



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

Centro Nacional de Investigación
y Desarrollo Tecnológico

Tesis de Maestría

Método para la detección de emociones mediante el
análisis de expresión facial en videos de evaluación
de la experiencia de usuario.

presentado por

Ing. Luis Alfredo Hernández Pérez

como requisito para la obtención del grado de
Maestro en Ciencias de la Computación

Director de tesis

Dr. Juan Gabriel González Serna

Codirectora de tesis

Dra. Andrea Magadan Salazar

Cuernavaca, Morelos, México. Diciembre de 2022.



Centro Nacional de Investigación y Desarrollo Tecnológico
Departamento de Ciencias Computacionales

Cuernavaca, Mor., 08/noviembre/2022

OFICIO No. DCC/089/2022

Asunto: Aceptación de documento de tesis
CENIDET-AC-004-M14-OFICIO

DR. CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial del C. LUIS ALFREDO HERNÁNDEZ PÉREZ, con número de control M20CE038, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado "MÉTODO PARA LA DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS DE EXPRESIÓN FACIAL EN VIDEOS DE EVALUACIÓN DE LA EXPERIENCIA DE USUARIO" y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

DR. JUAN GABRIEL GONZÁLEZ SERZA
Director de tesis

DRA. ANDREA MAGADÁN SALAZAR
Codirectora de Tesis

DR. DANTE MÚJICA VARGAS
Revisor 1

DR. NOÉ ALEJANDRO CASTRO SÁNCHEZ
Revisor 2

C.c.p. Depto. Servicios Escolares.
Expediente / Estudiante
JGGS/ibm



Interior Internado Palmira S/N, Col. Palmira, C. P. 62490, Cuernavaca, Morelos
Tel. 01 (777) 3627770, ext. 3201, e-mail: dcc@tecnm.mx tecnm.mx | cenidet.tecnm.mx



2022 Flores
Año de Magón
PRELACION DE LA REVOLUCIÓN MEXICANA



Centro Nacional de Investigación y Desarrollo Tecnológico
Subdirección Académica

Cuernavaca, Mor., **22/noviembre/2022**
No. De Oficio: **SAC/167/2022**
Asunto: **Autorización de
impresión de tesis**

**LUIS ALFREDO HERNÁNDEZ PÉREZ
CANDIDATO(A) AL GRADO DE MAESTRO(A) EN CIENCIAS
DE LA COMPUTACIÓN
P R E S E N T E**

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado "MÉTODO PARA LA DETECCIÓN DE EMOCIONES MEDIANTE EL ANÁLISIS DE EXPRESIÓN FACIAL EN VIDEOS DE EVALUACIÓN DE LA EXPERIENCIA DE USUARIO", ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

ATENTAMENTE
Excelencia en Educación Tecnológica®
"Educación Tecnológica al Servicio de México"



DR. CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO

C. c. p. Departamento de Ciencias Computacionales
Departamento de Servicios Escolares

CMAZ/RMA



Dedicatorias

Este trabajo está dedicado a mis padres, hermanos y amigos que siempre estuvieron ahí apoyándome en cada meta que me propongo y que sin dudarlo me brindaron su confianza y fortaleza para no sentirme solo y rendirme, mi familia y amigos son lo más importante y el motivo por el cual cada día me quiero superar y cumplir las metas que me propongo por ello dedico este trabajo a ellos.



Agradecimientos

Primeramente, a dios por permitirme lograr el sueño de tener un grado más profesionalmente.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico que me brindó durante mis estudios de maestría.

Al Tecnológico Nacional de México campus Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) por brindarme todas las herramientas necesarias para la culminación de esta tesis.

A mis padres y hermanos, por su apoyo incondicional, sus palabras de aliento en los buenos y malos momentos que me ayudan a ser un mejor ser humano y seguir creciendo como persona.

A mi abuelo Constantino Trinidad Pérez Santiago por el apoyo y consejos que me brindo desde los inicios de mi vida académica para seguir adelante con cada meta que me proponga.

A mi amigo Pedro Chulin Cruz por haberme dado la confianza que necesitaba para iniciar mis estudios de maestría.

A mi compañero de clases y sobre todo a mi gran amigo el cual considero como un hermano, Armando Miranda Molina que me brindo su amistad y apoyo durante la maestría y que sin conocerme me brindo su ayuda en ocasiones en las que la parte económica era un obstáculo para mí en la maestría, además no dudo en compartirme de sus conocimientos adquiridos para mejorar en partes de mi investigación.

A mis amigos Jhokenny Amaranta Bautista Alfonzo, Francisco Micheel Bautista Alfonzo, Yareni Lizbeth Bautista Alfonzo quienes me brindaron su apoyo alentándome cada día a lograr esta meta.

A mi asesor el Dr. Juan Gabriel González Serna que me guio durante toda la investigación y me brindo su apoyo y conocimientos para llevar a cabo un buen trabajo, que me motivo a seguir estudiando la maestría y no desistir, le agradezco mucho porque fue una parte muy importante para lograr obtener el grado.

A mi asesora la Dr. Andrea Magadan Salazar que me brindo sus conocimientos y experiencia en la investigación y poder desarrollar de manera exitosa esta investigación, agradezco también la paciencia que tuvo conmigo para transmitirme cada conocimiento que debía tener para lograr cada objetivo de la tesis.

A mi comité revisor el Dr. Dante Mujica Vargas y el Dr. Noe Alejandro Castro Sánchez por todo los consejos, observaciones y sugerencias que me brindaron en cada una de las presentaciones de avance que hicieron que mejorara cada vez y poder entregar una investigación de calidad y profesional.

Resumen

Las expresiones faciales resultan de gran importancia cuando se trata de la interacción entre personas, ya que revelan estados emocionales y cognitivos complejos los cuales son visualmente expresados por las personas (Arango et al., 2013; Fasel y Luetin, 2003). El rostro humano es considerado el principal sistema visual para mostrar las emociones, además de ser el área visual más importante y compleja de la comunicación no verbal y la que más se observa y analiza durante cualquier tipo de interacción social.

La experiencia de usuario (UX) se puede considerar como un atributo de calidad de un producto, proceso o servicio, definiendo si es fácil de adaptar, eficiente de usar y cómodo. La evaluación de la experiencia del usuario (UX por sus siglas en inglés) puede usarse como una herramienta para determinar el grado de aceptación o rechazo que experimenta un usuario cuando utiliza un producto, proceso o servicio. En el proceso de evaluación de la experiencia de usuario (UXE), una fuente de información muy relevante que está presente y que es difícil de alterar, es el rostro del usuario, ya que los cambios de estados emocionales de las personas, por lo general, hacen hincapié en la transmisión de un mensaje durante la interacción social y con ello se puede identificar de manera natural el impacto emocional que el usuario presenta durante la interacción con un producto digital. Este hecho ha motivado un gran número de investigaciones para que las máquinas puedan reconocer las emociones humanas. Hay varias aplicaciones de la computación afectiva para facilitar la interacción humano-computadora (A. Sunitha, et al, 2017).

Por ello en esta tesis se presenta la implementación de un algoritmo de aprendizaje automático con la capacidad de realizar el reconocimiento de emociones mediante el análisis de expresiones faciales con la finalidad de demostrar que el reconocimiento de emociones aporta información importante y complementaria en una evaluación de la UX integrando el algoritmo a una plataforma denominada UXLab. Para el análisis de características se extrajeron puntos de referencia del rostro utilizando un modelo pre-entrenado el cual localiza 68 puntos (x, y) y los sitúa en el rostro de los cuales se utilizaron solamente 17 para calcular distancias geométricas entre ciertos músculos del rostro tomando como referencia el Sistema de Codificación de Acción Facial para emociones (EMFACS por sus siglas en inglés) propuesta por Eckman, lo que permite describir las emociones. El sistema se entrenó con un repositorio propio y las pruebas de evaluación fueron con el conjunto de datos CK+ usado en la literatura, las pruebas de validación realizadas fueron tanto con imágenes estáticas como con videos generados en una evaluación de la experiencia de usuario. Los resultados obtenidos son del 84.52% de precisión, en el reconocimiento de 3 emociones Felicidad, Enojo, Sorpresa y la expresión Neutral.

Abstract.

Facial expressions are of great importance when it comes to the interaction between people, since they reveal complex emotional and cognitive states that are expressed visually by people (Arango et al., 2013; Fasel and Luetin, 2003). The human face is considered the main visual system to show emotions, as well as being the most important and complex visual area of non-verbal communication and the one that is most observed and analyzed during any type of social interaction.

User experience (UX) can be considered as a quality attribute of a product, process or service, defining whether it is easy to adapt, efficient to use and comfortable. User experience (UX) evaluation can be used as a tool to determine the degree of acceptance or rejection that a user experiences when using a product, process or service. In the user experience evaluation (UXE) process, a very relevant source of information that is present and difficult to alter is the user's face, since changes in people's emotional states usually, emphasize the transmission of a message during social interaction and thus the emotional impact that the user presents during the interaction with a digital product can be identified in a natural way. This fact has motivated a large number of investigations so that machines can recognize human emotions. There are several applications of affective computing to facilitate human-computer interaction (A. Sunitha, et al, 2017).

For this reason, this thesis presents the implementation of a machine learning algorithm with the ability to perform emotion recognition through the analysis of facial expressions in order to demonstrate that emotion recognition provides important and complementary information in an evaluation of the UX integrating the algorithm to a platform called UXLab. For the analysis of features, reference points were extracted from the face using a pre-trained model that locates 68 points (x, y) and locates them on the face, of which only 17 were used to calculate geometric distances between certain muscles of the face taking as reference the Facial Action Coding System for emotions (EMFACS for its acronym in English) proposed by Eckman, which allows describing emotions. The system was trained with its own repository and the evaluation tests were with the CK+ data set used in the literature, the validation tests carried out were both with static images and videos generated in an evaluation of the user experience. The results obtained have an accuracy of 84.52%, in the recognition of 3 emotions Happiness, Anger, Surprise and Neutral expression.

Contenido

| | |
|--|----|
| Capítulo 1 Introducción..... | 1 |
| 1.1 Descripción del problema | 2 |
| 1.2 Justificación | 4 |
| 1.3 Objetivo General..... | 5 |
| 1.3.1 Objetivos Específicos | 5 |
| 1.4 Alcances y limitaciones..... | 5 |
| 1.4.1 Alcances | 5 |
| 1.4.2 Limitaciones | 5 |
| 1.5 Antecedentes | 6 |
| 1.6 Estructura de la tesis | 7 |
| Capítulo 2 Marco teórico..... | 8 |
| 2.1 <i>Experiencia de usuario.</i> | 9 |
| 2.2 <i>Interacción Humano computadora (HCI).</i> | 9 |
| 2.3 <i>Estados cognitivos.</i> | 9 |
| 2.4 <i>Estados Afectivos (emocionales).</i> | 9 |
| 2.5 <i>Sensores fisiológicos</i> | 10 |
| 2.6 <i>Reconocimiento de expresiones faciales (FER)</i> | 11 |
| 2.7 <i>Reconocimiento facial</i> | 11 |
| 2.8 <i>Expresiones faciales</i> | 11 |
| 2.9 <i>Reconocimiento de rostro</i> | 12 |
| 2.10 <i>Pre procesamiento</i> | 13 |
| 2.11 <i>Segmentación</i> | 13 |
| 2.12 <i>Extracción de características.</i> | 13 |
| 2.13 <i>Entrenamiento</i> | 13 |
| 2.14 <i>Reconocimiento</i> | 14 |
| 2.15 <i>Puntos de referencia facial</i> | 14 |
| 2.16 <i>Sistema de codificación de acción facial.</i> | 14 |
| 2.17 <i>Sistema de codificación de acción facial para emociones.</i> | 15 |
| 2.18 <i>Unidades de Acción.</i> | 16 |

| | | |
|------------|---|----|
| 2.19 | <i>Aprendizaje automático.</i> | 16 |
| 2.20 | <i>Métricas de evaluación de algoritmos.</i> | 17 |
| 2.21 | <i>Matriz de confusión.</i> | 17 |
| 2.21.1 | Accuracy | 18 |
| 2.21.2 | Precision | 19 |
| 2.21.3 | Recall | 20 |
| 2.21.4 | F1-Score | 20 |
| 2.22 | <i>Máquina de Soporte Vectorial.</i> | 21 |
| 2.23 | <i>RandomForest.</i> | 22 |
| 2.24 | <i>K vecinos más cercanos</i> | 23 |
| Capítulo 3 | Estado del arte | 24 |
| 3.1 | <i>Estado del Arte de Detección de Rostro y Reconocimiento de Expresiones Faciales.</i> | 26 |
| 3.2 | <i>Estado del Arte de Reconocimiento de Emociones.</i> | 29 |
| 3.3 | <i>Resumen Comparativo de las Técnicas Utilizadas en el Estado del Arte para el Reconocimiento de Expresiones, Emociones Faciales y Estados Cognitivos.</i> | 40 |
| Capítulo 4 | Metodología de solución | 43 |
| 4.1 | Esquema de metodología de solución | 44 |
| 4.2 | Descripción de la Metodología de Solución | 44 |
| Capítulo 5 | Pruebas y resultados | 47 |
| 5.1 | experimentación Y Resultados | 48 |
| 5.1.1 | Repositorio de imágenes | 48 |
| 5.1.2 | Métricas para evaluar | 51 |
| 5.1.3 | Experimentación | 51 |
| 5.1.4 | Resultados | 92 |
| Capítulo 6 | Conclusión y trabajos futuros | 93 |
| 6.1 | Conclusión | 93 |
| 6.2 | Trabajos futuros | 95 |
| 6.2.1 | Unidades de acción para la detección de Interés | 96 |
| 6.2.2 | Unidades de acción para la detección del aburrimiento | 97 |
| 6.2.3 | Unidades de acción para la detección de la distracción | 98 |

| | | |
|-------|---|-----|
| 6.3 | Aportaciones | 99 |
| 6.3.1 | Procedimiento para la extracción de características con base geométrica. | 99 |
| 6.3.2 | Desarrollo e implementación del módulo de análisis facial para emociones en la plataforma UXLab. | 99 |
| 6.4 | Publicaciones | 100 |
| 7 | Referencias y bibliografía..... | 101 |
| 7.1 | Referencias..... | 101 |
| 7.2 | Bibliografías..... | 104 |

Índice de Figuras

| | | |
|-------------------|--|----|
| <i>Figura 2.1</i> | Clasificación de estados mentales (Soriano Terrazas, 2018)..... | 10 |
| <i>Figura 2.2</i> | Reconocimiento Facial (Lacort, J., 2017) | 11 |
| <i>Figura 2.3</i> | Expresiones faciales de emociones básicas, ejemplo de expresión facial de imágenes de la base de datos JAFFE. Cada columna representa las emociones de neutral, felicidad, enojo, tristeza, sorpresa, disgusto, y miedo respectivamente de izquierda a derecha (Ho An Kwang, Jin Chung Myung, 2009)..... | 12 |
| <i>Figura 2.4</i> | Representación de la matriz de confusión Shu et al (2018). | 18 |
| <i>Figura 2.5</i> | Representación de la métrica Accuracy Shu et al (2018) | 19 |
| <i>Figura 2.6</i> | Representación de la métrica Precision Shu et al (2018) | 19 |
| <i>Figura 2.7</i> | Representación de la métrica Recall Shu et al (2018) | 20 |
| <i>Figura 3.1</i> | Esquema de metodología de solución (Valderrama Cárdenas, 2019)..... | 27 |
| Figura 3.2 | Etapas primarias de un sistema de reconocimiento facial (Cadena Moreano J. A., et al, 2019)..... | 28 |
| <i>Figura 3.3</i> | Resultados obtenidos en porcentajes de precisión de reconocimiento de emociones (Yang, et al, 2017) | 30 |
| <i>Figura 3.4</i> | Resultados de reconocimiento facial de emociones mediante la metodología de extracción de características geométricas (Dewi, et al, 2018) | 33 |
| <i>Figura 3.5</i> | Metodología de solución propuesta por (Yi, 2019). El FESI incluye principalmente tres contenidos: el cálculo de la distancia euclidiana dentro del cuadro, el ajuste de la polilínea y el cálculo de la pendiente, la determinación de dos cuadros clave..... | 37 |
| <i>Figura 3.6</i> | Precisiones obtenidas con diferentes bases de datos (Yi, 2019) | 37 |
| <i>Figura 4.1</i> | Esquema de Metodología de solución | 44 |
| <i>Figura 5.1</i> | Rostros con las emociones de Felicidad, Enojo, Sorpresa y Neutral respectivamente del conjunto de datos CK+ Lucey, P (2010)..... | 49 |

| | |
|--|----|
| <i>Figura 5.2</i> Rostros con las emociones de Felicidad, Enojo, Sorpresa y Neutral respectivamente del conjunto de datos FER Sambare, M. (2020). | 50 |
| <i>Figura 5.3</i> Rostros con las emociones de Felicidad, Enojo, Sorpresa y Neutral respectivamente del conjunto de datos PROPIO | 51 |
| <i>Figura 5.4</i> Los 68 puntos de referencia faciales del conjunto de datos iBUG 300-W (ibug) | 52 |
| <i>Figura 5.5</i> Los 68 puntos de referencia faciales posicionados en el rostro | 53 |
| <i>Figura 5.6</i> Detección del rostro con diferentes poses y variabilidad de iluminación persona uno. | 54 |
| <i>Figura 5.7</i> Detección del rostro con diferentes poses y variabilidad de iluminación persona dos. | 54 |
| <i>Figura 5.8</i> Detección del rostro con diferentes poses y variabilidad de iluminación persona tres. | 54 |
| <i>Figura 5.9</i> localización de los puntos de referencia facial en el rostro de la persona uno con diferente orientación y variación de luminosidad en la imagen. | 55 |
| <i>Figura 5.10</i> localización de los puntos de referencia facial en el rostro de la persona dos con diferente orientación y variación de luminosidad en la imagen | 55 |
| <i>Figura 5.11</i> localización de los puntos de referencia facial en el rostro de la persona tres con diferente orientación y variación de luminosidad en la imagen. | 55 |
| <i>Figura 5.12</i> Esquema del sistema de clasificación de emociones | 56 |
| <i>Figura 5.13</i> Puntos faciales en un rostro con expresión de enojo del conjunto de datos CK+ | 57 |
| <i>Figura 5.14</i> Puntos [71, 72] del parpado superior e inferior del ojo izquierdo. | 61 |
| <i>Figura 5.15</i> Gráfico de resultado de las métricas de Precision y Recall de los 3 algoritmos. | 65 |
| <i>Figura 5.16</i> Gráfico de resultado de las métricas F1- Score de los 3 algoritmos. | 65 |
| <i>Figura 5.17</i> Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Felicidad..... | 66 |
| <i>Figura 5.18</i> Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino. | 67 |
| <i>Figura 5.19</i> Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de Felicidad..... | 68 |
| <i>Figura 5.20</i> Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino..... | 69 |
| <i>Figura 5.21</i> Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Enojo. | 69 |
| <i>Figura 5.22</i> Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino. | 70 |
| <i>Figura 5.23</i> Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de Enojo. | 71 |
| <i>Figura 5.24</i> Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino..... | 72 |



Figura 5.25 Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Sorpresa.....73

Figura 5.26 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.....74

Figura 5.27 Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de sorpresa.74

Figura 5.28 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino.....75

Figura 5.29 Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Neutral.....76

Figura 5.30 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.....77

Figura 5.31 Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de Neutral.....77

Figura 5.32 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino.....78

Figura 5.33 Grafico de promedio de reconocimiento de cada emoción resultado de las pruebas.79

Figura 5.34 Emoción de Felicidad y sus Unidades de Acción.80

Figura 5.35 Emoción de Enojo y sus Unidades de Acción.81

Figura 5.36 Emoción de Sorpresa y sus Unidades de Acción.82

Figura 5.37 Secuencia de video detectando la emoción Neutra.82

Figura 5.38 Modulo de Pos procesamiento de datos de seguimiento ocular de UXLab.....84

Figura 5.39 Modulo de pos-procesamiento de datos de análisis facial de UXLab84

Figura 5.40 Integración del Módulo de Análisis Facial al Reproductor de UXLab, visualización de los módulos en conjunto.....85

Figura 5.41 Visualización del Inicio de una sesión de evaluación de la UX en la plataforma UXLab para su análisis86

Figura 5.42 Visualización del Inicio de los datos de una sesión de evaluación de la UX en la plataforma UXLab para su análisis87

Figura 5.43 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro neutral88

Figura 5.44 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro alegre88

Figura 5.45 Visualización del Inicio de una sesión de evaluación de la UX en la plataforma UXLab para su análisis89

Figura 5.46 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro neutral.....90

Figura 5.47 Visualización del estado de ansiedad del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro alegre91

| | |
|---|----|
| <i>Figura 5.48 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro alegre.</i> | 91 |
| <i>Figura 5.49 Gráfico de comparación del trabajo de M. Alvares y el propuesto.</i> | 92 |
| <i>Figura 6.1 Visualización de los datos procesados en el grabador de UXLab de una sesión de evaluación UX.</i> | 95 |
| <i>Figura 6.2 Unidad de Acción 2 Levantador de cejas externo Noldus (2021)</i> | 96 |
| <i>Figura 6.3 AU 1. Levantador de cejas interior Noldus (2021)</i> | 96 |
| <i>Figura 6.4 AU 5. Elevador de tapa superior Noldus (2021)</i> | 97 |
| <i>Figura 6.5 AU 17. Mentonera Noldus (2021)</i> | 97 |
| <i>Figura 6.6 AU 23. Tensor de labios Noldus (2021)</i> | 97 |
| <i>Figura 6.7 AU 24. Labios presionados Noldus (2021)</i> | 98 |
| <i>Figura 6.8 AU 43. Ojos cerrados Noldus (2021)</i> | 98 |

Índice de tablas

| | |
|---|----|
| <i>Tabla 2.1 Clasificación de los estados afectivos (Soriano Terrazas, 2018)</i> | 10 |
| <i>Tabla 2.2 Unidades de Acción que conforman cada emoción en el EMFACS. P.E Group. (2016)</i> | 15 |
| <i>Tabla 3.1 Cadenas de búsqueda o palabras claves</i> | 25 |
| <i>Tabla 3.2 Selección de trabajos relacionados</i> | 25 |
| <i>Tabla 3.3 Comparación de precisión de identificación de emociones con otros clasificadores y el método propuesto por (Dewi, et al, 2018)</i> | 33 |
| <i>Tabla 3.4 total, de precisiones de todos los algoritmos probados (Kartali et al, 2018)</i> | 35 |
| <i>Tabla 3.5 Comparativa de identificación de emociones y estados cognitivos del estado del arte</i> | 40 |
| <i>Tabla 5.1 Características del Repositorio de imágenes propio</i> | 51 |
| <i>Tabla 5.2 Puntos relevantes para generar las distancias que nos ayudaran a identificar las 4 emociones.</i> | 57 |
| <i>Tabla 5.3 Relación entre los puntos faciales y las unidades de acción de la emoción de Felicidad</i> | 58 |
| <i>Tabla 5.4 Relación entre los puntos faciales y las unidades de acción de la emoción de Enojo</i> | 58 |
| <i>Tabla 5.5 Relación entre los puntos faciales y las unidades de acción de la emoción de Sorpresa</i> | 59 |
| <i>Tabla 5.6 Métricas del Algoritmo MSV</i> | 62 |
| <i>Tabla 5.7 Muestra que tomo el clasificador para realizar las pruebas</i> | 62 |
| <i>Tabla 5.8 Métricas del Algoritmo Random Forest</i> | 63 |
| <i>Tabla 5.9 Muestra que tomo el clasificador para realizar las pruebas</i> | 63 |
| <i>Tabla 5.10 Métricas del Algoritmo KNN</i> | 63 |
| <i>Tabla 5.11 Muestra que tomo el clasificador para realizar las pruebas</i> | 64 |
| <i>Tabla 5.12 Características de las personas que se les realizo las pruebas con el clasificador</i> | 66 |
| <i>Tabla 5.13 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de alegría</i> | 66 |

| | |
|--|----|
| Tabla 5.14 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de alegría..... | 68 |
| Tabla 5.15 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de enojo..... | 70 |
| Tabla 5.16 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de enojo..... | 71 |
| Tabla 5.17 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de sorpresa. | 73 |
| Tabla 5.18 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de sorpresa. | 75 |
| Tabla 5.19 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de neutral. | 76 |
| Tabla 5.20 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de neutral. | 78 |
| Tabla 5.21 Promedio de reconocimiento de cada emoción de los resultados de la prueba | 79 |
| Tabla 5.22 Porcentaje de reconocimiento de emociones en video. | 83 |
| Tabla 5.23 Resultados obtenidos por M. Alvares et al (2018) | 92 |
| Tabla 5.24 Resultados obtenidos en la investigación. | 92 |
| Tabla 6.1 Códigos de unidades de acción de la cabeza P.E Group (2016)..... | 98 |
| Tabla 6.2 Códigos de unidades de acción de los ojos P.E Group (2016)..... | 99 |

Capítulo 1

Introducción

La experiencia del usuario (UX) se puede considerar como un atributo de calidad de un producto, sistema o servicio digital, definiendo si es fácil de adaptar, eficiente de usar y cómodo. La evaluación de la calidad de la experiencia del usuario (QUX por sus siglas en inglés) puede usarse como una herramienta para determinar el grado de experiencia del usuario para un sistema o servicio.

La evaluación UX incluye todas las emociones, las creencias, las preferencias, las percepciones, las respuestas físicas, las psicológicas, los comportamientos y logros que ocurren antes, durante y después de su uso (Hussain, et al., 2018).

Durante el proceso de evaluación de la UX, se utilizan instrumentos de recolección de información convencionales, tales como cuestionarios o entrevistas después de que el usuario que realiza la evaluación interactúa con un producto, digital. Por lo que es difícil precisar la retroalimentación de los usuarios, ya que estos instrumentos dependen de la voluntad y de la habilidad del usuario para poder expresar cómo se sintió durante la interacción con el producto digital, una de las desventajas de realizar la evaluación de la UX con métodos convencionales es que los datos obtenidos pueden ser manipulados e interpretados de manera diferente por los evaluadores. Una alternativa que permite complementar estos instrumentos convencionales es el registro y procesamiento de datos fisiológicos y de comportamiento, por ejemplo, procesamiento de datos EEG, ECG y GSR, en conjunto con el análisis facial, el análisis de la postura, el seguimiento ocular, la pupilometría, análisis de voz, entre otros, los cuales, se pueden utilizar para identificar el estado emocional y/o cognitivo del usuario durante el proceso de evaluación UX.

La información fisiológica que se puede obtener de un usuario es amplia y heterogénea. Por ejemplo, se pueden registrar datos fisiológicos como el ritmo cardíaco (ECG), de conductividad en piel (GSR), la actividad bioeléctrica del cerebro (EEG), la actividad bioeléctrica en músculos (EMG), y también se pueden registrar datos biométricos como, el movimiento ocular (ET), las expresiones faciales y la postura del cuerpo, entre otras (García Pinzón 2018), con el registro y procesamiento de estos datos es posible identificar estados

emocionales y cognitivos durante el proceso de evaluación de la UX de un producto digital. De este modo se pueden analizar datos reales durante la interacción con un producto digital de los usuarios participantes en el proceso de evaluación de la experiencia de usuario, siendo más difíciles de alterar y más confiables para la evaluación UX.

El ser humano utiliza consciente e inconscientemente las expresiones faciales para mostrar sus estados emocionales. Sin embargo, el reconocimiento de la expresión facial presenta problemas técnicos desafiantes en el área de la visión por computadora por lo que se han desarrollado algoritmos que analizan imágenes de rostros humanos (Valderrama Cárdenas, 2019; Moreano, et al, 2019; y Sawhney Shreyak et al, 2019) capaces de identificar estados emocionales (Sunitha et al, 2017; Yang et al, 2017; Nwosu et al, 2017; Krestinskaya, y Pappachen, 2017; Liu Xiao, y Lee Kiju, 2018; M. Alvarez et al, 2018; Dewi et al, 2018; Mangala, y Prajwala, 2018; Kartali et al, 2018; Rabhi Yassine et al, 2018; Verma monu et al, 2018; y Yi, et al 2019) y estados cognitivos (Vázquez Rodríguez, 2017; Candra Kirana Karetika et al, 2018; H.S Gunavathi, y M. Siddappa, 2018; y Gama Velasco, 2018), otros trabajos se desarrollaron con un enfoque diferente, como la interacción humano computadora, la educación, la seguridad, la recuperación de imágenes y los videos de personas (Valderrama Cárdenas, 2019; Moreano, et al, 2019; y Sawhney Shreyak et al, 2019). Sin embargo, se encontraron pocos trabajos enfocados en el proceso de evaluación de la QUX de productos digitales (Md. Munim et al, 2017; y Liu Xiao, Lee Kiju, 2018).

Por lo consiguiente, el objetivo de este trabajo de investigación se centró en la evaluación de algoritmos para el procesamiento de imágenes de expresiones faciales que permitan identificar emociones básicas en usuarios que participan en un proceso de evaluación de la experiencia de usuario (UX por sus siglas en inglés), con el objetivo de aumentar el nivel de precisión de las técnicas de reconocimiento de estados emocionales y cognitivos mediante la fusión y correlación con datos fisiológicos multimodales y de comportamiento e implementarlo en una plataforma de la evaluación de la experiencia de usuario.

1.1 Descripción del problema

Durante el proceso de evaluación convencional de la experiencia de usuario el evaluador pone al usuario frente a la computadora o dispositivo para que pueda interactuar con el producto digital con el objetivo de obtener información del usuario sobre el grado de satisfacción que el producto le provocó durante la interacción, después de que el usuario termina la evaluación el evaluador aplica al usuario los instrumentos necesarios para obtener la información que requiere, esto se lleva a cabo mediante instrumentos convencionales como cuestionarios y/o entrevistas, por lo que la aplicación de este tipo de instrumentos muchas veces, dependen de la voluntad del usuario, entre las desventajas identificadas es fácil de alterar la información, son relativamente

simple y carente de evidencia de datos objetivos, no se puede aplicar durante la evaluación por lo que es difícil precisar la retroalimentación de los usuarios.

Esto hace que el análisis de la evaluación no sea confiable provocando que el producto digital salga a producción con deficiencias y en ocasiones generando pérdidas monetarias para la empresa.

Por lo anterior, existe la necesidad de encontrar una mejor manera para realizar la evaluación de la experiencia de usuario, ya que nuevos productos digitales van desarrollándose día a día y no es suficiente realizar una evaluación de la experiencia de usuario convencional ya que estos instrumentos no aportan al evaluador los datos adecuados para realizar el análisis y determinar el impacto de satisfacción que el usuario tiene con el producto.

Por ello se ha buscado la forma de mejorar el procesamiento de esta información y realizar una mejor evaluación de la experiencia de usuario que se adapte a las nuevas tecnologías, enfocadas en la evaluación de la experiencia de usuario que utilizan algoritmos de aprendizaje automático y aprendizaje profundo para clasificar datos fisiológicos como las señales de EEG, ECG, GSR y biométricos como el seguimiento ocular, con los cuales se obtienen información real de las impresiones que experimenta el usuario durante la interacción y permiten identificar estados cognitivos y emocionales que son inducidos durante la interacción con el producto digital; esta información es difícil de alterar y con la ventaja de que es obtenida durante la evaluación haciéndola más confiable.

Por otro lado, existe una parte del ser humano en donde se puede obtener información importante que permite conocer de manera más rápida el impacto emocional y el nivel de satisfacción del usuario con respecto al producto digital; se trata del rostro, que por medio del análisis de expresiones faciales se pueden clasificar las emociones, esta información puede ser relevante en la nueva forma de realizar la evaluación de la experiencia de usuario y poder aportar al evaluador información que complemente y contribuya a un mejor análisis de la evaluación y evitar que el producto salga a producción con deficiencias.

