos errores. Asimismo, el número de predicciones correctas e incorrectas se resumen con los valores de conteo resultantes y se desglosan por cada clase³⁶. Además, existen dos tipos de matrices de confusión en las tareas de aprendizaje automático, y se determina su uso en función del número de variables de clasificación planteadas, así se plantea matriz binaria y multivariable.

4.1.2.6.1 Matriz de confusión binaria

Es aquella que se según Mohajón³⁷ se emplea cuando se tiene un problema de clasificación que involucra únicamente dos resultados o categorías de clasificación, como se puede apreciar en la figura 4:

Figura 4: Matriz de confusión binaria

		Actual Values	
		Yes	No
Predicted Values	Yes	True Positive	False Positive
	No	False Negative	True Negative

Fuente: Datasource, 2020.

Donde:

- Falso positivo (False positive): El modelo predice como incorrecta una clase positiva.
- Falso negativo (False negative): El que el modelo predice incorrectamente la clase negativa cuando en realidad es positiva.
- Verdadero Positivo (True positive): Resultado en el que el modelo predice correctamente la clase positiva.
- Verdadero Negativo (TN): Resultado donde el modelo predice correctamente la clase negativa.

³⁶ DATASOURCE. Data Science. Machine learning. Python. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: <https://www.datasource.ai/es/data-science-articles/compreension-de-la-matriz-de-confusion-y-como-implementarla-en-python>

³⁷ MOHAJÓN, J. Towards Data Science. Matriz de confusión para su modelo de aprendizaje automático multiclase. [Sitio Web]. [Consultado el 19 de febrero de 2022]. Disponible en: <https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826>

Adicionalmente, a partir de la matriz de confusión, se generan medidas como lo son:

- Precisión³⁸ (Precision): Métrica que calcula la precisión de las predicciones del modelo con base a las etiquetas.
- Exhaustividad³⁹ (Recall): Calcula la recuperación de las predicciones del modelo con base a las etiquetas.
- Valor- F1⁴⁰ (f1-score) : Relación entre precisión y exhaustividad, en un rango de 0 a 1, siguiendo la fórmula :

$$F1 = 2 * \rightarrow \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

- Accuracy: Corresponde a la frecuencia con la que las predicciones coinciden con los labels o etiquetas preestablecidas⁴¹.

4.1.2.6.2 Matriz de confusión Multivariable

Se considera como aquella matriz de confusión compuesta por más de dos categorías de clasificación. Además, Mohajón⁴² explica a que, a diferencia de la matriz de confusión binaria, cada uno de las variables de falso positivo, falso negativo, verdadero positivo y verdadero negativo se determina por cada categoría y no solo por la totalidad de resultados de la matriz. A continuación, en la figura 5 se ejemplifica la estructura de la matriz de confusión multivariable.

Figura 5: Matriz de confusión multivariable

		True Class		
		Apple	Orange	Mango
Predicted Class	Apple	7	8	9
	Orange	1	2	3
	Mango	3	2	1

Fuente: Towards Data Science, 2020.

³⁸ TENSORFLOW. tf. keras. metrics. Precision. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/metrics/Precision

³⁹ TENSORFLOW. tf. keras. metrics. Recall. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/metrics/Recall

⁴⁰ TENSORFLOW. tf. keras. metrics. F1 Score. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/addons/api_docs/python/tfa/metrics/F1Score

⁴¹ KERAS. tf.compat.v1.metrics.precision [Sitio web]. [Consultado el 15 de marzo de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/compat/v1/metrics/accuracy

⁴² MOHAJON, J. Matriz de confusión para su modelo de aprendizaje automático multiclase. [Sitio Web]. [Consultado 19 de febrero de 2022]. Disponible en: <https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826>

Por ejemplo, a partir del caso presentado en la figura 5, se describe uno de los cálculos que se realiza por clase (Apple) teniendo en cuenta las fórmulas explicadas en el literal 4.1.2.2.12.1.

- Verdadero Positivo: 7
- Verdadero Negativo: $(2+3+2+1) = 8$
- Falso Positivo: $(8+9) = 17$
- Falso Negativo: $(1+3) = 4$

Precision, recall y f1-score para la clase apple

- Precision: $7/(7+17) = 0.29$
- Recall: $7/(7+4) = 0.64$
- F1-score: 0.40

Realizando así cada cálculo por cada una de las clases

4.1.2.7 Épocas (Epoch)

Es un corte arbitrario, que generalmente se define como una revisión de un conjunto de datos, utilizado para medir el entrenamiento de fases distintas, lo cual es de mucha utilidad para un registro y evaluación periódica. Cuando se utiliza `validation_data` o `validation_split` con el método de ajuste de los modelos Keras, la evaluación se ejecutará al final de cada época. Dentro de Keras, existe la posibilidad de añadir callbacks específicamente diseñados para ser ejecutados al final de una época⁴³.

4.2 MARCO CONCEPTUAL

4.2.1 Inteligencia Artificial (IA)

De acuerdo con Lasse Rouhiainen⁴⁴ en su libro Inteligencia Artificial, 101 cosas que debes saber sobre nuestro futuro, es la capacidad de las máquinas para realizar acciones que por lo general requieren de inteligencia humana, usando algoritmos, aprender de los datos y utilizar dicha información para realizar toma de decisiones, la IA tiene varias subramas en las que se destacan:

⁴³ KERAS. Getting started. Faq. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://keras.io/getting_started/faq/#what-do-sample-batch-and-epoch-mean

⁴⁴ ROUHIAINEN, Lasse. Inteligencia Artificial, 101 cosas que debes saber sobre nuestro futuro. Editorial Planeta S.A. [Sitio Web]. [Consultado el 19 de octubre de 2021]. Disponible en: https://static0planetadelibroscom.cdnstatics.com/libros_contenido_extra/40/39308_Inteligencia_artificial.pdf

4.2.2 Machine Learning:

Acuerdo al IBM, es un tipo de inteligencia artificial (IA), que permite que un sistema pueda aprender datos sin necesidad de programación explícita, sin ser un proceso fácil de realizar; se tiene que conforme a que el algoritmo procesa los datos puede dar como resultado modelos más precisos basados en experiencia, por lo que es necesario realizar un entrenamiento del mismo con un modelo de entrada que tendrá que dar una salida⁴⁵.

4.2.3 Deep learning:

En este tipo de machine learning, es uno de los algoritmos inspirados en el funcionamiento del cerebro humano basado en redes neuronales, procesando y aprendiendo grandes cantidades de datos, realizando tareas repetitivas ayudando a mejorar de forma gradual el resultado a través de “deep layers”, permitiendo un aprendizaje progresivo⁴⁶.

4.2.4 Aprendizaje no supervisado

También conocido como aprendizaje automático no supervisado, utiliza algoritmos de aprendizaje automático para analizar y agrupar conjuntos de datos no etiquetados. Estos algoritmos descubren patrones ocultos o agrupaciones de datos sin necesidad de intervención humana. Su capacidad para descubrir similitudes y diferencias en la información lo convierte en la solución ideal para el análisis exploratorio de datos, estrategias de venta cruzada, segmentación de clientes y reconocimiento de imágenes.⁴⁷

4.2.5 Procesamiento lenguaje natural

Este tipo de procesamiento facilita a las computadoras la interpretación del lenguaje de uso humano a uno comprensible, por ejemplo, permite a las computadoras la entrada e interpretación de texto, voz, imágenes, medición de sentimientos y determinación de importancia de los mensajes entre otros.⁴⁸

⁴⁵IBM. ¿Qué es Machine Learning?. [Consultado el 13 de octubre de 2021]. Disponible en: <https://www.ibm.com/co-es/analytics/machine-learning>

⁴⁶IBM. Deep Learning. [Sitio Web]. [Consultado 19 de octubre de 2021]. Disponible en <https://www.ibm.com/co-es/cloud/deep-learning>

⁴⁷ IBM. Centro de aprendizaje de la nube de IBM. ¿Qué es el aprendizaje no supervisado?. Aprendizaje sin supervisión. [Sitio Web]. [Consultado el 05 de octubre del 2021] Disponible en: <https://www.ibm.com/cloud/learn/unsupervised-learning>

⁴⁸ SAS. Software y soluciones de analítica. Procesamiento del lenguaje natural. [Consultado el 13 de octubre de 2021] Disponible en: https://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.html

4.2.6 Dataset

De acuerdo con la Asociación para la información científica y tecnológica, se tiene que un dataset puede entenderse como un conjunto de datos relacionados, tomados en condiciones similares que cumplen con criterios preestablecidos, por lo que es de esperarse que los datos tengan la misma estructura sintáctica⁴⁹.

4.2.7 EigenFace

Un EigenFace es aquel componente invariable que surge a partir de una serie de caras que representan vectores propios de la matriz cuadrada que contiene los valores que indican los grados de variación entre los datos a partir de variables aleatorias. Extraen, según IDIS50 la información facial relevante como los ojos, la nariz y los labios por medio de la variación estadística entre las imágenes.

4.2.8 Loadings

Pueden interpretarse como el peso/importancia que tiene cada variable en cada componente y, por lo tanto, ayudan a conocer que tipo de información recoge cada una de las componentes.⁵¹

4.2.9 Salud mental

Según el ministerio de salud y protección social Colombia, la salud mental es un estado que se expresa en la vida cotidiana a través de las diferentes interacciones humanas y en el comportamiento, permitiendo al individuo desplegar sus medios emocionales, cognitivos y mentales usados para trabajar, establecer relaciones y realizar contribuciones a la comunidad.⁵²

⁴⁹ ALLEN H. SACCHI, Simone, WICKETT, Karen. Definitions of dataset in the scientific and technical literature. [Sitio Web]. [Consultado 13 oct., 2021]. Disponible en: <https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/meet.14504701240>

⁵⁰ JPL. IDIS. Eigenface. [Sitio Web]. [Consultado el 08 de febrero del 2022]. Disponible en: <https://proyectoidis.org/eigenface/>

⁵¹ AMAT, J. Loadings. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado el 11 de octubre de 2021]. Disponible en https://www.cienciadedatos.net/documentos/35_principal_component_analysis

⁵² MINSALUD. Abecé sobre la salud mental, sus trastornos y estigma. [Sitio Web] [Consultado el 19 de octubre de 2021]. Disponible en:

4.2.10 Depresión

De acuerdo a la Organización Panamericana de la Salud (OPS), la depresión se entiende como una enfermedad grave que puede interferir con aspectos de la vida regular y capacidad para realizar actividades necesarias tales como trabajar, dormir, estudiar, comer, entre otros; a su vez, la misma puede ser causada por factores genéticos, biológicos, ambientales y psicológicos en su mayoría, siendo la gravedad, frecuencia y duración de los síntomas independiente para cada paciente⁵³.

4.2.11 Emociones

Se tienen actualmente diferentes definiciones de que es una emoción dependiendo del autor, las mismas según Russell involucran sentimientos, para Larsen y Ketelaar son respuestas emocionales que persisten a través del tiempo y contexto, de ahí que McCullough, Emmons, & Tsang expusieran que es posible volver a experimentar una emoción a lo largo de la vida, siendo para Gross, Sutton y Ketelaar, Lazarus y Malatesta, un aspecto completamente subjetivo del ser humano⁵⁴. De las emociones, se pueden destacar las propuestas por el psicólogo Paul Ekman quien es pionero en el estudio de las mismas y de las expresiones faciales, dando como resultado, que se contemplan 6 emociones básicas: alegría, ira, miedo, asco, sorpresa y tristeza⁵⁵.

4.2.12 Diagnóstico

De acuerdo con el Instituto Nacional del Cáncer, es un proceso por el cual se identifica una enfermedad o afección con base a signos o síntomas, por lo cual se pueden realizar diferentes tipos de análisis teniendo en cuenta también el historial médico y antecedentes familiares⁵⁶.

⁵³ OPS. Depresión. [Sitio Web]. [Consultado el 13 de octubre de 2021]. Disponible en: <https://www.paho.org/es/temas/depresion>

⁵⁴ KELTNER, D., & Lerner, J. S. Emotion. Handbook of Social Psychology. [Sitio Web]. [Consultado el 19 de octubre de 2021]. Disponible en : [doi:10.1002/9780470561119.socpsy001009](https://doi.org/10.1002/9780470561119.socpsy001009)

⁵⁵ COSSIN, Florencia Carla; Wanda Yanina Rubinstein; Daniel Gustavo Politis. How many are the basic emotions? Preliminary study in a sample of healthy old adults. [Sitio Web]. [Consultado el 13 de octubre de 2021]. Disponible en: <https://www.redalyc.org/journal/3691/369155966032/movil/>

⁵⁶ INSTITUTO NACIONAL DEL CÁNCER. Diagnóstico. [Sitio Web]. [Consultado el 19 octubre de 2021]. Disponible en: <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/diagnostico>

4.3 ESTADO DEL ARTE

Aya Hassouneh, A.M. Mutawa, M. Murugappan, de la Universidad de Kuwait, en la investigación titulada “Development of a Real Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods”⁵⁷ permiten ver el alcance que permite la clasificación de emociones; a partir, de etiquetas por reconocimiento facial junto con señales encefalográficas utilizando redes neurológicas convolucionales, la distinción entre pacientes físicamente discapacitados, infantes con autismo, personas sordas y aquellos que se encuentran postrados en cama, teniendo resultados de 87,27% a 99,81% de precisión. Adicionalmente, resaltan las características que tuvieron en cuenta dentro de la construcción de su dataset como emplear diferentes fondos, diferentes tonos de piel, diferentes tipos de iluminación y la categorización entre 6 emociones principales (felicidad, tristeza, enojo, miedo, disgusto y sorpresa).

En la investigación “Use of machine learning approach to predict depression in the elderly in China: A longitudinal study”⁵⁸ los autores desarrollan un algoritmo de aprendizaje automático, que basado en la información de personas de la tercera edad pueda determinar, según sus variables individuales de estilo de vida, estatus marital, enfermedades persistentes y las respuestas binarias de si o no que la persona determinaba en cada caso, obteniendo como resultado si la persona es vulnerable a presentar signos de depresión, esto con el fin de poder mitigarlo y contribuir a un tratamiento médico psicológico, de este modo se mejora la calidad de vida de las personas. Emplea algoritmos como el K-Nearest Neighbors (K-NN), logistic regression(LR) teniendo en cuenta que las salidas de su algoritmo siguen la distribución de Bernoulli, Support Vector Machine para la predicción y utilizando Deep Neural Networks con múltiples capas, medido por medio de precision, F1 Score, sensibilidad, entre otras, en donde los resultados estuvieron entre el 58%-75%

La investigación titulada “Automated facial video-based recognition of depression and anxiety symptom severity: cross-corpus validation”⁵⁹.