Las técnicas computacionales para el reconocimiento de expresiones faciales tienen muchas áreas de aplicación tales como la interacción humana computadora (HCI por sus siglas en inglés) que se utiliza en los robots sociales, la seguridad, la vigilancia y los juegos interactivos de computadoras. También se utiliza en la ciencia del comportamiento para proporcionar información social (origen, edad y género) y se utiliza en la medicina científica para la monitorización del dolor, tratamiento del retraso mental, la depresión y la ansiedad (A. Sunitha, et al, 2017).

Sin embargo, en el estado del arte se encontraron trabajos de investigación relacionada solo con el reconocimiento de emociones, pero enfocada a otro tipo de necesidad, solo uno de ellos está enfocado al reconocimiento de emociones durante el proceso de evaluación de la Experiencia de usuario (UX).

Por lo anterior, en este trabajo de investigación se presenta el resultado de la evaluación de algoritmos capaces de identificar estados emocionales y cognitivos mediante el análisis facial, pero con la diferencia de que se aplicaron durante el proceso de evaluación de la experiencia de usuario (UX) tomando como referencia los trabajos encontrados en el estado del arte para igualar o mejorar la precisión y que el costo computacional no sea tan elevado.

1.2 Justificación

Los humanos usan expresiones faciales para mostrar sus emociones. Sin embargo, el reconocimiento de la expresión facial es un problema desafiante en el área de visión por computadora, para enfrentar este problema se han desarrollado algoritmos capaces de detectar los estados emocionales y cognitivos, pero la mayoría lo han hecho con un enfoque diferente ya sea en aplicaciones como la interacción humano computadora, medicina (rehabilitación, terapia, asesoramiento, etc.), e-learning, entretenimiento, monitoreo de emociones, mercadeo, derecho (Kartali aneta, et al 2018; M. Alvarez et al, 2018; Dewi et al, 2018; Mangala, y Prajwala, 2018; Kartali et al, 2018; Rabhi Yassine et al, 2018; Verma monu et al, 2018; y Yi, et al 2019).

En la investigación del estado del arte se encontró un trabajo enfocado en el reconocimiento de emociones durante el proceso de evaluación de la experiencia de usuario de productos digitales el cual tiene por objetivo diseñar y desarrollar una herramienta de evaluación de la experiencia del usuario basada en la web utilizando la medición de las emociones del usuario. En este trabajo se propone un marco conceptual y se implementó un sistema usando Affdex SDK. Finalmente, se probó la herramienta con cuatro participantes (Munim Kazi Md, et al, 2017). Por lo que no se realizaron pruebas confiables.

Por lo anterior, la motivación para desarrollar esta investigación es complementar el proceso de la UX con el enfoque del reconocimiento de emociones mediante el análisis de expresiones faciales y que en conjunto con el análisis de datos fisiológicos obtenidos mediante sensores pueda determinar a un más el impacto emocional del usuario, evitando dejar en último plano la información importante que el rostro del usuario nos puede proporcionar, por lo que se realizó una evaluación y desarrollo de algoritmos capaces de identificar los estados emocionales básicos mediante la detección y análisis facial, durante el proceso de evaluación de la experiencia de usuario para incorporarlo al conjunto de datos multimodales y evaluar la precisión del reconocimiento de los estados emocionales y cognitivos en el proceso de la UX.

1.3 Objetivo General

Desarrollar y evaluar un algoritmo para la detección de tres emociones básicas, mediante el análisis de expresiones faciales en usuarios que realizan el proceso de evaluación de la experiencia de usuario de productos digitales.

1.3.1 Objetivos Específicos

- ❖ Evaluar la precisión de técnicas y algoritmos que mediante el análisis facial realicen el reconocimiento de tres emociones básicas como la felicidad, el enojo, sorpresa y el neutro.
- ❖ Implementar un algoritmo para identificar estados emocionales en el proceso de evaluación de la experiencia de usuario.
- ❖ Incorporar los algoritmos para el reconocimiento de estados emocionales en el módulo de análisis facial de la plataforma UXLab.

1.4 Alcances y limitaciones

1.4.1 Alcances

En este trabajo de investigación se evaluaron tres algoritmos de reconocimiento de emociones mediante el análisis de expresiones faciales encontrados en la literatura utilizando como dispositivo de entrada una cámara web convencional, durante el proceso de reconocimiento se pretende identificar dos emociones positivas (felicidad, y sorpresa), una emoción negativa (Enojo) y la emoción neutral. Posteriormente una vez identificado el mejor algoritmo se incorporó al módulo de análisis facial en la plataforma de evaluación de la experiencia de usuario llamado UXLab el cual se viene desarrollando desde el 2016 por alumnos de maestría y doctorado del TecNM/CENIDET.

1.4.2 Limitaciones

Dentro de las limitaciones que se tienen en esta investigación se encuentran las siguientes.

- No se realizarán nuevos algoritmos de reconocimiento facial.
- No se desarrollará ningún algoritmo de aprendizaje automático nuevo.
- No se usará ningún sensor de movimiento como Kinect
- No se trabajará con sensores fisiológicos.
- No se buscará optimizar ningún algoritmo de reconocimiento facial.

1.5 Antecedentes

Dentro del grupo de investigación de Sistemas Híbridos Inteligentes del TecNM/CENIDET se ha realizado varias investigaciones que abarcan diferentes áreas relacionadas a las Interacción Humano Computadora (HCI por sus siglas en inglés) y con el uso de sensores biométricos y fisiológicos, para realizar el proceso de evaluación de la Experiencia de Usuario (UX por sus siglas en inglés).

A continuación, se describe los trabajos de investigación realizados en el TecNM/CENIDET relacionados con esta investigación.

1.5.1. Método para integrar y sincronizar datos EEG y multimedia para su aplicación en el proceso de evaluación de la experiencia del usuario (Fouilloux quiroz, 2018).

En esta tesis de investigación se desarrolló una herramienta de software que captura y almacena información proveniente de múltiples dispositivos para registrar datos multimodales durante el proceso de evaluación de la experiencia de usuario (UX) de un producto digital, esta herramienta graba y sincroniza datos multimodales de un usuario, como: audio, video del rostro, video de la pantalla, y datos fisiológicos EEG.

En este trabajo se desarrolló la primera versión de la plataforma para la evaluación de la UX actualmente denominada UXLab el cual se usará en esta investigación.

1.5.2. Sistema embebido para analizar datos biométricos e identificar estados cognitivos y emocionales en procesos de evaluación de la experiencia del usuario (García pinzón, 2018).

En esta tesis se desarrolló un sistema que permite la detección de estados emocionales y cognitivos mediante el procesamiento de datos fisiológicos utilizando sensores fisiológicos como electrocardiograma (ECG), electroencefalograma (EEG) y respuesta galvánica en piel (GSR), estos sensores de bajo costo se diseñaron y calibraron para incorporarlos al sistema de grabación UXLab.

1.5.3. Algoritmo para analizar el movimiento ocular del usuario para identificar la correlación con estados mentales (Lagunés Ramírez, 2020).

En esta tesis de maestría en desarrollo se diseñó un método y algoritmo para analizar el movimiento ocular de un usuario durante el proceso de evaluación UX para obtener datos

relevantes y generar mapas de calor, de trazo de trayectorias y áreas de interés correlacionados con el estado cognitivo que experimenta durante el proceso de evaluación de un producto digital. Esta tesis propone la implementación del algoritmo en un módulo de visualización en la plataforma de evaluación de experiencia de usuario UXLab.

1.5.4. Máquina de aprendizaje para el reconocimiento de emociones mediante datos fisiológicos multimodales. (Morales Morante, 2020).

En esta tesis de maestría en desarrolló una metodología y algoritmos basados en máquina de aprendizaje para el reconocimiento de emociones básicas este algoritmo utiliza como entrada datos fisiológicos multimodales de electroencefalograma (EEG), electrocardiograma (ECG) y respuesta galvánica en piel (GSR) para su evaluación estos algoritmos de clasificación se incorporaron en la plataforma para evaluar la experiencia de usuario UXLab.

1.6 Estructura de la tesis.

El documento está conformado por siete capítulos, los cuales describen a fondo cada una de las etapas que se realizaron para el desarrollo de esta investigación. El capítulo I, presenta el análisis del problema de la investigación, la justificación, objetivos, alcances y limitaciones y los antecedentes que se tienen sobre esta tesis. El capítulo II, presenta el marco teórico en donde se describen todos los conceptos de que se necesitan conocer para que se tenga un mejor entendimiento de lo que en este documento se estará hablando, el capítulo III, presenta una recopilación de los trabajos realizados más relevantes que se relacionan con el tema de investigación y que sirvieron como referentes para poder solucionar la problemática de esta tesis, el capítulo IV describe la metodología de solución propuesta, la cual se basa en el estado del arte. En esta metodología se describe las fases que se llevaron a cabo para darle solución al problema de investigación, en el capítulo V, se describe la implementación del algoritmo de aprendizaje automático con la finalidad de poder clasificar las emociones que un usuario presenta durante la evaluación de la experiencia de usuario, en este mismo capítulo se presentan las pruebas y los resultados obtenidos con la implementación del algoritmo a la plataforma UXLab, en el capítulo VI, se presenta las conclusiones y los trabajos futuros de esta investigación, finalmente las referencias consultadas para la realización de esta tesis

Capítulo 2

Marco teórico

En este capítulo se describen los conceptos que se utilizan en el tema de investigación para un mejor entendimiento del tema. Además, se explica el funcionamiento de los diferentes algoritmos utilizados en el sistema desarrollado, así como las métricas para evaluar los resultados de los algoritmos utilizados para el desarrollo de este trabajo.

2.1 Experiencia de usuario.

La evaluación de la experiencia de usuario (UX por sus siglas en inglés) abarca todos los aspectos de la interacción del usuario final con la compañía, sus servicios y sus productos. La evaluación UX también puede usarse para formar la estructura e identificar los atributos de diseño de un sistema. Básicamente la evaluación UX incluyen un conjunto de procedimientos o métodos para evaluar la pertinencia frente a las características específicas del producto, proceso o servicio a evaluar (Kazi Md et. Al 2017).

2.2 Interacción Humano computadora (HCI).

El termino Interacción Humano Computadora es utilizado en el estudio del diseño y el uso de tecnología computacional enfocado en las interfaces aplicadas en la comunicación entre una persona y una computadora. Este enfoque implica cualquier herramienta que tenga como objetivo el dialogo abierto entre las dos entidades (Jacko,2012).

2.3 Estados cognitivos.

Un estado cognitivo es una condición neuropsicológica que presenta un sujeto durante la realización de una o más tareas cognitivas, las cuales contemplan los procesos necesarios para el cálculo, la concentración y la memoria, además de las funciones ejecutivas, de las cuales se consideran las siguientes:

- Concentración.
- Planificación de tareas.
- Resolución de problemas.
- Formular conceptos abstractos.
- Autoconciencia.

2.4 Estados Afectivos (emocionales).

Se entiende por estados afectivos a la condición neuropsicológica que presenta un sujeto. Los estados afectivos se clasifican en emociones y sentimientos. Los estados afectivos se diferencian principalmente por el tiempo de duración, los sentimientos tienen una duración

mayor que las emociones. En la tabla 2.1. se presenta la clasificación de los estados afectivos y algunos ejemplos que permiten diferenciarlos (Soriano Terrazas, 2018).

Tabla 2.1 Clasificación de los estados afectivos (Soriano Terrazas, 2018)

| Emociones | Sentimientos |
|-----------|---------------|
| Miedo | Vergüenza |
| Ira | Resentimiento |
| Tristeza | Amor |
| Alegría | Simpatía |
| Sorpresa | Orgullo |
| Asco | Celos |

Los estados emocionales se clasifican en dos tipologías: estados emocionales positivos y estados emocionales negativos. En la **¡Error! No se encuentra el origen de la referencia.** 2.1 se muestra una categorización de estados mentales según la neurociencia (Soriano Terrazas, 2018).

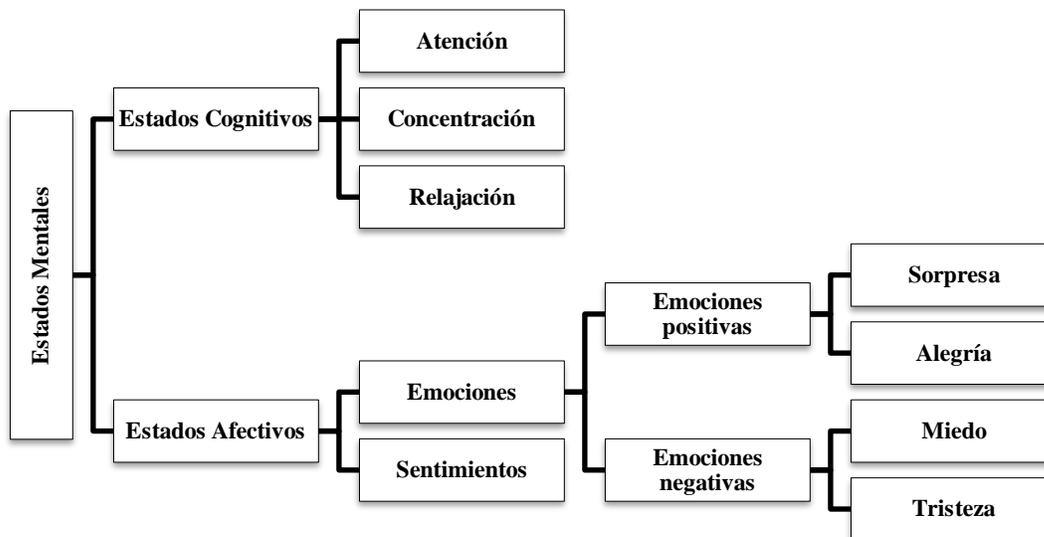


Figura 2.1 Clasificación de estados mentales (Soriano Terrazas, 2018)

2.5 Sensores fisiológicos

Los sensores fisiológicos permiten medir ciertas características de las funciones y mecanismos que tienen los seres vivos, por ejemplo, los sensores de electrocardiograma (ECG), respuesta galvánica en piel (GSR), los sensores de electroencefalograma (EEG), entre otros. Ayudan a los investigadores a obtener pistas del comportamiento humano cuando las personas se enfrentan a ciertos estímulos.

2.6 Reconocimiento de expresiones faciales (FER)

El reconocimiento facial de la expresión emocional es la capacidad de todos los individuos de reconocer formas básicas de expresión afectiva, la cual aparece en los rostros de las personas (Iglesias-Hoyos, et al.2016).

2.7 Reconocimiento facial

Método por el cual una cara humana es capturada para, posteriormente, ser utilizada con el fin de identificar a un individuo. Estos métodos miden y analizan la estructura y la forma de la cara, sus proporciones (distancia entre los ojos, nariz y boca; el contorno de las cuencas de los ojos; los lados de la boca; la localización de la nariz y la boca; etc.). En la verificación de identidad, el usuario debe situarse frente a una cámara. Para prevenir un posible ataque mediante el uso de una imagen en lugar del individuo real, estos sistemas, en ocasiones, esperan el parpadeo de los ojos, solicitan que el usuario sonría, haga algún gesto o utilizan otra técnica con la que se comprueba la temperatura de la cara: la termografía facial. La Figura 2.2 muestra un ejemplo de cómo un sistema realiza el reconocimiento facial (Simancas Ruesgas Borja, 2019)

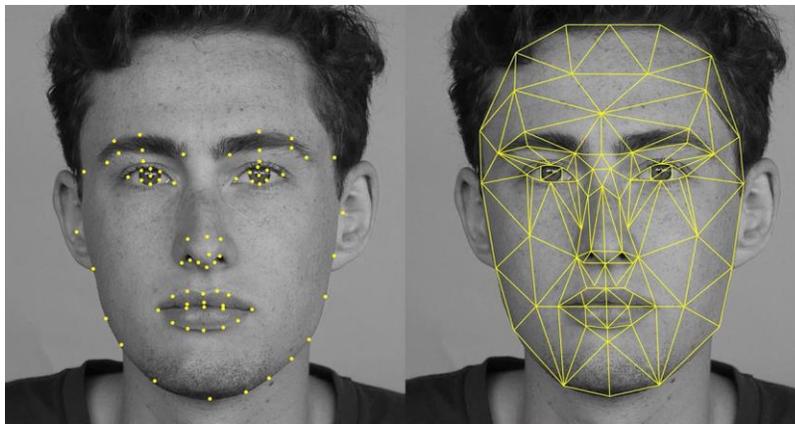


Figura 2.2 Reconocimiento Facial (Lacort, J., 2017)

2.8 Expresiones faciales

Las expresiones de la cara son probablemente la variable que más se observa para obtener información de las emociones de los interlocutores. Las expresiones faciales de estas emociones son comunes en todas las culturas, e innatas, es decir, se nace con la capacidad de emitirlas, no se

tiene que aprender a hacerlo (Murase H, Nayar S, 1995). En la Figura 2.3 se muestra las expresiones básicas.



Figura 2.3 Expresiones faciales de emociones básicas, ejemplo de expresión facial de imágenes de la base de datos JAFFE. Cada columna representa las emociones de neutral, felicidad, enojo, tristeza, sorpresa, disgusto, y miedo respectivamente de izquierda a derecha (Ho An Kwang, Jin Chung Myung, 2009).

2.9 Reconocimiento de rostro

El reconocimiento de rostros es un área del reconocimiento de patrones que ha sido investigada desde hace varios años. Durante todo este tiempo se han propuesto técnicas diferentes para la resolución de esta tarea, entre las que se pueden mencionar (Valvert Gamboa J.R, 2006).

- Igualamiento de plantillas
- Cálculo de Eigen Faces geométricas
- Métodos que utilizan redes neuronales

Independientemente de la técnica que sea implementada, se utiliza dos conjuntos de datos.

- ✓ El primero es utilizado para la etapa de aprendizaje, el cual es llamado conjunto de entrenamiento. El objetivo es que los Patrones que integran este conjunto sean lo más diferente posibles entre sí, y que, además, represente el problema, para poder tener un buen porcentaje de generalización.
- ✓ El segundo conjunto de patrones es llamado conjunto de pruebas, y es utilizado en la etapa de reconocimiento.

Por último, para reconocer un rostro debe realizar los siguientes pasos generales:

- ✓ Preprocesamiento de la imagen
- ✓ Extracción de características
- ✓ Entrenamiento
- ✓ Reconocimiento

A continuación, se define y describe cada uno de ellos

2.10 Pre procesamiento

El objetivo del preprocesamiento es conseguir una mejora en la calidad y/o la apariencia de la imagen original. Resalta ciertas características de una imagen: bordes, contrastes y, se ocultan o eliminan otras, por ejemplo, el ruido. El preprocesamiento es una etapa necesaria para otras fases posteriores de análisis tales como: segmentación, extracción de características, reconocimiento, e interpretación o calificación. Los principales tipos de preprocesamiento son el realzado y filtrado (Vázquez Rodríguez C, 2017).

2.11 Segmentación

La segmentación es un proceso por el cual se extrae de la imagen cierta información subyacente para su posterior uso. La segmentación está basada en dos principios fundamentales: discontinuidad y similitud (Vázquez Rodríguez C, 2017).

La segmentación orientada a bordes (discontinuidad) y orientada a regiones (similitud). Un borde es una línea que separa dos regiones por tanto de diferentes propiedades por ello se le conoce como discontinuidad. Una región en general es un área de la imagen en la que sus píxeles poseen propiedades similares ya sea de intensidad, de color u otras. Tanto la detección de bordes como la de regiones implican una manipulación de la imagen original, que supone en definitiva una transformación de la imagen original de forma que los valores de los píxeles originales son modificados mediante ciertas funciones (Vázquez Rodríguez C, 2017).

2.12 Extracción de características.

Una vez segmentada la imagen, es necesario obtener las características de las distintas regiones de interés. La extracción de características basadas en el contenido de la imagen, tanto en la parte de creación como de recuperación del banco de datos trabaja a partir de imágenes a color; para realizarlo es necesario extraer patrones que permitan representarlo y describirlo ya sea por su color, tamaño, posición, entre otros. Para hacer una buena descripción de características debe tratarse de una descripción única y completa por lo que no deberá presentar ambigüedades (Vázquez Rodríguez C, 2017).

2.13 Entrenamiento.

El entrenamiento básicamente consiste en utilizar alguna forma de aprendizaje que le permita al sistema aprender rostros que formaran el conjunto de entrenamiento. El tipo de entrenamiento a utilizar dependerá de la metodología utilizada para el reconocimiento.

2.14 Reconocimiento

La última etapa, en un sistema de reconocimiento de rostros es la etapa de reconocimiento. Consiste en alimentar al sistema con imágenes de rostros diferentes a las utilizadas durante el entrenamiento, esperando obtener como resultado, alguna forma de codificación que nos permita saber de qué rostro se trata.

2.15 Puntos de referencia facial

La detección de puntos de referencia faciales se define como una tarea para detectar puntos de referencia clave en la cara y rastrearlos (siendo resistente a deformaciones faciales rígidas y no rígidas debido a movimientos de la cabeza y expresiones faciales). Los puntos de referencia faciales se utilizan para localizar y representar regiones destacadas de la cara, como ojos, cejas, nariz, mandíbula, boca, etc. (Rosebrock, A. 2021; Mallick, S. 2021)

Es una técnica que se ha utilizado en aplicaciones como alineación facial, estimación de la postura de la cabeza, intercambio de caras, detección de parpadeo, detección de somnolencia, detección de expresiones faciales, etc. (Rosebrock, A. 2021; Mallick, S. 2021)

En el contexto de los puntos de referencia faciales, es necesario detectar las estructuras faciales importantes en la cara utilizando métodos de predicción de forma. La detección de puntos de referencia faciales implica dos pasos (Rosebrock, A. 2021; Mallick, S. 2021):

1. Localización del rostro detectado en la imagen.
2. Detección de estructuras faciales clave en la cara.
 - Nariz
 - Mandíbulas
 - Ojos izquierdo y derecho
 - Ceja izquierda y derecha
 - Boca

2.16 Sistema de codificación de acción facial.

El Sistema de Codificación de Acción Facial (FACS) es un sistema integral basado en la anatomía para describir todos los movimientos faciales perceptibles visualmente. Descompone las expresiones faciales en componentes individuales del movimiento muscular, llamados Unidades de Acción (AU) P.E Group. (2016).

FACS se utiliza en muchos entornos personales y profesionales diferentes. A menudo se utiliza en diversos entornos científicos para la investigación. También lo utilizan animadores e informáticos interesados en el reconocimiento facial P.E Group. (2016).

El sistema de codificación de acción facial está definido como un estándar para la medición de las expresiones faciales, el psicólogo Paul Ekman y Wallace V. Friesen adoptaron una taxonomía de los movimientos faciales humanos por su apariencia en la cara denominado Sistema de Codificación de Acción Facial (FACS por sus siglas en inglés) publicaron un manual en 1978 y fue revisado por última vez en el año 2002 P.E Group. (2016).

ACS también puede permitir una mayor conciencia y sensibilidad a los comportamientos faciales sutiles. Tales habilidades son útiles para psicoterapeutas, entrevistadores y cualquier persona que trabaje en comunicaciones P.E Group. (2016).

Usando FACS, los humanos pueden codificar manualmente casi cualquier expresión facial anatómicamente posible, descomponiéndola en las unidades de acción específicas (AU) y sus segmentos temporales que produjeron la expresión. Como las AU son independientes de cualquier interpretación, se pueden usar para cualquier proceso de toma de decisiones de orden superior, incluido el reconocimiento de emociones básicas o comandos preprogramados para un entorno inteligente.

2.17 Sistema de codificación de acción facial para emociones.

De acuerdo con el EMFACS (Sistema de codificación de acciones faciales emocionales) que es un acrónimo de “Emotion FACS” este se basa en la utilización del Sistema de Codificación de la Acción Facial para detectar las emociones a través de la expresión de la cara. EMFACS (Emotion FACS) es una aplicación selectiva de la calificación FACS, en la que los codificadores solo califican el comportamiento que probablemente tenga un significado emocional. En este sistema sólo se consideran aquellas unidades de acción que se asocian a determinadas emociones, las conocidas como emociones básicas. Dichas emociones son la alegría, la tristeza, la sorpresa, el miedo, la ira, el disgusto y el desprecio. P.E Group. (2016)

Este sistema fue desarrollado por Paul Ekman y Wallace V. Friesen en la década de los 80 P.E Group. (2016), con la misión de simplificar la puntuación de las unidades de acción de la cara, centrándose únicamente en los signos de la emoción como se ve en la Tabla 2.2

Tabla 2.2 Unidades de Acción que conforman cada emoción en el EMFACS. P.E Group. (2016)

| Nombre | Unidad de acción |
|----------|---------------------------|
| Sorpresa | AU 1+2+5+26 |
| Miedo | AU 1+2+4+5+20+25, 26 o 27 |
| Alegría | AU 6+12 |

| | |
|-----------|--|
| Tristeza | AU 1+4+15 con o sin 54+64 |
| Asco | AU 9+17 |
| Ira | AU 4+5+7+17+23 |
| Desprecio | AU R12/L12 o R14/L14. Las letras R o L indican el lado de la cara donde se produce la acción muscular, siendo esta Left (izquierda) o Right (derecha). |

2.18 Unidades de Acción.

Cada movimiento facial observable recibe el nombre de Unidad de Acción (AU por sus siglas en inglés “Action Units”) así la composición de la expresión de una emoción puede ser dividida en una o varias AU para las distintas representaciones de una o varias emociones. La forma más eficaz para medir las emociones en el rostro humano es mediante algún sistema de codificación de acción facial con el cual se pueda relacionar el movimiento de ciertas regiones del rostro con algún identificador AU y que los algoritmos puedan aprender de ello Paul Ekman, et al. (2002).

2.19 Aprendizaje automático.

El aprendizaje automático (Machine Learning ML) es un área de la inteligencia artificial (IA) que se centra en desarrollar sistemas que aprenden, o mejoran el rendimiento, en función de los datos que consumen. Inteligencia artificial es un término amplio que se refiere a sistemas o máquinas que imitan la inteligencia humana. Se suele mencionar al aprendizaje automático y a la IA en las mismas conversaciones, y los términos a veces se usan indistintamente, pero no significan lo mismo. Un aspecto importante para destacar es que, aunque todo aprendizaje autónomo es IA, no toda IA es aprendizaje autónomo (Oracle, 2022).

Los algoritmos son los motores que impulsan el aprendizaje automático. En general, actualmente se utilizan dos tipos de algoritmos de aprendizaje: aprendizaje supervisado y aprendizaje no supervisado. La diferencia entre ellos se define por cómo cada uno aprende acerca de los datos para hacer predicciones (Oracle, 2022).

- **Aprendizaje automático supervisado:** Los algoritmos supervisados de aprendizaje autónomo son los más utilizados. Con este modelo, un científico de datos actúa como guía y enseña al algoritmo las conclusiones que debe hacer. Al igual que un niño que aprende a identificar las frutas al memorizarlas con un libro de imágenes, en el aprendizaje supervisado, el algoritmo se capacita mediante un conjunto de datos que ya está etiquetado y tiene un resultado predefinido (Oracle, 2022).
- **Aprendizaje autónomo no supervisado:** El aprendizaje autónomo no supervisado utiliza un enfoque más independiente, en el que una computadora aprende a identificar procesos y

patrones complejos sin que un ser humano proporcione una guía cercana y constante. El aprendizaje autónomo no supervisado implica la capacitación basada en datos que no tiene etiquetas o un resultado específico definido (Oracle, 2022).

Para continuar con la analogía de la enseñanza infantil, el aprendizaje autónomo no supervisado es similar a un niño que aprende a identificar frutas mediante la observación de colores y patrones, en lugar de memorizar los nombres con la ayuda de un maestro. El niño buscaría similitudes entre las imágenes y las separaría en grupos, asignando a cada grupo su propia etiqueta nueva. Los ejemplos de algoritmos de aprendizaje autónomo no supervisados incluyen el agrupamiento de k-means, el análisis de componentes principales e independientes y las leyes de asociación (Oracle, 2022).

Los ejemplos de aprendizaje autónomo supervisado incluyen algoritmos tales como regresión lineal y logística, clasificación multiclase y máquinas de vectores de soporte (Oracle, 2022).

2.20 Métricas de evaluación de algoritmos.

Las métricas para evaluar los algoritmos de clasificación se describen en las siguientes secciones:

La validación de resultados es la etapa final en el proceso de reconocimiento emocional que consiste en la evaluación y comparación de los modelos programados a partir de pruebas de entrenamiento y predicción con el repositorio desarrollado. El resultado de estas pruebas suele ser una matriz de confusión, la cual es utilizada para el cálculo de métricas especializadas en la medición de diferentes aspectos del modelo en cuestión Shu et al (2018).

2.21 Matriz de confusión.

La matriz de confusión es una herramienta estadística para la interpretación de los resultados obtenidos por un clasificador (shu et al., 2018). Ésta se describe en cuestión de cuatro valores:

- Verdaderos positivos. número de predicciones correctas de clase positiva
- Falsos Negativos. número de predicciones incorrectas de clase negativa
- Falsos Positivos. número de predicciones incorrectas de clase positiva
- Verdaderos negativos. número de predicciones correctas de clase negativa

En la figura 2.4 se puede visualizar este concepto en forma de conjuntos.



Figura 2.4 Representación de la matriz de confusión Shu et al (2018).

2.21.1 Accuracy

En español Exactitud, es una forma de medir la frecuencia con la que el algoritmo clasifica un conjunto de datos correctamente. La precisión es el número de predicciones correctamente clasificadas de todas las predicciones. Más formalmente, se define como el número de verdaderos positivos y verdaderos negativos dividido por el número de verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos (shu et al., 2018). La fórmula 1 describe su cálculo y la figura 5 su representación en conjuntos.

$$Accuracy = \frac{VP+VN}{VP+VN+FP+FN} \quad (1)$$

Donde:

VP: es el número de predicciones correctas de clase positiva.

VN: es el número de predicciones incorrectas de clase negativa.

FP: es el número de predicciones incorrectas de clase positiva.

FN: es el número de predicciones correctas de clase negativa.

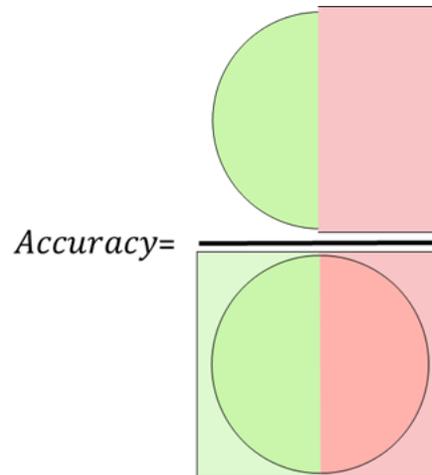


Figura 2.5 Representación de la métrica Accuracy Shu et al (2018)

2.21.2 Precision

En español precisión, es una métrica que cuantifica el número de predicciones positivas correctas realizadas. Se calcula como la proporción de verdaderos positivos, dividida por el número total de ejemplos positivos predichos (shu et al., 2018). La fórmula 2 describe su cálculo y la figura 6 su representación en conjuntos.

$$Precision = \frac{VP}{VP + FP} \quad (2)$$

Donde:

VP: es el número de predicciones correctas de clase positiva.

FP: es el número de predicciones incorrectas de clase positiva.