⁵⁷ HASSOUNEH, Aya. MUTAWA, A. MURUGAPPAN, M. Universidad de Kuwait. Development of a Real Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods. [Sitio Web]. [Consultado 27 de enero de 2022]. Disponible en: <https://www.sciencedirect.com/ucatalica.basesdedatosezproxy.com/science/article/pii/S235291482030201X>

⁵⁸ SU, Dai. ZHANG, Xingyu. HE, Kevin. CHEN, Yingchun. Use of machine learning approach to predict depression in the elderly in China: A longitudinal study. ScienceDirect. Volume 282, 1 March 2021, Pages 289-298

⁵⁹ PAMPOUCHIDOU, A. PEDIADITIS, M. KAZANTZAKI, E. SFAKIANAKIS, S. APOSTOLAKI, A. ARGYRAKI, K. MANOUSOS, D. MERIAUDEAU, F. MARIAS, K. YANG, F. TSIKNAKIS, M. BASTA, M. VGONTZAS, A. SIMOS, P. Machine Vision Applications. Automated facial video-based recognition of depression and anxiety symptom severity: cross-corpus validation. [Sitio Web]. [Consultado el 13 de octubre de 2021]. Disponible en: 10.1007/s00138-020-01080-7

Describe la detección de emociones precisa como uno de los signos no verbales de depresión que permiten por medio de su análisis, ser un método mínimamente intrusivo para monitorear la sintomatología depresiva. Además, detallan la importancia de establecer las herramientas que se diseñan para contribuir al fin anteriormente expuesto, como un método de soporte de los sistemas que contribuyen en la toma de decisiones de los profesionales de la salud que llevan a cabo las sesiones terapéuticas de manera remota. El método que exponen tiene como salida el reconocimiento de pacientes que tienen o no síntomas depresivos, logrando obtener resultados por encima del 79% en las medidas de desempeño, a partir de implementaciones de técnicas de machine learning como Active appearance model with Support Vector Machines, Principal components analysis (PCA) para la reducción de dimensionalidad de las entradas, resaltando el uso de datasets previamente clasificados médicamente entre pacientes depresivos y pacientes sanos.

Hyen Byeon y Chang Kwak, del departamento de control e instrumentación ingenieril de la universidad Chosun en Corea, desarrollaron la investigación titulada “Facial Expression Recognition Using 3D Convolutional Neural Network”⁶⁰, en la cual desarrollan un método automático para el reconocimiento de expresiones faciales basado en vídeo para aplicación en HRI (Human-Robot Interaction), enfocado en seis emociones principales propuestas por Ekman: felicidad, tristeza, enojo, sorpresa, disgusto y miedo. Además, describen que dado al tiempo en el que tarda una persona en cambiar de emoción, denota la necesidad de resguardar y hacer el estudio de cada uno de los rostros a partir de imágenes estáticas y vídeos con tiempos base. Dentro de su propuesta, desarrollan una red convolucional que permite leer imágenes en tres dimensiones donde resaltan el potencial de emplear técnicas como Principal Components Analysis (PCA), Multilinear Principal Component Analysis (MPCA) y Tensor-based Multilinear Principal Component Analysis (TMPCA).

Wang Qingxiang, Yang Huanxin, Yanhong Yu en su investigación, “expresión facial para la detección de depresión en pacientes chinos”⁶¹. Constatan la urgencia de la necesidad de herramientas auxiliares de detección temprana de la depresión a partir de las emociones presentadas por Ekman y la precisión del 80,77% que el método que desarrollaron presentó. Rescatan la utilización de técnicas como FACS, la cual permite la detección automática de caras en

⁶⁰ BYEON, Young-Hyen. KWAK, Keun-Chang. Facial Expression Recognition Using 3D Convolutional Neural Network. [Sitio Web]. [Consultado el 7 septiembre de 2021]. Disponible en:

<https://www.sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S235291482030201X>

⁶¹ WANG, Qingxiang. YANG, Huanxin. YU, Yanhong. Facial Expression video analysis for depression detection in Chinese patients. [Sitio Web]. [Consultado el 23 de agosto de 2021]. disponible en:

<https://www.sciencedirect.com/science/article/abs/pii/S1047320318302761?via%3Dihub>

videos a partir de la extracción de figuras geométricas y de técnicas de clasificación como Radial Basis Function (RBF); así como las siguientes medidas de desempeño: accuracy, recall(sensitivity), positive predictive value (precision) and F1 score para evaluar la efectividad del método en cuanto a la detección de la depresión.

Fahima TabassumaMd.Imdadul Islama, RisalaTasin KhanbM.R. Aminc, en la investigación titulada “Human face recognition with combination of DWT and machine learning”⁶² proponen la unión de varas técnicas ya empleadas en el reconocimiento facial con el fin de generar una detección facial más confiable y precisa, siendo reconocido el rostro en 2D. Además, se expone el hallazgo en el cual se afirma que el método combinado que emplea redes convolucionales, linear discriminant analysis, principal component analysis, discrete wavelet transform en un sistema difuso radica su nivel de precisión en la calidad de la imagen y la diversidad del dataset conformado, obteniendo resultados que se encuentran en el rango entre 89,56% a 93,34% de precisión.

En el estudio titulado " Facial recognition using Haar cascade and LBP classifiers" realizado por integrantes del departamento de ciencias de la computación y del instituto de tecnología y gerencia de China⁶³. Presenta el análisis comparativo de dos métodos de reconocimiento facial, Haar cascade y Local Binary Pattern, en donde explican diferentes modos de realizar la detección facial y exponen que la precisión del método de Haar Cascade es mejor en comparación con la del otro método; pero resaltan a su vez, que esta emplea mayor tiempo de ejecución versus lo que tarda en ejecutarse el método de LBPH.

4.4 HERRAMIENTAS UTILIZADAS

A continuación, se especifican las características de las herramientas tanto de software como de hardware empleadas en el desarrollo del proyecto

⁶² TABASSUMA. ISLAMA, TASIN, AMINC. ScienceDirect. Human face recognition with combination of DWT and machine learning. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S1319157819309395>

⁶³ SHETTY, BHOMIKA, DEEKSHA, JEEVAN REBEIRO, RAMYASHREE. ScienceDirect. KeAi Chinese Roots Global Impact. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://doi.org/10.1016/j.glt.2021.08.044>

4.4.1 Software

4.4.1.1 Google Colab

Es un producto de Google Research, el cual permite escribir y ejecutar código Python en el navegador, está pensado para tareas de aprendizaje automático, análisis de datos y educación, adicionalmente, ofrece sin costo uso individual de GPU⁶⁴.

4.4.1.2 Sublime Text

Según Python⁶⁵ Es un excelente editor de texto que permite manejar un entorno de programación para Python facilitando el manejo de varios archivos relacionados con el proyecto. Sus características de resaltado de sintaxis, autocompletado y búsqueda entre otras permiten solventar las necesidades que surgen en la codificación, así como la compilación a partir de licencia gratuita.

4.4.1.3 Google Drive

Servicio gratuito de Google que proporciona opciones de almacenamiento flexible que pueden ser administradas fácilmente por los usuarios, permitiendo uso compartido y colaboración en los mismos⁶⁶. Dentro del paquete de Google drive, se encuentran incluidos los servicios de:

4.4.1.3.1 Hojas de cálculo

Diseñado para organizaciones ágiles, permiten el trabajo colaborativo y la compatibilidad con archivos y sistemas externos como Excel, trabajando con bases de datos y proporcionando diversas operaciones por celdas⁶⁷.

⁶⁴GOOGLE RESEARCH. Colaborarity. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://research.google.com/colaboratory/intl/es/faq.html>

⁶⁵ PYTHON. Miscelánea. Configurar Sublime Text para programar en Python. [Sitio Web]. [Consultado 12 de febrero de 2022]. Disponible en: <https://www.programaenpython.com/miscelanea/configurar-sublime-text-para-programar-en-python/>

⁶⁶GOOGLE WORKSPACE. Drive. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://workspace.google.com/intl/es-419/products/drive/>

⁶⁷ GOOGLE Workspace. Hojas de cálculo . [Consultado el 10 de mayo de 2022]. Disponible en: <https://workspace.google.com/intl/es-419/products/sheets/>

4.4.1.3.2 Documentos

Producto de Google que permite la edición de documentos de texto en forma colaborativa, permitiendo la exportación en formatos .docx, .pdf, .odt, .rtf, .txt o .html. Adicionalmente, proporciona de forma gratuita un historial ilimitado de versiones del documento⁶⁸.

4.4.1.4 Github

Ofrece diversos productos tanto gratuitos como de pago para clasificar y colaborar con desarrollo de código, en la versión gratuita se cuenta con ilimitadas colaboraciones de código, incluido el soporte de la comunidad, autenticación en 2 factores y 500 MB de almacenamiento⁶⁹.

4.4.1.5 Anaconda

Es un repositorio que cuenta con más de 8000 paquetes de aprendizaje automático y ciencia de datos de código abierto, que lo hace un sistema de gestión y paquetes de código que se puede ejecutar en Windows, macOS y Linux, además de que instala y ejecuta, también actualiza los paquetes y sus dependencias, principalmente está diseñado para Python, pero puede empaquetar y distribuir en cualquier idioma⁷⁰.

4.4.1.6 Python 3.10.4

Intérprete del lenguaje de programación Python, distinguido por ser la versión principal más reciente que contiene nuevas funciones y optimizaciones. Adicionalmente corrige la regresión introducida por BPO 46968 y contiene la consola que permite desempacar las librerías necesarias para compilar con el lenguaje de Python⁷¹

⁶⁸ GOOGLE Workspace. Documentos . [Consultado el 10 de mayo de 2022]. Disponible en: <https://workspace.google.com/intl/es-419/products/docs/>

⁶⁹ GITHUB Docs. Productos de GitHub. [Consultado el 10 de mayo de 2022]. Disponible en: <https://docs.github.com/es/get-started/learning-about-github/githubs-products>

⁷⁰ ANACONDA. Products.[Sitio web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://www.anaconda.com/products/distribution>

⁷¹ PYTHON. Pitón 3.10.4. [Sitio Web]. [Consultado 28 de marzo de 2022]. Disponible en: <https://www.python.org/downloads/release/python-3104/>

4.4.2 Hardware

4.4.2.1 Computador 1

- Marca: *HP*
- Modelo: *Pavilion -15- cw0003la*
- Tipo de Sistema: *x64*
- Procesador: *AMD Ryzen (™) 5 2 procesadores principales y 4 lógicos*
- CPU: *2 GHz*
- RAM: *8 GB*
- Disco: *512 GB SSD*

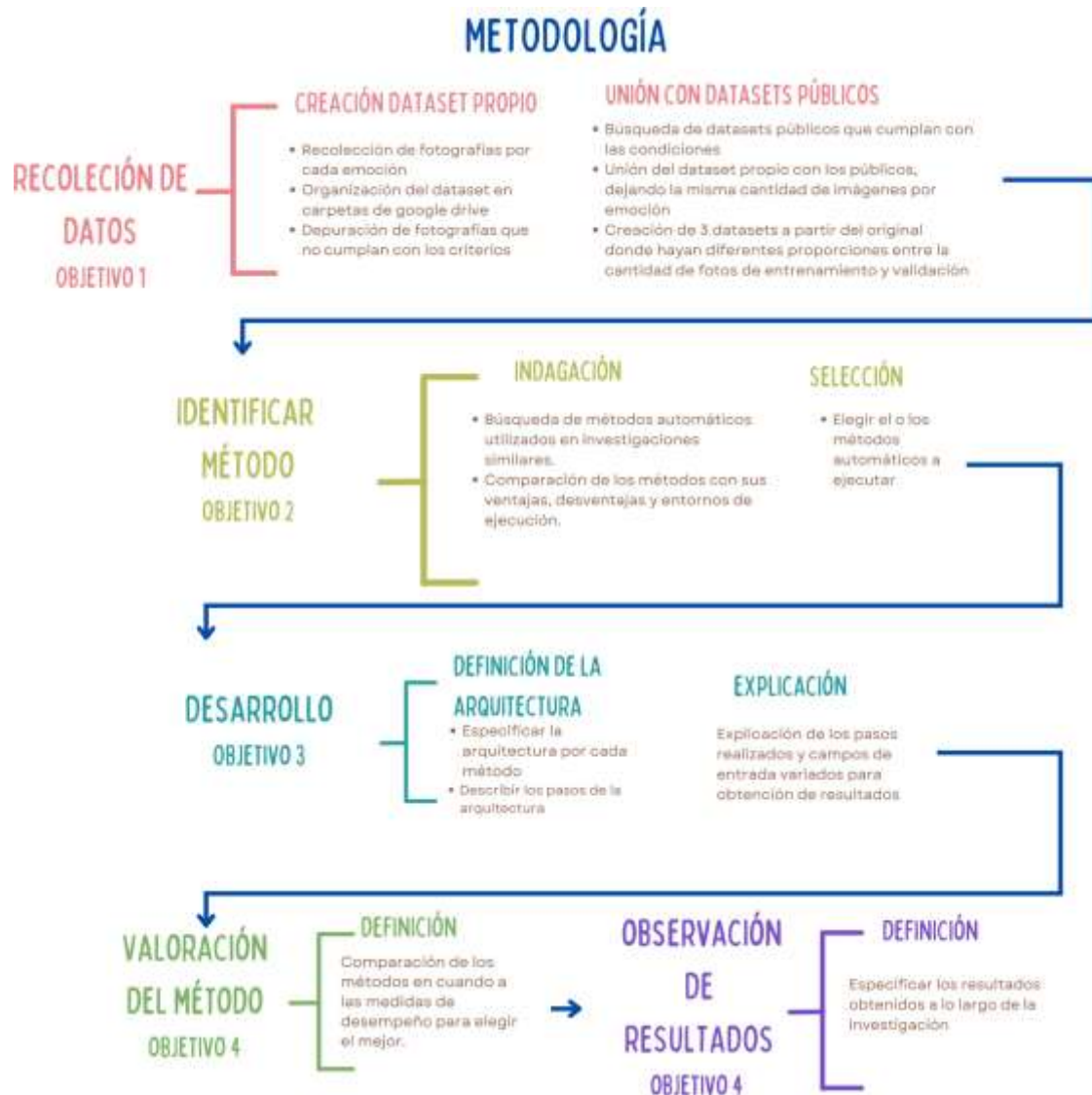
4.4.2.2 Computador 2

- Marca: *Acer*
- Modelo: *Aspire E5-476*
- Tipo de Sistema: *x64*
- Procesador: *Intel(R) Core (™) i5 7th Gen - 2 procesadores principales y 4 lógicos*
- CPU: *2.50GHz*
- RAM: *12GB*
- Disco: *250 GB SSD*

5. METODOLOGÍA

Para la metodología del presente trabajo de investigación, se propone una metodología definida en cinco fases de desarrollo como se evidencia en la figura 6 Metodología propuesta:

Figura 6: Metodología propuesta



Fuente: Elaboración propia

Con base a lo anterior, se procede a especificar cada una de las fases a realizar:

5.1 FASE 1: RECOLECCIÓN DE DATOS

Esta fase se centra en la construcción de un dataset, a partir de una muestra de personas entre los 18 y 29 años, quienes después de firmar el consentimiento para el uso de su imagen que fue previamente revisado por la doctora Eliana Ivette Ortiz Garzón de parte de la facultad de psicología de la Universidad Católica de Colombia para esta investigación, realizaron el envío de 6 fotos de sus rostros representando las emociones propuestas por Ekman : ira, asco, miedo, alegría, tristeza y sorpresa. Adicionalmente, con el fin de aumentar la cantidad de imágenes en el dataset, se indaga en la red con el fin de encontrar dataset públicos con características similares con el objetivo de unirlos al dataset propio y obtener un balanceo en el mismo, es decir, que todas las emociones presenten la misma cantidad de fotografías.

5.2 FASE 2: IDENTIFICAR MÉTODO

Con el fin de seleccionar los modelos existentes y comúnmente utilizados en investigaciones similares, se realiza una amplia investigación, describiendo las ventajas, desventajas y entorno de ejecución de cada modelo, para poder realizar una selección de modelos para su posterior puesta en marcha con el fin de elegir el modelo que mejor porcentaje de certidumbre brinde.

5.3 FASE 3: DESARROLLO

Teniendo en cuenta, los métodos seleccionados de la fase anterior, se procede a especificar los componentes de la arquitectura que cada uno de ellos emplea, detallando las características del manejo de los datos por cada una de las fases del proceso, las variables, técnicas y procedimientos necesarios para la obtención de las etiquetas de clasificación resultantes de los rostros de entrada según las seis emociones; así como los valores de las métricas precisión, recall, f1-score, support y accuracy de los métodos generados por cada muestra del dataset (dataset7030, dataset7525, dataset8020) junto con el estudio del ajuste de parámetros realizado según las técnicas que cada uno de ellos emplea.

5.4 FASE 4: VALORACIÓN DEL MÉTODO

En esta fase, se realiza la comparación de los resultados obtenidos en el proceso anterior junto con el análisis de los mismos, así como la determinación del algoritmo con mejor desempeño que obtiene a las salidas esperadas.

5.5 FASE 5: OBSERVACIÓN DE RESULTADOS

Por último, se realiza el estudio y descripción de los resultados obtenidos durante todo el proceso de investigación, la determinación de los aportes y oportunidades de trabajos futuros que surgen a partir de los análisis realizados

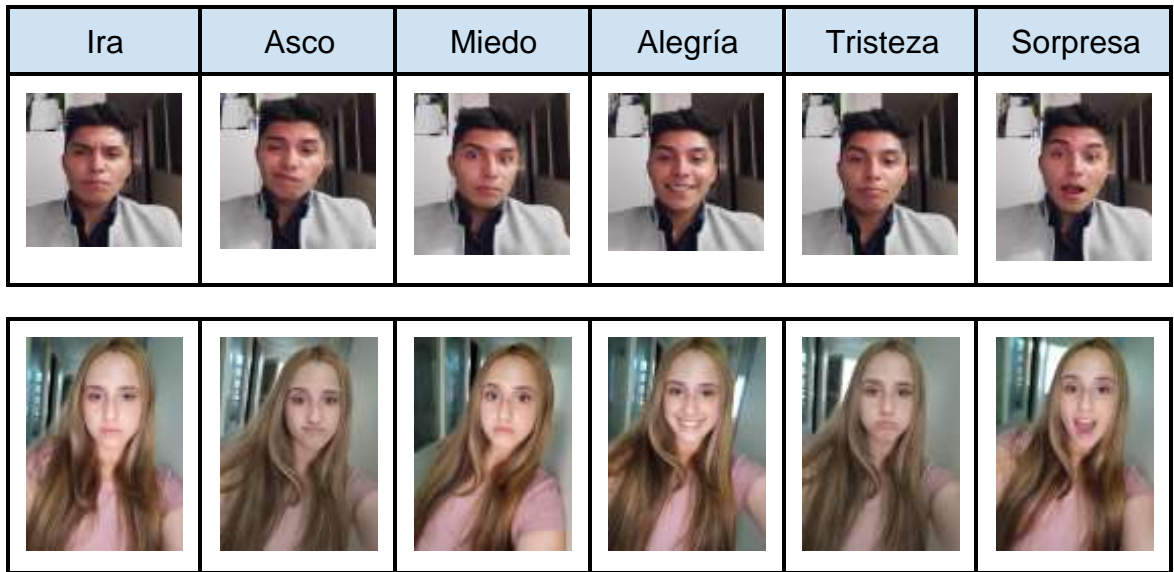
6. CONSTRUCCIÓN DEL DATASET

La construcción del dataset de entrenamiento y validación, contó con la participación voluntaria de 150 adultos jóvenes entre los 18 a 29 años, quienes después de firmar el consentimiento para el uso de su imagen que fue previamente revisado por la doctora Eliana Ivette Ortiz Garzón de parte de la facultad de psicología de la Universidad Católica de Colombia para esta investigación, realizaron el envío de 6 fotos por persona, en donde se evidenciaban las emociones propuestas por Ekman: ira, asco, miedo, alegría, tristeza y sorpresa.

Para las mismas, se solicitó evidenciar el rostro en una vista frontal, sin taparlo con algún objeto como tapabocas, cabello, brazos entre otros. Además, se permitió el uso de gafas debido a que el método realizado, permite el reconocimiento facial con presencia o no de este objeto; donde conviene resaltar, que para el 2019 según la OMS⁷², alrededor de 2200 millones de personas en el mundo presentan deficiencia visual o ceguera, por lo que a su vez hacía importante que la persona al momento de ser examinada por el sistema, tuviera la posibilidad de en caso de tenerlas, hacer uso de sus gafas de forma habitual, sin que el reconocimiento de sus expresiones faciales se vean alteradas por la falta de uso de las mismas. A continuación, se ejemplifica en la tabla 1 el tipo de fotos recibidas por emoción:

⁷² OMS. La OMS presenta el primer Informe mundial sobre la visión. [Consultado el 05 de marzo de 2022]. Disponible en : <https://www.who.int/es/news/item/08-10-2019-who-launches-first-world-report-on-vision>

Tabla 1. Ejemplos de fotos recibidas para el dataset



Fuente: Elaboración propia

Dando así un total de 900 imágenes recogidas de manera directa; posterior a la recepción de las fotografías se realizó la división de las mismas por carpetas en la herramienta de uso gratuito Google Drive para cada emoción, teniendo la separación de las fotografías se realizó una depuración del material recibido, descartando las fotografías que no corresponden con las indicaciones dadas, eran borrosas o no tenían una expresión acorde a la emoción solicitada y ubicándolas según correspondía a la expresión realizada. De este modo el dataset quedó compuesto por 880 fotografías a color, distribuidas como se muestra en la tabla 2:

Tabla 2. Consolidado de fotos aceptadas para el dataset propio.

Emoción	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Total
Fotografías	165	124	144	158	160	129	880

Fuente: Elaboración propia

Con el fin de poder aumentar la cantidad de imágenes en el dataset y de este modo hacer que el método se pudiera entrenar y validar más veces se procedió a realizar consultas para encontrar datasets en a web, con condiciones similares, siendo indiferente la cantidad de imágenes que tuvieran o si se encontraban en escalas de grises o a color, para ello se hizo uso de Dataset Search, una herramienta gratuita de Google, en la que se facilita la búsqueda de datasets por palabras, permitiendo filtrar por últimas actualizaciones, formato de descarga, derechos de uso, tema y si es gratuito o no.

Para esta investigación, se decidió utilizar datasets recientes, en una ventana de tiempo de 3 años, debido a que la calidad de estas imágenes responde a las necesidades de las redes neuronales, convolucionales, adicional que fuera gratuito, con el fin de reducir costos, por consiguiente se realizó la elección de los datasets: Fer-2013⁷³ y Face expression recognition dataset⁷⁴ de Kaggle y Kanade Dataset CK⁷⁵, de estos datasets, se eligieron aleatoriamente la cantidad de fotografías necesarias para completar un total de 1320 por cada emoción, con el fin de unirlo al dataset realizado previamente, a continuación en la tabla 3 se especifica la cantidad de imágenes extraídas de cada uno:

Tabla 3. Consolidado de imágenes extraídas de datasets externos

Dataset/ Emociones	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Total de fotos extraídas
Fer-2013	700	700	700	700	700	700	4200
Face expression recognition	400	400	400	400	400	400	2400
Kanade Dataset CK	55	96	76	62	60	91	440

Fuente: Elaboración propia

Realizando una mezcla aleatoria de los dataset se obtuvo como resultado 1320 imágenes por emoción con el fin de que todas las emociones se entrenan y validaran con la misma cantidad de data, siendo 7920 imágenes en total para el dataset.

Para tener diferentes proporciones entre las imágenes destinadas entrenamiento y validación, se realizaron 3 datasets dividiendo de forma aleatoria las imágenes dejando las proporciones más comúnmente utilizadas para este tipo de investigaciones: 80% y 20%, 75% y 25%, 70% y 30% de entrenamiento y validación respectivamente, a continuación en la tabla 4 se presenta el consolidado de la composición de cada uno de los datasets anteriormente mencionados:

⁷³ MANAS SAMBARE. Fer-13. [Consultado el 03 de marzo del 2022]. Disponible en: <https://www.kaggle.com/datasets/msambare/fer2013>

⁷⁴ JONATHA OENIX. Face expression recognition dataset. [Consultado el 03 de marzo del 2022]. Disponible en: <https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset>

⁷⁵ Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

Tabla 4. Conformación de los 3 diferentes datasets

Nombre	Porcentaje entrenamiento	Cantidad de fotos por emoción de entrenamiento	Porcentaje validación	Cantidad de Fotos por emoción de validación
Dataset7030	70%	924	30%	396
Dataset7525	75%	990	25%	330
Dataset8020	80%	1056	20%	264

Autor : Elaboración propia

Partiendo de lo explicado anteriormente, sobre la importancia de estudio de reconocimiento de emociones en personas con problemas visuales, que manejan accesorios como las gafas; se muestra la cantidad de fotos de personas que las utilizaron durante el proceso de toma de sus fotos por cada categoría, teniendo en cuenta la conformación expuesta en la Tabla 4 para los tres datasets:

Tabla 5. Fotografías de personas con Gafas incluidas por categoría en el dataset 7030

Dataset7030	GAFAS			
	Train		Validation	
	Angry	54	Angry	24
	Disgust	50	Disgust	25
	Fear	63	Fear	16
	Happy	92	Happy	31
	Sad	35	Sad	11
	Surprise	54	Surprise	20

Autor: Elaboración propia

En la tabla 5, se presenta la cantidad de fotos que se obtuvieron de personas que implementaron el uso de las gafas dentro de la muestra 7030 por cada una de las emociones, a partir de la cual, se tiene presente la que la categoría que registra mayor cantidad de fotos con gafas es la categoría "Happy" tanto en train como en validación.

Tabla 6. Fotografías de personas con Gafas incluidas por categoría en el dataset 7525

Dataset7525	GAFAS			
	Train		Validation	
	Ira	69	Ira	22
	Asco	60	Asco	25
	Miedo	51	Miedo	11
	Alegría	83	Alegría	25
	Tristeza	29	Tristeza	12
	Sorpresa	51	Sorpresa	24

Autor: Elaboración propia

En la tabla 6, se presenta la cantidad de fotos que se obtuvieron de personas que implementaron el uso de las gafas dentro de la muestra 7525 por cada una de las emociones, donde para la subdivisión de esta muestra en train la categoría que más cuenta con fotos de la la emoción con gafas es “Happy” y para el de validación es “Happy” y “Disgust”

Tabla 7. Fotografías de personas con Gafas incluidas por categoría en el dataset 8020

Dataset8020	GAFAS			
	Train		Validation	
	Ira	85	Ira	18
	Asco	56	Asco	19
	Miedo	50	Miedo	12
	Alegría	106	Alegría	23
	Tristeza	36	Tristeza	7
	Sorpresa	68	Sorpresa	21

Autor: Elaboración propia

En la tabla 7, se presenta la cantidad de fotos que se obtuvieron de personas que implementaron el uso de las gafas dentro de la muestra 8020 por cada una de las emociones, donde para la subdivisión de esta muestra en train la categoría qué más cuenta con fotos de la la emoción con gafas es “Happy” y para el de validación es “Happy” y “Surprise”.

7. SELECCIÓN DE MÉTODO AUTOMÁTICO

Para la selección del método, se tuvo en cuenta las limitaciones que se tenían para la ejecución de los mismos, como era la ausencia de GPU local apta para el procesamiento de imágenes. Para ello, se realizó la investigación correspondiente de algunos métodos usados comúnmente en este tipo de investigaciones, encontrando las ventajas y desventajas de uso, para realizar una selección conforme a las necesidades de la investigación, por lo cual, en la tabla 8 se encuentra el consolidado de lo mismo:

Tabla 8. Métodos comúnmente utilizados

Método	¿Cuándo se usa?	Ventajas	Desventajas	Entorno de ejecución
Secuencial Keras ⁷⁶	Para utilización de diversas capas, en la que para cada salida se obtiene un tensor como resultado	-Una vez construido, se comporta como un modelo funcional de forma más simple, para entrada da una salida. -Permite agregar gran cantidad de capas de entrenamiento	-Necesita la unión consecuente de las capas de entrenamiento para no fallar -No permite un entrenamiento cíclico.	Entornos donde sea posible ejecutar Python : Google Colab, Jupyter, Anaconda.
Modelo Funcional Keras ⁷⁷	Para usos similares a Secuencial, sin embargo, de forma cíclica, modelos con capas compartidas y de múltiples entradas y salidas	-Permite manejar modelos no lineales -Valida el modelo mientras lo define	-No admite arquitecturas de tipo dinámico -Puede necesitar varios métodos personalizados	Entornos donde sea posible ejecutar Python : Google Colab, Jupyter, Anaconda.