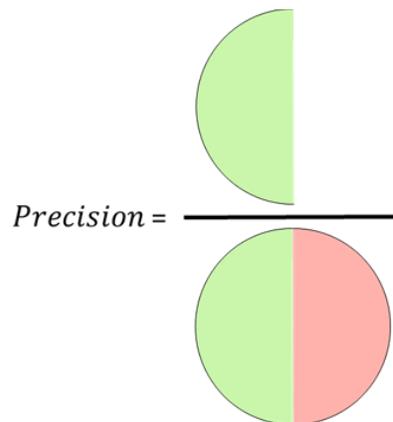


Figura 2.6 Representación de la métrica Precision Shu et al (2018)

2.21.3 Recall

En español recuerdo, es una métrica que cuantifica el número de verdaderos positivos acertados a partir de todas las predicciones positivas que podrían haberse realizado (shu et al., 2018). A diferencia de la precisión que solo comenta las predicciones positivas correctas de todas las predicciones positivas, el recuerdo proporciona un indicador de las predicciones positivas perdidas. De esta forma, Recall proporciona una noción de la cobertura total de la clase positiva. La fórmula 3 describe su cálculo y la figura 7 su representación en conjuntos.

$$Recall = \frac{VP}{VP+FN} \quad (3)$$

Donde:

VP: es el número de predicciones correctas de clase positiva.

FN: es el número de predicciones correctas de clase negativa.

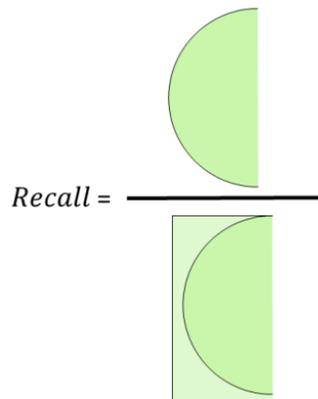


Figura 2.7 Representación de la métrica Recall Shu et al (2018)

2.21.4 F1-Score

En español Valor-F o también conocido como f1, es una métrica ampliamente utilizada debido a que es la medida predilecta para resumir el rendimiento de un modelo (shu et al., 2018). Por sí solo, ni la precisión ni el *recall* validan totalmente el modelo. Se puede tener una precisión excelente con un *recall* terrible y viceversa. Por lo que F-1/F-measure proporciona una forma de combinar el rendimiento de la *precision* y *recall* en una sola métrica. La fórmula 4 describe su cálculo.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

Una vez formalizados los conceptos de las métricas que se aplicaran para evaluar cada algoritmo y los repositorios de imágenes para el entrenamiento y prueba de los algoritmos, se describe el proceso de experimentación.

2.22 Máquina de Soporte Vectorial

Máquinas de Vector Soporte (Vector Support Machines, SVMs) es un algoritmo de clasificación y regresión desarrollado en la década de los 90, dentro del campo de la ciencia computacional. Aunque inicialmente se desarrolló como un método de clasificación binaria, su aplicación se ha extendido a problemas de clasificación múltiple y regresión. SVMs ha resultado ser uno de los mejores clasificadores para un amplio abanico de situaciones, por lo que se considera uno de los referentes dentro del ámbito de aprendizaje estadístico y machine learning. Las Máquinas de Vector Soporte se fundamentan en el Maximal Margin Classifier, que, a su vez, se basa en el concepto de hiperplano, Joaquín A.R. (2021).

Hiperplano y Maximal Margin Classifier

En un espacio p-dimensional, un hiperplano se define como un subespacio plano y afín de dimensiones p-1. El término afín significa que el subespacio no tiene por qué pasar por el origen. En un espacio de dos dimensiones, el hiperplano es un subespacio de 1 dimensión, es decir, una recta. En un espacio tridimensional, un hiperplano es un subespacio de dos dimensiones, un plano convencional. Para dimensiones $p > 3$ no es intuitivo visualizar un hiperplano, pero el concepto de subespacio con p-1 dimensiones se mantiene Joaquín A.R. (2021).

La definición matemática de un hiperplano es bastante simple. En el caso de dos dimensiones, el hiperplano se describe acorde a la ecuación de una recta ver ecuación 2.1:

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 = 0 \quad (2.1)$$

Intuitivamente la SVM es un modelo que representa a los puntos de la muestra en el espacio, separando las clases en dos espacios lo más amplios posibles mediante un hiperplano de separación, el cual es definido como el vector entre los dos puntos de las dos clases más cercanas y a este vector se le llama vector soporte. Cuando una nueva muestra se pone en correspondencia con dicho modelo, en función del espacio al que pertenece, entonces puede ser clasificada a una u otra clase. Una buena separación entre las clases permitirá una clasificación correcta. Como en la mayoría de los métodos de clasificación supervisada, los datos de entrada son vistos como un vector p-dimensional (una lista ordenada de p números). La SVM busca un hiperplano que

separe de forma óptima a los puntos de una clase con respecto a otra, que previamente han podido ser proyectados en un espacio de dimensionalidad superior. En este concepto de separación óptima es donde reside la característica fundamental de las SVM, buscando que el hiperplano tenga la máxima distancia con los puntos más cercanos a él. De esta manera los puntos del vector que están de un lado del hiperplano se etiquetan con una categoría y los que se encuentran del otro lado, se etiquetan con otra categoría. A la variable predictora se le llama atributo y a los atributos utilizados para definir el hiperplano se les llaman características. La elección de la representación más adecuada del universo estudiado se lleva a cabo mediante un proceso llamado selección de características. Al vector formado por los puntos más cercanos al hiperplano se le llama vector soporte Camacho, J. A. (2020).

2.23 *RandomForest*

Un modelo Random Forest está formado por un conjunto de árboles de decisión individuales, cada uno entrenado con una muestra ligeramente distinta de los datos de entrenamiento generada mediante bootstrapping). La predicción de una nueva observación se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo. Muchos métodos predictivos generan modelos globales en los que una única ecuación se aplica a todo el espacio muestral. Cuando el caso de uso implica múltiples predictores, que interactúan entre ellos de forma compleja y no lineal, es muy difícil encontrar un único modelo global que sea capaz de reflejar la relación entre las variables. Los métodos estadísticos y de machine learning basados en árboles engloban a un conjunto de técnicas supervisadas no paramétricas que consiguen segmentar el espacio de los predictores en regiones simples, dentro de las cuales es más sencillo manejar las interacciones. Es esta característica la que les proporciona gran parte de su potencial. Los métodos basados en árboles se han convertido en uno de los referentes dentro del ámbito predictivo debido a los buenos resultados que generan en problemas muy diversos. Joaquín A.R. (2020).

Un modelo Random Forest está formado por un conjunto (ensemble) de árboles de decisión individuales, cada uno entrenado con una muestra aleatoria extraída de los datos de entrenamiento originales mediante bootstrapping). Esto implica que cada árbol se entrena con unos datos ligeramente distintos. En cada árbol individual, las observaciones se van distribuyendo por bifurcaciones (nodos) generando la estructura del árbol hasta alcanzar un nodo terminal. La predicción de una nueva observación se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo Joaquín A.R. (2020).

2.24 *K* vecinos más cercanos

El algoritmo KNN (K Nearest Neighbors) es un algoritmo de Machine Learning muy conocido, debido su simplicidad, ya que es muy fácil de entender y utilizar. KNN pertenece al conjunto de algoritmos de Aprendizaje Supervisado *Díaz, R. (2022)*.

El algoritmo KNN (K Nearest Neighbors) o en español, K Vecinos Más Cercanos, trata de buscar los *K* puntos más cercanos a un punto concreto para poder inferir su valor. Este algoritmo pertenece al conjunto de técnicas del aprendizaje automático supervisado, y puede ser utilizado tanto para problemas de clasificación, como de regresión. El algoritmo KNN se encuentra bajo la categoría de métodos denominados Lazy Learning (Aprendizaje Vago). Se llaman así, ya que este tipo de algoritmos no entrena un modelo, es decir, no se optimizan unos pesos, sino que simplemente compara como de parecidos son los puntos que conocemos, con los puntos nuevos, para obtener así una predicción *Díaz, R. (2022)*.

Capítulo 3

Estado del arte

En este apartado se presenta los trabajos que investigan las principales técnicas de reconocimiento de emociones, así como los algoritmos de clasificación que emplean para la identificación tanto de características faciales como las emociones.

Para la búsqueda y análisis del estado del arte se aplicó el Método de Revisión Sistemática de la Literatura (RSL) el cual facilita la identificación de los trabajos relacionados con el tema de investigación por lo que se plantean preguntas las cuales se pretenden responder con cada uno de los artículos elegidos Manterola, C., et al. (2013).

Para aplicar el método de RSL se parte de la descripción del problema el cual se describe en la sección 1.1 de este documento, posteriormente se continua con la formulación de las preguntas de investigación, después con la selección de fuentes para esto se emplearon las cadenas de búsquedas descritas en la tabla 3.1 y los siguientes buscadores de artículos mencionadas en la tabla 3.2.

Formulación de preguntas de investigación

- ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?
- ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales?
- ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial?
- ¿Qué técnicas de extracción de características existen para la clasificación de emociones?

Tabla 3.1 Cadenas de búsqueda o palabras claves

| Cadena | Cadenas Generales de búsqueda |
|----------|--|
| A | Facial recognition algorithms. |
| B | Facial recognition techniques. |
| C | Recognition of emotions. |
| D | Facial expressions |
| E | Facial Action Code System |
| F | Facial Expression Recognition |
| G | Facial Recognition Of Cognitive States |

Tabla 3.2 Selección de trabajos relacionados

| Repositorios | Cadena Utilizada | Estudios Encontrados | Estudios Candidatos | Estudios Seleccionados |
|---|---------------------|----------------------|---------------------|------------------------|
| ACADEMIA | A, B, C, D, E, F, G | 18 | 3 | 2 |
| Google Scholar | A, B, C, D, E, F, G | 26 | 4 | 2 |
| IEEE explorer | A, B, C, D, E, F, G | 80 | 18 | 13 |
| Repositorio de tesis TecNM/CENIDET | A, B, C, D, E, F, G | 4 | 4 | 3 |
| Total | | 128 | 29 | 20 |

Los criterios de inclusión y exclusión utilizados fueron los siguientes:

Inclusión:

- Sólo se seleccionaron los artículos que fueron publicados en el periodo 2017 – 2020.
- Sólo se seleccionaron aquellos artículos que responden a las preguntas de investigación, para lo cual se analiza el título, el resumen, las palabras clave, la introducción y las conclusiones.
- Solo se seleccionan los artículos de revistas indizadas en Scopus y JCR.
- Solo se seleccionan artículos de congresos internacionales indizados en Scopus.

Exclusión

- Se excluyen aquellos artículos con años anteriores a 2017.
- Se excluyen aquellos trabajos que no responden a las preguntas de investigación y esto se obtuvo al analizar su título, resumen, introducción y conclusiones.
- Se excluyen aquellos artículos que no aparecen en ningún índice.

Para dar respuesta a las preguntas de investigación se eligieron los siguientes trabajos de investigación ya que se consideró que son los que trabajan en la identificación de las emociones básicas de neutral, felicidad, enojo, tristeza, sorpresa, disgusto, y miedo.

El estado del arte lo dividiremos en dos secciones:

- 1) Estado del arte de Detección de rostro y reconocimiento de expresiones faciales.
- 2) Estado del Arte de Reconocimiento de emociones

3.1 Estado del Arte de Detección de Rostro y Reconocimiento de Expresiones Faciales.

3.1.1.Reconocimiento Automático del Rostro para Verificación de Identidad para Evaluación en Línea (Valderrama Cárdenas, 2019).

En esta tesis se desarrolló un sistema de visión artificial que monitorea, en tiempo real a los aspirantes que presentan el examen de admisión, mientras presentan dicho examen en línea; esto a través de capturas del rostro de los aspirantes para verificar su identidad; se pretende comprobar que la persona que se registró en el sistema con su credencial del Instituto Nacional

Electoral (INE) para presentar el examen, sea la misma que se encuentra frente a la computadora durante la ejecución del mismo.

Si bien esta tesis no responde ninguna de las preguntas de investigación que se plantearon, pero se tomó en cuenta ya que nos proporciona un algoritmo que es capaz de identificar el rostro de una persona en tiempo real por lo que nos será de gran utilidad en el proceso de identificación de rostro que implementaremos en nuestra metodología de solución así también tomaremos como referencia el esquema de su metodología de solución el cual se muestra en la figura 3.1.



Figura 3.1 Esquema de metodología de solución (Valderrama Cárdenas, 2019)

3.1.2. *Facial Recognition Techniques Using SVM: A Comparative Analysis* (Moreano, et al, 2019)

Este artículo presenta una revisión literaria del reconocimiento facial en 2D. El enfoque está en los resultados obtenidos por algunos investigadores con la aplicación de técnicas de extracción, clasificadores de patrones, bases de datos y sus respectivos porcentajes de eficiencia obtenidos. El objetivo es determinar técnicas eficientes que permitan un proceso óptimo de reconocimiento facial 2D, basado en la calidad de bases de datos, extractores de características y clasificadores de patrones.

Este trabajo no responde alguna de la preguntas de investigación pero proporciona una comparativa de las principales técnicas de reconocimiento facial dando información sobre la precisión del reconocimiento de cada técnica y de cómo en combinación con otras técnicas estos pueden mejorar el nivel de reconocimiento facial, el análisis que el autor realiza en este trabajo llega a la conclusión de que son dos las técnicas de clasificación de características que obtienen mejores resultados los cuales son el vector de características discriminante independiente (IDCV por sus siglas en inglés) en combinación con la máquina de soporte vectorial (SVM por sus siglas en inglés) obteniendo una precisión de reconocimiento del 98.92%, siguiendo en segundo lugar la red de propagación hacia atrás (BNP por sus siglas en inglés) en combinación con la

función de base radial (RBF por sus siglas en inglés) obteniendo una precisión de 98.88% y en tercer lugar la combinación de la máquina de soporte vectorial (SVM) con el análisis discriminatorio no paramétrico (NDA por sus siglas en inglés) con un 98% de precisión de reconocimiento por lo que se tomaran las técnicas para evaluarlas y poder elegir una de ellas para incorporarla en la metodología de solución, además de que proporciona una idea de qué técnica utilizar para el reconocimiento facial y cómo sería las etapas de un sistema de reconocimiento facial ya que también proporciona un esquema que se muestra en la figura 3.2, el cual se adaptara para la metodología de solución en el apartado de reconocimiento del rostro.

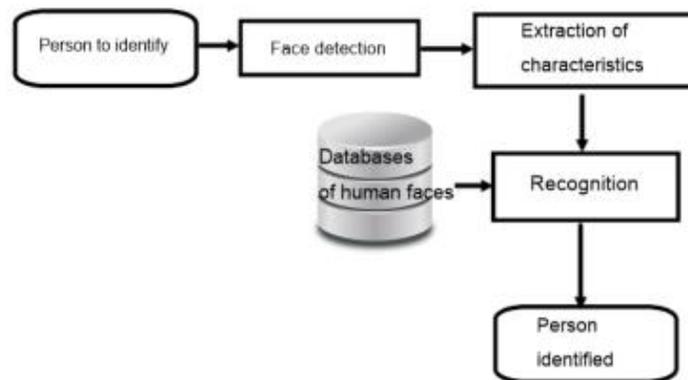


Figura 3.2 Etapa primaria de un sistema de reconocimiento facial (Cadena Moreano J. A, et al, 2019)

3.1.3. Real-Time Smart Attendance System Using Face Recognition Techniques (Sawhney Shreyak et al, 2019)

En este trabajo se propone un modelo para implementar un sistema automatizado de gestión de asistencia para estudiantes de una clase mediante el uso de la técnica de reconocimiento facial, mediante el uso de valores de la cara, análisis de componentes principales (PCA) y red neuronal convolucional (CNN).

Si bien este trabajo no responde del todo a las preguntas de investigación nos propone técnicas de reconocimiento facial las cuales se pueden poner a prueba y evaluarlas para determinar si se utilizaran en la parte de detección de rostro de la metodología de solución.

3.2 *Estado del Arte de Reconocimiento de Emociones.*

3.2.1. *Recognition Of Facial Emotions Based On Sparse Coding (Sunitha, et al, 2017)*

El objetivo principal del presente trabajo fue desarrollar un sistema de reconocimiento de emociones basado en la cara capaz de manejar variaciones en la pose facial, la iluminación y la obstrucción parcial del rostro. En otras palabras, el sistema tiene como objetivo representar y reconocer con firmeza las expresiones faciales en situaciones de la vida real.

En este trabajo se presenta una metodología llamada Extreme Sparse Learning (ESL), que puede incluir mutuamente un conjunto de premisas y un modelo de agrupación no lineal. El enfoque propuesto consolida la fuerza discriminativa de Extreme Learning Machine (ELM) con la propiedad de reproducción de una representación pequeña para potenciar la disposición exacta cuando se dan signos ruidosos e información manchada registrada en entornos comunes.

Este trabajo menciona que el método propuesto tiene un alto nivel de reconocimiento de emociones, pero no hace referencia el nivel de precisión logrado y qué emociones está reconociendo, por otra parte menciona que el costo computacional es considerablemente alto, además comenta que existen técnicas de identificación de expresiones faciales que son capaces de identificarlas en secuencias de video las cuales son, Hidden Markov Models (HMM), Dynamic Bayesian Networks (DBN), por ello de este artículo se analizarán estas técnicas para determinar la que más precisión y menos costo computacional proporcione.

3.2.2. *An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment (Yang, ET AL, 2017)*

El propósito de este estudio fue introducir un método basado en el reconocimiento facial para identificar la comprensión de los estudiantes de todo el proceso de aprendizaje a distancia. Este estudio propone un modelo de reconocimiento de emociones de aprendizaje, que consta de tres etapas: extracción de características, característica de subconjunto y clasificador de emociones. Se utiliza el método de Haar Cascades para detectar la imagen de entrada, una cara, como base para la extracción de ojos y boca, y luego a través de la detección de bordes Sobel para obtener el valor característico. A través del entrenamiento del clasificador de la Red Neural, se obtienen seis tipos de diferentes categorías emocionales (tristeza, felicidad, sorpresa, miedo, ira y asco).

Este trabajo responde claramente las siguientes preguntas de investigación que se plantearon en la revisión sistemática de literatura las cuales son ¿Existen algoritmos capaces de

identificar emociones mediante el reconocimiento facial?, ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales? Y ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? Ya que este trabajo proporciona una técnica de reconocimiento de emociones llamada HAAR Cascade, la cual en conjunto con una red neuronal identificas 6 emociones básicas en secuencias de video con solo utilizar la región de los ojos y la boca, obteniendo los resultados que se muestran en la **figura 3.3** para cada emoción, esto muestra que se considera evaluar este algoritmo y verificar si da los mismos resultados con la base de datos que se utilizara para esta investigación y poder adaptarla en el proceso de la metodología de solución.

| | Sad | Surprise | Happy | Anger | Disgust | Fear |
|--------------------------|------------|-----------------|--------------|--------------|----------------|-------------|
| Proposed solution | 78.54 | 93.26 | 95.25 | 91.22 | 84.32 | 82.58 |
| Current solution | 76 | 87.72 | 94 | 87.66 | 82.76 | 79.73 |

Figura 3.3 Resultados obtenidos en porcentajes de precisión de reconocimiento de emociones (Yang, et al, 2017)

3.2.3. Deep Convolutional Neural Network for Facial Expression Recognition Using Facial Parts (Nwosu, et al, 2017)

Este artículo propone el diseño de un sistema de reconocimiento de expresión facial (FER) basado en una red neuronal convolucional profunda mediante el uso de partes faciales. El método propuesto utiliza una red neuronal convolucional de dos canales en la cual las partes faciales (FP) se usan como entrada a la primera capa convolucional, los ojos extraídos se usan como entrada al primer canal mientras que la boca es la entrada al segundo canal. La información de ambos canales converge en una capa totalmente conectada que se usa para obtener información global de estas características locales y luego se usa para la clasificación.

En respuesta de la pregunta de investigación ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?, se seleccionó este artículo porque la solución que propone obtiene un alto nivel de precisión con su técnica de clasificación la región de los ojos y de la boca eso da la idea de que se puede utilizar solo estas regiones para poder identificar las emociones y reducir el tiempo de procesamiento de la imagen, además menciona la utilización del algoritmo Viola-Jones, que es un algoritmo robusto capaz de procesar imágenes extremadamente rápido para situaciones en tiempo real el cual es capaz de detectar la región de la cara independientemente de la variación en el tamaño, el fondo, el brillo y la transformación espacial de la imagen de entrada sin formato por lo que se pretende hacer uso de este algoritmo para la etapa de preprocesamiento del rostro.

3.2.4. Facial Emotion Recognition Using Min-Max Similarity Classifier (Krestinskaya, Pappachen, 2017)

En este artículo, se presenta un algoritmo de reconocimiento de emociones faciales eficiente y directo para reducir el problema de la falta de coincidencia de píxeles entre clases durante la clasificación. El método propuesto incluye la aplicación de la normalización de píxeles para eliminar las compensaciones de intensidad seguidas de una métrica Min-Max en un clasificador vecino más cercano que es capaz de suprimir los valores atípicos de características. Los resultados indican una mejora en el rendimiento de reconocimiento de 92.85% a 98.57% para el método de clasificación Min-Max propuesto cuando se prueba en la base de datos JAFFE.

Se eligió este artículo porque da respuesta a las preguntas de investigación siguientes, ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?, ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales? Y ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? La precisión que obtuvieron es considerablemente alta por lo que se pondrá a prueba el algoritmo propuesto para poder asegurar que puede ayudar a dar solución al tema de tesis.

3.2.5. Towards Developing A Tool for UX Evaluation Using Facial Expression (Md. Munim et al, 2017)

En este artículo, se propuso un sistema para evaluar la experiencia del usuario a partir de expresiones faciales usando Affdex SDK. También se realiza un estudio de evaluación para evaluar la efectividad y la eficiencia del sistema propuesto.

El trabajo realizado en este trabajo da una idea de lo que se pretende realizar en esta tesis, aunque al autor le falta mucho que mejorar en su propuesta aporta indicadores a tomar en cuenta, así como también responde la siguiente pregunta de investigación ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?, ya que menciona el algoritmo utilizado para la detección de emociones que el empleo para su investigación.

3.2.6. Optimized Facial Emotion Recognition Technique for Assessing User Experience (Liu Xiao, Lee Kiju, 2018)

Este trabajo presenta una nueva técnica de optimización en el procesamiento de imágenes para el reconocimiento de emociones basado en la expresión facial. El método combina dos filtros de preprocesamiento (prefiltros), es decir, filtro de brillo, contraste y filtro de extracción

de bordes, con aprendizaje basado en la red neuronal convolucional (CNN) y máquina de vector soporte (SVM) para la clasificación de emociones.

Este artículo responde las siguientes preguntas de investigación: ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?, ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales?, ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? En el cual habla de una combinación de la técnica de extracción de características llamada Máquina de Soporte Vectorial (SVM por sus siglas en inglés) y una Red Neuronal Convolucional (CNN por sus siglas en inglés) con 4 capas ocultas que le permiten al autor ajustar los parámetros para poder obtener una mejor precisión en la identificación de las emociones por lo que manejaron una gran cantidad de datos para el entrenamiento, validación y prueba del algoritmo esto da a entender que este tipo de algoritmos demanda un alto costo computacional y da una idea de cómo poder combinar técnicas y optimizarlas para poder reducir el costo computacional.

3.2.7. Facial Emotion Recognition: A Comparison of Different Landmark-Based Classifiers (M. Alvarez et al, 2018)

En este trabajo, después de aplicar algoritmos comunes de detección de rostros y alineación al conjunto de datos de Cohn-Kanade, se alimentó un grupo de puntos de referencia etiquetados por la emoción a diferentes clasificadores para comparar sus resultados. El clasificador perceptrón multicapa mostró el mejor rendimiento con una precisión promedio del 89%

Se eligió este trabajo como parte del estado del arte ya que responde las siguientes preguntas de investigación: ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?, ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales? y ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? Ya que este trabajo presenta una comparación de diferentes clasificadores y extractores de características para identificar emociones a través de análisis facial estos algoritmos de clasificación son Navie Bayes, Support Vector Machine, Decision Tree, Random Forest, AdaBoost y, finalmente, un multicapa Perceptron por lo que se tomaran estas técnicas para evaluarlas y probarlas con la base de datos que utilizan y otras que se han obtenido en los trabajos elegidos anteriormente y determinar cuál de ella da un mejor resultado y poder implementar la parte de clasificación en nuestra metodología de solución.

3.2.8. Geometric Facial Components Feature Extraction for Facial Expression Recognition
(Dewi, et al, 2018)

En este trabajo proponen un nuevo método de extracción de características geométricas que aplica técnicas de cálculo simples para componentes faciales para garantizar la robustez de cada variación de pose. A diferencia de otras características que requieren más esfuerzos en un proceso de transformación, el método propuesto funciona de manera eficiente directamente sobre una base de píxeles, **la figura 3.4** y **la tabla 3.3** muestran los resultados obtenidos.

En respuesta a la pregunta de investigación ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? Que se planteó, este trabajo propone un nuevo método de extracción de características faciales para identificar emociones por lo que se tomará el método propuesto como referencia para la metodología de solución de esta investigación ya que como menciona el autor que la característica geométrica es robusta a la variación de la posición de la cara, la escala, el tamaño y la orientación de la cabeza con el cual se considera que dará mejores resultados en nuestra metodología propuesta.



Figura 3.4 Resultados de reconocimiento facial de emociones mediante la metodología de extracción de características geométricas (Dewi, et al, 2018)

Tabla 3.3 Comparación de precisión de identificación de emociones con otros clasificadores y el método propuesto por (Dewi, et al, 2018)

| No | Classifier | Accuracy Rate |
|----|------------|---------------|
| 1 | FCM | 80.71% |
| 2 | FIS | 72% |
| 3 | SVM | 93.10% |
| 4 | our method | 93.67% |

3.2.9. Facial Expression Recognition by Calculating Euclidian Distance for Eigen Faces Using PCA (Mangala, Prajwala, 2018)

El sistema propuesto se basa en el cálculo de la distancia euclidiana para las caras de Eigen. En este enfoque se consideraron las siete emociones básicas que se clasifican como feliz, enojo, tristeza, miedo y asco, sorpresa, neutral. Se consideraron 50 imágenes de expresión facial son como entrenadas por Eigen Faces. Finalmente, las Caras Eigen presentadas se comparan con la imagen de prueba. En este enfoque, el reconocimiento facial se realiza mediante el "Análisis de componentes principales (PCA)". El resultado experimental proporciona tasas de reconocimiento del 98.5% para cada emoción.

Este trabajo responde tres de las preguntas de investigación planteadas las cuales son ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial?, ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales? Y ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? Ya que en este artículo propone la técnica de extracción de características llamada Análisis de Componentes Principales (PCA por sus siglas en inglés) que en combinación con una red neuronal hacen posible el reconocimiento de 7 emociones por lo que consideramos que la técnica PCA está entre los mejores extractores de características faciales haciendo que se tome en cuenta para el preprocesamiento de la imagen de la metodología de solución.

3.2.10. Real-Time Algorithms for Facial Emotion Recognition: A Comparison of Different Approaches (Kartali, et al, 2018)

En este artículo, se presenta una comparación de cinco enfoques diferentes para el reconocimiento de emociones en tiempo real de cuatro emociones básicas (felicidad, tristeza, ira y miedo) a partir de imágenes faciales. Se comparó tres enfoques de aprendizaje profundo basados en, redes neuronales convolucionales (CNN) y dos enfoques convencionales para la clasificación de las características del histograma de gradientes orientados (HOG): 1) AlexNet CNN, 2) solución comercial Affdex CNN, 3) FER- CNN, 4) Máquina de vectores de soporte (SVM) de características HOG, 5) Red neuronal artificial Perceptron multicapa (MLP) de características HOG en la tabla 3.4 se presenta el resultado que cada técnica obtuvo.

Este trabajo da respuesta a las siguientes preguntas de investigación, ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento facial? ¿Cuál es el nivel de

precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales?, ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial?, a pesar que responde estas preguntas todos los algoritmos y técnicas que utilizan se basan en, redes neuronales las cuales demandan un costo computacional alto además de que las pruebas realizadas son en un ambiente muy controlado por lo que los resultados fueron altos y lo que se pretende en esta investigación es analizar algoritmos que no demanden un costo computacional alto.

Tabla 3.4 total, de precisiones de todos los algoritmos probados (Kartali et al, 2018)

| Algoritmo de Reconocimiento de Emociones Faciales | Precisión Total [%] |
|--|--------------------------------|
| Affdex CNN | 85,05 |
| AlexNet CNN afinado | 76,64 |
| Clasificación SVM de características HOG | 63,55 |
| Clasificación MLP de características HOG | 56,07 |
| FER-CNN | 55,14 |

3.2.11. A Real-Time Emotion Recognition System for Disabled Persons (Rabhi Yassine et al, 2018)

Este documento propone un sistema de reconocimiento de emociones integrado en tiempo real diseñado para que una silla de ruedas eléctrica detecte, explote y evalúe el estado emocional de un usuario mayor o un usuario con algún deterioro cognitivo. Se emplea un conjunto de varias técnicas para la detección de rostros, la extracción de rasgos faciales y la clasificación de emociones como HOG.

En respuesta a las siguientes preguntas de investigación, ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? El autor realizó la detección de rostros, utilizando una combinación de, la pirámide de imagen mejor conocida como pirámide gaussiana el cual es simplemente para extraer características mientras se reduce el ruido, histograma de gradiente orientado (HOG por sus siglas en ingles) y un clasificador lineal junto con un esquema de detección de ventana deslizante además combino diferentes clasificadores lineales de acuerdo a lo que quiso extraer y así tener un tiempo de procesamiento y costo computacional moderado, en su investigación solo se enfocó en 3 de las 7 emociones básicas las cuales son ira, miedo y sorpresa se tomara como referencia la idea de su metodología para adaptarla en la metodología de solución.

3.2.12. Quest: Quadrilateral Senary Bit Pattern for Facial Expression Recognition (Verma monu et al, 2018)

Este documento proporciona un nuevo descriptor de características denominado Patrón de bits senarios cuadriláteros para el reconocimiento de expresiones faciales. El patrón QUEST codificó los cambios de intensidad al enfatizar la relación entre los píxeles vecinos y de referencia al dividirlos en dos cuadriláteros en el vecindario local. Por lo tanto, los bordes de gradiente resultantes revelan la información de variación de transición, que mejora la tasa de clasificación al discriminar las clases de expresión. Además, también mejora la capacidad del descriptor para hacer frente a las variaciones del punto de vista y los cambios de iluminación. El patrón QUEST genera un código compacto de seis bits, que mejora la eficiencia del sistema de reconocimiento de expresión facial con más discriminación. La efectividad del método propuesto se evalúa mediante la realización de varios experimentos en cuatro conjuntos de datos de referencia: MMI, GEMEP-FERA, OULU-CASIA e ISED.

Este artículo da respuesta a la siguiente pregunta de investigación ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? Por lo que se analizará y comparará con otras técnicas para determinar si realmente podrá servir en el proceso de extracción de características faciales para poder identificar las emociones, además de que se utilizará las bases de datos de emociones que el autor utiliza para realizar las pruebas y entrenamientos de los algoritmos que se elijan.

3.2.13. Facial Expression Recognition of Intercepted Video Sequences Based on Feature Point Movement Trend and Feature Block Texture Variation (Yi, et al 2019)

Este documento propone un Marco de referencia del Reconocimiento de Expresiones Faciales para las secuencias de video interceptadas mediante el uso de la tendencia de movimiento de puntos de características y variación de textura de bloques de características. En primer lugar, los puntos característicos están marcados por el Modelo de Apariencia Activa (AAM por sus siglas en inglés) y se seleccionan los 24 más representativos. En segundo lugar, la secuencia de la expresión facial se intercepta en el video de la cara determinando dos cuadros clave cuyas intensidades emocionales son mínimas y máximas, respectivamente. En tercer lugar, se ajusta la curva de tendencia que representa las variaciones de distancia euclidiana entre dos puntos de entidad seleccionados, y se calculan las pendientes de puntos específicos en la curva de tendencia. Finalmente, combinando el conjunto de pendientes que está compuesto por las pendientes calculadas con la diferencia de textura de bloques de características propuesta (FBTD) que se refiere a la variación de textura del parche facial, la característica de expresión final se forma y se ingresa a la red neuronal de convolución unidimensional (1DCNN por sus siglas en inglés) para el Reconocimiento de Expresiones Faciales (FER por sus siglas en inglés) **la figura 3.5** muestra los resultados obtenidos para cada emoción detectada.