⁷⁶ TENSORFLOW. El modelo secuencial. [Sitio Web]. [Citado el 25 de marzo de 2022]. Disponible en: https://www.tensorflow.org/guide/keras/sequential_model?hl=es-419

⁷⁷ TENSORFLOW. La API funcional "Keras" en TensorFlow. [Citado el 25 de marzo de 2022]. Disponible en: <https://www.tensorflow.org/guide/keras/functional?hl=es-419>

DCT por bloques ⁷⁸	Uso para reconocimiento facial.	-Reconocimiento facial según patrones comunes del rostro humano.	-Todas las imágenes del dataset deben estar alineadas, recortadas y con el mismo tamaño y ubicación de la persona - Previo a la aplicación del algoritmo, se debe conocer la ubicación de los ojos y alinear según criterio	-Matlab
Locality Preserving Projections LPP ⁷⁹	Usado para hacer predicciones con base a proyecciones proporcionadas .	-Agrupa los datos según su similitud dejando vecinos, con características similares a los lados. -Preserva al máximo el modelo inicial que crea, es decir, es difícil de corromper.	-Dificultad para hacer uso y recuperación de datos originales. -No es posible el manejo con múltiple cantidad de datos, ya que bajará significativamente su capacidad de relación.	Matlab, entornos donde sea posible ejecutar Python : Google Colab, Jupyter, Anaconda.
EigenFaces en OpenCV ⁸⁰	Cuando se tiene una gran dimensión de datos y se quiere hacer un reconocimiento a partir del conjunto más pequeño de variables no correlacionadas	Encuentra aquella combinación lineal que permite maximizar la variación total de los datos Tiene mejores resultados en el reconocimiento por medio de diferentes intensidades lumínicas	No toma en cuenta la realización de las combinaciones lineales por clase Requiere que las imágenes de entrada sean siempre en escala de grises y que tengan las mismas dimensiones	Anaconda Google Colab Python Local (Cualquier IDE de Python, Java y C++) Multiplataforma (para los sistemas operativos GNU/Linux, Mac OS X, Windows y Android, y

⁷⁸ ROGER GIMEO. Estudio de técnicas de reconocimiento facial. [Citado el 25 de marzo de 2022]. Disponible en:

https://upcommons.upc.edu/bitstream/handle/2099.1/9782/PFC_RogerGimeno.pdf

⁷⁹XIAOFEI He, PARTHA Niyogi . Locality Preserving Projections. [Sitio Web]. [Citado el 25 de marzo de 2022]. Disponible en:

<https://proceedings.neurips.cc/paper/2003/file/d69116f8b0140cdeb1f99a4d5096ffe4-Paper.pdf>

⁸⁰ OPENCV. Face Analysis. cv::face::EigenFaceRecognizer Class Reference. [Sitio Web]. [Citado el 28 de marzo de 2022]. Disponible en: https://docs.opencv.org/4.2.0/dd/d7c/classcv_1_1face_1_1EigenFaceRecognizer.html

				para diversas arquitecturas de hardware como x86, x64 (PC), ARM (celulares y Raspberry Pi)
FisherFaces en OpenCV ⁸¹	Cuando se requiere hacer el reconocimiento entre diferentes clases, agrupando estrechamente por clase y generando más distancia de las demás clases distintas	Permite tener en cuenta más información discriminativa que se pudo haber perdido cuando se emplea PCA	Requiere que las imágenes de entrada sean siempre en escala de grises y que tengan las mismas dimensiones Aprende de una matriz específica por clase, por lo que las características diferenciales entre las demás clases no son tan notorias (Ej la intensidad luminica)	Anaconda GoogleColab Python Local (Cualquier IDE de Python, Java y C++) Multiplataforma(para los sistemas operativos GNU/Linux, Mac OS X, Windows y Android, y para diversas arquitecturas de hardware como x86, x64 (PC), ARM (celulares y Raspberry Pi)
LBPH en OpenCV ⁸²	Cuando no se cuenta con un dataset balanceado y se requiera realizar el reconocimiento a partir de pocas imágenes	Permite evaluar la imagen no como todo un vector de gran dimensión, sino únicamente describir las características locales por cada imagen comparando así cada píxel con su vecino y si su intensidad es mayor lo guarda como 1 de lo contrario lo guarda como 0	Requiere que las imágenes de entrada sean siempre en escala de grises y que tengan las mismas dimensiones Tiene en cuenta solo un área pequeña a partir de un radio fijo	Anaconda Google Colab Python Local (Cualquier IDE de Python, Java y C++) Multiplataforma(para los sistemas operativos GNU/Linux, Mac OS X, Windows y Android, y para diversas arquitecturas de hardware como x86, x64 (PC), ARM (celulares y Raspberry Pi)

Autor: Elaboración propia

⁸¹ OPENCV. Face Analysis. cv::face::FisherFaceRecognizer Class Reference. [Sitio Web]. [Citado el 28 de marzo de 2022]. Disponible en: https://docs.opencv.org/4.2.0/d2/de9/classcv_1_1face_1_1FisherFaceRecognizer.html

⁸² OPENCV. Face Analysis..cv::face::LBPHFaceRecognizer Class Reference. [Sitio Web]. [Citado el 28 de marzo de 2022]. Disponible en: https://docs.opencv.org/4.2.0/df/d25/classcv_1_1face_1_1LBPHFaceRecognizer.html

Con el fin de realizar la ejecución de los métodos que cumplieran con las características necesitadas, se eligió secuencial de Keras, por la posibilidad de composición de diferentes capas, permite generar un modelo robusto pero a su vez de fácil manejo y entendimiento, adicional, como es posible desarrollarlo en Google Colab, el mismo es desarrolla en la GPU remota de forma gratuita, lo que reduce el tiempo de ejecución del mismo, de este modo, se pueden generar diversas pruebas, cambiando algunos parámetros de ejecución, como lo es el dataset y la cantidad de ciclos de desarrollo, mejor conocidos como épocas o en inglés Epochs.

Por otro lado, dado a los modelos que vienen preconfigurados en la librería de OpenCV, principalmente diseñados para procesadores Intel, que permite no sobrecargar el procesamiento del dispositivo, dado a que la configuración de la red neuronal se hace internamente con los parámetros de las características de se requieren para el reconocimiento de emociones. Se realizó el estudio a través de aquellos que no requieran de configuración de GPU local, EigenFaces, FisherFaces y LBPH, donde se tuvo en cuenta que para la creación de cada método el ajuste los parámetros sobre los cuales se realiza elreconocimiento facial de entrada, con el fin de aplicar el mismo formato para todas las imágenes, y el ajuste de los parámetros de reconocimiento, que serán una parte determinante para la determinación de resultados.

8. DISEÑO DE LOS MÉTODOS

A continuación, para el diseño de los métodos se seleccionaron las siguientes técnicas: Secuencial con KERAS, Secuencial con OpenCV, de los cuales se describen sus arquitecturas a continuación:

8.1 MÉTODO SECUENCIAL CON KERAS

Para la ejecución de este método se utilizó una arquitectura, la cual se encuentra dividida en 3 partes, para cada una de sus etapas como se visualiza en la figura 7, las cuales se detallan más adelante.

Figura 7: Arquitectura método secuencial

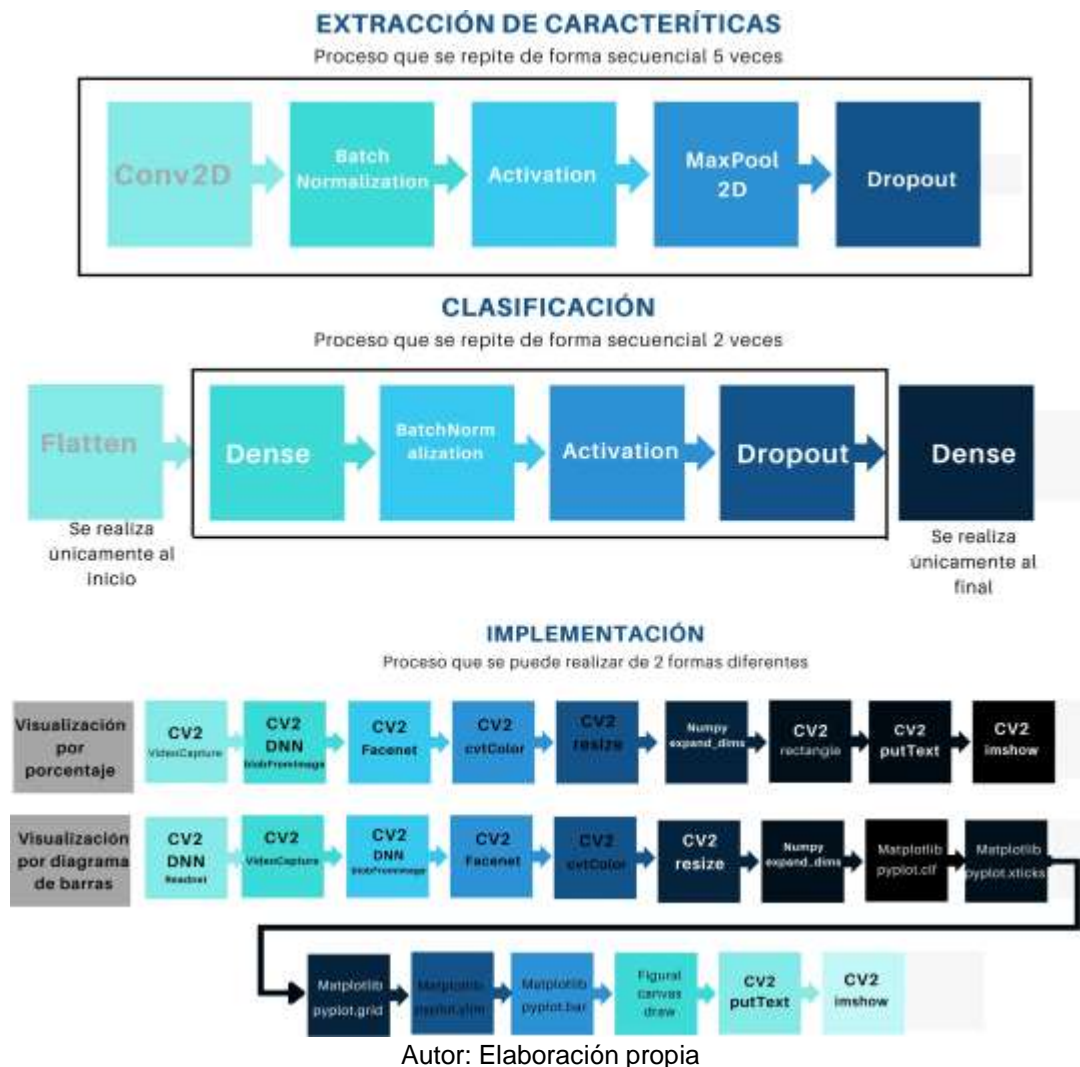


Elaboración propia

Para cada una de estas fases: extracción de características, clasificación e implementación de esta técnica, se define con más detalle cada una de estas etapas en la figura 8.

Los cuales se seleccionaron considerando, las diferentes librerías y procesos de Python tanto para su implementación en Google Colab, como para la distribución libre de Anaconda.

Figura 8: Definición de las etapas de la arquitectura del método secuencial



A continuación, se procede a especificar el proceso realizado en las tres etapas de esta arquitectura:

8.1.1 Extracción de características

En esta etapa la red neuronal convolucional, se centra en la extracción de características, haciendo uso del módulo Sequential() de keras, en específico las capas y procesos : Conv2D, BatchNormalization, Activation, MxPool2D y Dropout; por medio de la misma se realiza un proceso de entrenamiento del algoritmo, en donde se especifican la emociones, se normaliza la salida de la capa, posteriormente se transforma la entrada de tal forma que el

modelo empieza a crear predicciones sobre lo aprendido, más adelante se realiza la escalación de la imagen en lotes de 2x2 para su clasificación por píxeles y finalmente se generan diferentes nodos de salida con el fin de reducir el sobre ajuste.

8.1.2 Clasificación

En esta fase el secuencial de KERAS que se implementó en Google Colab, debido a la ventaja del uso de GPU gratuito, haciendo implementación de las capas y procesos de python : Flatten, Dense, Batch Normalization, Activation y Dropout, inicialmente se aplica una conversión de dimensión a la fotografía con el fin de poder realizar una estimación de acuerdo a los píxeles encontrados, posteriormente se realiza una normalización de la imagen y se genera la clasificación de las emociones, para finalmente realizar una evaluación de su precisión al momento de clasificar.

8.1.3 Implementación

Esta última fase se realiza en Python en la distribución de Anaconda, instalado previamente en el dispositivo. Se da uso de diferentes librerías de python, en donde inicialmente se especifica el arreglo de las emociones a utilizar (ira, alegría, asco, tristeza, sorpresa, miedo); por medio de CV2 realiza el reconocimiento del rostro, que se pasó a escala de grises y se redimensiona a 48 x 48, con el fin de que tenga las mismas condiciones de entrenamiento y validación del modelo, para posteriormente realizar la comparación de la expresión facial con lo registrado en el modelo, se fija un umbral del 40% de aproximación para determinar el acercamiento a cada emoción contrastado con el modelo entrenado; permitiendo de esta manera 2 tipos de visualizaciones diferentes, las cuales se especifican a continuación:

8.1.4 Porcentual

Para este tipo de visualización, elige la emoción que mayor porcentaje de acercamiento presente, finalmente realiza un rectángulo alrededor del rostro con CV2, mostrando en pantalla la emoción y el porcentaje de aproximación a la misma, a continuación, en la figura 9 se ejemplifica la visualización porcentual:

Figura 9: Ejemplo visualización porcentual

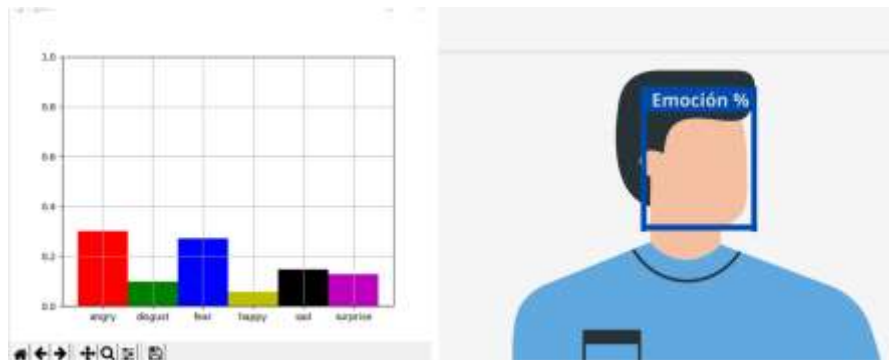


Autor: Elaboración propia

8.1.5 Diagrama de barras

En este tipo de visualización, con matplotlib, realiza la ilustración de un diagrama de barras en donde especifica el acercamiento a cada emoción, indicando el valor de cada una y el nombre de la misma, a continuación, en la figura 10 se ejemplifica la visualización por diagrama de barras:

Figura 10: Ejemplo visualización porcentual

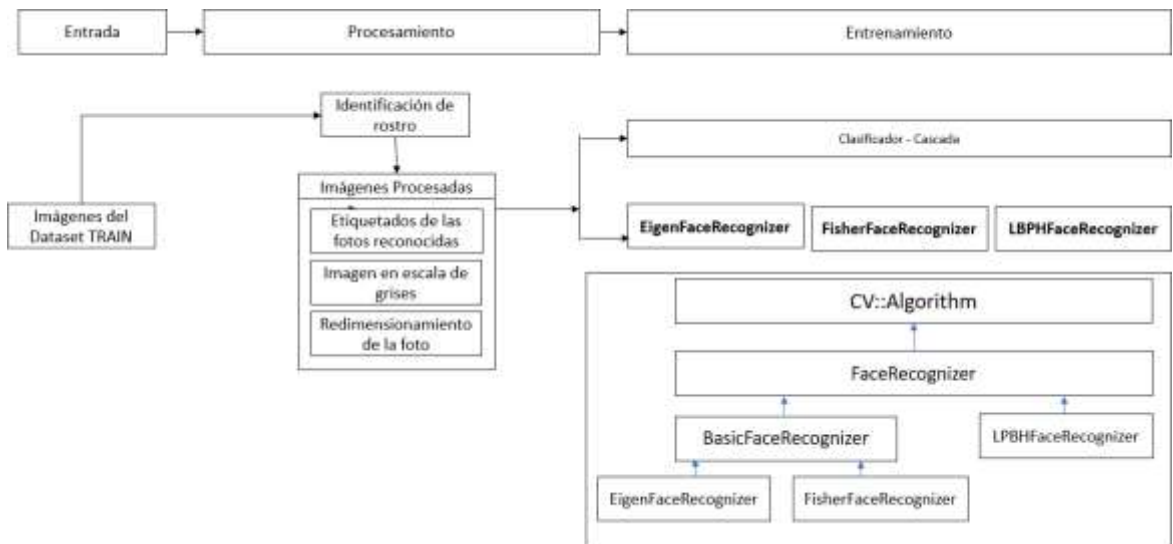


Autor: Elaboración propia

8.2 MÉTODOS CON OPENCV

Por otro lado, se seleccionaron tres técnicas de extracción de características: EigenFaces, FisherFaces y LBPH, las cuales se implementaron en la librería de OpenCV y siguen la siguiente arquitectura compuesta por 3 fases principales (Entrada, procesamiento y entrenamiento), como se describe en la figura 11:

Figura 11: Arquitectura de procesos para los métodos generados por OpenCV



Autor: Elaboración propia

Como se observa en la figura 11, los tres métodos empleados requieren de un mismo formato de imágenes que se obtiene a partir de la fase de entrada y procesamiento, adicionalmente, que los tres métodos requieren de las fases de CV:Algorithm y CV::face:FaceRecognizer y que tanto el método de EigenFaces como el de FisherFaces conllevan una capa adicional de procesamiento respecto a LBPH

8.2.1 Entrada

En la fase de entrada, se realiza la configuración de lectura de las imágenes designadas de entrenamiento, a partir del uso de las funcionalidades de la librería “os” de Python, así como la inicialización del arreglo que contiene las categorías de clasificación.

8.2.2 Procesamiento

En esta fase, se tuvo en cuenta las condiciones que los tres métodos (EigenFaces, FisherFaces y LBPH) requieren validar para el manejo de los datos dentro de las configuraciones internas que cada uno conlleva; de esta forma, a cada una de las imágenes de entrada se le realizó el redimensionamiento y la conversión a escala de grises. Además, con implementación del modelo pre entrenado de haarcascade_frontalface_default.xml y el estudio de diferentes parámetros para el reconocimiento facial se realizó el recorte de las imágenes respecto al momento en el cual detectara un rostro que pasaría a ser el dataset sobre el cual se entrenaría posteriormente cada uno de los modelos.

8.2.3 Entrenamiento

8.2.3.1 cv::face::EigenFace

Este método implementa parámetros internos, como el número de componentes, que corresponde al número de Eigenfaces que se determina para realizar el análisis de componentes principales (PCA), el umbral que corresponde al límite definido para la predicción, los eigenvalores correspondientes a los vectores en donde se puede hacer la transformada del vector ordenados de forma descendente, la media calculada de los datos de entrenamiento, las proyecciones y las categorías entre las cuales se reconocerá un rostro que se le deberán enviar al método. Además, es importante resaltar que este método asume que los datos de entrenamiento y testing ya están balanceados, se encuentran transformados en escala de grises y tienen el mismo tamaño, dado a que en caso de que no lo estén envía una excepción sobre los datos de entrada

8.2.3.2 cv::face::FisherFace

FisherFaces cuenta con la diferencia que implementa el análisis discriminante lineal por medio de Fisher's Linear Discriminant (FLD) que busca encontrar el vector que maximiza la separación entre las clases de los datos de entrada para de esta manera encontrar las diferencias más notorias entre cada clase. Además, una vez el método determina el discriminante lineal, maximiza la distancia de relación entre las clases y minimiza la varianza de los datos analizados dentro de cada una de las categorías. Así como el método anteriormente explicado, asume que las entradas se encuentran en escalas grises, en el mismo tamaño y que el dataset se encuentra balanceado, con los mismos parámetros que recibe el método de EigenFaces.

8.2.3.3 cv::face::LBPH

La técnica de extracción de características Local Binary Patterns, emplea parámetros como el radio, con el cual se construye el patrón circular binario local que además define dentro de que radio intentará obtener la mayor cantidad de datos de la imagen respecto al centro, la variable neighbors, determina la cantidad de puntos mínimos con los cuales se construye el patrón circular binario partiendo de un mínimo definido de 8 puesto que entre más puntos se empleen para la construcción del círculo mayor será el costo computacional. Adicionalmente se especifican dos parámetros que corresponden al número de celdas, de dimensiones del vector, el umbral de predicción y las categorías para clasificación; trabaja con entradas de imágenes en escala de grises como único requerimiento de formato para estos archivos de entrada.

8.2.3.4 cv::face::BasicFaceRecognizer

Realiza el procesamiento principal de los eigenvalores y los eigenvectores acomodándolos según se requiera por el método que se está procesando (EigenFaces o FisherFaces). Además, obtiene el número de componentes, el umbral y las proyecciones a partir de los parámetros encontrados por cada método.

8.2.3.5 cv::face::FaceRecognizer

Establece un constructor virtual en el cual se reciben los parámetros necesarios para todos los métodos. Estableciendo de esta manera, el número de componentes y umbral para el caso de EigenFaces y FisherFaces, y la mejor distancia entre los puntos mínimos establecidos dentro del círculo para el caso de LBPH.

8.2.3.6 cv::face::Algorithm

En esta capa se realizan el balanceo de los algoritmos realizados, además establece la clase hash para los datos guardados de imágenes, se establecen las clases que marcarán la identificación de las emociones a partir de los datos configurados.

8.3 Ajuste de los parámetros de los métodos seleccionados

8.3.1 Método secuencial de Keras

Para la ejecución del método y se eligió el entorno de ejecución Google Colab, debido a que de forma gratuita se podía realizar uso de máquinas virtuales con GPU en las que se podía realizar el proceso de forma más rápida, sin importar las especificaciones del computador o si contaba con tarjeta gráfica, ya que Google Colab no consume los recursos de la máquina y se puede ejecutar desde otros dispositivos como celulares y tablets.

Con el fin de encontrar los mejores resultados posibles del modelo, se variaron las épocas desde 10 hasta 150, de este modo, se seleccionaron 11 pruebas con diferentes muestreos del dataset, con el fin de poder evaluar cual dataset brindaba mejores resultados, como se indica en la tabla 9.

Tabla 9. Parámetros de ejecución método secuencial

Pruebas	DataSet	Cantidad de épocas
1	Dataset7030	30
2	Dataset7030	60
3	Dataset7030	90
4	Dataset7525	30
5	Dataset7525	60
6	Dataset7525	90
7	Dataset8020	30
8	Dataset8020	60
9	Dataset8020	90
10	Dataset8020	120
11	Dataset8020	150

Autor: Elaboración propia

8.3.2 Ajuste de métodos con OpenCV

En este caso, se realizó la evaluación en primera instancia de los parámetros de las dimensiones de la imagen que tendría en cuenta el modelo clasificador de haarcascade, a partir de una submuestra que contó con el 5% del dataset (396 imágenes) empleando los parámetros que se especifican en la tabla 10 partiendo del mínimo establecido en las investigaciones realizadas para el estado del arte (30) y el máximo establecido de(950)

Tabla 10. Parámetros de ejecución modelo haarcascade

Prueba	minSize	maxSize
1	100	1000
2	30	1300
3	30	950

Autor: Elaboración propia

A partir de los cuales se analizó según la cantidad de imágenes que entraban, cuantas reescribía por categoría y en cuantas escribía una imagen que no contara con un rostro. Posteriormente, se realizó el análisis de los métodos deEigenFaces, FisherFaces y LBPH, realizando la variación del parámetro que permite determinar el rango sobre el cual se almacenará información sobre los rostros y estudiando elcomportamiento de esta variación por cada una de las muestras según se detalla en la tabla 11.

Tabla 11. Parámetros de ejecución modelos diseñados con OPENCV

Parámetro	Método	Muestra
minNeighbors = 8	LBPH	7030
minNeighbors = 5	LBPH	7030
minNeighbors = 3	LBPH	7030
minNeighbors = 8	LBPH	7525
minNeighbors = 5	LBPH	7525
minNeighbors = 3	LBPH	7525
minNeighbors = 8	LBPH	8020
minNeighbors = 5	LBPH	8020
minNeighbors = 3	LBPH	8020
minNeighbors = 8	EigenFaces	7030

minNeighbors = 5	EigenFaces	7030
minNeighbors = 3	EigenFaces	7030
minNeighbors = 8	EigenFaces	7525
minNeighbors = 5	EigenFaces	7525
minNeighbors = 3	EigenFaces	7525
minNeighbors = 8	EigenFaces	8020
minNeighbors = 5	EigenFaces	8020
minNeighbors = 3	EigenFaces	8020
minNeighbors = 8	FisherFaces	7030
minNeighbors = 5	FisherFaces	7030
minNeighbors = 3	FisherFaces	7030
minNeighbors = 8	FisherFaces	7525
minNeighbors = 5	FisherFaces	7525
minNeighbors = 3	FisherFaces	7525
minNeighbors = 8	FisherFaces	8020
minNeighbors = 5	FisherFaces	8020
minNeighbors = 3	FisherFaces	8020

Autor: Elaboración propia

9. EVALUACIÓN DE MEDIDAS DE DESEMPEÑO

A continuación, se presentan los resultados de las medidas de desempeño del método secuencial de Keras y con los métodos secuenciales implementados por OpenCV:

9.1 MÉTODO SECUENCIAL DE KERAS

9.1.1 Matriz de confusión de las pruebas

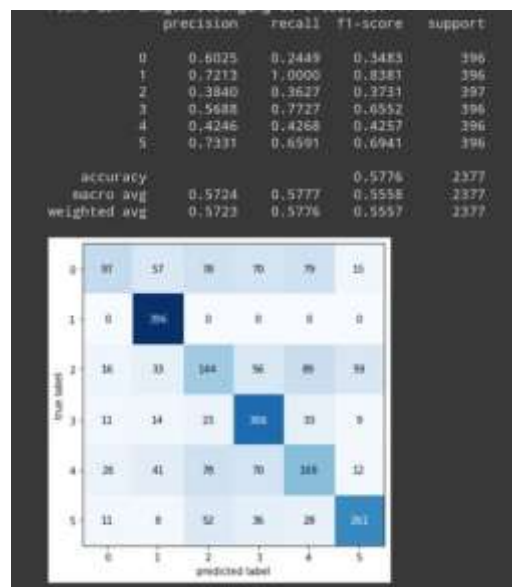
Para la evaluación de las medidas de desempeño del método secuencial de keras, con los diferentes parámetros indicados en el numeral 8.1.4 se realizó

una matriz de confusión para cada una de las pruebas, en donde en el eje X se encuentra “predicted label”, que hace referencia a la clasificación que realizó el algoritmo, y en el eje Y “true label”, que indica la emoción por la cual debía ser clasificado, adicional, se encuentra la representación de emociones de 0 a 5 siendo ira, asco, miedo, alegría, tristeza, sorpresa respectivamente, para cada una, se visualiza: precision, exhaustividad (recall), valor-f1 y soporte (support), haciendo este último referencia a la cantidad de fotografías que fueron clasificadas por el algoritmo. A continuación, se presentan las matrices de confusión correspondientes a la prueba con mejor accuracy para cada uno de los 3 datasets generados: Dataset 7030, Dataset7525 y Dataset8020:

9.1.2 Matriz de confusión Dataset7030

Los mejores resultados obtenidos con el Dataset7030 corresponden a la prueba 3, donde se ejecutó a 90 épocas; debido a que, aunque se aumenta el número de épocas de 60 a 90 en las pruebas un accuracy similar de aproximadamente 55%, lo que indica que el modelo no va a tender a evolucionar y por convergencia será su exactitud en promedio. Adicionalmente, como se observa en la figura 12, se evidencia la alta capacidad de detección correcta de las emociones 1 y 5 correspondientes a asco y sorpresa, ya que la precisión de las mismas se encuentra por encima de un 70%, del mismo modo, es visible en el mapa de calor, por su color más oscuro, que la emoción 1, obtuvo verdadero positivo en el total de las 396 a evaluar, de ahí que su re-call sea de 1. Por otro lado, la emoción 2 correspondiente a miedo, obtuvo los peores resultados, ya que tuvo una precisión del 38%, evidenciado también en su mapa de calor, por lo cual, se entiende que esta última emoción perjudicó de forma significativa el accuracy total de la prueba 3. Por otro lado, con respecto al mapa de calor general de la matriz, se puede observar que para las emociones 0 y 2, correspondientes a ira y miedo, el sistema no ha realizado clasificaciones correctas, ya que en las filas de las mismas se pueden observar tonos azules tenues en las 6 opciones de clasificación posibles.

Figura 12: Matriz de confusión prueba 3 con el Dataset7030 y 90 épocas

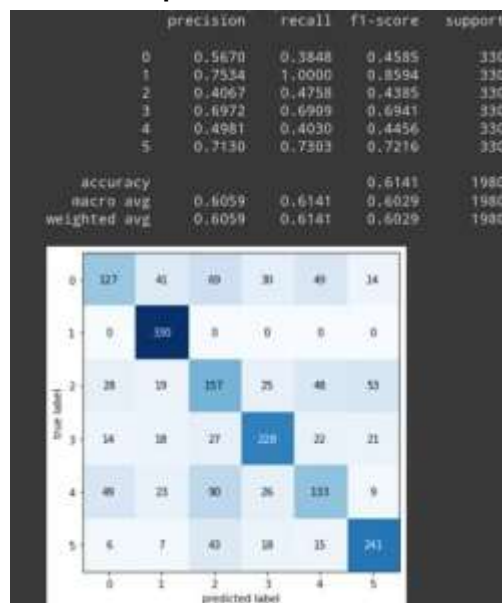


Autor: Elaboración propia

9.1.3 Matriz de confusión Dataset7525

Para los resultados obtenidos en las pruebas ejecutadas con el Dataset7525 se evidencia un incremento en la precisión de detección de emociones proporcional al aumento de épocas de ejecución, siendo la prueba 6 la que mejores resultados obtuvo, teniendo un 61,41% de accuracy total. Adicionalmente, como se puede observar en la figura 13, se tiene que al igual que con el Dataset7030, las mejores métricas de precisión las obtienen las emociones 1 y 5, correspondientes a asco y sorpresa, estando por encima del 70%, por otro lado, las peores métricas las presenta de nuevo la emoción 2, correspondiente a miedo, ya que tiene un 40,67% de precisión, esto se evidencia en el mapa de calor de dicha emoción ya que en su línea horizontal, presenta colores suaves en los 6 campos posibles, ya que de 330 imágenes, solo 157 fueron clasificadas de forma correcta, sin embargo, una situación similar se presenta en la emoción 0, correspondiente a ira, ya que de las 330 imágenes, solo obtuvo 127 en verdadero positivo.

Figura 13: Matriz de confusión prueba 6 con el Dataset7525 y 90 épocas



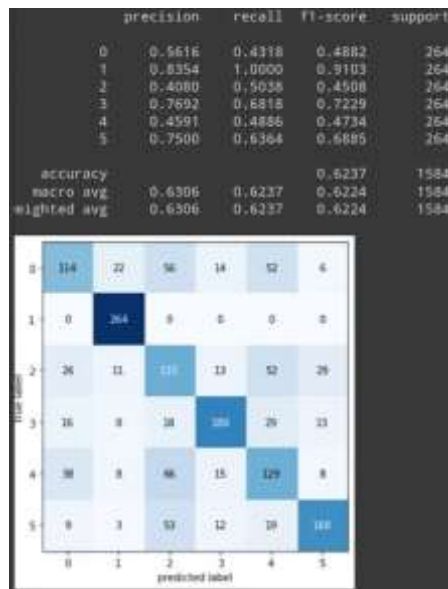
Autor: Elaboración propia

9.1.4 Matriz de confusión Dataset8020

Con respecto a las pruebas ejecutadas con el Dataset8020, se obtuvieron en general mejores resultados en comparación a los otros 2 datasets, por lo cual, se realizaron 2 pruebas adicionales, en donde se aumentó la cantidad de épocas, probando con 120 y 150; sin embargo, se pudo evidenciar una convergencia del accuracy de reconocimiento de emociones, de este modo se logró constatar que los mejores resultados se obtuvieron en la prueba 9, donde se alcanzó una accuracy global del 62,37%.

Como se evidencia en la figura 14, las emociones que mejor métrica de precisión obtuvieron fueron las 1, 3 y 5, correspondientes a asco, alegría y sorpresa, la primera con más de un 80% de precisión y las 2 últimas con más del 75%. Adicionalmente, se puede observar en general un mejor mapa de calor con respecto a las pruebas aplicadas en los otros datasets, ya que vemos que la diagonal principal del mismo presenta los tonos más oscuros, de modo tal, que en la mayor parte de la ejecución del método bajo las condiciones de la prueba 9 se obtuvieron verdaderos positivos.

Figura 14: Matriz de confusión prueba 9 con el Dataset8020 y 90 épocas



Autor: Elaboración propia

Con base a los anterior, se genera la tabla 12, en donde se agrupan los resultados en medida de porcentaje, de precisión para cada emoción por las 11 pruebas realizadas con los 3 datasets generados, adicional, el accuracy de cada prueba:

Tabla 12. Consolidado de porcentaje de precisión por cada emoción por prueba con método secuencial con Keras

Prueba	Ira (%)	Asco (%)	Miedo(%)	Alegría (%)	Tristeza (%)	Sorpresa (%)	Exactitud (Accuracy)
1	57,14	34,49	23,36	46,77	31,46	73,51	41,73
2	55,50	67,79	42,72	60,89	41,63	64,89	57,55
3	60,25	72,13	38,40	56,88	42,46	73,31	57,76
4	42,86	44,70	26,96	61,92	31,73	65,88	47,12
5	46,96	56,89	44,00	77,61	49,02	64,66	57,17
6	56,70	75,34	40,67	69,72	49,81	71,30	61,41
7	34,92	41,18	35,05	62,13	32,63	83,75	46,40

8	61,42	81,15	41,74	75,70	37,72	75,80	60,04
9	56,16	83,54	40,80	76,92	45,91	75,00	62,37
10	53,92	74,16	44,93	68,21	48,50	67,16	61,17
11	53,59	78,34	42,31	66,12	44,83	71,37	60,67

Fuente : Elaboración propia

9.2 MÉTODOS DE OPENCV

9.2.1 Observaciones de estudio de parámetros para clasificador Haarcascade

Partiendo de los parámetros que se plantearon en la sección 8.3.2 se realizó la variación dentro de la submuestra constituida por el 5% del dataset, equivalente a 396 imágenes para el análisis y determinación de los parámetros que se emplearán dentro del modelo de clasificación, como se muestra en la tabla 13:

Tabla 13. Tabulación de resultados según parámetros definidos para el clasificador (mínimo tamaño: 100 y máximo tamaño:1000) aplicado en la submuestra

minSize: 100 maxSize:1000			
Emoción	Imágenes Existentes	Imágenes re escritas apartir de detección	Detecciones erróneas
Ira	396	224	13
Asco	396	227	17
Miedo	396	214	37
Alegría	396	250	8
Tristeza	396	172	12
Sorpresa	396	260	23
TOTAL			110

Fuente : Elaboración propia

A partir de la tabla 13, se analiza el comportamiento del clasificador haarcascade, donde se rescata en primera medida, que el porcentaje de detecciones erróneas llevadas a cabo con los parámetros de mínimo tamaño de 100 y máximo tamaño de 1000 estuvo por debajo del 50% (del 2% al 9%) respecto al total de imágenes de entrada analizadas (396). Además, la categoría con menor cantidad de detecciones erróneas fue la categoría “Alegría”.

Tabla 14. Tabulación de resultados según parámetros definidos para el clasificador (mínimo tamaño: 30 y máximo tamaño:1300) aplicado en dentro de la submuestra

minSize: 30 maxSize:1300			
Emoción	Imágenes Existentes	Imágenes re escritas a partir de detección	Detecciones erróneas
Ira	396	225	15
Asco	396	228	19
Miedo	396	214	34
Alegría	396	250	7
Tristeza	396	173	12
Sorpresa	396	260	23
TOTAL			110

Fuente: Elaboración propia

A partir de la tabla 14, se analiza el comportamiento del clasificador haarcascade, donde se rescata en primera medida, que el porcentaje de detecciones erróneas llevadas a cabo con los parámetros de mínimo tamaño de 30 y máximo tamaño de 1300 estuvo por debajo del 50% (del 2% al 8%) respecto al total de imágenes de entrada analizadas (396). Además, la categoría con menor cantidad de detecciones erróneas fue la categoría “Alegría”.

Tabla 15. Tabulación de resultados según parámetros definidos para el clasificador (mínimo tamaño: 30 y máximo tamaño:950) aplicado en la submuestra

minSize: 30 maxSize:950			
Emoción	Imágenes Existentes	Imágenes re escritas a partir de detección	Detecciones erróneas
Ira	396	225	14
Asco	396	228	17
Miedo	396	214	25
Alegría	396	250	7
Tristeza	396	173	9
Sorpresa	396	261	24
TOTAL			96

Fuente: Elaboración propia

A partir de la tabla 15, se analiza el comportamiento del clasificador haarcascade, donde se rescata en primera medida, que el porcentaje de detecciones erróneas llevadas a cabo con los parámetros de mínimo tamaño de 30 y máximo tamaño de 950 estuvo por debajo del 50% (del 2% al 6%) respecto al total de imágenes de entrada analizadas (396). Además, la categoría con menor cantidad de detecciones erróneas fue la categoría “Alegría”.

Por los análisis anteriormente expuestos, se puede evidenciar que, aunque no muestra una diferencia significativa, los parámetros con menor cantidad de detecciones erróneas (minSize:30 maxSize:950) serán los empleados como parámetros de entrada para el entrenamiento de los tres métodos seleccionados para el análisis por parte de OpenCV. Adicionalmente, fue posible evidenciar que, a pesar de la variación de parámetros llevada a cabo, la categoría con menor cantidad de detecciones erróneas fue “Alegría”; que, a su vez, fue una de las categorías con mayor cantidad de imágenes reescritas junto con la categoría “Sorpresa”.

9.2.2 Matriz de confusión de las pruebas

Para la evaluación de las medidas de desempeño de los diferentes parámetros indicados en el numeral 7.2.4.2 se realizó una matriz de confusión organizada de la misma forma que se indicó para la matriz de confusión de las pruebas con el método de Keras, pero esta vez por cada uno de los métodos desarrollados de OpenCV, sin embargo de todas las pruebas realizadas, en el presente documento se detallan únicamente los resultados de la métrica de exactitud como se presenta en las tablas 16, 17, 18, 19, 20, 21, 22, 23 y 24 con los resultados categorizados por cada emoción, prueba y por Método secuencial de OPENCV desarrollado (EigenFaces, FisherFaces y LBPH) según las 3 subdivisiones muestrales del dataset:

Tabla 16. Consolidado de exactitud por cada emoción y prueba en dataset7030 por FisherFaces

FISHERFACES				Dataset7030			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.1039	0.1250	0.1887	0.2083	0.0962	0.1667	0.1497
Prueba con minNeighbors = 10	0.1633	0.1905	0.2000	0.2708	0.0789	0.2000	0.1858
Prueba con minNeighbors = 5	0.1282	0.2464	0.1538	0.1825	0.1010	0.2500	0.1695
Prueba con minNeighbors = 3	0.1481	0.2544	0.1462	0.1852	0.1067	0.2262	0.1743

Fuente: Elaboración propia

La tabla 16, especifica el consolidado de resultados del método Fisherfaces evaluado a partir del dataset 7030 con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de 10.

Tabla 17. Consolidado de exactitud por cada emoción y prueba en dataset7525 por FisherFaces

FISHERFACES				Dataset7525			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.1250	0.1379	0.1628	0.2615	0.0541	0.2909	0.1877
Prueba con minNeighbors= 10	0.1613	0.1111	0.1154	0.3200	0.0435	0.2683	0.2011
Prueba con minNeighbors= 5	0.1591	0.2203	0.1667	0.2391	0.0843	0.3297	0.2057
Prueba con minNeighbors= 3	0.1611	0.2766	0.1683	0.2294	0.1020	0.2925	0.2044

Fuente : Elaboración propia

La tabla 17, especifica el consolidado de resultados del métodoFisherfaces evaluado a partir del dataset 7525, con la variación deparámetro de vecinos y mostrando la mejor estadística de precisiónempleando el método con el parámetro de 5.

Tabla 18. Consolidado de exactitud por cada emoción y prueba en dataset8020 por FisherFaces

FISHERFACES				Dataset8020			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.2558	0.1154	0.0882	0.1837	0.2143	0.2826	0.1991
Prueba con minNeighbors= 10	0.3103	0.0541	0.0769	0.1613	0.2353	0.2500	0.1830
Prueba con minNeighbors= 5	0.2289	0.2807	0.1029	0.1829	0.1509	0.2921	0.2106
Prueba con minNeighbors= 3	0.2258	0.2708	0.1020	0.1812	0.1395	0.3400	0.2197

Fuente : Elaboración propia

La tabla 18, especifica el consolidado de resultados del método Fisherfaces evaluado a partir del dataset 8020, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de 3.

Demostrando a partir de lo anterior que para la evaluación del método de Fisherfaces el análisis que evidencia los mejores resultados de rendimiento es aquel que se entrenó a partir del 80% del dataset y se evaluó a partir del 20% restante. Además, se puede resaltar que el método muestra mejores resultados cuando se disminuye el valor de la variable de vecinos cercanos.

Tabla 19. Consolidado de exactitud por cada emoción y prueba en dataset7030 por EigenFaces

EIGENFACES				Dataset7030			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.3016	0.5063	0.2857	0.3934	0.1818	0.4694	0.3772
Prueba con minNeighbors= 10	0.3514	0.5091	0.1944	0.4091	0.1538	0.4286	0.3628
Prueba con minNeighbors= 5	0.2783	0.5714	0.2809	0.3363	0.2459	0.4857	0.3909
Prueba con minNeighbors= 3	0.2597	0.5437	0.2609	0.3297	0.2143	0.4882	0.3761

Fuente : Elaboración propia

La tabla 19, especifica el consolidado de resultados del método EigenFaces evaluado a partir del dataset 7030, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de vecinos= 5.

Tabla 20. Consolidado de exactitud por cada emoción y prueba en dataset7525 por EigenFaces

EIGENFACES				Dataset7525			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors= 8	0.2708	0.5231	0.2609	0.3333	0.2414	0.4390	0.3610
Prueba con minNeighbors= 10	0.3226	0.5000	0.2000	0.3667	0.1667	0.3478	0.3333
Prueba con minNeighbors= 5	0.2529	0.5902	0.2597	0.2796	0.2545	0.4396	0.3695
Prueba con minNeighbors= 3	0.2467	0.5644	0.2623	0.2867	0.2088	0.4752	0.3645

Fuente : Elaboración propia

La tabla 20, especifica el consolidado de resultados del método EigenFaces evaluado a partir del dataset 7525, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de vecinos= 5.

Tabla 21. Consolidado de exactitud por cada emoción y prueba en dataset8020 por EigenFaces

EIGENFACES				Dataset8020			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors= 8	0.2683	0.5102	0.3125	0.3182	0.2727	0.4211	0.3628
Prueba con minNeighbors= 10	0.3200	0.5312	0.2222	0.3226	0.1765	0.3333	0.3333
Prueba con minNeighbors= 5	0.2754	0.5979	0.2982	0.2674	0.2927	0.4146	0.3773
Prueba con minNeighbors= 3	0.2807	0.5660	0.2644	0.2836	0.2286	0.4531	0.3714

Fuente : Elaboración propia

La tabla 21, especifica el consolidado de resultados del método EigenFaces evaluado a partir del dataset 8020, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de vecinos= 5.

Considerando los análisis realizados por los resultados de evaluación de los métodos de EigenFaces, se obtiene que la mejor implementación de este se da cuando se usa el parámetro de vecinos mínimos = 5, además, se pudo evidenciar que a pesar de los ensayos realizados con las diferentes muestras del dataset, la mayoría de los resultados que se obtuvieron están entre el 33% al 37%.