Con la información obtenida de este artículo se da respuesta a las siguientes preguntas de investigación: ¿Existen algoritmos capaces de identificar emociones mediante el reconocimiento

facial?, ¿Cuál es el nivel de precisión obtenida actualmente en los algoritmos en la identificación de estados emocionales? Y ¿Qué técnicas existen actualmente para la identificación de emociones mediante el reconocimiento facial? Los autores proponen una metodología de reconocimiento de emociones mediante el reconocimiento facial por medio de secuencias de video en el cual utilizan la técnica de diferencia de textura de bloque de características (FBTD por sus siglas en ingles) y una red neuronal convolucional para realizar el reconocimiento, por lo que se utilizará la estructura de la metodología para poder construir una similar ya que los datos de entrada que se utilizara en esta investigación serán secuencias de video capturadas desde una video cámara. **La figura 3.6** representa la metodología utilizada por los autores la cual da una idea de cómo abordar la parte del desarrollo del módulo de reconocimiento facial de emociones que se pretende implementar.

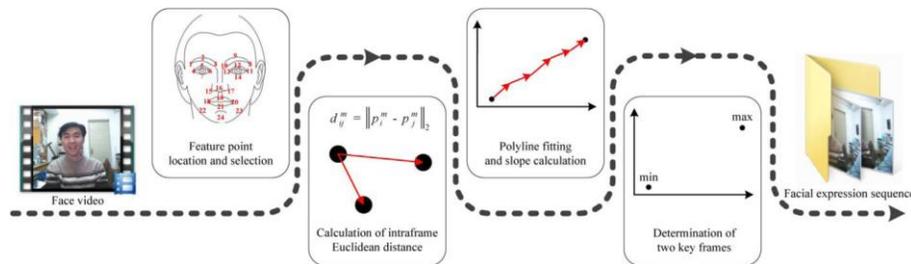


Figura 3.5 Metodología de solución propuesta por (Yi, 2019). El FESI incluye principalmente tres contenidos: el cálculo de la distancia euclidiana dentro del cuadro, el ajuste de la polilínea y el cálculo de la pendiente, la determinación de dos cuadros clave.

Precision, recall, and F1-measure for the databases.

| | BHU Facial Expression Database | | | MMI Facial Expression Database | | | Combination of two databases | | |
|-----------|--------------------------------|-------------|-------------|--------------------------------|-------------|-------------|------------------------------|-------------|-------------|
| | Precision | Recall | F1-measure | Precision | Recall | F1-measure | Precision | Recall | F1-measure |
| Happiness | 97.9 | 99.2 | 98.5 | 99.1 | 99.6 | 99.3 | 99.8 | 100 | 99.9 |
| Sadness | 95.1 | 92.2 | 93.6 | 95.6 | 92.8 | 94.2 | 96.9 | 93.6 | 95.2 |
| Fear | 92.8 | 93.4 | 93.1 | 96.4 | 94.4 | 95.4 | 96.2 | 95.4 | 95.8 |
| Disgust | 91.3 | 92.9 | 92.1 | 92.0 | 96.4 | 94.1 | 92.3 | 96.4 | 94.3 |
| Surprise | 98.8 | 98.2 | 98.5 | 99.6 | 99.6 | 99.6 | 100 | 100 | 100 |
| Anger | 95.2 | 95.2 | 95.2 | 96.8 | 96.5 | 96.6 | 97.0 | 96.6 | 96.8 |

Figura 3.6 Precisiones obtenidas con diferentes bases de datos (Yi, 2019)

3.2.14. Valoración Automática de la Motivación del Estudiante Mediante Expresiones Faciales (Vázquez Rodríguez, 2017)

Este trabajo propone utilizar Kinect y la librería *facetracking* para crear un sistema que localice automáticamente el rostro y realizar una interpretación de sus elementos para detectar la motivación (interés) en una actividad. La evaluación de las pruebas de este sistema obtuvo que, para casos donde el sistema determinó que los sujetos estaban prestando atención y realmente estaban prestando atención fue de 89.79% y para los casos donde el sistema evaluó que los sujetos de prueba no prestaron atención y en realidad no lo hicieron fue del 90.72%.

En este trabajo además de realizar la detección de emociones mediante videos también realiza la clasificación de estados de interés y desinterés, ya que las técnicas que utilizó el autor para poder identificar estos estados en los alumnos dieron resultados alentadores por lo que se evaluarán estas técnicas y determinar si se utilizarán para complementar la solución de la investigación si bien los estados de interés y desinterés no son temas de interés en esta investigación podría ser utilizado para trabajos futuros de esta tesis.

3.2.15. Facial Emotion Recognition Based On Viola-Jones Algorithm In The Learning Environment (Candra Kirana Karetika et al, 2018)

Este artículo Presenta un sistema de reconocimiento de emociones faciales basado en el algoritmo Viola-Jones en el entorno de aprendizaje. Básicamente, el algoritmo Viola-Jones es un algoritmo de detección de rostros. Sin embargo, utilizaron características basadas en la cara para detectar la cara y reconocer la emoción, por lo tanto, aplicamos la función rectangular y el algoritmo AdaBoost en cascada, que son el concepto principal del algoritmo Viola-Jones en ambos procesos.

Este artículo es interesante ya que menciona que el método propuesto identifica emociones y al igual que el artículo mencionado en la sección 3.2.15 realiza la clasificación de los estados de interés y aburrimiento, Entonces se infiere que algunas técnicas como lo es el algoritmo viola-jones que es para el reconocimiento de emociones podrían ser utilizadas para el reconocimiento no solo de emociones, la precisión de este algoritmo no es muy alta y fue probado bajo un ambiente controlado dando resultados en accuracy, precision, recall, y el tiempo que consume son de 74%, 73%, 76% y 15 segundo por cuadro respectivamente por lo que se evaluara para determinar si se utilizara este algoritmo.

3.2.16. Towards Cognitive State Detection Using Facial Expression And Hand-Over-Face Gesture (H.S Gunavathi, M. Siddappa, 2018)

Este trabajo presenta un enfoque automático para la detección de estados cognitivos. Dividiendo el problema en tres tareas: 1) Codificación y clasificación de descriptores de traspaso 2) Detección de cuatro expresiones faciales (feliz, triste, neutral y sorpresa) utilizando puntos de referencia faciales y 3) Combinación de expresiones faciales y gestos con la mano sobre la cara para el reconocimiento de los estados cognitivos (interesado, concentrado, inseguro, aburrido y feliz). El enfoque propuesto reconoce los estados cognitivos con una precisión promedio de 90.51%

y los resultados justifican que la fusión de características multimodales funciona mejor en comparación con la clasificación unimodal.

La elección de este artículo fue porque es de gran importancia para el tema de investigación ya que en él se identifican estados emocionales que es en lo que se centrará esta tesis además de que el autor realiza la clasificación de estados cognitivos el cual no es tema de interés en esta investigación, pero aportaría mucha información para trabajos futuros, este trabajo responde todas las preguntas de investigación por lo que es el trabajo que más aportará a la solución de la investigación.

3.2.17. Desarrollo de un Sistema Interactivo en Visión Artificial para la Rehabilitación del Tobillo (Gama Velasco, 2018)

En esta tesis se Desarrolló un Sistema inteligente de rehabilitación de tobillo conformado por un rehabilitador de tobillo de 1 grado de libertad, un juego serio y un sistema de visión artificial que realiza el reconocimiento de expresiones faciales para conocer la motivación o molestia del paciente con el equipo de rehabilitación.

El Sistema Desarrollado propone la detección de tres expresiones faciales diferentes a las básicas, que ayuda a monitorear al paciente, de tal forma que cuente con información sobre la rehabilitación.

A pesar de que este trabajo menciona la detección de expresiones faciales no básicas realmente lo que identifica son estados cognitivos (Dolor, motivación o concentración y desmotivación o distracción) durante el proceso de rehabilitación del paciente, por lo que se eligió como parte del estado del arte ya que propone una herramienta (librería facetracking) que es la que identifica el rostro y la evaluación de tres algoritmos de clasificación de expresiones faciales en donde el que tuvo mayor precisión de reconocimiento fue el algoritmo J48, el autor utilizó el dispositivo kinect para poder identificar y extraer características del rostro. Lo que se pretende es evaluar el algoritmo sin usar el dispositivo Kinect para determinar si nos será de utilidad en la solución de la investigación.

En la tabla 3.5, se extrae la información relevante de cada uno de los trabajos que se estudiaron en el estado del arte con la finalidad de tener un mejor entendimiento de los algoritmos, bases de datos, técnicas, precisión, software y hardware que se necesitaron para realizar la clasificación de emociones y el reconocimiento del rostro con esto podemos ver que la mayoría de los algoritmos requiere de un alto rendimiento ya que la extracción de características es a base de apariencia por lo que el coste computacional es alto, además de que en algunos trabajos los autores generaron sus propios repositorios creando un ambiente controlado en la ejecución de sus pruebas, la diferencia de esta investigación es que la extracción de

características tiene una base geométrica y las pruebas se realizaron en un ambiente no controlado.

3.3 Resumen Comparativo de las Técnicas Utilizadas en el Estado del Arte para el Reconocimiento de Expresiones, Emociones Faciales y Estados Cognitivos.

Tabla 3.5 Comparativa de identificación de emociones y estados cognitivos del estado del arte

| Trabajo | Tiempo Real | Algoritmo | Precisión | Costo Computacional | Emociones | Dataset |
|---------------------------------|-------------|--|---|---------------------|--|--|
| Sunitha, et al, 2017 | si | Extreme Sparse Learning | No lo menciona | alto | No menciona cuales | No menciona cual |
| Yang, et al, 2017 | si | Haar Cascades Red Neuronal | 87.5% | -- | Felicidad, Tristeza, sorpresa, Enojo, Asco, miedo | JAFFE 213 imágenes (cada imagen: 256 * 256 píxeles) de rostros de mujeres japonesas. |
| Nwosu, et al, 2017 | si | (CNN) red neuronal convolucional como clasificador y el algoritmo Viola-Jones como identificador de rostro | 98% para sorpresa, disgusto, felicidad y tristeza. 97% en la detección de Enojado, Neutral y Miedo | Medio | Felicidad, Tristeza, Sorpresa, Enojo, Disgusto, Miedo, Neutral | Expresión Facial Femenina Japonesa (JAFFE) y Extendido Cohn-Kanada (CK +) |
| Krestinska ya, Pappachen , 2017 | No | clasificador de similitud Min-Max | 98.57% | Baja | Felicidad Tristeza sorpresa enojo disgusto miedo Neutral | JAFFE |
| Md. Munim et al, 2017 | Si | Affdex SDK | No lo menciona | -- | Compromiso Valencia Desprecio Sorpresa Ira Tristeza Asco Miedo alegría | Propia |
| Vázquez Rodríguez, 2017 | Si | Árbol binario | 89.79% atención 90.72% distracción | -- | Atención y distracción | Propia |
| Liu Xiao, Lee Kiju, 2018 | No | red neuronal convolucional (CNN) y máquina de vectores de soporte (SVM) filtro de extracción de | 98.19% | Modo rado | Enojado Asco Miedo Feliz Triste Sorpresa | FER2013 de Google |

| | | bordes (EEF) | | | neutral | |
|---|----|--|--|----------------|--|--|
| M. Alvarez et al, 2018 | No | Histograma de gradientes orientados (HOG) Navie Bayes Support Vector Machine Decision Tree Random Forest AdaBoost Perceptron Multicapa | Navie Bayes 82.05% Support Vector Machine 38.46% Decision Tree 76.92% Random Forest 82.05% AdaBoost 38.46% Perceptron Multicapa 88.03% | -- | Ira Asco Desprecio Miedo Alegría tristeza sorpresa | Cohn-Kanade |
| Dewi, et al, 2018 | No | extracción de características de componentes faciales geométricos | 93.67% | -- | Enojado Feliz Asco Miedo Sorpresa tristeza | CK + |
| Mangala, Prajwala, 2018 | No | Análisis de componentes principales (PCA) | 98.5% | Bajo | Feliz Enojo Tristeza Miedo Asco Sorpresa neutral | Eigen Faces |
| Kartali, et al, 2018 | Si | Affdex CNN AlexNet CNN afinado Clasificación SVM de características HOG Clasificación MLP de características HOG FER-CNN | Affdex CNN AlexNet 85.05 % CNN afinado 76.64% Clasificación SVM de características HOG 63.55% Clasificación MLP de características HOG 56.07% FER-CNN 55.14% | -- | felicidad, tristeza, ira y miedo | Cohn-Kanade extendido (CK +) |
| Rabhi Yassine et al, 2018 | Si | Pirámide de imagen en combinación con Histograma de gradientes orientados PCA y SVM | 73% | -- | felicidad, tristeza, ira, sorpresa, miedo y asco. | CK |
| Verma monu et al, 2018 | si | Patrón de bits senarios cuadriláteros QUEST Support Vector Machine (SVM) | 86.2% | Medio | felicidad, tristeza, ira, sorpresa, miedo y asco. | MMI, GEMEP-FERA, OULUCASIA e ISED |
| Candra Kirana Karetika et al, 2018 | Si | Viola-Jones | 74% | No lo menciona | Interés y aburrido | conjunto de datos de aprendizaje de la Universidad de Malang |
| H.S Gunavathi, M. | Si | patrón binario local extendido robusto mediano (MRELBP) | 90.51% | Modo | Emociones: neutral, triste, feliz y sorpresa | corpus Cam3D CK + |

| | | | | | | |
|------------------------------------|----|---|--|----------------|--|---|
| Siddappa, 2018 | | histograma de características del cuadro Viola y Jones Clasificador basado en reglas SVM multiclase | | | Estados cognitivos: interesado, aburrido, inseguro, feliz y reflexivo. | |
| Gama Velasco, 2018 | Si | J48 árbol de decisión Naive Bayes SVM Maquina de soporte vectorial | 95.74% | No lo menciona | Concentración, distracción y dolor | propia |
| Yi, et al 2019 | Si | diferencia de textura de bloque de características propuesta (FBTD) y red neuronal de convolución unidimensional (1DCNN) | 97% | Alto | felicidad, tristeza, miedo, asco, sorpresa e ira | base de datos de expresión facial de la Universidad de Beihang (BHU), la base de datos de expresión facial MMI y la combinación de dos bases de datos |
| Valderrama Cárdenas, 2019 | Si | Capa de foto receptores con retinex multiescala FFRM algoritmo de preprocesamiento. SVM Redes Bayecianas MLP Random Foprest | 96.08% | Bajo | NA | FEI FACES94 Propia |
| Moreano, et al, 2019 | No | EIGENFACE LDA PCA-SVM DCV-NN KDCV-NN DCV-RBF KDCV-RBF IDCV-SVM | EIGENFACE 93,87 % LDA 93,8 % PCA-SVM 95,56 % DCV-NN 96,67 % KDCV-NN 97,75 % DCV-RBF 97,77 % KDCV-RBF 97,65 % IDCV-SVM 98,92 % | alto | NA | Gavab ORL BDD COMPUESTA UMIST AR YALE JAFPE |
| Sawhney Shreyak et al, 2019 | Si | análisis de componentes principales (PCA) y red neuronal convolucional (CNN). | No lo menciona | bajo | NA | propia |

Capítulo 4

Metodología de solución

En este capítulo se detalla la metodología de solución propuesta para resolver la problemática identificada en esta investigación. Con base en el estudio del estado del arte se diseñó una metodología que permite implementar un algoritmo de aprendizaje automático capaz de clasificar las emociones mediante el análisis de expresiones faciales durante la evaluación de la experiencia de usuario, esta metodología consta de 7 fases, cada una es fundamental para la realización de las demás, la primera fase de la metodología corresponde a la Selección de algoritmos de detección de rostro y clasificación de emociones, la segunda fase es el diseño o adquisición de una base de datos de imágenes del rostro con emociones, la tercera fase se trata de la evaluación de los algoritmos, la fase cuatro corresponde a la implementación del algoritmo al módulo de análisis facial de la plataforma UXLab, la fase cinco corresponde a la realización de pruebas funcionales y de integración, en la fase seis se lleva a cabo la corrección de errores de implementación del algoritmo, la siete correspondió a la Validación del algoritmo en UXLab para realizar el análisis de emociones durante la evaluación de la experiencia de usuario.

4.1 Esquema de metodología de solución

La figura 4.1 presenta la propuesta de solución del problema a resolver en esta investigación, se presentará como esquema y posteriormente se describirá cada proceso.

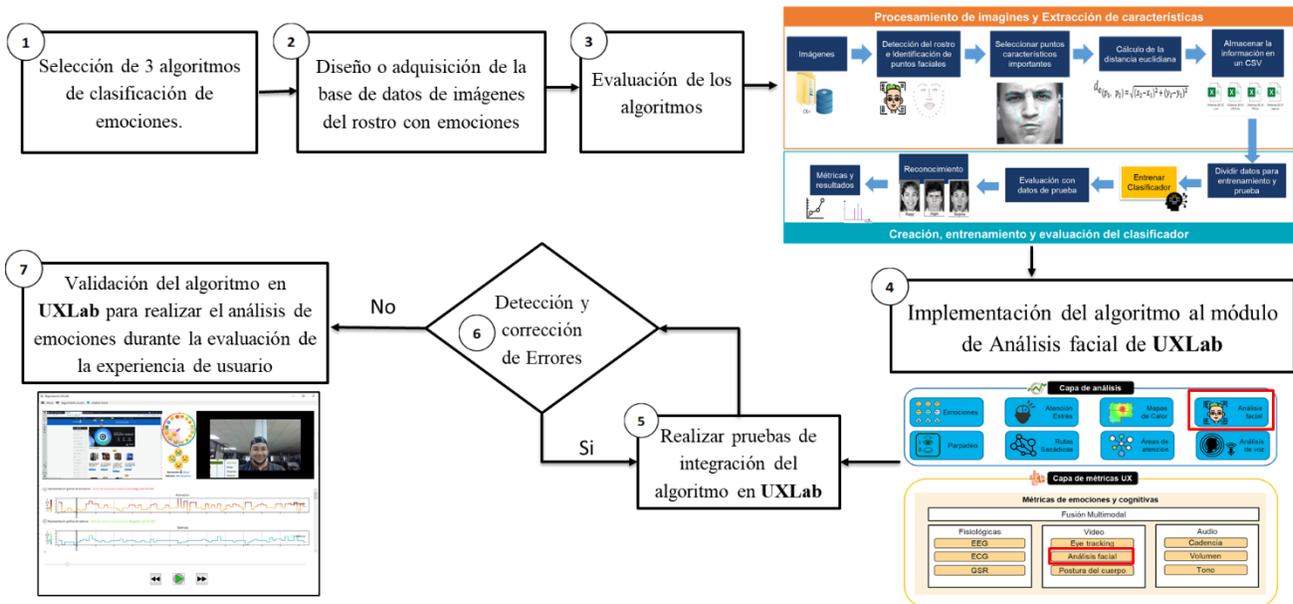


Figura 4.1 Esquema de Metodología de solución

4.2 Descripción de la Metodología de Solución

Fase 1. Selección de tres algoritmos de clasificación de emociones

En esta fase se realizó el análisis del estado del arte para identificar los estudios previos que se han realizado referente al tema de investigación para ir recolectando información útil como algoritmos, bases de datos y técnicas de reconocimiento facial, emociones y estados cognitivos los cuales nos ayudaran a dar solución a la tesis.

De acuerdo con el análisis del estado del arte se tomó un modelo pre entrenado para el reconocimiento del rostro y tres algoritmos de clasificación de aprendizaje automático; para el análisis facial, en este caso se utilizó, el modelo pre-entrenado de la librería Dlib el cual detecta el rostro en una imagen y predice 68 puntos de referencia facial, utilizado por (Valderrama Cárdenas, 2019, M. Alvarez et al. 2018), en la identificación de emociones se evaluaron los algoritmos de (1) *Maquina de Soporte Vectorial (SVM)* propuesta por (M. Alvarez et al. 2018, H.S Gunavathi, M. Siddappa, 2018), (2) *Bosques aleatorios* (M. Alvarez et al. 2018) y (3) *K-*

vecinos más cercanos (M. Alvarez et al. 2018) ya que con estos algoritmos los autores pudieron detectar emociones mediante el rostro extrayendo características mediante una técnica con base geométrica, dichos algoritmos se evaluaron seleccionando el que mejores resultados proporcionó para implementarlo en el módulo de análisis facial de la plataforma UXLab, haciendo las adecuaciones necesarias para someterlos a pruebas e identificar errores y lograr estabilidad en el sistema y la mayor precisión posible en reconocimiento de emociones en el proceso de evaluación UX.

Fase 2. Diseño o adquisición de una base de datos de rostros con emociones.

El análisis del estado del arte permitió identificar las mejores bases de datos de imágenes, de expresiones faciales, emociones y estados cognitivos los cuales permitieron entrenar los algoritmos para obtener la mayor precisión posible de identificación, por otra parte, en esta fase también se conformó una base de datos propia para poder identificar mejor las emociones y estados cognitivos ya que los rasgos faciales de las bases de datos son de otra región etnográfica y la intensidad con la que expresan las emociones varía considerablemente, se conformó el repositorio para poder obtener mejor resultado en la identificación de emociones y estados cognitivos relacionados con las características etnográficas de México. Las mejores bases de datos identificadas son JAFFE (Yang, et al, 2017), Cohn-Kanade extendido (CK+) (Kartali, et al, 2018), Eigen Faces (Mangala, Prajwala, 2018), corpus Cam3D (H.S Gunavathi, M. Siddappa, 2018), FEI y FACES94 (Valderrama Cárdenas, 2019) que son las más amplias y las que más se utilizan en el reconocimiento tanto de rostros como de emociones y estados cognitivo.

Fase 3. Evaluación de los Algoritmos.

Consistió en evaluar los tres algoritmos seleccionados en el estado del arte y entrenarlos con las bases de datos seleccionadas y creadas para determinar cuál de éstos proporciona el mejor resultado en reconocimiento de las emociones, el mejor de éstos fue el que se utilizó para implementarse en el módulo de análisis facial de UXLab y poder realizar las pruebas durante la evaluación de la experiencia de usuario.

Fase 4. Implementación del algoritmo al módulo de análisis facial de la plataforma UXLab.

Consistió en implementar algoritmos capaces de poder identificar expresiones faciales, emociones y estados cognitivos mediante el reconocimiento y análisis facial enfocada en el proceso de evaluación de la experiencia de usuario (UX), por lo que se incluyó en la plataforma UXLab para realizar las pruebas pertinentes y validación del algoritmo de reconocimiento facial.

Fase 5. Realizar pruebas de integración del algoritmo en UXLab.

En esta fase se realizaron las pruebas con el algoritmo que mejor desempeño de reconocimiento proporcionó, se realizaron los protocolos de pruebas necesarios para obtener

pruebas significativas que indique que realmente el algoritmo está realizando bien la clasificación y dar solución a la problemática planteada, se realizaron pruebas individuales con el algoritmo y se realizaron pruebas de integración en la plataforma de UXLab.

Fase 6. Detección y corrección de Errores.

En esta fase se identificaron los errores generados durante las pruebas realizadas en la fase cinco, que implicaron pruebas de integración, errores de funcionamiento del módulo de reconocimiento facial en la plataforma UXLab y los errores en el proceso de reconocimiento de emociones mediante el análisis de expresiones faciales. Posteriormente se implementó la solución para corregir adecuando y optimizando los algoritmos para una mejor precisión y realizar las pruebas nuevamente en caso de encontrar fallas o inconsistencias.

Fase 7. Validación del algoritmo en UXLab para realizar el análisis de emociones durante la evaluación de la experiencia de usuario.

Se revisaron los resultados obtenidos en las pruebas, mediante análisis de probabilidad y la elaboración de gráficas de la información. Se midió la relación de los estados emocionales y cognitivos con las experiencias mostradas a los usuarios.

Se elaboró la documentación de las pruebas realizadas y finalmente se incorporó a la plataforma UXLab una vez que ya no se presentó ningún problema en la integración al sistema.

Capítulo 5

Pruebas y resultados.

En la literatura se proponen técnicas y algoritmos capaces de identificar emociones y estados cognitivos mediante sistemas de reconocimiento de expresiones faciales los cuales varían en su nivel de precisión. La diferencia en la exactitud se debe a que existen factores que impiden una buena identificación de las características del rostro como lo son: la oclusión, el ruido, la luminosidad de la imagen, pose u orientación del rostro, etc. Otro factor que influye en el reconocimiento de emociones son los repositorios de imágenes que se utilizan para el aprendizaje de los algoritmos de clasificación ya que son adquiridas, generalmente, en ambientes controlados, diferentes a los utilizados en los sistemas reales. Y si se desea un alto nivel de precisión por ejemplo con arquitecturas de aprendizaje profundo, implica tener un alto costo computacional por la cantidad de imágenes para ser entrenados y tener un óptimo rendimiento.

Según la disponibilidad de los datos, los trabajos publicados como los de Sunitha A, et al. (2017), Nwosu, et al. (2017), Krestinskaya, et al. (2017), Liu xiao, et al. (2018), M. Alvarez, et al. (2018), Dewi, et al. (2018), Mangal, et al. (2018), Kartali, et al. (2018), Rabbi, et al. (2018), Cadena, et al. (2019), Sawhney, et al. (2019), Lacort, et al. (2017), Hussain, et al. (2018), Bahera, et al. (2020), jain, et al. (2018), sharma, et al. (2016), [17] James Garrett, Jesse (2011) y [19] Yang, D et al (2017) se clasifican en general en dos categorías. La primera categoría es el método basado en la apariencia que utiliza la modalidad de textura y exploran las diferencias de expresión en el espacio de píxeles. Considerando la imagen de la textura facial completa como la característica expresiva, algunos investigadores mapean el vector de alta dimensión en el subespacio de baja dimensión aplicando un algoritmo de reducción de dimensionalidad, como el análisis de componentes principales (PCA) Rabbi, et al. (2018), y Sawhney, et al. (2019). Algunos otros investigadores seleccionan algunas áreas faciales representativas como objetos de análisis y proponen muchos métodos de extracción de características locales para calcular descriptores locales de estas áreas. Los métodos más populares incluyen Haar Cascades Yang, D et al (2017), filtro de extracción de bordes (EEF) Liu xiao, et al. (2018) e Histograma de gradientes orientados (HOG) Rabbi, et al. (2018). Aunque el método basado en la apariencia puede capturar la información detallada y sutil de la expresión facial, las características son muy sensibles a la luminancia, la variación de la

postura de la cabeza y la oclusión. La segunda categoría se puede considerar como un método de base geométrica. Este método generalmente necesita marcar un punto de característica facial cuyos movimientos correspondientes pueden ayudar a capturar las características expresivas como los empleados en M. Alvarez, et al. (2018), Dewi, et al. (2018) y Mangal, et al. (2018).

Por lo anterior, el objetivo de esta tesis fue el desarrollo de un método para el reconocimiento de tres emociones mediante el análisis de expresión facial implementando un algoritmo de clasificación basado en máquinas de aprendizaje aplicando la extracción de características con un método basado en el cálculo de distancias geométricas para aplicarlo en videos de rostros grabados durante el proceso de evaluación de la experiencia de usuario (UXE por sus siglas en inglés) el objetivo es obtener información relacionada con los aspectos emocionales que experimenta el usuario durante la interacción con la interfaz de usuario de un producto digital (software) esta información es de utilidad para los evaluadores de la UX ya que les permite identificar la polaridad emocional que experimenta el usuario durante el periodo de la interacción.

En este capítulo se presentan los resultados obtenidos de las pruebas realizadas para evaluar los métodos de, preprocesamiento de imágenes, localización del rostro y la clasificación de emociones. En el preprocesamiento de imágenes solo se realiza la redimensión de las imágenes y la conversión a escala de grises, posteriormente se realiza la detección del rostro en las imágenes utilizando la librería Dlib D. King (2015) con el fin de emplear un método con base al cálculo de distancias geométrica de los puntos de referencias faciales para la extracción de características.

Para la clasificación de las emociones se compararon tres clasificadores que son: Máquina de Soporte Vectorial, Random Forest y K vecinos más cercanos. Los resultados son evaluados mediante las métricas de: Accuracy, Precision, Recall y F1-Score. Además, se presentan las pruebas de integración del clasificador en el módulo de análisis facial de la plataforma de UXLab que es una plataforma software desarrollada en CENIDET.

5.1 experimentación Y Resultados.

Para realizar las pruebas se identificaron y evaluaron los repositorios de imágenes que se utilizaron para el entrenamiento y prueba de los algoritmos de clasificación, también se describen las métricas que se evaluaron para una mejor comprensión de la visualización de los resultados.

5.1.1 Repositorio de imágenes

En la fase dos de la metodología de solución propuesta en la figura 4.1 del capítulo 4, se realizó la búsqueda de repositorios de imágenes emocionales, como resultado se seleccionaron

tres repositorios, dos tomados de la literatura: 1) Cohn-Kanade extendido (CK+) para evaluar y validar los tres clasificadores de emociones, 2) FER2013 que en conjunto con 3) Un repositorio propio sirvió para el entrenamiento de los clasificadores. A continuación, se describen los diferentes repositorios de imágenes.

5.1.1.1 Cohn-Kanade extendido (CK+)

Este repositorio es una extensión de la base de datos de imágenes de Cohn-Kanade Lucey, P (2010) con mejoras en el etiquetado de las emociones, contiene 593 secuencias de video de 123 personas con expresiones faciales codificadas por el Sistema de Codificación de Acción Facial (FACS por sus siglas en inglés) Lucey, P (2010), todas las secuencias son seguidas por el modelo de apariencia activa (AAM por sus siglas en inglés) Lucey, P (2010), con 68 puntos de referencia. Cada secuencia incluye desde la expresión neutra hasta la máxima expresión donde, 0 = neutral, 1 = enojo, 2 = desprecio, 3 = disgusto, 4 = miedo, 5 = feliz, 6 = tristeza y 7 es la emoción máxima que corresponde a Sorpresa quedando así las siete emociones básicas determinadas por Ekman P.E Group (2016), tomadas entre julio de 2010 y septiembre de 2010. Todas las caras están representadas principalmente por personas de entre 18 y 60 años, el 69% eran mujeres, el 81%, euroamericanos, el 13% afroamericanos y el 6% de otros grupos. Previo a las pruebas se instruyó a los participantes para que realizaran una serie de 23 demostraciones faciales; estos incluían unidades de acción única y combinaciones de unidades de acción. Cada exhibición comenzaba y terminaba en una cara neutral con las excepciones señaladas. Las secuencias de imágenes para vistas frontales y vistas de 30 grados se digitalizaron en matrices de píxeles de 640x490 o 640x480 con valores de color de escala de grises de 8 bits o de 24 bits. en la figura 5.1 se muestra un ejemplo de las secuencias de este repositorio de imágenes.



Figura 5.1 Rostros con las emociones de Felicidad, Enojo, Sorpresa y Neutral respectivamente del conjunto de datos CK+ Lucey, P (2010).

5.1.1.2 FER2013

FER2013 Sambare, M. (2020) se conforma de imágenes de rostros en escala de gris de 48x48 píxeles. Los rostros se capturaron automáticamente para que esté más o menos centrado y ocupe aproximadamente la misma cantidad de espacio en cada imagen. Cada rostro se clasifica en

función de la emoción que se muestra en la expresión facial en una de siete categorías (0 = Enojado, 1 = Disgusto, 2 = Miedo, 3 = Feliz, 4 = Triste, 5 = Sorpresa, 6 = Neutral) con diferentes poses. El conjunto de formación consta de 28.709 ejemplos, pero el conjunto de pruebas público consta de 3.589 ejemplos, en la figura 5.2 se muestra un ejemplo de cada una de las emociones de este repositorio.