Tabla 22. Consolidado de exactitud por cada emoción y prueba en dataset7030 por LBPH

LBPH				Dataset7030			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.3617	0.5227	0.2203	0.4561	0.1522	0.4595	0.3772
Prueba con minNeighbors= 10	0.3871	0.4918	0.1944	0.5122	0.1471	0.4348	0.3761
Prueba con minNeighbors= 5	0.3947	0.5674	0.2872	0.4312	0.2353	0.5128	0.4223
Prueba con minNeighbors= 3	0.3750	0.5900	0.3066	0.4162	0.1954	0.5473	0.4221

Fuente : Elaboración propia

La tabla 22, especifica el consolidado de resultados del método LBPH evaluado a partir del dataset 7030, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de vecinos= 5.

Tabla 23. Consolidado de exactitud por cada emoción y prueba en dataset7525 por LBPH

LBPH				Dataset7525			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.3421	0.5385	0.2391	0.4783	0.1795	0.4667	0.3935
Prueba con minNeighbors= 10	0.4000	0.5000	0.1935	0.5312	0.1667	0.3529	0.3757
Prueba con minNeighbors= 5	0.3846	0.5667	0.3239	0.4176	0.2619	0.5156	0.4305
Prueba con minNeighbors= 3	0.3704	0.5963	0.3333	0.4225	0.2258	0.5656	0.4334

Fuente : Elaboración propia

La tabla 23, especifica el consolidado de resultados del método LBPH evaluado a partir del dataset 7525, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de vecinos= 3.

Tabla 24. Consolidado de exactitud por cada emoción y prueba en dataset8020 por LBPH

LBPH				Dataset8020			
Prueba	Ira	Asco	Miedo	Alegría	Tristeza	Sorpresa	Exactitud (Accuracy)
Prueba con minNeighbors = 8	0.3667	0.5593	0.2500	0.5000	0.2143	0.5200	0.4159
Prueba con minNeighbors= 10	0.3810	0.5250	0.2000	0.5357	0.1905	0.3846	0.3856
Prueba con minNeighbors= 5	0.4423	0.5932	0.3333	0.4474	0.2769	0.5273	0.4537
Prueba con minNeighbors= 3	0.4205	0.6069	0.3500	0.4505	0.2353	0.5446	0.4480

Fuente: Elaboración propia

La tabla 24, especifica el consolidado de resultados del método LBPH evaluado a partir del dataset 8020, con la variación de parámetro de vecinos y mostrando la mejor estadística de precisión empleando el método con el parámetro de vecinos= 5.

Partiendo de los resultados de los métodos de LBPH generados por cada muestra del dataset, se obtiene mejor precisión en la clasificación de las categorías que implementan los parámetros de valores más pequeños como 3 y 5. Además, sus resultados al encontrarse dentro del rango de 37% a 45% demuestran que las técnicas empleadas por el método de Local Binary pattern permiten clasificar las diferentes categorías con mayor exactitud en comparación con los anteriores métodos presentados.

10.RESULTADOS Y ANÁLISIS DE RESULTADOS

Para este trabajo de investigación, se realizó la construcción de un dataset propio a partir de 6 fotografías recolectadas, una por cada una de las emociones propuestas por Ekman a 150 personas que se encuentran en el rango de edad entre 18 y 29 años, esto a partir de las estadísticas encontradas durante la investigación de la problemática, adicionalmente se realizó un incremento de datos, con 3 datasets de la web, consolidando un total 7920 imágenes, compuestas por 1320 imágenes por cada emoción (alegría, tristeza, miedo, enojo, disgusto y sorpresa), posteriormente, se realizó el muestreo de 3 datasets, en donde se varió la proporción de cantidad de fotografías destinadas para el entrenamiento y clasificación en: 80% y 20%, 75% y 25%, 70% y 30% respectivamente, además, se designan los nombres de los muestreos del dataset de la siguiente manera: Dataset7030, Dataset7525 y Dataset8020 agregando a esto, se tuvo en cuenta la cantidad de personas que emplearon accesorios como las gafas como se detalla en las tablas 5, 6 y 7 del capítulo 5, dado que para este estudio se analizó en un principio como una variable condicionante, para el reconocimiento de emociones que posteriormente durante el desarrollo se evidenció que no tuvo resultados desfavorables por esta condición.

A partir de la construcción del Dataset, se realizó una búsqueda en la literatura acerca de diferentes métodos de detección comúnmente utilizados en investigaciones similares, teniendo como resultado 7 diferentes métodos, para los cuales se tuvieron en cuenta sus ventajas, desventajas y entorno de desarrollo con base a lo cual se seleccionaron los métodos secuenciales de Keras y OpenCV; el primero, debido a que su manejo por capas permitía el uso de diferentes librerías de forma secuencial, con lo que permitió la interacción de las mismas, con el fin de que el modelo fuera más robusto, adicionalmente, la facilidad de implementación en Google Colab, facilitó su ejecución ya que no fue necesario contar con una GPU local, y el segundo, la librería OpenCV, que dentro de sus más recientes versiones, permite realizar 3 técnicas diferentes de extracción de características (EigenFaces, FisherFaces y LBPH) por medio de los constructores virtuales que permiten el desarrollo de métodos secuenciales empleando cada una de las técnicas que a su vez, permiten optimizar tiempos de ejecución, así como la posibilidad de realizar el procesamiento asociado a procesadores Intel y no depender de la GPU.

Dentro de la fase de desarrollo y evaluación de los métodos estudiados, fue posible encontrar que los dos métodos principales, Keras y OpenCV, aplicados en la investigación, demostraron un mejor rendimiento sobre los modelos entrenados a partir del Dataset8020. Así mismo, a partir de las pruebas y métodos expuestos en la sección 8.2.2, empleando las técnicas de extracción de características dentro de los métodos secuenciales desarrollados por la librería OpenCV (EigenFaces, Fisherfaces y LBPH) , se pudo comprobar que entre menor es el parámetro de distancia entre los puntos vecinos se obtienen mejores resultados por parte del método entrenado. Adicionalmente, se pudo

evidenciar que de los tres métodos que se evaluaron empleando OpenCV, el que presentó mejor desempeño fue el método que emplea las técnicas de extracción de características de LBPH entrenado a partir de la muestra del Dataset8020. Sin embargo, a pesar de que en comparación con el método secuencial de Keras, tardaron menos tiempo de entrenamiento por cada uno de ellos, los resultados de las métricas indicaron un desempeño inferior al evidenciado por el otro método como se evidencia y detalla a continuación en la tabla 25:

Tabla 25. Comparación de precisión entre las mejores pruebas de los métodos secuenciales de Keras y OpenCV

Emoción	Método secuencial de Keras Prueba 9	Método secuencial de OpenCV por LBPH Prueba minNeighbors=5
Ira	56,16	44,23
Asco	83,54	59,32
Miedo	40,80	33,33
Felicidad	76,92	44,74
Tristeza	45,91	27,69
Sorpresa	75,00	52,73
Exactitud (Accuracy)	62,37	45,37

Fuente: Elaboración propia

Teniendo en cuenta las matrices de confusión del método secuencial de Keras, para el cual, el mejor resultado se obtuvo cuando estaba involucrado el Dataset8020, en específico en la prueba 9, presentando el mayor accuracy entre el total de pruebas realizadas en el presente trabajo de investigación con los diferentes métodos. Como se observa en la tabla 26, dicha prueba tiene un total de 62,37% de accuracy, correspondiente al porcentaje de casos que el modelo acertó, es decir, de 1584 casos de validación, el modelo clasificó de forma correcta 988; por otro lado el 50% de las emociones evaluadas presentan una precisión de detección mayor al 70%, y ninguna se encuentra por debajo de una 40%, esto corresponde a la relación entre falsos positivos y verdaderos positivos que el modelo detectó; adicional a esto en cuanto a exhaustividad, solo la emoción asco, presentó un total perfecto de 100%, por lo cual, de las 264 imágenes a evaluar clasificó todas de forma correcta, mientras que las otras emociones se encuentran todas por encima de un 42%; en cuanto a la relación entre la precisión y exhaustividad, es decir, el valor-f1, para la emoción asco presenta un 91,03%, siendo la emoción que mejores métricas presentó, por otro lado el 50% de las emociones se encuentran por encima de

un 65% de valor-f1 y el restante no está por debajo de un 45%, presentando en general valores aceptables. Sin embargo, las emociones miedo y tristeza, presentan métricas bajas con respecto al resto de emociones, por lo que influyeron de forma negativa en el accuracy general de la prueba.

Tabla 26. Resultados obtenidos en la prueba 9 del método secuencial de Keras

Emoción/ Métrica	Precisión (%)	Exhaustividad (%)	Valor - F1 (%)	Accuracy (%)
Ira	56,16	43,18	48,82	62,37
Asco	83,54	100	91,03	62,37
Miedo	40,80	50,38	45,08	62,37
Alegría	76,92	68,18	72,29	62,37
Tristeza	45,91	48,86	47,34	62,37
Sorpresa	75	63,64	68,85	62,37

Autor: Elaboración propia

Bajo la prueba 9, se realizó la fase de implementación, obteniendo como resultado 2 formas diferentes de evidenciar el reconocimiento facial de emociones: De forma porcentual y con diagrama de barras, cabe resaltar que, en la visualización porcentual, es posible la detección de emociones para 1 o más personas, lo que puede ser de utilidad en terapias grupales, como se evidencia en la tabla 27:

Tabla 27. Tipos de visualización de detección de emociones

Visualización porcentual	Visualización por diagrama de barras
	

Autor: Elaboración propia

Por lo anterior, el mejor método planteado para los fines propuestos de clasificación de emociones propuestas por Ekman, es el método secuencial de Keras. Cuya arquitectura se especifica en el capítulo 7.1 y se añade el hallazgo 90 épocas, que se deben implementar como parámetro que expone las mejores medidas de desempeño, para el cual es posible garantizar su validez según los resultados favorables de 62% sobre el total, a partir de todas características anteriormente expuestas sobre cada uno de los elementos (dataset, arquitectura, desarrollo y evaluación) que se evaluaron para el diseño de un método automático para el apoyo en la identificación de emociones en tratamientos psicológicos de depresión en adultos jóvenes a partir de técnicas de aprendizaje automático.

CONCLUSIONES

Dentro de esta investigación, se realizaron diferentes actividades investigativas explicadas a lo largo de este proyecto, que permitieron identificar los elementos que deben considerarse para la implementación de un método automático que identifique las emociones faciales planteadas por Ekman en adultos jóvenes, como la construcción de un Dataset a partir de las características pertinentes a la población de estudio y variables relacionadas a ella, la arquitectura del método, el desarrollo y su evaluación en donde se destacan los siguientes hallazgos:

- De los conjuntos de datos generados, el que mejor resultados obtuvo fue el Dataset8020 que tiene 80% de imágenes de entrenamiento y un 20% de validación, debido a que al tener un mayor porcentaje de imágenes para el entrenamiento facilitó que el modelo tuviera una mejor precisión al momento de identificar emociones.
- El método secuencial de Keras, presenta una mejor exactitud con un 62% al momento de clasificar emociones faciales en tiempo real frente a los métodos de la librería OpenCV que tienen un 45%.
- Los resultados obtenidos en la matriz de confusión y de las medidas precisión, exactitud y exhaustividad demostraron tener una capacidad de acierto por encima del 62% en el proceso de clasificación de las seis emociones establecidas por Ekman.
- Mediante la implementación de este algoritmo es posible realizar la detección de las seis emociones establecidas por Ekman, en tiempo real de una o más personas sin tener limitaciones respecto al uso de accesorios como las gafas.

TRABAJO FUTURO

Esta investigación tiene gran impacto en el proceso de construcción del conocimiento, tanto en el área de psicología como en el área de ingeniería; ya que permite dar paso a la búsqueda de proyectos que complementen lo realizado y se constituyan como base que sea de gran utilidad en los diagnósticos de los pacientes que asisten a terapia virtual psicológica en la Unidad de Servicios Psicológicos de la Universidad Católica de Colombia.

Considerando lo anterior, se proyectan trabajos encaminados a desarrollar el ciclo de vida del software partiendo del método automático escogido, que a partir de las estadísticas de las emociones encontradas en los pacientes, entregadas por la lectura del método, se organicen de forma que sean de entendimiento para el profesional de la salud, y pueda ejecutarlo por medio de una interfaz que le permita hacer uso del método automático sin tener que interactuar directamente con el código.

Por otro lado, a partir de la identificación correcta de las emociones investigadas en el presente trabajo, se prevé la necesidad futura por identificar microexpresiones de los pacientes que permitan detallar aún más el diagnóstico y se le entregue más información al profesional de la salud mental de manera que se complemente y contribuya el diagnóstico emitido.

Finalmente, se prevé que este trabajo de grado tiene implicaciones positivas y genera insumos para el desarrollo de plataformas de otras entidades que presten servicios psicológicos favoreciendo en la precisión de más diagnósticos psicológicos que se emitan en establecimientos no solo en la Universidad Católica de Colombia sino también ajenos a ella.

BIBLIOGRAFÍA

AMAT, J. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado 11 de octubre de 2021]. Disponible en: https://www.cienciadedatos.net/documentos/35_principal_component_analysis

AMAT, J. Eigenvalores. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado el 11 de octubre de 2021]. Disponible en https://www.cienciadedatos.net/documentos/35_principal_component_analysis

AMAT, J. Eigenvectores. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado el 11 de octubre de 2021]. Disponible en https://www.cienciadedatos.net/documentos/35_principal_component_analysis

AMAT, J. Loadings. Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE. [Sitio Web]. [Consultado el 11 de octubre de 2021]. Disponible en https://www.cienciadedatos.net/documentos/35_principal_component_analysis

ALLEN H. SACCHI, Simone, WICKETT, Karen. Definitions of dataset in the scientific and technical literature. [Sitio Web]. [Consultado 13 oct., 2021]. Disponible en: <https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/meet.14504701240>

ANACONDA. imutils 0.5.4. [Sitio Web]. [Consultado el 03 de febrero de 2022]. Disponible en: <https://anaconda.org/conda-forge/imutils>

ANACONDA. Products.[Sitio web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://www.anaconda.com/products/distribution>

ANALYTICSLANE. Visualización de datos en Python con Seaborn. [Sitio Web]. Consultado el 07 de febrero de 2022]. Disponible en: <https://www.analyticslane.com/2018/07/20/visualizacion-de-datos-con->

APRENDER DE MACHINE LEARNING. ¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador. [Consultado el 03 de agosto de 2021]. Disponible en: <https://www.aprendemachinelearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>

BYEON, Young-Hyen. KWAK, Keun-Chang. Facial Expression Recognition Using 3D Convolutional Neural Network. [Sitio Web]. [Consultado el 7 septiembre de 2021]. Disponible en: <https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S235291482030201X>

COSSIN, Florencia Carla; Wanda Yanina Rubinstein; Daniel Gustavo Politis. How many are the basic emotions? Preliminary study in a sample of healthy old adults. [Sitio Web]. [Consultado el 13 de octubre de 2021]. Disponible en: <https://www.redalyc.org/journal/3691/369155966032/movil/>

COSSIN, Florencia. RUBINSTEIN, Wanda. POLITIS, Daniel. Universidad de Buenos Aires. ¿Cuántas son las emociones básicas ? Estudio preliminar en una muestra de adultos mayores sanos. En el Anuario de Investigaciones, vol. XXIV, pp. 253-257, 2017.