- Enojado con 958 imágenes.
- Disgusto con 111 imágenes.
- Miedo con 1024 imágenes.
- Feliz con 1774 imágenes.
- Triste con 1247 imágenes.
- Sorpresa con 831 imágenes.
- Neutral con 1233 imágenes.



Figura 5.2 Rostros con las emociones de Felicidad, Enojo, Sorpresa y Neutral respectivamente del conjunto de datos FER Sambare, M. (2020).

5.1.1.3 Repositorio Propio

Este repositorio de imágenes propio se creó para entrenar y evaluar los algoritmos de clasificación de emociones, utilizando rostros de personas mexicanas, ya que los repositorios antes mencionados contienen imágenes con rostros de otras regiones etnográficas, por lo que al entrenar con dichos repositorios los resultados obtenidos no fueron los esperados. Por lo anterior surgió la necesidad de crear un repositorio propio que permitió obtener mejores resultados, el repositorio contiene 532 imágenes de rostros en formato RGB con un tamaño de 968 X 720 píxeles de los cuales se seleccionaron 432 imágenes para realizar el entrenamiento y las pruebas; para que el repositorio estuviera balanceado entre las clases, se clasificaron en 4 emociones, las imágenes fueron tomadas del 1 de mayo del 2021 en ambientes naturales y el resto fueron solicitadas a personas y conocidos, cada conjunto de imágenes se almacenó en cuatro repositorios que corresponden a cada emoción como se describe a continuación ver Tabla 5.1.

Tabla 5.1 Características del Repositorio de imágenes propio

| Etiqueta | Emoción | Cantidad | Hombres | Mujeres | Edad |
|----------|----------|----------|---------|---------|---------------|
| 0 | Alegría | 149 | 59% | 49% | Entre 27 y 40 |
| 1 | Enojo | 115 | 54% | 46% | Entre 27 y 40 |
| 2 | Sorpresa | 119 | 47% | 53% | Entre 27 y 40 |
| 3 | Neutral | 149 | 64% | 36% | Entre 27 y 40 |

La figura 5.3 muestra un ejemplo de cada emoción de las imágenes que contiene el repositorio propio.



Figura 5.3 Rostros con las emociones de Felicidad, Enojo, Sorpresa y Neutral respectivamente del conjunto de datos PROPIO

5.1.2 Métricas para evaluar.

La validación de resultados es la etapa final en el proceso de reconocimiento emocional que consiste en la evaluación y comparación de los modelos programados a partir de pruebas de entrenamiento y predicción con el repositorio desarrollado. El resultado de estas pruebas suele ser una matriz de confusión, la cual es utilizada para el cálculo de métricas especializadas en la medición de diferentes aspectos del modelo en cuestión Shu et al (2018), las métricas que se evaluaron para los algoritmos son las que más se utilizan en la literatura, de acuerdo con la naturaleza de los datos y al tipo de algoritmo utilizado en este caso se utilizó un algoritmo de aprendizaje automático del tipo supervisado por lo que las métricas a evaluar son Precisión, Recall y F1-Score, que se describieron en el sección 1.27 del capítulo 2.

5.1.3 Experimentación.

En esta sección se describen las pruebas que se realizaron para evaluar los algoritmos de clasificación, se divide en tres secciones la primera describe la detección del rostro en las imágenes utilizando la librería Dlib D. King(2015) se realizaron pruebas con la finalidad de verificar la robustez en la detección del rostro ante variabilidad de iluminación y pose ya que esta librería permite colocar puntos de referencia en el rostro para poder realizar el cálculo de distancias geométricas la cual sirvió para la extracción de características para relacionarlas con

las unidades de acción propuestas por Ekman P.E Group (2016), la segunda sección corresponde a la evaluación de los algoritmos de clasificación y la tercera sección presenta el resultado de las pruebas de integración del algoritmo que presentó mejores resultados a la plataforma de Evaluación de la Experiencia de Usuario UXLab.

5.1.3.1 Detección de rostro con el algoritmo Histograma de Gradientes Orientado un conjunto de árboles de regresión en cascada de Dlib.

La biblioteca Dlib nos proporciona un modelo pre-entrenado, el cual es una implementación del artículo *Alineación de cara de un milisegundo con un conjunto de árboles de regresión de Kazemi y Sullivan (2014)*, localiza el rostro en una imagen y estima la ubicación de 68 coordenadas (x, y) que se asignan a estructuras faciales mediante un conjunto de árboles de regresión en cascada. En la figura 5.4 se muestran las 68 coordenadas.

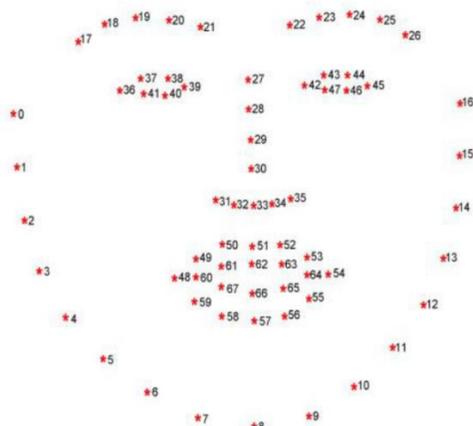


Figura 5.4 Los 68 puntos de referencia faciales del conjunto de datos iBUG 300-W (ibug)

El algoritmo 1 muestra el pseudocódigo para estimar los puntos de referencia en la cara desarrollado por Sullivan, 2014 e implementado por D. King (2015) en su librería Dlib.

Algoritmo 1 Entrenamiento de árboles de regresión para localización de puntos faciales

Entrada: Tener datos de entrenamiento $\{(I_{\pi_i}, \hat{S}_i^{(t)}, \Delta S_i^{(t)})\}_{i=1}^N$ y el aprendizaje tasa (factor de contracción) $0 < v < 1$

Salida: $r_t(I, \hat{S}^{(t)}) = f_k(I, \hat{S}^{(t)})$

1. Inicializar

$$f_0(I, \hat{S}^{(t)}) = \arg \min_{\gamma \in \mathbb{R}^{2p}} \sum_{i=1}^N \|\Delta S_i^{(t)} - \gamma\|^2$$

2. **para** $k = 1, \dots, K$: **hacer**

(a). asignar

para $i = 1, \dots, N$ **hacer**

$$r_{ik} = \Delta \hat{S}^{(t)} - f_{k-1}(I_{\pi i}, \hat{S}_i^{(t)})$$

(b). Ajuste un árbol de regresión a los objetivos r_{ik} dando una función de regresión débil $gk(I, \hat{S}^{(t)})$.

(c). Actualizar

$$f_k(I, \hat{S}^{(t)}) = f_{k-1}(I, \hat{S}^{(t)}) + v gk(I, \hat{S}^{(t)})$$

Fin para

Fin para

3. **Salida** $r_t(I, \hat{S}^{(t)}) = f_k(I, \hat{S}^{(t)})$

Las ubicaciones de los rasgos faciales se pueden evaluar a través de los diferentes conjuntos de puntos [punto de inicio, punto final]:

- Ojo izquierdo: puntos [42, 47]
- Boca: puntos [48, 67]
- Ceja izquierda: puntos [22, 26]
- Nariz: puntos [27, 34]
- Ceja derecha: puntos [17, 21]
- Ojo derecho: puntos [36, 41]
- Línea de la mandíbula: puntos [0, 16]

Para que el algoritmo detector de Dlib D. King (2015) realice la predicción de los puntos, toma como parámetro de entrada la imagen en escala de grises obteniendo las coordenadas x, y en píxeles de cada punto, mediante un bucle que va iterando punto por punto del predictor. En la figura 5.4 se puede observar cómo se posicionan los puntos del predictor. La figura 5.5 muestra los 68 puntos localizados en un rostro con una expresión de felicidad.

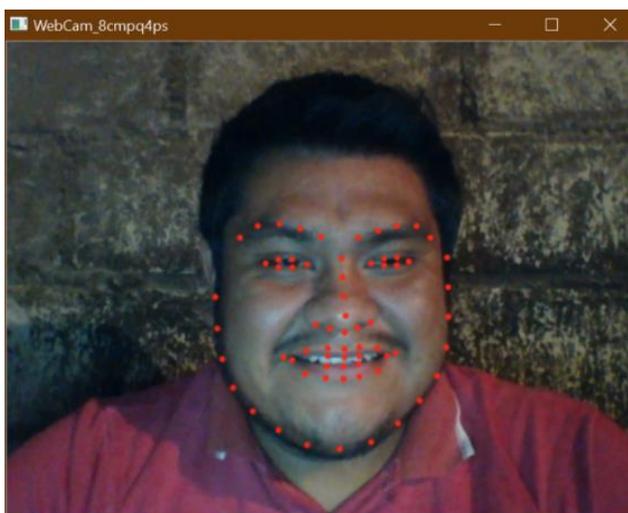


Figura 5.5 Los 68 puntos de referencia faciales posicionados en el rostro que presenta la emoción de felicidad

Se realizaron pruebas con tres personas con variación de iluminación y poses con la finalidad de poder determinar la limitante que la librería tiene en la detección del rostro, las figura 5.6 a la 5.8 muestran las pruebas realizadas en la detección del rostro.



Figura 5.6 Detección del rostro con diferentes poses y variabilidad de iluminación persona uno.



Figura 5.7 Detección del rostro con diferentes poses y variabilidad de iluminación persona dos.



Figura 5.8 Detección del rostro con diferentes poses y variabilidad de iluminación persona tres.

Como se puede observar en la figura 5.6 y 5.7 se logró detectar el rostro aun cuando la luminosidad de la imagen es baja y la pose mientras no sea totalmente de perfil y se pierda alguna región del rostro como los ojos, boca, nariz o barbilla la librería es capaz de detectar el rostro en una imagen, la primera imagen de la figura 5.8 muestra en qué punto ya no se puede detectar el rostro; cuando la cabeza de la persona está completamente de perfil.

De la misma forma se realizaron pruebas para verificar si los puntos de referencia facial se localizaban correctamente en presencia de variación de pose y luminosidad de las imágenes, las figuras 5.9 a la 5.11 muestran los resultados de las pruebas.



Figura 5.9 localización de los puntos de referencia facial en el rostro de la persona uno con diferente orientación y variación de luminosidad en la imagen.



Figura 5.10 localización de los puntos de referencia facial en el rostro de la persona dos con diferente orientación y variación de luminosidad en la imagen



Figura 5.11 localización de los puntos de referencia facial en el rostro de la persona tres con diferente orientación y variación de luminosidad en la imagen.

Como se puede observar, la variación de luminosidad no impide que los puntos de referencia sean localizados en el rostro, el inconveniente es la variación de pose ya que entre más perfilada esté la cabeza de la persona hay cierta desalineación de los puntos de referencia con respecto al rostro, la pose óptima para que los puntos de referencia se acoplen en el rostro es la pose frontal, estas pruebas permitieron restringir la variación de pose a la hora de realizar la extracción de características.

Después de realizar las pruebas con la librería Dlib se realizó la evaluación de los algoritmos, para esta investigación se evaluaron tres: 1) *Máquina de Soporte Vectorial*, 2) *Random Forest* y 3) *K vecinos más cercanos*, que se identificaron como los más utilizados en la revisión de la literatura relacionada, tomando como premisa que se requiere un algoritmo que ofrezca el menor costo computacional posible, el método para la extracción de características propuesto es un método de base geométrica que utiliza puntos de referencia facial para calcular distancias entre los puntos que se necesitan para identificar las emociones, estos puntos fueron seleccionados con

base al sistema de codificación de acción facial para emociones propuesta por Ekman P.E Group (2016), relacionándola con cada unidad de acción que interviene en cada emoción a detectar.

5.1.3.2 Evaluación de los algoritmos.

Para evaluar los tres algoritmos, se implementó un procedimiento para crear el clasificador de emociones, el procedimiento está dividido en dos partes, cada una consta de 5 fases, la primera parte incluye el proceso para la extracción de características con el método de base geométrica propuesto y el preprocesamiento de las imágenes, la segunda parte corresponde a la creación, entrenamiento, prueba y validación del clasificador como se puede ver en la figura 5.12

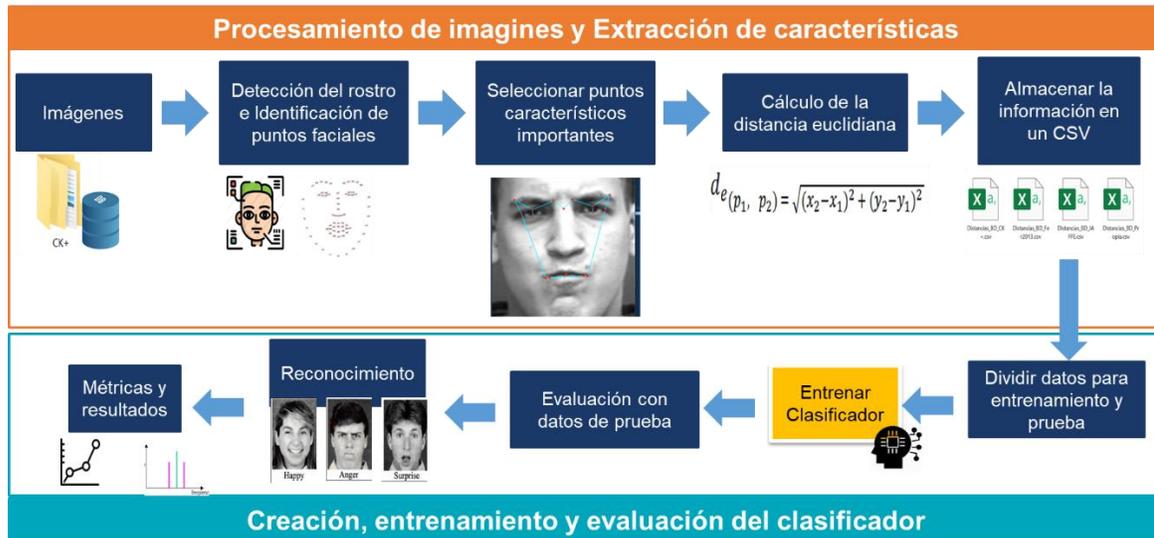


Figura 5.12 Esquema del sistema de clasificación de emociones

Para un mejor entendimiento de cómo se realizó la evaluación de los algoritmos se describirá la primera parte del procedimiento antes mencionado ya que a partir del procesamiento de los repositorios de imágenes se generan los conjuntos de datos que contendrán las distancias calculadas de los puntos de referencia facial para realizar el entrenamiento y pruebas de los clasificadores, posteriormente se presentarán los resultados de las métricas a evaluar para cada clasificador.

5.1.3.2.1 Procesamiento de imágenes y extracción de características.

La primera fase de este procedimiento corresponde a los datos de entrada, en este caso, las imágenes del rostro, se utilizaron dos repositorios, uno para el entrenamiento y otro para realizar

las pruebas, cada repositorio consta de 432 imágenes con rostros que presentan expresiones faciales de cuatro emociones básicas, que son alegría, enojo, sorpresa y neutral, el primer repositorio *Cohn-Kanade extendido* más conocido como *CK+* (Lucey, P, 2010) se seleccionó de la literatura ya que es utilizado en varios trabajos para el reconocimiento de emociones, este repositorio se utilizó para realizar las pruebas, el segundo repositorio es propio que se combinó con el repositorio de FER2013 que se utiliza en la literatura, esto para igualar las imágenes para cada clase y no se tenga problemas de desbalanceo a la hora de entrenar el algoritmo.

La segunda fase corresponde a la detección del rostro y localización de los puntos de referencia, como se mencionó anteriormente, se utilizó el modelo pre-entrenado de la librería Dlib el cual predice 68 puntos faciales de siete regiones del rostro, ojo izquierdo, ojo derecho, ceja izquierda, ceja derecha, boca, nariz y la línea de la mandíbula, para que el modelo proporcione estas coordenadas, fue necesario convertir las imágenes a escala de grises y redimensionarlas a un tamaño de 300 X 300 píxeles para que todas las imágenes tengan el mismo tamaño.

La tercera fase fue seleccionar los puntos de referencia facial relevantes, de los 68 puntos solo se tomaron en cuenta aquellos que sirven para calcular las distancias necesarias que permitan identificar las emociones y posteriormente guardar esas distancias en un archivo CSV quedando como puntos importantes los siguientes 17 ver Tabla 5.2.

Tabla 5.2 Puntos relevantes para generar las distancias que nos ayudaran a identificar las 4 emociones.

| No. | Región del rostro | Puntos o coordenadas |
|-----|-------------------|--------------------------|
| 1 | Ojo izquierdo | [45, 71,72] |
| 2 | Boca | [60, 64, 62, 66, 48, 54] |
| 3 | Ceja izquierda | [22, 25] |
| 4 | Nariz | [27] |
| 5 | Ceja derecha | [18,21] |
| 6 | Ojo derecho | [36, 69,70] |

La figura 5.13 muestra los 17 puntos seleccionados localizados en un rostro con la expresión de enojo.

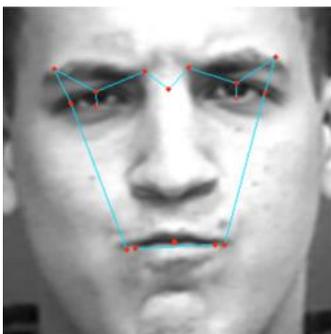


Figura 5.13 Puntos faciales en un rostro con expresión de enojo del conjunto de datos CK+ (Lucey, P, 2010)

La determinación de selección de estos 17 puntos fue con base al sistema de codificación de acción facial para emociones propuesta por Ekman el cual proporciona las unidades de acción que intervienen en cada emoción, unidades de acción se refiere a que Ekman le asignó un identificador con las siglas AU y un número a los movimientos de cada uno de los músculos del rostro que corresponde a cada emoción, por ejemplo para la emoción de felicidad o alegría intervienen las unidades de acción AU6 y AU12 la primera corresponde al levantamiento de las mejillas y la segunda corresponde a la contracción de las comisuras de la boca, en esas regiones se encuentran los puntos 36, 45, 48, 60, 62, 64 y 66, de esta forma fue como se relacionó cada punto con las unidades de acción.

Las tablas 5.3 a la 5.5 muestran la relación de las unidades de acción con los puntos de referencia facial seleccionados para la emoción de felicidad, enojo y sorpresa.

Tabla 5.3 Relación entre los puntos faciales y las unidades de acción de la emoción de Felicidad

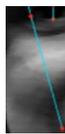
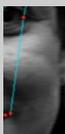
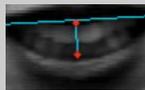
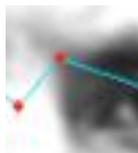
| No | Combinación | Puntos | Imagen |
|------------|---|-----------|---|
| AU6 y AU12 | | | |
| 1 | Long_mejillas_elevadas_derecha | [36,48] |  |
| 2 | Long_mejillas_elevadas_izquierda | [45, 64] |  |
| 3 | Long_comisura_interior_izquierda_y_derecha_de_la_boca | [60, 64] |  |
| 4 | Long_labio_interior_inferior_y_superior_de_la_boca | [62, 66] |  |

Tabla 5.4 Relación entre los puntos faciales y las unidades de acción de la emoción de Enojo

| No | Combinación | Puntos | Imagen |
|----------------------|--|-----------|---|
| AU4, AU5, AU7 y AU23 | | | |
| 1 | Long_comisura_derecha_de_ceja_izquierda_inicio_tabique | [22, 27] |  |

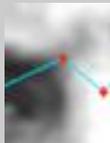
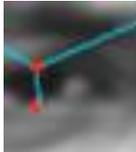
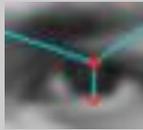
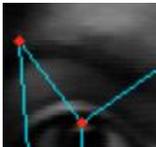
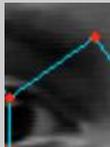
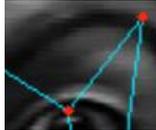
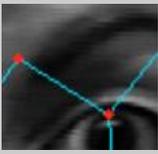
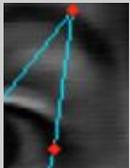
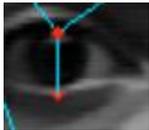
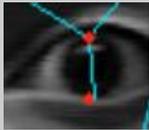
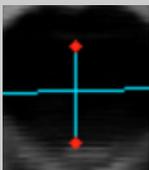
| | | | |
|---|---|-----------|--|
| 2 | Long_comisura_izquierda_de_ceja_derecha_inicio_de_tabique | [21, 27] |  |
| 3 | Long_Elevador_de_parpado_superior_de_Ojo_Derecho | [69, 70] |  |
| 4 | Long_Elevador_de_parpado_superior_de_Ojo_Izquierdo | [71, 72] |  |
| 5 | Long_comisura_interior_izquierda_y_derecha_de_la_boca | [60, 64] |  |
| 6 | Long_labio_interior_inferior_y_superior_de_la_boca | [62, 66] |  |

Tabla 5.5 Relación entre los puntos faciales y las unidades de acción de la emoción de Sorpresa

| No | Combinación | Puntos | Imagen |
|---------------------|--|-----------|---|
| AU1, AU2, AU5 y AU6 | | | |
| 1 | Long_comisura_derecha_de_Elevador_de_ceja_interior_derecha | [18,69] |  |
| 2 | Long_comisura_izquierda_de_Elevador_de_ceja_interior_derecha | [21, 69] |  |
| 3 | Long_comisura_izquierda_de_Elevador_de_ceja_interior_izquierda | [25, 71] |  |

| | | | |
|----|---|-----------|---|
| 4 | Long_comisura_derecha_de_Elevador_de_ceja_interior_izquierda | [22, 71] |  |
| 5 | Long_comisura_derecha_de_Elevador_de_ceja_interior_comisura_derecha_ojo_derecho | [18, 36] |  |
| 6 | Long_comisura_izq_de_Elevador_de_ceja_interior_comisura_izquierda_ojo_izquierdo | [25, 45] |  |
| 7 | Long_Elevador_de_parpado_superior_de_Ojo_Derecho | [69, 70] |  |
| 8 | Long_Elevador_de_parpado_superior_de_Ojo_Izquierdo | [71, 72] |  |
| 9 | Long_comisura_interior_izquierda_y_derecha_de_la_boca | [60, 64] |  |
| 10 | Long_labio_interior_inferior_y_superior_de_la_boca | [62, 66] |  |

La cuarta fase corresponde al cálculo de distancias entre los puntos seleccionados, la cual, se llevó a cabo mediante la ecuación de la distancia euclidiana, la ecuación 5.1 muestra la expresión matemática del cálculo.

$$d_e(p_1, p_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (5.1)$$

Donde:

$d_e(p_1, p_2)$ es la distancia que se pretende calcular de un punto característico del rostro a otro punto característico del rostro.

- x_1 es la coordenada X en pixeles del punto uno.
- x_2 es la coordenada X en pixeles del punto dos.
- y_1 es la coordenada Y en pixeles del punto uno.
- y_2 es la coordenada Y en pixeles del punto dos.

Por ejemplo, para calcular la longitud del parpado superior e inferior del ojo izquierdo se tiene que calcular la distancia entre el punto 71 y 72 como se ve en la figura 5.14.

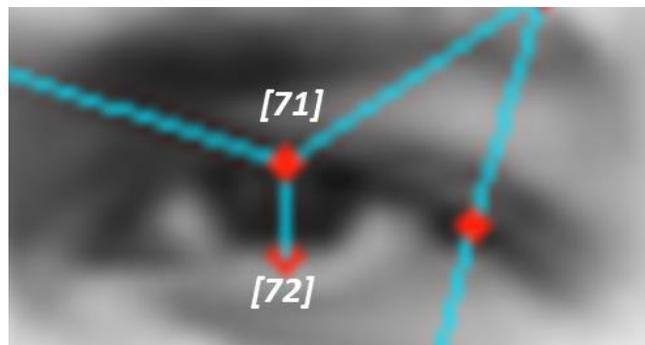


Figura 5.14 Puntos [71, 72] del parpado superior e inferior del ojo izquierdo.

La quinta fase es la parte de almacenamiento de las distancias calculadas en un archivo CSV, el cual se utilizó como datos de entrada para el entrenamiento y prueba del clasificador.

A continuación, se presentan las pruebas realizadas para cada algoritmo de clasificación evaluado, así como los resultados de las métricas.

De acuerdo con la revisión de la literatura, se seleccionó un algoritmo para la clasificación que se ha utilizado en el reconocimiento de emociones con base geométrica, como se describe en M. Alvares et al (2018), Mangala et al (2018) y Dewi et al (2018). El algoritmo se denomina Máquina de Vector Soporte o más conocido en inglés como Support Vector Machine SVM Joaquín A.R. (2021). Es un algoritmo de clasificación y regresión desarrollado en la década de los 90, dentro del campo de la ciencia computacional. Aunque inicialmente se desarrolló como un método de clasificación binaria, su aplicación se ha extendido a problemas de clasificación múltiple y regresión. El algoritmo SVM se considera uno de los mejores clasificadores para un amplio abanico de situaciones, por lo que se considera uno de los referentes dentro del ámbito de aprendizaje estadístico y de máquinas de aprendizaje.

Para iniciar con el entrenamiento del algoritmo de clasificación, se dividió el conjunto de datos propio combinado con el conjunto de datos FER2013, para cada clase se tomó 108

imágenes para evitar un desbalanceo de clases en el repositorio teniendo un total de 432 imágenes de este repositorio se tomó el 80% para entrenar los algoritmos, posteriormente se establecieron los hiperparámetros óptimos con los cuales se logró obtener un 84.52% de precisión en el reconocimiento. También se realizó el entrenamiento y experimentación de dos algoritmos más, RandomForest y KNN, se entrenaron con el mismo conjunto de datos que el primer algoritmo y se eligieron los hiperparámetros óptimos para cada uno.

Posteriormente al aprendizaje, se evaluó el modelo, en esta prueba se consideró el conjunto de datos CK+. De este repositorio se tomó el 20% de los datos para probar el clasificador.

5.1.3.2.2 Máquina de Soporte Vectorial.

Con el algoritmo de máquina de soporte vectorial se obtuvo un Accuracy del 84.52 % de reconocimiento, la tabla 5.6 muestra los porcentajes de cada una de las métricas evaluadas para este algoritmo.

Tabla 5.6 Métricas del Algoritmo MSV.

| Emoción | Precisión | Recall | F1-Score |
|------------------|-----------|--------|-------------|
| Felicidad | 0.96 | 0.92 | 0.94 |
| Enojo | 0.71 | 0.62 | 0.67 |
| Sorpresa | 0.91 | 1.00 | 0.95 |
| Neutral | 0.76 | 0.79 | 0.78 |

Como se puede observar en la tabla 5.6, las emociones mejor identificadas fueron felicidad y sorpresa, que son las emociones que tienen características muy diferentes a las demás, por otra parte, las emociones de enojo y neutral la clasificación no fue muy buena debido a que estas dos emociones son muy parecidas a la hora de que las personas expresan dicha emoción, la tabla 5.7 muestra la distribución de la muestra y los resultados en la clasificación que el clasificador tomó para realizar las pruebas.

Tabla 5.7 Muestra que tomo el clasificador para realizar las pruebas

| Emoción | Personas | Aciertos | Equivocaciones | Porcentaje reconocimiento de la muestra. |
|-----------|----------|----------|----------------|--|
| Felicidad | 24 | 22 | 2 | 92% |
| Enojo | 16 | 10 | 6 | 62% |
| Sorpresa | 20 | 20 | 0 | 100% |
| Neutral | 24 | 19 | 5 | 79% |

5.1.3.2.3 Random Forest

Con el algoritmo de Random Forest se obtuvo un Accuracy del 82.14 % de reconocimiento, la tabla 5.8 muestra los porcentajes de cada una de las métricas evaluadas para este algoritmo.

Tabla 5.8 Métricas del Algoritmo Random Forest.

| Emoción | Precisión | Recall | F1-Score |
|------------------|-----------|--------|-------------|
| Felicidad | 1.00 | 0.88 | 0.93 |
| Enojo | 0.57 | 0.75 | 0.65 |
| Sorpresa | 0.97 | 0.95 | 0.97 |
| Neutral | 1.00 | 0.74 | 0.72 |

Como se puede observar en la tabla 5.8, las emociones mejor identificadas fueron felicidad y Sorpresa que son las emociones que tienen características muy diferentes a las demás, por otra parte, las emociones de enojo y neutral la clasificación no fue muy buena debido a que estas dos emociones son muy parecidas cuando las personas expresan dichas emociones, la tabla 5.9 muestra la distribución de la muestra y los resultados en la clasificación que el clasificador tomo para realizar las pruebas.

Tabla 5.9 Muestra que tomo el clasificador para realizar las pruebas

| Emoción | Personas | Aciertos | Equivocaciones | Porcentaje reconocimiento de la muestra. |
|-----------|----------|----------|----------------|--|
| Felicidad | 24 | 21 | 3 | 88% |
| Enojo | 16 | 12 | 4 | 75% |
| Sorpresa | 20 | 19 | 1 | 95% |
| Neutral | 24 | 17 | 7 | 74% |

5.1.3.2.4 KNN

Con el algoritmo de KNN se obtuvo un Accuracy del 84.52 % de reconocimiento, la tabla 5.10 muestra los porcentajes de cada una de las métricas evaluadas para este algoritmo.

Tabla 5.10 Métricas del Algoritmo KNN.

| Emoción | Precisión | Recall | F1-Score |
|------------------|-----------|--------|-------------|
| Felicidad | 0.96 | 0.96 | 0.96 |
| Enojo | 0.82 | 0.56 | 0.67 |
| Sorpresa | 1.00 | 0.85 | 0.92 |
| Neutral | 0.69 | 0.92 | 0.79 |

Como se puede observar en la tabla 5.10, en esta prueba las emociones mejor identificadas fueron felicidad y sorpresa, por otra parte, la clasificación de la emoción de enojo no fue buena debido a que esta emoción es muy parecida a la emoción neutral y la emoción de sorpresa en esta prueba no fue tan buena como en las anteriores, la tabla 5.11 muestra la distribución de la muestra y los resultados en la clasificación que el clasificador tomo para realizar las pruebas.

Tabla 5.11 Muestra que tomo el clasificador para realizar las pruebas

| Emoción | Personas | Aciertos | Equivocaciones | Porcentaje reconocimiento de la muestra. |
|-----------|----------|----------|----------------|--|
| Felicidad | 24 | 23 | 1 | 96% |
| Enojo | 16 | 9 | 7 | 67% |
| Sorpresa | 20 | 17 | 3 | 92% |
| Neutral | 24 | 22 | 2 | 79% |

Como se puede ver en las tablas, se obtuvieron buenos resultados con los tres algoritmos, pero, aunque se obtuvieron buenos resultados en el Accuracy en las métricas se puede identificar el mejor algoritmo en la clasificación de las cuatro emociones. De los tres el que mejor resultados presentó en las métricas fue el clasificador con Máquina de Soporte Vectorial realizando un buen reconocimiento en cada emoción como se ve en las figuras 5.15 y 5.16 las métricas de *Precision-Recall* y *F1-Score* respectivamente, también podemos ver que las emociones mejor clasificadas son la *alegría* y la *sorpresa* ya que son difíciles de confundir por las características propias de estas emociones, en cambio las emociones de *enojo* y *neutral* son más fáciles de ser confundidas por la forma en que son expresadas por las personas.

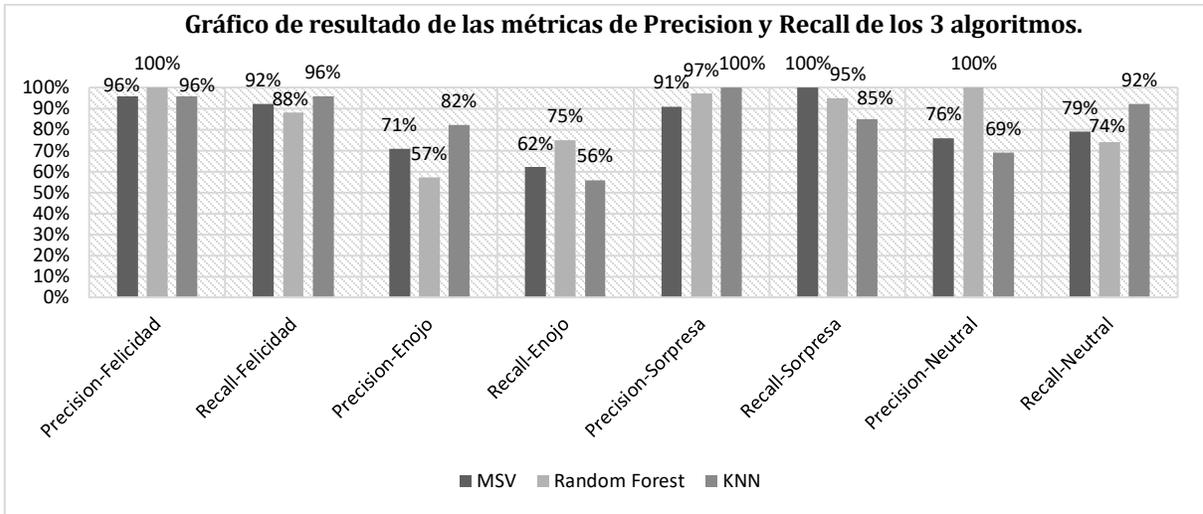


Figura 5.15 Gráfico de resultado de las métricas de Precisión y Recall de los 3 algoritmos.

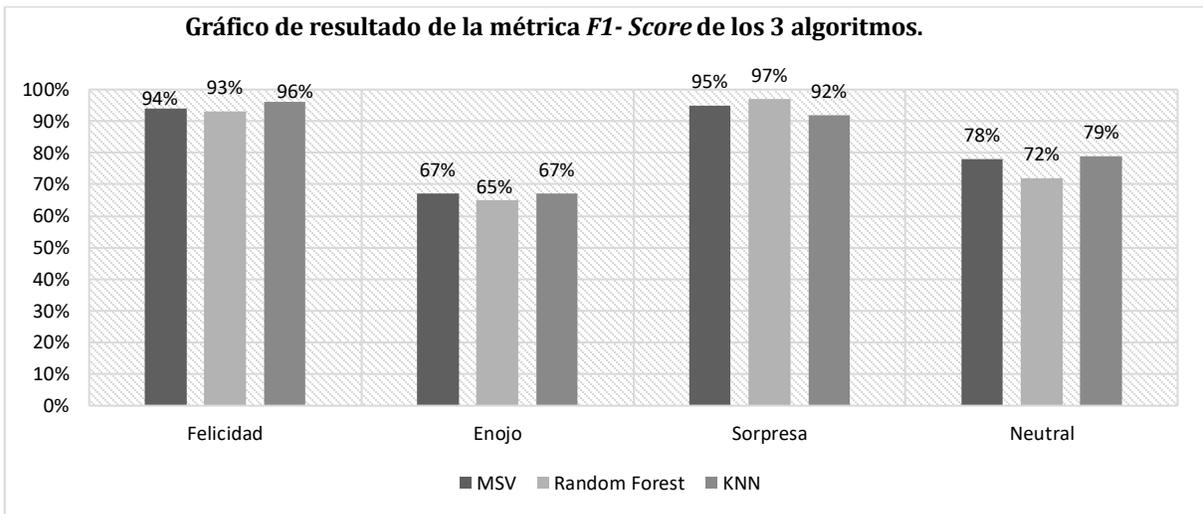


Figura 5.16 Gráfico de resultado de las métricas F1-Score de los 3 algoritmos.

5.1.3.2.5 Pruebas con imágenes estáticas

Para complementar la evaluación de los algoritmos, se realizaron pruebas con imágenes estáticas, diferentes a las de entrenamiento, implementando el algoritmo de Máquina de Soporte Vectorial el cual dio el mejor resultado, obteniendo un buen porcentaje de reconocimiento para cada emoción.

Para realizar las pruebas se tomaron 10 personas por cada emoción, la tabla 5.12 muestra las características de las personas.

Tabla 5.12 Características de las personas que se les realizó las pruebas con el clasificador

| Emoción | Personas | Sexo | | Edad | Pose |
|-----------|----------|------|---|--------------------|---------|
| | | H | M | | |
| Felicidad | 10 | 5 | 5 | Entre 20 y 30 años | Frontal |
| Enojo | 10 | 5 | 5 | Entre 20 y 30 años | Frontal |
| Sorpresa | 10 | 5 | 5 | Entre 20 y 30 años | Frontal |
| Neutral | 10 | 5 | 5 | Entre 20 y 30 años | Frontal |

La figura 5.17, muestra los resultados de la clasificación con las personas de sexo masculino mediante un gráfico de barras de la emoción de alegría, con el clasificador a simple vista se puede ver que los resultados fueron favorables por lo que en la tabla 5.13 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.18 se representa por medio de un gráfico.

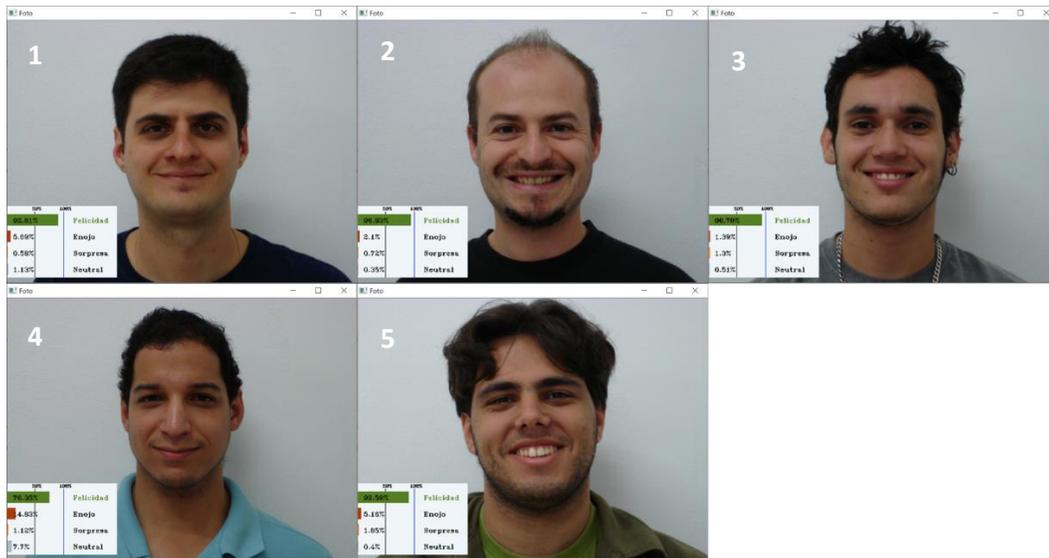


Figura 5.17 Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Felicidad.

Tabla 5.13 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de alegría

| Persona | Emoción | Porcentaje |
|---------|------------------|---------------|
| 1 | Felicidad | 92.61% |
| | Enojo | 5.69% |
| | Sorpresa | 0.58% |
| | Neutral | 1.13% |
| 2 | Felicidad | 96.83% |
| | Enojo | 2.1% |
| | Sorpresa | 0.72% |
| | Neutral | 0.35% |
| 3 | Felicidad | 96.79% |
| | Enojo | 1.39% |
| | Sorpresa | 1.3% |
| | Neutral | 0.51% |

| | | |
|---|------------------|---------------|
| 4 | Felicidad | 76.35% |
| | Enojo | 14.83% |
| | Sorpresa | 1.12% |
| | Neutral | 7.7% |
| 5 | Felicidad | 92.59% |
| | Enojo | 5.16% |
| | Sorpresa | 1.85% |
| | Neutral | 0.4% |

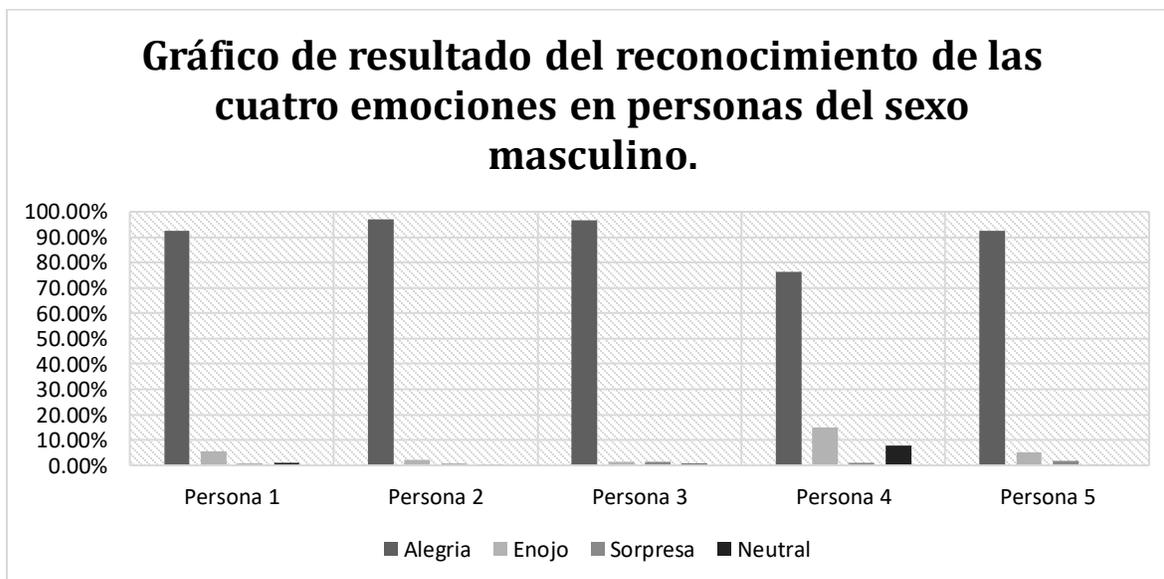


Figura 5.18 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.

La figura 5.19, muestra los resultados de la clasificación con las personas de sexo femenino mediante un gráfico de barras de la emoción de alegría con el clasificador a simple vista se puede ver que los resultados fueron favorables por lo que en la tabla 5.14 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.20 se representa por medio de un gráfico.



Figura 5.19 Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de Felicidad.

Tabla 5.14 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de alegría

| Persona | Emoción | Porcentaje |
|---------|------------------|---------------|
| 1 | Felicidad | 96.86% |
| | Enojo | 1.03% |
| | Sorpresa | 1.40% |
| | Neutral | 0.63% |
| 2 | Felicidad | 88.76% |
| | Enojo | 1.9% |
| | Sorpresa | 8.64% |
| | Neutral | 0.71% |
| 3 | Felicidad | 94.39% |
| | Enojo | 1.38% |
| | Sorpresa | 3.73% |
| | Neutral | 0.5% |
| 4 | Felicidad | 89.62% |
| | Enojo | 4.99% |
| | Sorpresa | 2.2% |
| | Neutral | 3.19% |
| 5 | Felicidad | 94.56% |
| | Enojo | 3.75% |
| | Sorpresa | 1.31% |
| | Neutral | 0.39% |

Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino.

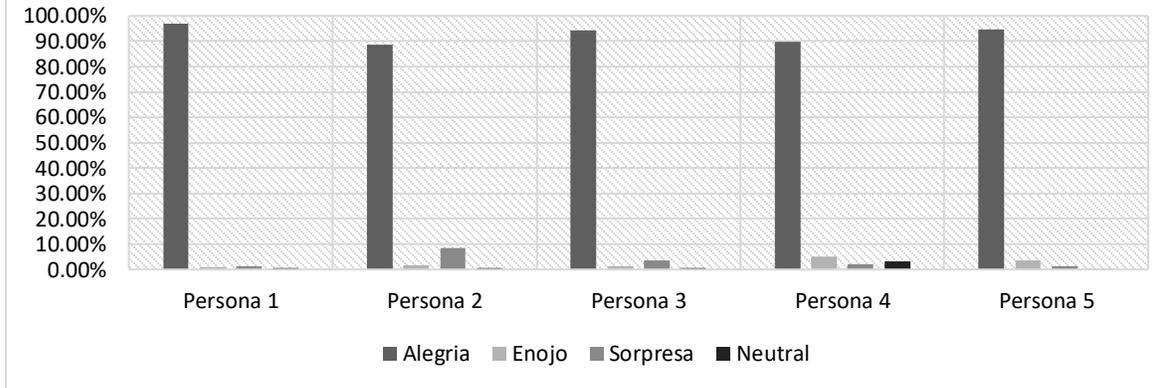


Figura 5.20 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino

La figura 5.21, muestra los resultados de la clasificación con las personas de sexo masculino mediante un gráfico de barras de la emoción de enojo con el clasificador a simple vista se puede ver que los resultados fueron favorables por lo que en la tabla 5.15 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.22 se representa por medio de un gráfico.

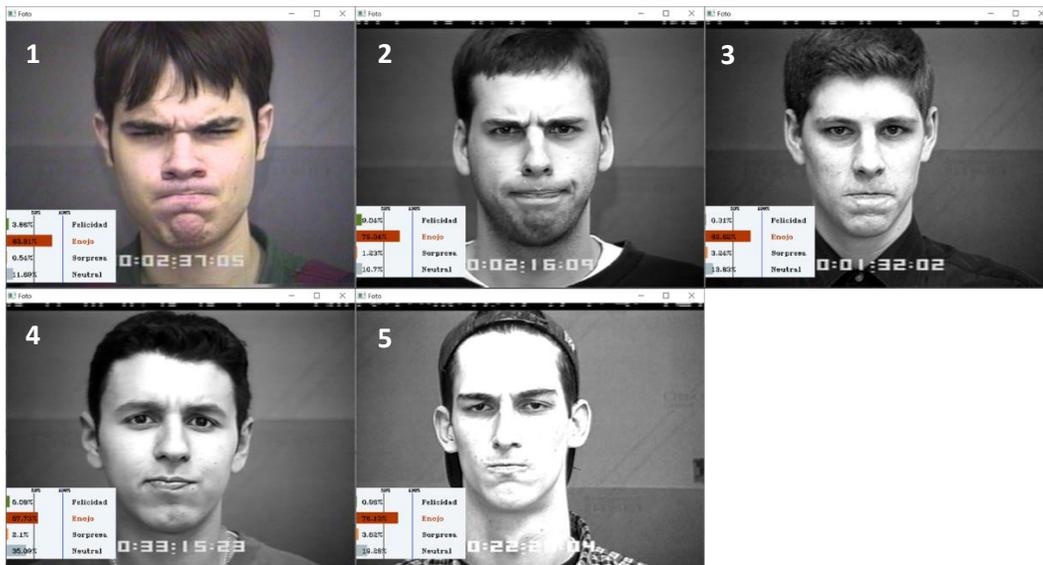


Figura 5.21 Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Enojo.

Tabla 5.15 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de enojo

| Persona | Emoción | Porcentaje |
|---------|--------------|---------------|
| 1 | Felicidad | 3.86% |
| | Enojo | 83.91% |
| | Sorpresa | 0.54% |
| | Neutral | 11.69% |
| 2 | Felicidad | 9.04% |
| | Enojo | 79.04% |
| | Sorpresa | 1.23% |
| 3 | Felicidad | 0.31% |
| | Enojo | 82.62% |
| | Sorpresa | 3.24% |
| 4 | Felicidad | 5.09% |
| | Enojo | 57.73% |
| | Sorpresa | 2.1% |
| 5 | Felicidad | 0.96% |
| | Enojo | 76.13% |
| | Sorpresa | 3.62% |
| | Neutral | 19.26% |

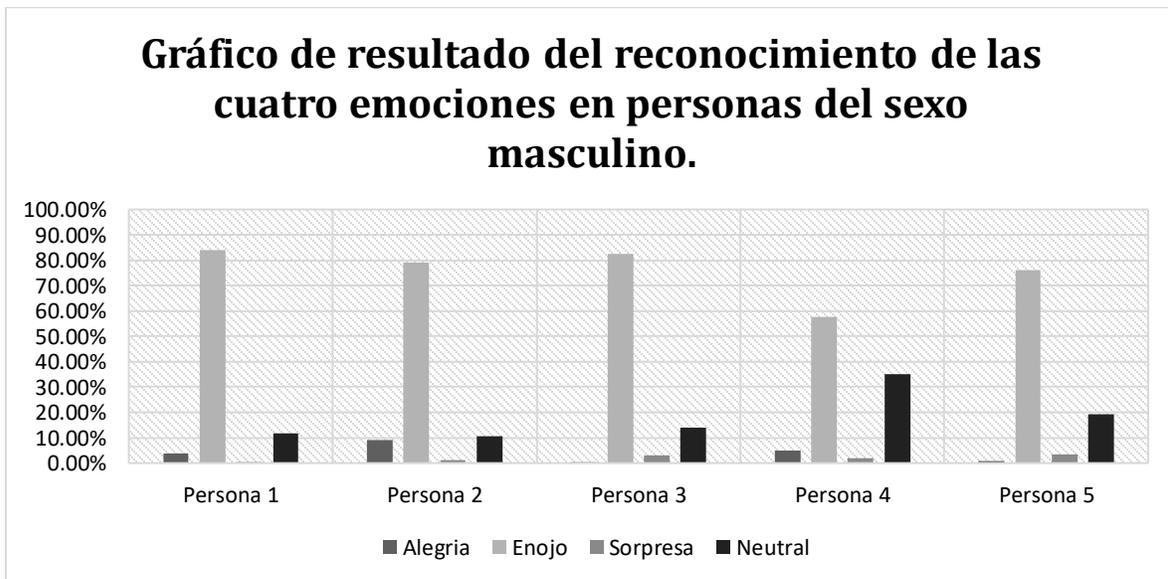


Figura 5.22 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.

La figura 5.23, muestra los resultados de la clasificación con las personas de sexo femenino mediante un gráfico de barras de la emoción de enojo con el clasificador a simple vista se puede ver que los resultados fueron aceptables, sin embargo, en el caso de la persona 3 su expresión fue clasificada como neutral esto sucede cuando la persona no tiene tan pronunciada la emoción que

se quiere expresar por lo que en la tabla 5.16 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.24 se representa por medio de un gráfico.



Figura 5.23 Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de Enojo.

Tabla 5.16 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de enojo.

| Persona | Emoción | Porcentaje |
|---------|--------------|---------------|
| 1 | Felicidad | 5.12% |
| | Enojo | 62.7% |
| | Sorpresa | 1.31% |
| | Neutral | 30.87% |
| 2 | Felicidad | 0.09% |
| | Enojo | 92.06% |
| | Sorpresa | 2.1% |
| 3 | Felicidad | 3.98% |
| | Enojo | 12.69% |
| | Sorpresa | 6.63% |
| 4 | Felicidad | 0.83% |
| | Enojo | 74.28% |
| | Sorpresa | 1.25% |
| 5 | Felicidad | 0.49% |
| | Enojo | 71.32% |
| | Sorpresa | 27.09% |
| | Neutral | 1.1% |

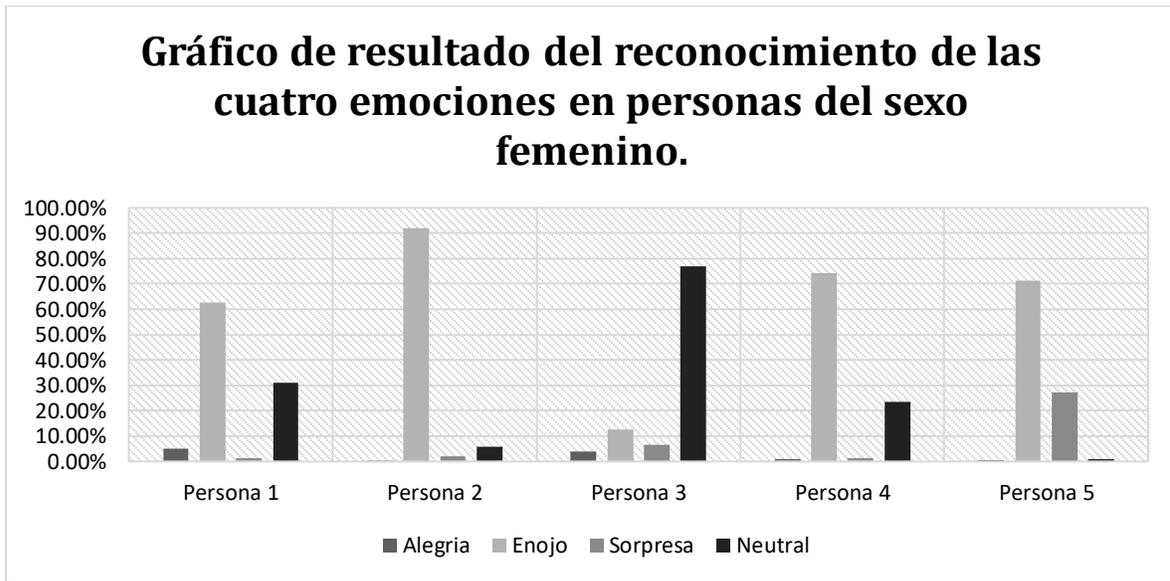


Figura 5.24 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino.

La figura 5.25, muestra los resultados de la clasificación con las personas de sexo masculino mediante un gráfico de barras de la emoción de sorpresa con el clasificador a simple vista se puede ver que los resultados fueron favorables, pero en la prueba de la persona 4 en su expresión emoción presento características similares a la emoción neutral, también podemos observar que la persona 5 alcanzo un 99.09% de reconocimiento ya que las características de la emoción están bien pronunciadas en su rostro, por lo que en la tabla 5.17 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.26 se representa por medio de un gráfico.



Figura 5.25 Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Sorpresa.

Tabla 5.17 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de sorpresa.

| Persona | Emoción | Porcentaje |
|---------|-----------------|---------------|
| 1 | Felicidad | 0.42% |
| | Enojo | 4.29% |
| | Sorpresa | 91.29% |
| | Neutral | 4.0% |
| 2 | Felicidad | 1.58% |
| | Enojo | 2.66% |
| | Sorpresa | 94.64% |
| | Neutral | 1.12% |
| 3 | Felicidad | 0.65% |
| | Enojo | 7.26% |
| | Sorpresa | 88.66% |
| | Neutral | 3.44% |
| 4 | Felicidad | 0.96% |
| | Enojo | 4.19% |
| | Sorpresa | 67.54% |
| | Neutral | 27.32% |
| 5 | Felicidad | 0.14% |
| | Enojo | 0.53% |
| | Sorpresa | 99.09% |
| | Neutral | 0.25% |

Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.

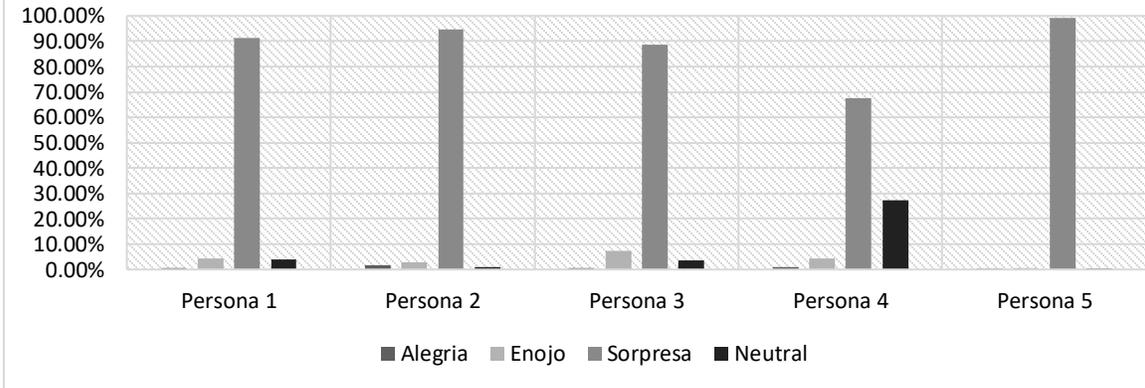


Figura 5.26 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.

La figura 5.27, muestra los resultados de la clasificación con las personas de sexo femenino mediante un gráfico de barras de la emoción de sorpresa con el clasificador a simple vista se puede ver que los resultados fueron favorables, pero en el caso de las personas 3 y 4 tuvieron una ligera confusión con otras emociones como enojo y neutral respectivamente como se puede ver cada persona tiene diferentes formas de expresar cada emoción, por lo que en la tabla 5.18 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.28 se representa por medio de un gráfico.



Figura 5.27 Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de sorpresa.

Tabla 5.18 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de sorpresa.

| Persona | Emoción | Porcentaje |
|---------|-----------------|---------------|
| 1 | Felicidad | 0.17% |
| | Enojo | 0.83% |
| | Sorpresa | 98.75% |
| | Neutral | 0.25% |
| 2 | Felicidad | 0.21% |
| | Enojo | 0.9% |
| | Sorpresa | 98.84% |
| | Neutral | 0.06% |
| 3 | Felicidad | 1.33% |
| | Enojo | 18.23% |
| | Sorpresa | 73.65% |
| | Neutral | 6.79% |
| 4 | Felicidad | 1.16% |
| | Enojo | 1.36% |
| | Sorpresa | 83.25% |
| | Neutral | 14.23% |
| 5 | Felicidad | 0.53% |
| | Enojo | 7.31% |
| | Sorpresa | 90.41% |
| | Neutral | 1.75% |

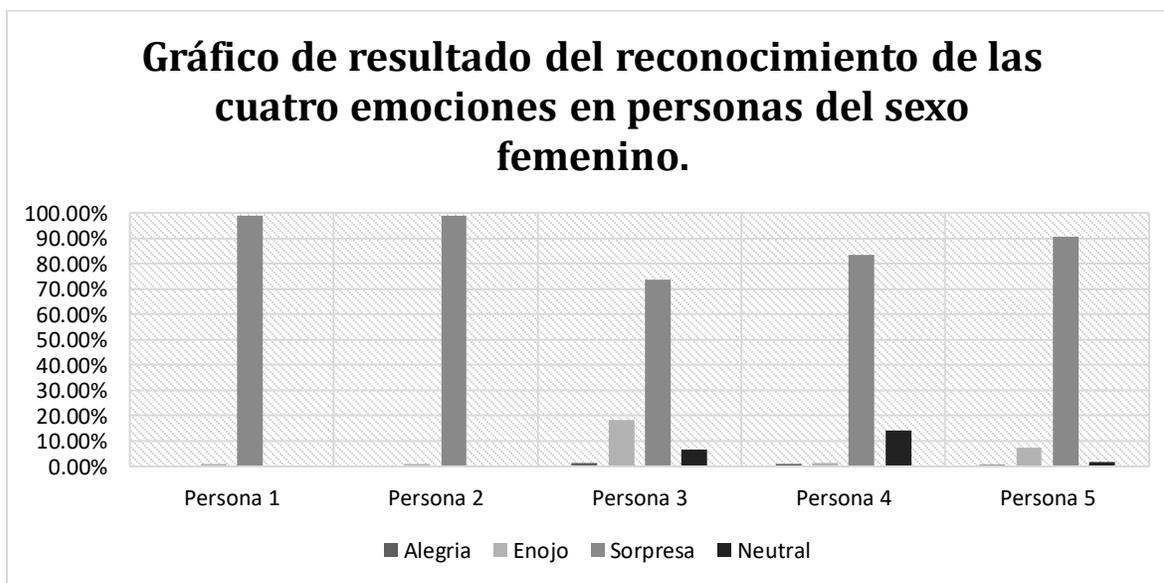


Figura 5.28 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino.

La figura 5.29, muestra los resultados de la clasificación de la emoción neutra de las personas de sexo masculino mediante un gráfico de barras, como se mencionó en la sección de evaluación de los algoritmos el clasificador obtuvo un porcentaje de precisión baja en la identificación de dicha emoción por lo que se ve reflejado en estas pruebas, por otra parte el porcentaje supera el

50% por lo que los resultados son aceptables, en la tabla 5.19 se muestran los resultados obtenidos de esta prueba mediante porcentajes y en la figura 5.30 se representa por medio de un gráfico.



Figura 5.29 Pruebas de reconocimiento de emociones con personas del sexo masculino de la emoción de Neutral.

Tabla 5.19 Porcentajes de reconocimiento de las emociones en personas del sexo masculino con la expresión facial de neutral.

| Persona | Emoción | Porcentaje |
|---------|----------------|---------------|
| 1 | Felicidad | 1.52% |
| | Enojo | 24.38% |
| | Sorpresa | 4.45% |
| | Neutral | 69.65% |
| 2 | Felicidad | 0.57% |
| | Enojo | 17.18% |
| | Sorpresa | 12.92% |
| | Neutral | 69.34% |
| 3 | Felicidad | 3.32% |
| | Enojo | 34.86% |
| | Sorpresa | 3.91% |
| | Neutral | 57.92% |
| 4 | Felicidad | 1.03% |
| | Enojo | 13.34% |
| | Sorpresa | 7.89% |
| | Neutral | 77.73% |
| 5 | Felicidad | 4.07% |
| | Enojo | 13.15% |
| | Sorpresa | 7.49% |
| | Neutral | 75.29% |

Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.

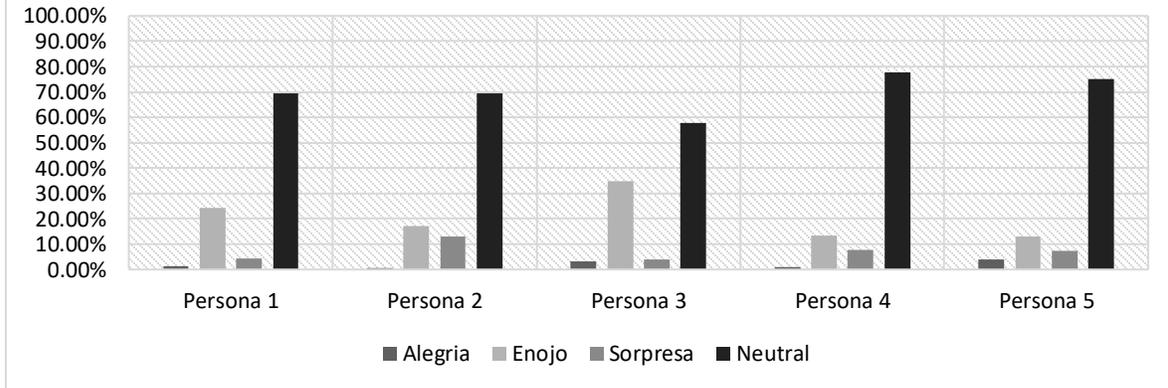


Figura 5.30 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo masculino.

La figura 5.31, muestra los resultados de la clasificación con las personas de sexo femenino mediante un gráfico de barras de la emoción de neutral con el clasificador, como se puede ver en las imágenes en el caso de la persona 3 fue clasificada como enojada, como se les mencionó en la evaluación de los algoritmos la emoción de enojo y neutral tienden a confundirse en ocasiones ya que algunas personas expresan el enojo poniéndose serios como si estuvieran neutros, por lo que en la tabla 5.20 se muestran los resultados obtenidos en el reconocimiento de la emoción de esta prueba mediante porcentajes y en la figura 5.32 se representa por medio de un gráfico.



Figura 5.31 Pruebas de reconocimiento de emociones con personas del sexo Femenino de la emoción de Neutral.