DATASOURCE. Data Science. Machine learning. Python. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: <https://www.datasource.ai/es/data-science-articles/compreension-de-la-matriz-de-confusion-y-como-implementarla-en-python>

GITHUB Docs. Productos de GitHub. [Consultado el 10 de mayo de 2022]. Disponible en: <https://docs.github.com/es/get-started/learning-about-github/githubs-products>

GOOGLE RESEARCH. Colaborarity. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://research.google.com/colaboratory/intl/es/faq.html>

GOOGLE Workspace. Documentos . [Consultado el 10 de mayo de 2022]. Disponible en: <https://workspace.google.com/intl/es-419/products/docs/>

GOOGLE WORKSPACE. Drive. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://workspace.google.com/intl/es-419/products/drive/>

GOOGLE Workspace. Hojas de cálculo . [Consultado el 10 de mayo de 2022]. Disponible en: <https://workspace.google.com/intl/es-419/products/sheets/>

GUZMAN, Juan, TAMAYO, Alejandro. Ansiedad y depresión, 'verdugos' de los jóvenes en la pandemia. [Consultado el 30 de agosto de 2021]. [Sitio Web]. Disponible en: <https://www.javeriana.edu.co/pesquisa/ansiedad-depresion-jovenes-pandemia/>

HASSOUNEH, Aya. MUTAWA, A. MURUGAPPAN, M. Universidad de Kuwait. Development of a Real Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods. [Sitio Web]. [Consultado 27 de enero de 2022]. [Sitio Web]. Disponible en: <https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S235291482030201X>

IBM. Centro de aprendizaje de la nube de IBM. ¿Qué es el aprendizaje no supervisado?. Aprendizaje sin supervisión. [Sitio Web]. [Consultado el 05 de

octubre del 2021] Disponible en:
<https://www.ibm.com/cloud/learn/unsupervised-learning>

IBM. ¿Qué es Machine Learning?. [Consultado el 13 de octubre de 2021].
Disponible en:<https://www.ibm.com/co-es/analytics/machine-learning>

IBM. Deep Learning. [Sitio Web]. [Consultado 19 de octubre de 2021]. Disponible
en <https://www.ibm.com/co-es/cloud/deep-learning>

INSTITUTO NACIONAL DEL CÁNCER. Diagnóstico. [Sitio Web]. [Consultado
el 19 octubre de 2021]. Disponible en:
[https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/d
ef/diagnostico](https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/diagnostico)

JONATHA OENIX. Face expression recognition dataset. [Consultado el 03 de
marzo del 2022]. Disponible en:
[https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dat
aset](https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset)

JPL. IDIS. Eigenface. [Sitio Web]. [Consultado el 08 de febrero del 2022].
Disponible en: <https://proyectoidis.org/eigenface/>

Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial
expression analysis. Proceedings of the Fourth IEEE International Conference
on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

KELTNER, D., & Lerner, J. S. Emotion. Handbook of Social Psychology. [Sitio
Web]. [Consultado el 19 de octubre de 2021]. Disponible en :
doi:10.1002/9780470561119.socpsy001009

KERAS. Deep learning for humans. [Sitio Web] [Consultado el 05 de febrero de
2022]. Disponible en: <https://keras.io/>

KERAS. Dense layer. [Sitio Web]. [Consultado el 05 de febrero de 2022].
Disponible en: https://keras.io/api/layers/core_layers/dense/

KERAS. Dropout layer. [Sitio Web]. [Consultado el 05 de febrero de 2022].
Disponible en: https://keras.io/api/layers/regularization_layers/dropout/

KERAS. Flatten layer. [Sitio Web]. [Consultado el 05 de febrero de 2022].
Disponible en: https://keras.io/api/layers/reshaping_layers/flatten/

KERAS. Getting started. Faq. [Sitio web]. [Consultado el 05 de febrero de 2022].
Disponible en:
https://keras.io/getting_started/faq/#what-do-sample-batch-and-epoch-mean

KERAS. MaxPooling2D layer. [Sitio Web]. [Consultado el 05 de febrero de
2022]. Disponible en: https://keras.io/api/layers/pooling_layers/max_pooling2d/

KERAS. tf.compat.v1.metrics.precisión [Sitio web]. [Consultado el 15 de marzo de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/compat/v1/metrics/accuracy

LÓPEZ, David. SANZ, Natalia. Mente a Mente. Terapia para la depresión. [Sitio Web]. [Consultado el 06 de agosto de 2021]. Disponible en: <https://www.menteamente.com/terapia-depresion>

MACHUCA Fernando. ¿Qué es Python? El lenguaje de programación más popular para aprender en 2021.[Sitio Web]. [Consultado el 19 de octubre de 2021]. Disponible en: <https://www.crehana.com/cl/blog/desarrollo-web/que-es-python/>

MANAS SAMBARE. Fer-13. [Consultado el 03 de marzo del 2022]. Disponible en: <https://www.kaggle.com/datasets/msambare/fer2013>

Matplotlib. Matplotlib: Visualization with Python.[Sitio Web] [Consultado el 05 de febrero de 2022]. Disponible en: <https://matplotlib.org/>

MINSALUD. Abecé sobre la salud mental, sus trastornos y estigma. [Sitio Web] [Consultado el 19 de octubre de 2021]. Disponible en:

MINSALUD. Depresión: hablemos. [Sitio Web] [Consultado el 26 de julio de 2021]. Disponible en: <https://www.minsalud.gov.co/Paginas/Depresi%C3%B3n-hablemos.aspx>

MOHAJÓN, J. Towards Data Science. Matriz de confusión para su modelo de aprendizaje automático multiclase. [Sitio Web]. [Consultado el 19 de febrero de 2022]. Disponible en: <https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826>

NATIONAL INSTITUTE OF MENTAL HEALTH . Depresión. [Sitio Web]. EEUU. [Consultado el 18 de junio de 2021]. Disponible en: <https://www.nimh.nih.gov/health/publications/espanol/depression-sp>

NUMPY. NumPy documentation. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: <https://numpy.org/doc/stable/>

OMS. La OMS presenta el primer Informe mundial sobre la visión. [Consultado el 05 de marzo de 2022]. Disponible en : <https://www.who.int/es/news/item/08-10-2019-who-launches-first-world-report-on-vision>

OPENCV. Face Analysis. cv::face::EigenFaceRecognizer Class Reference. [Sitio Web]. [Citado el 28 de marzo de 2022]. Disponible en: https://docs.opencv.org/4.2.0/dd/d7c/classcv_1_1face_1_1EigenFaceRecognizer.html

OPENCV. Face Analysis. cv::face::FisherFaceRecognizer Class Reference. [Sitio Web]. [Citado el 28 de marzo de 2022]. Disponible en: https://docs.opencv.org/4.2.0/d2/de9/classcv_1_1face_1_1FisherFaceRecognizer.html

OPENCV. Face Analysis..cv::face::LBPHFaceRecognizer Class Reference. [Sitio Web]. [Citado el 28 de marzo de 2022]. Disponible en: https://docs.opencv.org/4.2.0/df/d25/classcv_1_1face_1_1LBPHFaceRecognizer.html

OPENCV. Introducción. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en : <https://docs.opencv.org/4.5.5/d1/dfb/intro.html>

OPS. Depresión. [Sitio Web]. [Consultado el 13 de octubre de 2021]. Disponible en: <https://www.paho.org/es/temas/depresion>

ORGANIZACIÓN MUNDIAL DE LA SALUD, OMS. Depresión. [Consultado el 18 de agosto de 2021]. Disponible en: <http://www.who.int/topics/depression/es/>

OSPINA, Alexis. PULIDO, Mateo. Comparar tres métodos convencionales de reconocimiento facial y confrontar el de mejor resultado frente a computación cognitiva. [Sitio Web]. Disponible en: https://repositorio.itm.edu.co/bitstream/handle/20.500.12622/1706/Rep_Itn_pre_Ospina.pdf?sequence=1&isAllowed=y

PAMPOUCHIDOU, A. PEDIADITIS, M. KAZANTZAKI, E. SFAKIANAKIS, S. APOSTOLAKI, A. ARGYRAKI, K. MANOUSOS, D. MERIAUDEAU, F. MARIAS, K. YANG, F. TSIKNAKIS, M. BASTA, M. VGONTZAS, A. SIMOS, P. Machine Vision Applications. Automated facial video-based recognition of depression and anxiety symptom severity: cross-corpus validation. [Sitio Web]. [Consultado el 13 de octubre de 2021]. Disponible en: 10.1007/s00138-020-01080-7

PERIÓDICO EL TIEMPO. La depresión, un mal que crece en el mundo. [Consultado el 18 de mayo de 2022]. Disponible en: <https://www.eltiempo.com/archivo/documento/CMS-16826965>

PINZÓN, Jorge. Bienestar Colsanitas. Suicidio en Colombia: alerta amarilla. [Sitio Web]. [Consultado el 05 de agosto de 2021]. Disponible en: <https://www.bienestarcolsanitas.com/articulo/suicidio-en-colombia-alerta-amarilla.html>

PYTHON. Documentation. Biblioteca Estándar. Interfaces misceláneas del sistema operativo. [Sitio Web]. [Consultado el 28 de enero del 2022]. Disponible en: <https://docs.python.org/es/3.10/library/os.html>

PYTHON. Miscelánea. Configurar Sublime Text para programar en Python. [Sitio Web]. [Consultado 12 de febrero de 2022]. Disponible en: <https://www.programaenpython.com/miscelanea/configurar-sublime-text-para-programar-en-python/>

PYTHON. Pitón 3.10.4. [Sitio Web]. [Consultado 28 de marzo de 2022]. Disponible en: <https://www.python.org/downloads/release/python-3104/>

ROGER GIMEO. Estudio de técnicas de reconocimiento facial. [Citado el 25 de marzo de 2022]. Disponible en: https://upcommons.upc.edu/bitstream/handle/2099.1/9782/PFC_RogerGimeno.pdf

ROUHIAINEN, Lasse. Inteligencia Artificial, 101 cosas que debes saber sobre nuestro futuro. Editorial Planeta S.A. [Sitio Web]. [Consultado el 19 de octubre de 2021]. Disponible en: https://static0planetadelibroscom.cdnstatics.com/libros_contenido_extra/40/39308_Inteligencia_artificial.pdf

SAS. Software y soluciones de analítica. Procesamiento del lenguaje natural. [Consultado el 13 de octubre de 2021] Disponible en: https://www.sas.com/es_co/insights/analytics/what-is-natural-language-processing-nlp.html

SAUNDERS, Adam. BBVA. El impacto de la tecnología en el crecimiento y el empleo. [Consultado el 25 de noviembre de 2021]. Disponible en: <https://www.bbvaopenmind.com/articulos/el-impacto-de-la-tecnologia-en-el-crecimiento-y-el-empleo>

SCIKIT-LEARN. Machine learning en Python. [Sitio Web]. [Consultado 04 de marzo de 2022]. Disponible en: <https://scikit-learn.org/stable/>

SEMANA S: A.: Salud Mental. El mundo triste de la depresión. [Sitio Web]. [Consultado el 05 de agosto de 2021]. Disponible en: <https://www.semana.com/vida-moderna/articulo/depresion-en-paises-violentos/364663-3/>

SHETTY, BHOOMIKA, DEEKSHA, JEEVAN REBEIRO, RAMYASHREE. ScienceDirect. KeAi Chinese Roots Global Impact. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://doi.org/10.1016/j.gltp.2021.08.044>

SU, Dai. ZHANG, Xingyu. HE, Kevin. CHEN, Yingchun. Use of machine learning approach to predict depression in the elderly in China: A longitudinal study. ScienceDirect. Volume 282, 1 March 2021, Pages 289-298

TABASSUMA. ISLAMA, TASIN, AMINC. ScienceDirect. Human face recognition with combination of DWT and machine learning. [Sitio Web]. [Consultado el 10 de mayo de 2022]. Disponible en: <https://www-sciencedirect-com.ucatolica.basesdedatosezproxy.com/science/article/pii/S1319157819309395>

TENSORFLOW. El modelo secuencial. [Sitio Web]. [Citado el 25 de marzo de 2022]. Disponible en: https://www.tensorflow.org/guide/keras/sequential_model?hl=es-419

TENSORFLOW. Introducción a Tensores. [Sitio Web]. [Consultado el 03 de febrero]. Disponible en : <https://www.tensorflow.org/guide/tensor>

TENSORFLOW. La API funcional "Keras" en TensorFlow. [Citado el 25 de marzo de 2022]. Disponible en: <https://www.tensorflow.org/guide/keras/functional?hl=es-419>

TENSORFLOW. tf. keras. metrics. F1 Score. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/addons/api_docs/python/tfa/metrics/F1Score

TENSORFLOW. tf. keras. metrics. Precision. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/metrics/Precision

TENSORFLOW. tf. keras. metrics. Recall. [Sitio web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/metrics/Recall

TENSORFLOW. tf.keras.layers.BatchNormalization.[Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/layers/BatchNormalization

TENSORFLOW. tf.keras.layers.Conv2D. [Sitio Web] [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Conv2D

TENSORFLOW. Tf.tensor. [Sitio Web]. [Consultado el 05 de febrero de 2021]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/Tensor

TENSORFLOW.. tf.keras.layers.Activation. [Sitio Web]. [Consultado el 05 de febrero de 2022]. Disponible en: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Activation

UNIVERSIDAD DE ALCALÁ. Pandas: Herramienta básica para el data science en Python. [Sitio Web]. [Consultado el 07 de febrero de 2022]. Disponible en: <https://www.master-data-scientist.com/pandas-herramienta-data-science/>

UNZUETA, D. Towards Data Science. Fundamentos del aprendizaje automático. Discriminante lineal de Fisher: explicación intuitiva. [Sitio Web]. [Consultado 22 de noviembre de 2021]. Disponible en: <https://towardsdatascience.com/fishers-linear-discriminant-intuitively-explained-52a1ba79e1bb>

WANG, Qingxiang. YANG, Huanxin. YU, Yanhong. Facial Expression video analysis for depression detection in Chinese patients. [Sitio Web]. [Consultado el 23 de agosto de 2021]. disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S1047320318302761?via%3Dihub>

XIAOFEI He, PARTHA Niyogi . Locality Preserving Projections. [Sitio Web].
[Citado el 25 de marzo de 2022]. Disponible en:
<https://proceedings.neurips.cc/paper/2003/file/d69116f8b0140cdeb1f99a4d5096ffe4-Paper.pdf>