Tabla 5.20 Porcentajes de reconocimiento de las emociones en personas del sexo femenino con la expresión facial de neutral.

| Persona | Emoción | Porcentaje |
|---------|----------------|---------------|
| 1 | Felicidad | 6.22% |
| | Enojo | 10.92% |
| | Sorpresa | 4.93% |
| | Neutral | 77.93% |
| 2 | Felicidad | 3.87% |
| | Enojo | 21.16% |
| | Sorpresa | 5.04% |
| | Neutral | 69.93% |
| 3 | Felicidad | 2.43% |
| | Enojo | 51.66% |
| | Sorpresa | 2.05% |
| | Neutral | 43.85% |
| 4 | Felicidad | 6.52% |
| | Enojo | 2.85% |
| | Sorpresa | 8.32% |
| | Neutral | 82.31% |
| 5 | Felicidad | 3.39% |
| | Enojo | 6.22% |
| | Sorpresa | 4.37% |
| | Neutral | 86.02% |

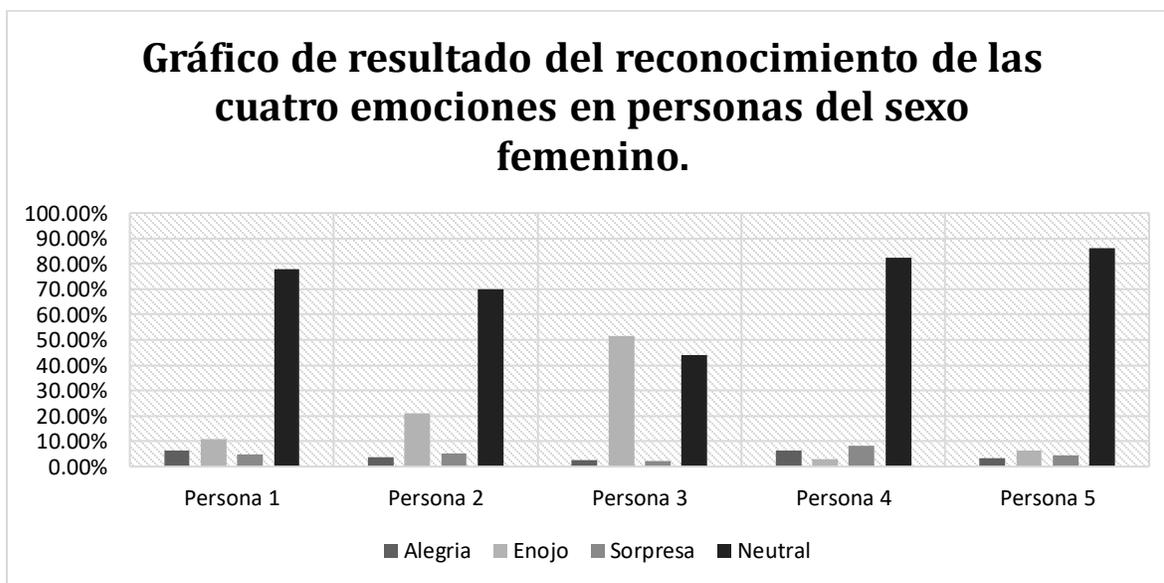


Figura 5.32 Gráfico de resultado del reconocimiento de las cuatro emociones en personas del sexo femenino.

Como se pudo ver en las pruebas el clasificador pudo realizar una clasificación aceptable, los resultados indican que para la clasificación el sexo de la persona no es un factor que afecte en el análisis, también se puede observar que el mayor problema es la pose del rostro y la falta de

expresión de algunas personas que hacen que el clasificador se confunda en ciertas emociones que comparten similitud de características como la emoción neutra y enojo.

Por otro lado, las emociones mejor clasificadas son la felicidad y la sorpresa ya que estas dos emociones tienen características muy específicas que las hacen diferenciarse de las demás emociones, por ejemplo, para una persona alegre siempre tiende a contraer las comisuras de la boca y en ocasiones a tener cierta apertura de ella y para la emoción de sorpresa generalmente se define por abrir los párpados superior e inferior, tener una notable apertura de la boca y las cejas elevadas por tal motivo estas dos emociones son las que mejor pudo clasificar el algoritmo. La tabla 5.21 y la figura 5.33, muestran el promedio de reconocimiento para cada emoción en las pruebas realizadas.

Tabla 5.21 Promedio de reconocimiento de cada emoción de los resultados de la prueba

| Emoción | Personas | Sexo | | Promedio de reconocimiento |
|-----------|----------|------|---|----------------------------|
| | | H | M | |
| Felicidad | 10 | 5 | 5 | 91.94% |
| Enojo | 10 | 5 | 5 | 69.25% |
| Sorpresa | 10 | 5 | 5 | 88.61% |
| Neutral | 10 | 5 | 5 | 71.00% |

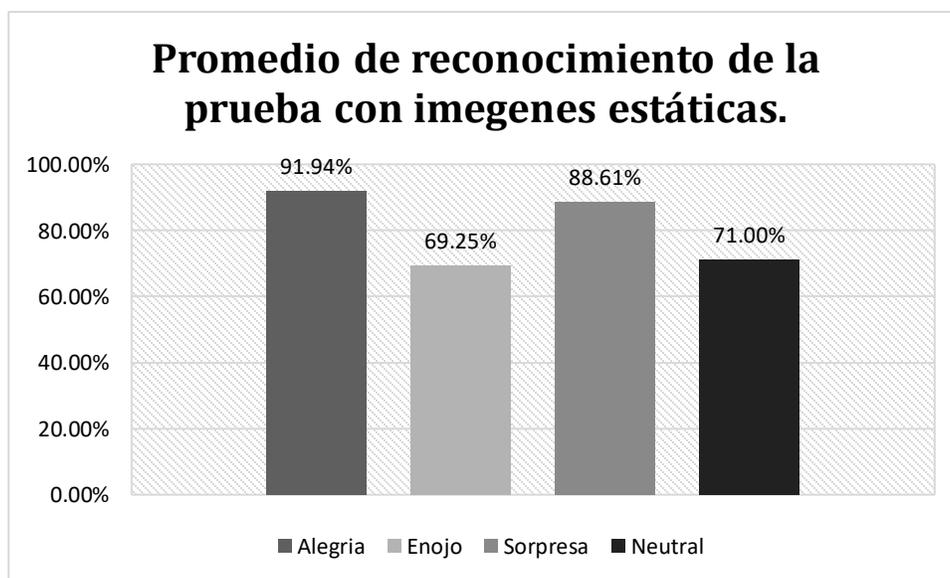


Figura 5.33 Grafico de promedio de reconocimiento de cada emoción resultado de las pruebas

5.1.3.2.6 Pruebas con videos

Se realizaron pruebas con el procesamiento de video implementando el algoritmo de Máquina de Soporte Vectorial el cual dio el mejor resultado en las pruebas anteriores, obteniendo un buen porcentaje de reconocimiento para cada emoción los resultados de clasificación en un video con el porcentaje de probabilidad de reconocimiento de cada emoción.

Para la realización de las pruebas se tomaron como datos de entrada videos del rostro de 6 personas que fueron participantes de una evaluación de la experiencia de usuario generados por la plataforma UXLab, los cuales durante la sesión estuvieron interactuando con diferentes productos digitales (páginas web) y también les fueron inducidas las emociones mediante imágenes generadas por Sistema Internacional de Imágenes Afectivas (IAPS por sus siglas en inglés) Lang, P.J, et al(2008) las cuales provocan en la persona cierta carga emocional estos videos fueron grabados de noviembre a diciembre del 2020.

La figura 5.34, muestra los resultados en video obtenidos para la emoción de felicidad el gráfico indica el porcentaje de reconocimiento.

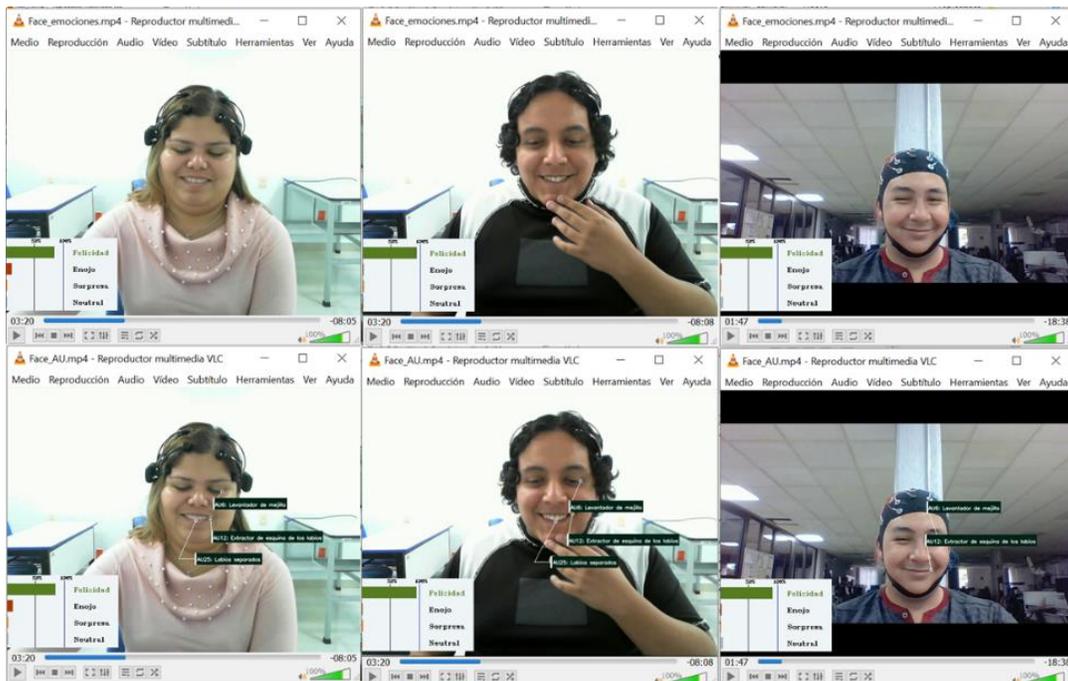


Figura 5.34 Emoción de Felicidad y sus Unidades de Acción.

La figura 5.35, muestra los resultados en video obtenidos para la emoción de enojo el gráfico indica porcentaje de reconocimiento.

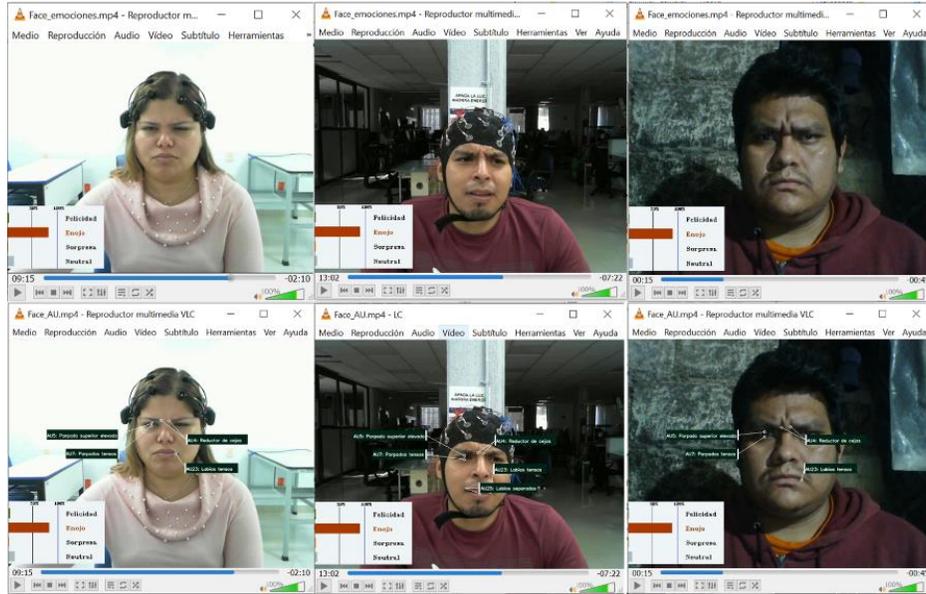


Figura 5.35 Emoción de Enojo y sus Unidades de Acción.

Como en los videos no se hizo la inducción de la emoción de sorpresa y los participantes no expresaron esa emoción a lo largo de la sesión se decidió realizar la prueba con videos de dos personas en donde si presentaban la emoción de sorpresa con la finalidad de verificar si el clasificador realizaba de manera correcta la identificación de esta emoción.

La figura 5.36, muestra los resultados en video obtenidos para la emoción de Sorpresa el gráfico indica porcentaje de reconocimiento.

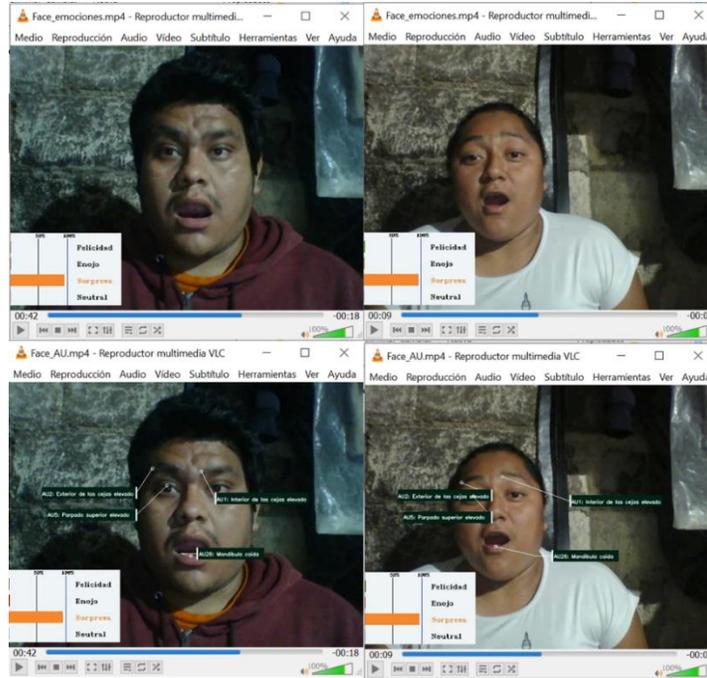


Figura 5.36 Emoción de Sorpresa y sus Unidades de Acción.

La figura 5.37, muestra los resultados en video obtenidos para la emoción de Neutral el gráfico indica porcentaje de reconocimiento

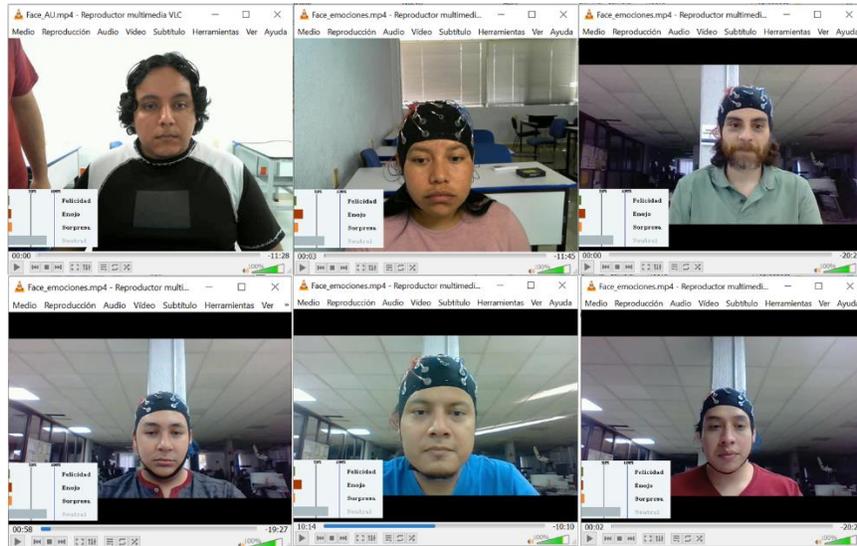


Figura 5.37 Secuencia de video detectando la emoción Neutra.

Los resultados obtenidos en las pruebas con videos, el clasificador se desempeñó de manera adecuada clasificando las emociones con un buen porcentaje para cada emoción en donde la emoción de felicidad el reconocimiento es del 95%, para la emoción de enojo el porcentaje de reconocimiento está entre es del 98%, para le emoción de sorpresa el porcentaje de reconocimiento es del 91% y para la emoción neutra es del 88% aunque en las pruebas se tuvo

buenos resultados el principal problema para poder clasificar mucho mejor cada emoción es la pose del rostro de la persona ya que en video es más notorio porque la persona todo el tiempo de la sesión se encuentra en constante movimiento por eso el clasificador tiene una variación en el porcentaje de reconocimiento, la Tabla 5.22 muestra el porcentaje de reconocimiento que alcanza el clasificador a la hora de identificar las emociones.

Tabla 5.22 Porcentaje de reconocimiento de emociones en video.

| Emoción | Precisión |
|------------------|------------------|
| Felicidad | 95.43% |
| Enojo | 98.6 % |
| Sorpresa | 91.46% |
| Neutral | 88.00% |

5.1.3.2.7 Pruebas de integración en UXLab.

Como parte de la experimentación se realizaron las pruebas de integración del algoritmo de análisis facial a la plataforma UXLab, durante el proceso de integración se presentaron algunos problemas que se solucionaron sobre la marcha, con las pruebas realizadas se determinó que el algoritmo no provoca ninguna inconsistencia con los demás módulos de UXLab por lo que el módulo de análisis facial para la clasificación de emociones queda concluido.

UXLab tiene varios módulos de procesamiento que se pueden ejecutar individualmente mediante un script se manda a ejecutar cada módulo que se requiera, para procesar la información, el clasificador fue integrado al script para poder realizar el análisis facial de emociones de manera individual permitiéndole al usuario realizar el proceso en el momento que lo requiera en las figuras 5.38 y 5.39 se muestra la ejecución del script para el procesamiento de seguimiento ocular y el de análisis facial.

```
C:\WINDOWS\system32\cmd.exe - PostProcessScript.py -d C:\uxlab\recordings\USR1\2020-11-17\002 -p eyetracking face
C:\>cd uxlab
C:\uxlab>PostProcessScript.py -d C:\uxlab\recordings\USR1\2020-11-17\002 -p eyetracking face
.....: Comienza procesamiento de eyetracking ::::::...
Buscando video de pantalla...
Se encontro video de la pantalla
Ejecutando...
C:\uxlab\recordings\USR1\2020-11-17\002
//Video_display.mp4
--Generando video de puntos--
test: C:\uxlab\recordings\USR1\2020-11-17\002//Video_display.mp4
--Video de puntos terminado--
--Fin--
--Generando mapa de calor--
Moviepy - Building video C:\uxlab\recordings\USR1\2020-11-17\002\VideoFull_Heatmap.mp4.
Moviepy - Writing video C:\uxlab\recordings\USR1\2020-11-17\002\VideoFull_Heatmap.mp4
t: 1%|█          | 411/29122 [01:37<2:35:05, 3.09it/s, now=None]
```

Figura 5.38 Modulo de Pos procesamiento de datos de seguimiento ocular de UXLab

```
C:\WINDOWS\system32\cmd.exe
.....: Inicia el procesamiento de los Datos ::::::...
Reconocimiento de Emociones y Unidades de Accion...
Buscando video de WebCam...
Se encontro video de la WebCam
Iniciando Analisis de Unidades de Accion...
Por favor espere mientras termina el analisis...
Analisisi Finalizado
Iniciando el proceso de Clasificacion de Emociones...
Por favor espere mientras termina de realizarse la clasificacion...
Analisisi Finalizado
C:\uxlab>
```

Figura 5.39 Modulo de pos-procesamiento de datos de análisis facial de UXLab

Una vez que el proceso termina son generados los videos con el análisis facial que contendrá la clasificación de las emociones para posteriormente poder visualizarlo en el módulo de reproducción de la plataforma como se puede ver en la figura 5.40.

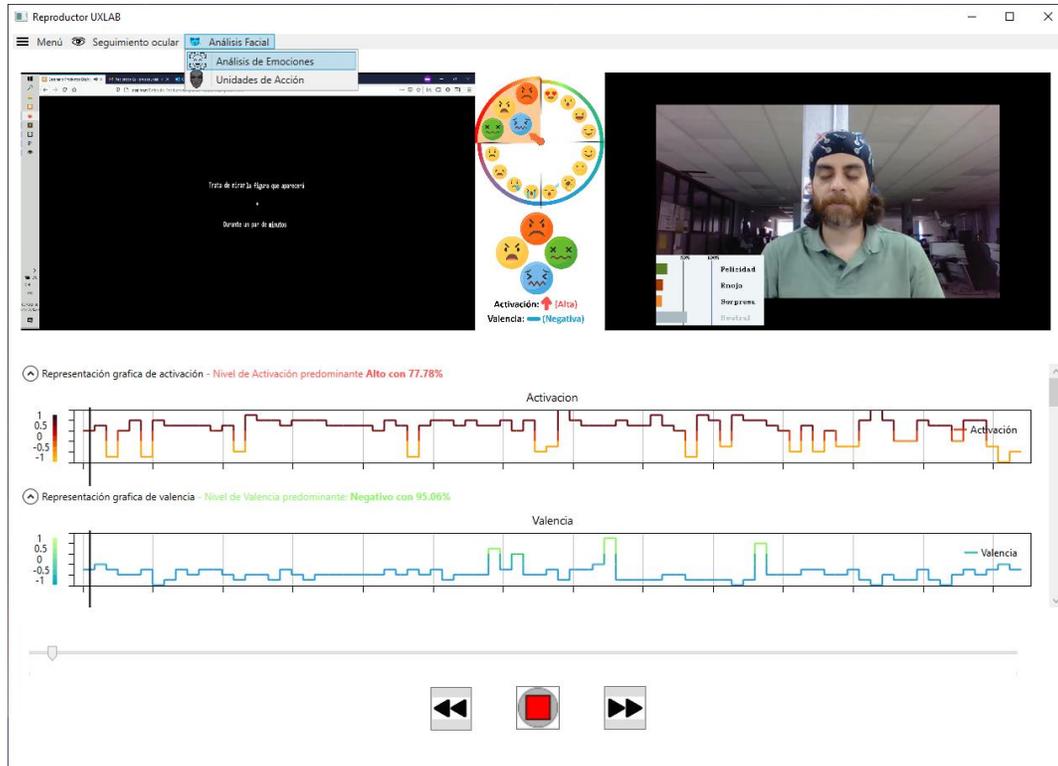


Figura 5.40 Integración del Módulo de Análisis Facial al Reproductor de UXLab, visualización de los módulos en conjunto

Una vez realizadas las pruebas de integración y con la finalidad de demostrar que esta investigación será de gran utilidad en la evaluación de la experiencia de usuario se realizaron pruebas en conjunto con todos los datos que se generan en una sesión de interacción con un producto digital durante la evaluación de la UX, en donde podemos observar como el módulo de análisis facial complementa con información importante la evaluación para que el evaluador pueda realizar un mejor análisis sobre algún producto.

Como se puede ver en la figura 5.41, en esta sesión de evaluación los datos de señales fisiológicas clasifican el estado del usuario con una activación alta y una valencia negativa esto nos quiere decir que la persona está presentando un estado de ansiedad, pero el análisis de emociones nos indica que el usuario expresa la emoción de felicidad esto es razonable ya que es el comienzo de la evaluación con ello podemos observar que el análisis de emociones complementa a la información que UXLab le proporciona al evaluador.

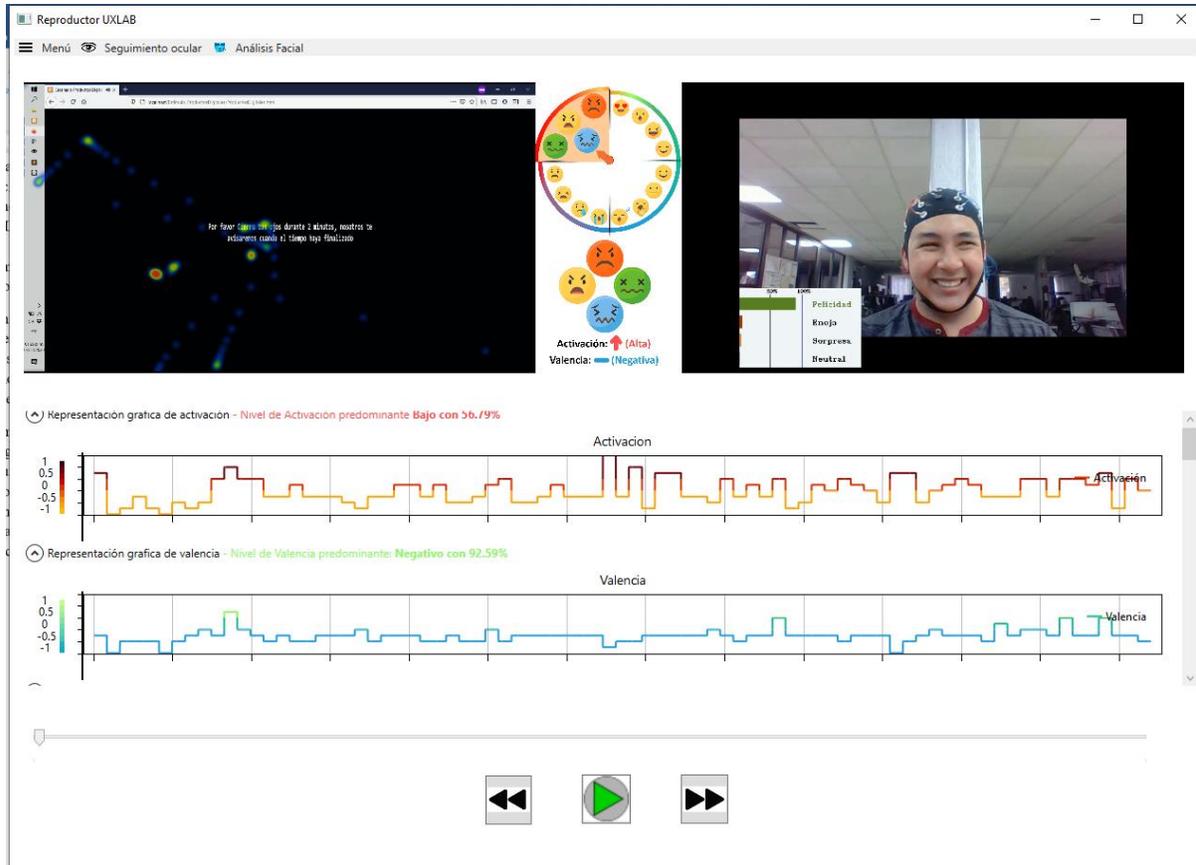


Figura 5.41 Visualización del Inicio de una sesión de evaluación de la UX en la plataforma UXLab para su análisis

En la figura 5.42, podemos ver otro caso de como el análisis facial de emociones realizado en esta investigación complementa la información para que el evaluador de la UX pueda realizar un mejor análisis de la evaluación, en este caso los datos fisiológicos indican en el modelo valencia – activación que la persona pasa por un estado de tristeza por lo que para el evaluador seria confuso poder interpretar de manera correcta esta información pero el módulo de análisis facial nos indica que la persona expresa la emoción de felicidad dándole más facilidad al evaluador de realizar un mejor análisis.



Figura 5.42 Visualización del Inicio de los datos de una sesión de evaluación de la UX en la plataforma UXLab para su análisis

La figura 5.43 y 5.44 se muestra la transición de la emoción neutra a la emoción de felicidad del participante que durante la evaluación presenta un estado de alegría de acuerdo al indicador del modelo valencia – activación, con esto el evaluador puede saber que el usuario tuvo una buena experiencia en ese lapso de interacción con el producto ya que la emoción de neutral también puede traducirse como concentración y segundos después se expresa esa satisfacción con una sonrisa en donde el análisis facial lo clasifica como alegría.



Figura 5.43 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro neutral



Figura 5.44 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro alegre

Como se puede ver en la figura 5.45, en esta sesión de evaluación de otro usuario los datos de señales fisiológicas clasifican el estado del usuario con una activación alta y una valencia negativa esto nos quiere decir que la persona está presentando un estado de ansiedad, pero el análisis de emociones nos indica que el usuario expresa la emoción de felicidad esto es razonable ya que es el comienzo de la evaluación, minutos después como se ve en la figura 5.46 podemos observar que el estado del usuario pasa a alegre y también se puede observar que está concentrado ya que presenta un rostro neutral y fijo con ello podemos determinar que el análisis de emociones complementa a la información que UXLab le proporciona al evaluador.



Figura 5.45 Visualización del Inicio de una sesión de evaluación de la UX en la plataforma UXLab para su análisis



Figura 5.46 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro neutral

La figura 5.47 y 5.48 nos muestran la transición de la emoción neutra a la emoción de felicidad del participante que durante la evaluación presenta un estado de ansiedad y segundos después pasa a un estado de alegría de acuerdo al indicador del modelo valencia – activación, con esto el evaluador puede saber que el usuario tuvo una buena experiencia en ese lapso de interacción con el producto ya que la emoción de neutral también puede traducirse como concentración y la emoción de alegría se expresa como una buena satisfacción por medio de una sonrisa, con esta prueba podemos determinar que el algoritmo implementado en esta investigación realmente aporta información que complementa a los otros análisis para que un evaluador pueda realizar un buen análisis de la evaluación de la experiencia de usuario de algún producto digital.



Figura 5.47 Visualización del estado de ansiedad del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro alegre

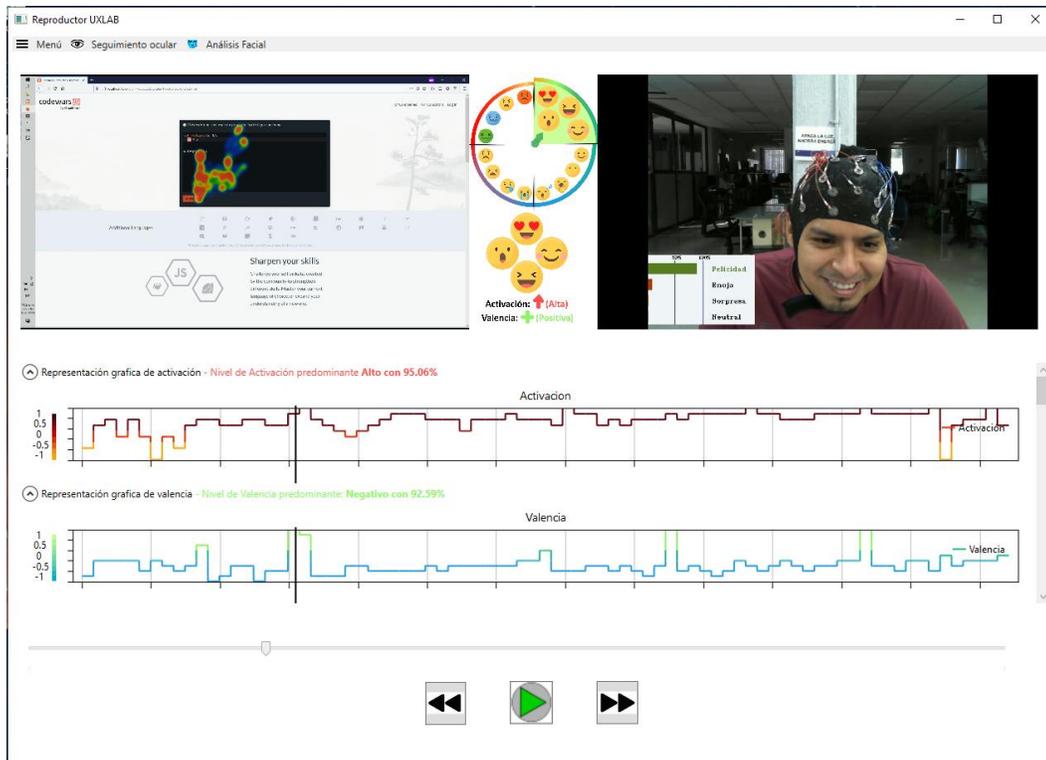


Figura 5.48 Visualización del estado de alegría del usuario durante una sesión de evaluación de la UX en la plataforma UXLab donde el usuario presenta un rostro alegre.

5.1.4 Resultados.

Se realizó una comparación con el trabajo de M. Álvarez y el realizado, M. Álvarez realizó una comparación de diferentes clasificadores utilizando un método geométrico para la extracción de características utilizando 68 puntos de referencia facial, con el trabajo que se realizó en esta investigación se superó los resultados del autor, utilizando Máquina de Soporte Vectorial y 17 puntos de referencia facial como se puede ver en la tabla 5.23 y 5.24 y en la figura 5.49 en forma de gráfica.

Tabla 5.23 Resultados obtenidos por M. Alvares et al (2018) .

| Algoritmos | Precisión | Puntos Faciales |
|-----------------------|-----------|-----------------|
| Naive Bayes | 52.13 % | 68 |
| SVM | 58.97 % | 68 |
| Decision Tree | 52.99 % | 68 |
| Random Forest | 59.83 % | 68 |
| AdaBoost | 36.75 % | 68 |
| Multilayer Perceptron | 67.52 % | 68 |

Tabla 5.24 Resultados obtenidos en la investigación.

| Algoritmos | Precisión | Puntos Faciales |
|---------------|-----------|-----------------|
| SVM | 84.52% | 17 |
| Random Forest | 82.14% | 17 |
| KNN | 84.52% | 17 |

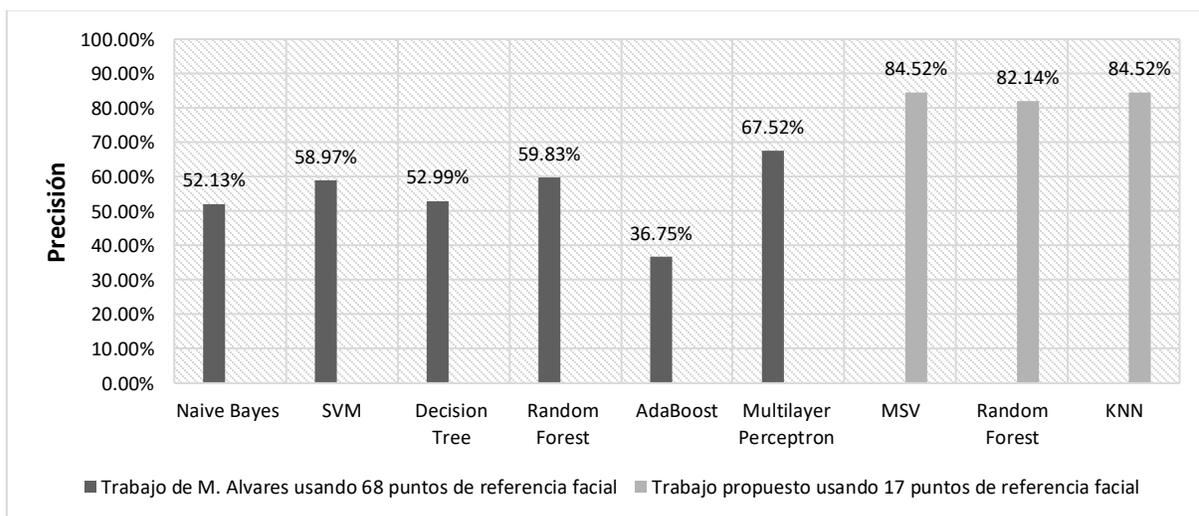


Figura 5.49 Gráfico de comparación del trabajo de M. Alvares y el propuesto.

Capítulo 6

Conclusión y trabajos futuros.

6.1 Conclusión

En esta investigación se presentó el desarrollo de un sistema de reconocimiento de emociones, mediante el análisis de expresiones faciales tomando algunos puntos de referencia del rostro. El reconocimiento de emociones mediante el análisis facial con base geométrica y un clasificador de aprendizaje automático como lo es Máquina de Soporte Vectorial como se demostró da buenos resultados y el costo computacional no es alto. El sistema se evaluó con dos conjuntos de imágenes y presenta un buen desempeño. Sin embargo, aún hay trabajo para mejorar la precisión de reconocimiento que es del 84.52 %, para mejorar estos resultados se puede trabajar con la mejora del procesamiento de las imágenes, aplicando filtros de iluminación y contraste, realizando una alineación del rostro, obtener imágenes con una mayor nitidez y que los rostros tengan más pronunciadas las expresiones para cada emoción y aumentar el número de imágenes para cada clase. La relación que se hizo con los puntos de referencia facial y el sistema de codificación de acción facial de emociones para la extracción de características permitió que la clasificación fuera menos costosa computacionalmente y nos acerca más a la realidad en la clasificación de las emociones por computadora ya que de esta forma es como un humano identifica las emociones.

Por otra parte, también fue posible demostrar que el análisis facial para la clasificación de emociones durante la evaluación de la experiencia de usuario proporciona al evaluador de la UX información valiosa con respecto a lo que el usuario siente durante la interacción con el producto digital ya que en conjunto con los datos multimodales (EEG, ECG, GSR y Eyetracking) se puede realizar un mejor análisis de cada evaluación y cuando la clasificación de los datos fisiológicos y seguimiento ocular no proporcionan información clara el análisis de emociones podrá complementar la información ya que por medio de las expresiones del rostro del usuario se

puede notar fácilmente lo que el usuario va experimentando durante la interacción con el producto digital.

En el proceso de una evaluación de la UX, el evaluador es una parte fundamental ya que es el encargado de realizar el análisis de todos los datos que se obtuvieron del usuario durante la interacción con el producto digital, en UXLab una vez que el usuario termine la evaluación la grabación tanto de la pantalla de la computadora como del rostro y los datos fisiológicos se almacenan y se procesan mediante algoritmos de aprendizaje automático y profundo generando nueva información que le servirá al evaluador para realizar el análisis de manera más fácil, el evaluador utiliza la información generada ya procesada y los visualiza en el reproductor de la plataforma UXLab en ella se visualiza de manera ordenada cada una de las secciones que le corresponde a cada información en ese momento es donde el evaluador empieza a realizar el análisis visualizando y relacionando cada información con lo que el usuario interactuó durante la evaluación, es decir el evaluador va visualizando el video de la interacción del usuario con el producto digital y a la par visualiza la información procesada de los datos multimodales (EEG, ECG, GSR y Eyetracking) y de análisis facial para poder determinar las emociones y satisfacción que el usuario presento en cada momento con la interacción durante la evaluación como se ve en la figura 6.1, aquí el evaluador tiene que visualizar la sección de video de la pantalla y con el control de reproducción ir deteniendo el proceso en el tiempo que se desee analizar cómo podemos ver el usuario estaba interactuando con una página web en donde tenía que rellenar un formulario podemos ver el mapa de calor que nos indica la intensidad de interés que el usuario está poniendo en cada parte del formulario entre más rojo sea más interés le presta el usuario a cada componente, el evaluador podrá determinar fácilmente que el usuario sintió una buena satisfacción al rellenar el formulario ya que el indicador de valencia-activación nos indica que el usuario está presentando alegría teniendo una valencia alta y una activación positiva, por otro lado los datos fisiológicos nos indican la emoción que el usuario está presentando, otra información muy importante y que fue la que se desarrolló en esta investigación es el análisis facial de emociones el cual comprueba y complementa el análisis y nos indica que en efecto el usuario está feliz o alegre realizando esa actividad con el producto digital demostrándolo por medio de su expresión que es una sonrisa y dándonos la información de que emoción está presentando en ese momento el usuario con este análisis el evaluador puede determinar que el módulo de este producto digital tuvo una buena satisfacción para el usuario y que le fue fácil utilizarlo e interactuar con el producto con este proceso el evaluador podrá determinar que secciones del producto se tienen que mejorar o en su caso eliminar para que el producto no salga con deficiencias y haya pérdidas monetarias.

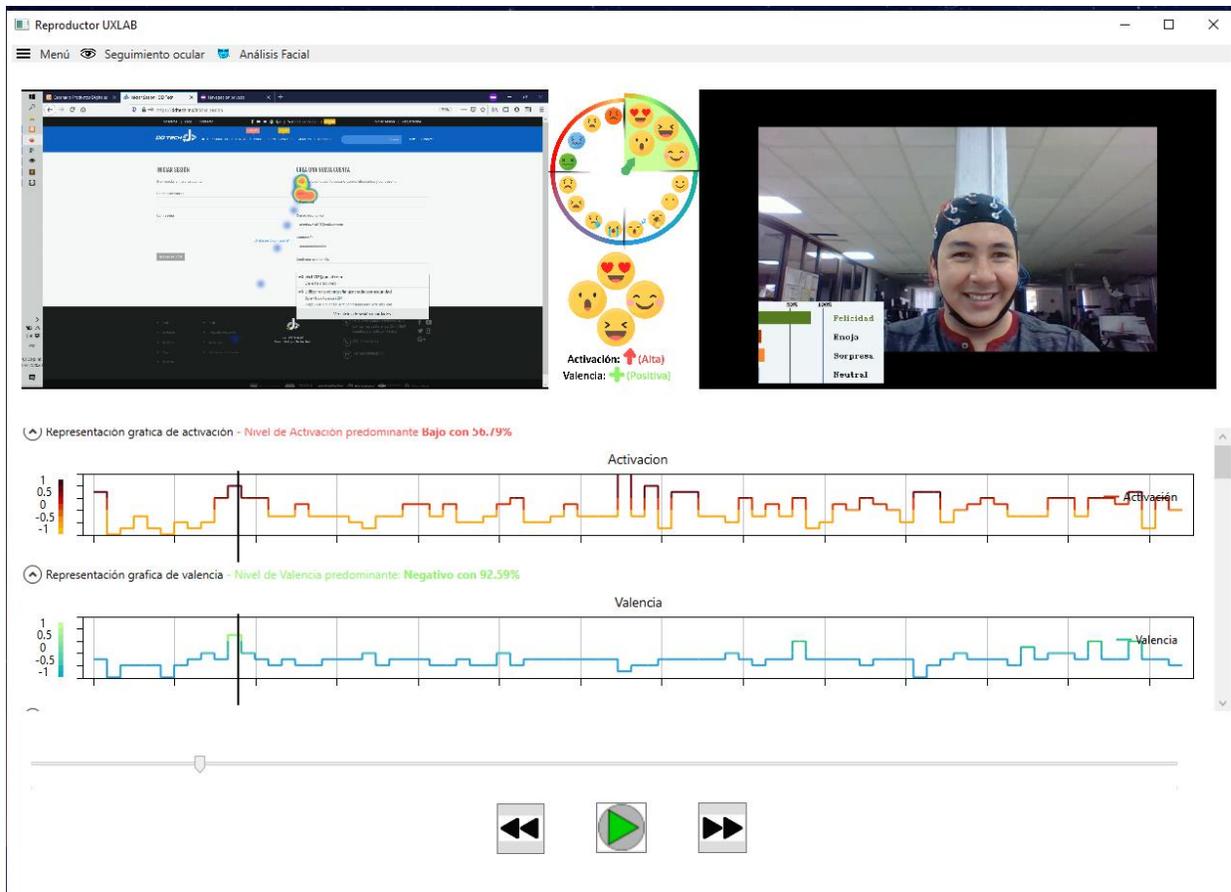


Figura 6.1 Visualización de los datos procesados en el grabador de UXLab de una sesión de evaluación UX.

6.2 Trabajos futuros

Por otra parte, también se pretendía identificar tres estados cognitivos, pero por las limitaciones de tiempo y falta de repositorios de imágenes de estos estados mentales no fue posible realizar dichas pruebas, por lo que se propone para abordar este tema como trabajo futuro, la propuesta es la identificación de las unidades de acción que Ekman propone para la detección del interés y aburrimiento, para la distracción se propone utilizar las unidades de acción que corresponden al movimiento de la cabeza y el movimiento ocular, el método para la extracción de características y el algoritmo de clasificación sería el mismo que se usó para la detección de emociones realizándole las adecuaciones necesarias para las nuevas características.

El sistema de codificación de acción facial desarrollado por Ekman Paul Ekman, et al. (2002). es extenso ya que no solo incluye las unidades de acción para las emociones, sino que también se encuentran otras unidades de acción dentro de ellas se encuentran las unidades de acción de los movimientos de la cabeza y los ojos que serán tomados para identificar la distracción y también se encuentran las que intervienen en la identificación del interés y

aburrimiento, a continuación se presentan las unidades de acción que se deben de tomar en cuenta para poder identificar los estados de interés, aburrimiento y distracción.

6.2.1 Unidades de acción para la detección de Interés.

AU 2. Levantador de cejas externo: Contribuye a las emociones sorpresa y miedo, y al interés de la actitud afectiva. Frontalis (pars lateralis) es el músculo facial subyacente ver figura 5.50. *Noldus (2021)*



Figura 6.2 Unidad de Acción 2 Levantador de cejas externo Noldus (2021)

AU 1. Levantador de cejas interior: Contribuye a las emociones tristeza, sorpresa y miedo, y al interés de la actitud afectiva. Base muscular: frontalis (pars medialis) ver figura 5.51. *Noldus (2021)*

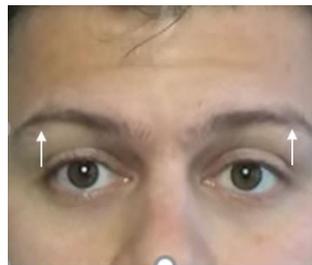


Figura 6.3 AU 1. Levantador de cejas interior Noldus (2021)

AU 5. Elevador de tapa superior: Contribuye a la sorpresa, el miedo y la ira, y al interés. Base muscular: levator palpebrae superioris y músculo tarsal superior ver figura 5.52. *Noldus (2021)*

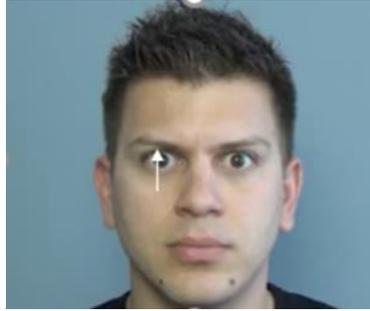


Figura 6.4 AU 5. Elevador de tapa superior Noldus (2021)

AU 17. Mentonera: Esta Unidad de Acción contribuye al interés y confusión de las actitudes afectivas. El músculo facial subyacente es mentalis ver figura 5.53. *Noldus (2021)*

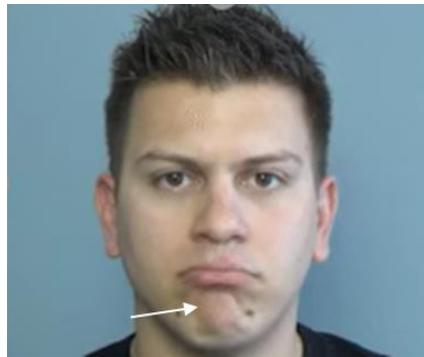


Figura 6.5 AU 17. Mentonera Noldus (2021)

6.2.2 Unidades de acción para la detección del aburrimiento.

AU 23. Tensor de labios: Contribuye a la emoción de la ira, a la confusión y al aburrimiento de las actitudes afectivas. Base muscular: orbicularis oris ver figura 5.54. *Noldus (2021)*



Figura 6.6 AU 23. Tensor de labios Noldus (2021)

AU 24. Labios presionados: Esta Unidad de Acción contribuye al aburrimiento de la actitud afectiva. El músculo facial subyacente es orbicularis oris ver figura 5.55. *Noldus (2021)*



Figura 6.7 AU 24. Labios presionados Noldus (2021)

AU 43. Ojos cerrados: Contribuye a la actitud afectiva el aburrimiento. La base muscular consiste en la relajación del Levator palpebrae superioris ver figura 5.56. *Noldus (2021)*

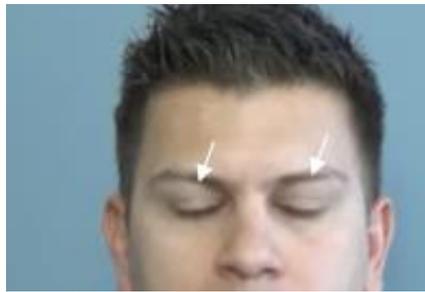


Figura 6.8 AU 43. Ojos cerrados Noldus (2021)

6.2.3 Unidades de acción para la detección de la distracción.

Para la identificación del estado de distracción se tomará como características las unidades de acción que corresponden al movimiento de la cabeza y a las unidades de acción que corresponden al movimiento de los ojos, la tabla 5.25 y 5.26 muestran estas unidades de acción. *P.E Group (2016)*

Tabla 6.1 Códigos de unidades de acción de la cabeza P.E Group (2016)

| Unidad de acción | Nombre de acción |
|------------------|---|
| 51 | Girar cabeza a la izquierda |
| 52 | Girar cabeza a la derecha |
| 53 | Alzar la cabeza |
| 54 | Bajar la cabeza |
| 55 | Inclinación de la cabeza hacia la derecha |
| M55 | Inclinación de la cabeza a la izquierda |
| 56 | Leve inclinación |

Tabla 6.2 Códigos de unidades de acción de los ojos P.E Group (2016)

| Unidad de acción | Nombre de acción |
|------------------|-------------------------------|
| 61 | Mover ojos hacia la izquierda |
| 62 | Mover ojos a la derecha |
| 63 | Ojos hacia arriba |
| 64 | Ojos hacia abajo |

Con estas unidades de acción se propone realizar la extracción de características con base en el calcula de distancias geométricas como se realizó con la clasificación de las emociones, lo que faltaría sería un repositorio de imágenes para cada estado cognitivo y realizar el procesamiento y el entrenamiento del algoritmo.

6.3 Aportaciones

6.3.1 Procedimiento para la extracción de características con base geométrica.

Para poder extraer las características del rostro en esta investigación se desarrolló un procedimiento para procesar las imágenes e ir extrayendo los puntos de referencias necesarios para poder obtener un conjunto de datos con el que se entrenó el algoritmo propuesto, en este procedimiento se realizó la detección del rostro en una imagen y por medio de la localización de puntos de referencia esto tomando como referencia las unidades de acción que intervienen en cada una de las emociones a clasificar las cuales se describen en el sistema de codificación de acción facial para emociones que Ekman desarrollo en su investigación Paul Ekman, et al. (2002).

de este conjunto de puntos se realizó el cálculo de distancias entre ellos mediante la distancia euclidiana, teniendo como resultado un conjunto de datos con 17 distancias que fueron las que describieron mejor cada emoción, el esquema de seguimiento de este procedimiento se puede observar en la primera parte de la figura 5.12 del capítulo 5 y el pseudocódigo se puede observar en el algoritmo 2 del mismo capítulo.

6.3.2 Desarrollo e implementación del módulo de análisis facial para emociones en la plataforma UXLab.

El módulo de reconocimiento emocional mediante el análisis facial implementado en UXLab es una de las mejoras más importantes en la herramienta. Permite utilizar las grabaciones del rostro del usuario para clasificar emociones mediante las expresiones que el usuario presenta durante una evaluación de la UX. Los resultados son fáciles de interpretar ya que se presenta un

gráfico donde indica la emoción que se está detectando en ese momento y se puede ver el porcentaje de detección de cada emoción, por otro lado también se puede visualizar las unidades de acción que intervienen encada una de las emociones que se detectan durante la interacción del usuario con el producto dando información importante al evaluador de la UX para tener un mejor análisis de la evaluación en conjunto con los demás datos que UXLab ya procesa.

6.4 Publicaciones

Durante la elaboración de esta investigación se realizaron tres publicaciones las cuales fueron temas sobre las etapas de obtención de los resultados a continuación se mencionan los temas y los congresos en los que se presentaron.

- Artículo titulado “*RECONOCIMIENTO DE EMOCIONES MEDIANTE EL ANÁLISIS DE EXPRESIONES FACIALES*”, presentado para la Escuela de Inteligencia Computacional y Robótica 2021, llevada a cabo del 21 al 25 de junio del 2021.
- Artículo titulado "*RECONOCIMIENTO DE EMOCIONES MEDIANTE EL ANÁLISIS DE EXPRESIONES FACIALES*", presentado en el marco de la 7a jornada de ciencia y tecnología aplicada, celebrado del 17 al 19 de noviembre de 2021, en el TECNM/CENIDET, publicado en el Vol. 4, Núm. 2, Julio - diciembre 2021 de la revista de la Jornada de Ciencia y Tecnología Aplicada con ISBN en trámite, pag. 62-67.
- Artículo titulado “Recognition of emotions by analyzing facial expressions from user experience evaluation videos”, presented in 2nd International Conference on advancing Knowledge from Multidisciplinary Perspectives in Engineering & Technology (ICAKMPET-2021) organized by Institute for Engineering Research and Publication (IFERP) held on 22nd & 23rd December 2021. Plicated in International Journal of Mechanical Engineering Vol. 7 No. 1 January, 2022, With ISBN: 0974-5823, page: 6325 – 6333.

7 Referencias y bibliografía

7.1 Referencias

- Sunitha A, P. Ajay Kumar Reddy, S.Nanda Kishore³, G.N Kodanda Ramaiah (2017), *Recognition of Facial Emotions Based on Sparse Coding*. *Journal of Engineering Research and Application*, ISSN: 2248-9622.
- Yang, D, Alsadoon, A., P.W.C, Prasad, Singh, A, K, Elchouemi, A. (2017, Dic), *International Conference on Smart Computing and Communications*. Kurukshetra, India. *An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment Networks*. <https://doi.org/10.1016/j.procs.2017.12.003>
- Nwosu, L, Wang, H., Lu, J, Unwala,I, Yang, Zhang, T.(2017, Septiembre). *deep convolutional neural network for facial expression recognition using facial parts*. DOI 10.1109/DASC-PICom-DataCom-CyberSciTec.2017.213
- Krestinskaya, et al. (2017, septiembre 13-16). *Conferencia Internacional sobre Avances en Informática, Comunicaciones e Informática (ICACCI) Udupi, India, Facial Emotion Recognition using Min-Max Similarity Classifier*. **IEEE**. 10.1109/ICACCI.2017.8125932
- MD. Munim, et al. (2017, Dicember 7-9). *3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, Towards Developing a Tool for UX Evaluation Using Facial Expression*. **IEEE**. 10.1109 / EICT.2017.8275227
- Vázquez Rodríguez, C. A. (2017). *Valoración automática de la motivación y atención del estudiante mediante expresiones faciales (tesis de maestría)*. Morelos, México: CENIDET.
- Liu Xiao, Lee Kiju. (2018, August 15-17). *IEEE Games, Entertainment, Media Conference (GEM), Galway, Irlanda, Optimized Facial Emotion Recognition Technique for Assessing User Experience*.**IEEE**, 10.1109 / GEM.2018.8516518
- M. Alvarez V, et al. (2018, August 22-24). *Conferencia Internacional sobre Investigación en Inteligencia y Computación en Ingeniería (RICE), San Salvador, El Salvador, Facial Emotion Recognition: A Comparison of Different Landmark-Based Classifiers*. **IEEE**. 10.1109 / ARROZ.2018.8509048
- Dewi, Y.L, Widyanio, M., Basaruddin, T. (2018). *Geometric facial components feature extraction for facial expression recognition*.
- Mangal, Divya., Prajwala.(2018, Abril 3-5). *International Conference on Communication and Signal Processing, Facial expression recognition by calculating euclidian distance for eigen faces using pca*.
- Kartali, et al. (2018, Nobiembre 20-21).*14th Symposium on Neural Network and Aplication (NEUREL), Belgrade, serbia real-time algorithms for facial emotion recognition: a comparison of different approaches*.**IEEE**.

- Rabhi Yassine, et al. (2018, March 21-24). *4a Conferencia internacional sobre tecnologías avanzadas para el procesamiento de señales e imágenes (ATSIP)*, Susa, Túnez, *A Real-time Emotion Recognition System for disabled persons*. **IEEE**. 10.1109 / ATSIP.2018.8364339
- Verma monu, et al. (2018, October 7-10). *IEEE International Conference on Systems, Man, and Cybernetics, Miyazaki, Japón, Japón QUEST:Quadrilateral Senary bit Pattern for Facial Expression Recognition*. **IEEE**. 0.1109 / SMC.2018.00260
- Candra Kirana Karetika, et al. (2018). *International Seminar on Application for Technology of Information and Communication (iSemantic), Malang, Indonesia Facial Emotion Recognition based on Viola-Jones Algorithm in the Learning Environment*. **IEEE**.
- H. S Gunavathi, M. Siddappa, et al. (2018, May 18-19). *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT-2018), Bangalore, India, India Towards Cognitive State Detection using Facial Expression and Hand-over-Face Gesture*. **IEEE**. 10.1109/ RTEICT42901.2018.9012547
- Gama Velasco, A. K. (2018). *Desarrollo de un Sistema interactivo basado en visión artificial para la rehabilitación del tobillo (tesis de maestría)*. Morelos, México: CENIDET.
- Yi, J, Chen, A., Cai, Z, Sima, Y, Zhou, Wu, Xingyu .(2019, Mayo). *Facial expression recognition of intercepted video sequences based on feature point movement trend and feature block texture variation*. <https://doi.org/10.1016/j.asoc.2019.105540>
- Valderrama Cárdenas, W. (2019). *Reconocimiento automático del rostro para verificación de identidad para evaluación en línea (tesis de maestría)*. Morelos, México: CENIDET.
- Cadena Moreano, J.A, La Serna Palomino, N., Llano Casa, A, C. (2019, September). *A Facial recognition technique using SVM: A comparative analysis*. e-ISSN: 1390-6542 / p-ISSN: 1390-9363.
- Sawhney Shreyak, et al. (2019, January 10-11). *IX Conferencia Internacional sobre Cloud Computing, Data Science & Engineering (Confluence), Noida, India, India Real-Time Smart Attendance System using Face Recognition Techniques*. **IEEE**. 10.1109/ CONFLUENCE.2019.8776934
- Lacort, J. (21 agosto 2017). *Las claves de los sistemas de reconocimiento facial: ¿cuál es su verdadero nivel de seguridad?* Recuperado el 2 de junio del 2020 de <https://www.xataka.com/seguridad/las-claves-de-los-sistemas-de-reconocimiento-facial-cual-es-su-verdadero-nivel-de-seguri>
- Burns, E. (2021, September 6). *Aprendizaje profundo (deep learning)*. ComputerWeekly.es; TechTarget. <https://www.computerweekly.com/es/definicion/Aprendizaje-profundo-deep-learning>
- Shu L, Xie J, Yang M, Li Z, Li Z, Liao D, Xu X, Yang X. *A Review of Emotion Recognition Using Physiological Signals*. *Sensors (Basel)*. 2018 Jun 28;18(7):2074. doi: 10.3390/s18072074
- Behera Ardhendu, et al (2020, abril), *International Journal of Artificial Intelligence in Education. Associating Facial Expressions and Upper-Body Gestures with Learning Tasks for Enhancing Intelligent Tutoring Systems*, <https://doi.org/10.1007/s40593-020-00195-2>
- Manterola, C., et al. (2013). *Revisiones sistemáticas de la literatura. Qué se debe saber acerca de ellas*. *Cirugía Española*, 91(3), 149-155. <https://doi.org/10.1016/j.ciresp.2011.07.009>

- López Caballero, V. (2018). *Revisión sistemática de la literatura del impacto del Elearning en el lugar de trabajo*. CENIDET, Morelos, México.
- D. King, "dlib C++ Library," 2015. [Online]. Available: [Www.Dlib.Net](http://www.Dlib.Net).
- Noldus (2021) *What Facial Action Units look like*, <https://www.noldus.com/applications/facial-action-coding-system>
- Jeovanny Soriano Terrazas (2018). *Metodología para caracterizar e inducir estados mentales a través de realidad virtual inmersiva e interfaz cerebro computadora (Tesis de maestría)*. CENIDET, Morelos, México.
- Joaquín A.R. (2021) *Máquinas de Vector Soporte (SVM) con Python available under a Attribution 4.0 International (CC BY 4.0) at* <https://www.cienciadedatos.net/documentos/py24-svm-python.html>
- Camacho, J. A. (2020, 28 marzo). *Support Vector Regression (SVR)*. JacobSoft. https://www.jacobsoft.com.mx/es_mx/support-vector-regression/
- Joaquín A.R. (2020) *Random Forest python. (s. f.)*. https://www.cienciadedatos.net/documentos/py08_random_forest_python.html
- Díaz, R. (2022, 18 noviembre). *Algoritmo KNN – cómo funciona y ejemplos en Python. The Machine Learners*. <https://www.themachinelearners.com/algoritmo-knn/>
- Jain Udit, et al. (2018, mayo), *2do Congreso Internacional de Tendencias en Electrónica e Informática (ICOEI)*. Tirunelveli, India. *Analysis of Face Detection and Recognition Algorithms using Viola Jones Algorithm with PCA and LDA*. 10.1109 / ICOEI.2018.8553811
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). *The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010)*, San Francisco, USA, 94-101.
- Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba. *Coding Facial Expressions with Gabor Wavelets*, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998). <http://doi.org/10.1109/AFGR.1998.670949> Open access content available at: <https://zenodo.org/record/3430156>
- P.E Group " *Facial Action Coding System - Paul Ekman Group, LLC.*," 2016 [Online]. Available: <https://www.paulekman.com/facial-action-coding-system/>
- Paul Ekman, et al. (2002). *Facial Action Coding System The Manual*. EEUA. Research Nexus division of Network Information Research Corporation. ISBN: 0-931835-01-1
- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System (FACS) [Database record]*. APA PsycTests. <https://doi.org/10.1037/t27734-000>
- Rosebrock, A. (2021, 3 julio). *Facial landmarks with dlib, OpenCV, and Python*. PyImageSearch. <https://pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/>
- Mallick, S. (2021, 5 mayo). *Facial Landmark Detection | LearnOpenCV #*. LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow with examples and tutorials. <https://learnopencv.com/facial-landmark-detection/>

- Sambare, M. (2020, Julio 19). FER-2013 Learn facial expressions from an image. Version 1. <https://www.kaggle.com/msambare/fer2013/metadata>
- Sharma S., et al. (2016, mayo 25-27). Conferencia Internacional sobre Tecnologías Avanzadas de Computación y Control de Comunicaciones (ICACCCT). Ramanathapuram, India, FAREC - CNN Based Efficient Face Recognition Technique using Dlib. IEEE. 10.1109 / ICACCCT.2016.7831628
- ¿Qué es el aprendizaje autónomo? (n.d.). Oracle.com. Retrieved February 14, 2022, from <https://www.oracle.com/mx/data-science/machine-learning/what-is-machine-learning/>

7.2 Bibliografías

- Fouilloux Quiroz, D. E. (2018). Método para integrar y sincronización de datos EEG y multimedia para su aplicación en un ambiente de evaluación para la experiencia del usuario (tesis de maestría). Morelos, México: CENIDET.
- García Pinzón, G. A. (2018). Sistema embebido para analizar datos biométricos e identificar estados cognitivos y emocionales en procesos de evaluación de la experiencia del usuario (tesis de maestría). Morelos, México: CENIDET.
- Lagunés Ramírez, D. A. (2020). Algoritmo para analizar el movimiento ocular del usuario para identificar la correlación con estados mentales (tesis de maestría en desarrollo). Morelos, México: CENIDET.
- Morales Morante, J. A. (2018). Máquina de aprendizaje para el reconocimiento de emociones mediante datos fisiológicos multimodales (tesis de maestría en desarrollo). Morelos, México: CENIDET
- Hussain, J., Ali Khan, W., Hur, T., Bilal, H., Bang, J., Hassan, A. . . . Lee, S. (2018, Mayo 18). A Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation. *Sensors*, 18(5). doi:10.3390/s18051622.
- Michael D. Robinson, Edward R. Watkins, Eddie Harmon-Jones. *Handbook of cognition and emotion*, p. cm. Includes bibliographical references and index. ISBN 978-1-4625-0999-7
- James Garrett, Jesse (2011) *The Elements of User Experience: User-Centered Design for the Web and Beyond*, Second Edition, ISBN 13: 978-0-321-68368-7