



MÓDULO PARA LA DETECCIÓN DE ESTADOS EMOCIONALES A TRAVÉS DE EXPRESIONES FACIALES EN EL AULA TECNOLÓGICA XAUCE ATCNEA

MODULE FOR THE DETECTION OF EMOTIONAL STATES THROUGH FACIAL EXPRESSIONS IN THE TECHNOLOGICAL CLASSROOM XAUCE ATCNEA

Ing. Glenda Piñero Muñoz

gmunoz@uci.cu

Universidad de las Ciencias Informáticas, Cuba

Ing. Jesús Abraham Kindelán Gómez

jakindelan@uci.cu

Universidad de las Ciencias Informáticas, Cuba

Ing. William Simón Ramírez

wsimon@uci.cu

Universidad de las Ciencias Informáticas, Cuba

Resumen

Esta investigación está enmarcada en el proceso de desarrollo de un módulo de reconocimiento de expresiones faciales capaz de reconocer estados emocionales tales como felicidad, ira, tristeza, disgusto, sorpresa, miedo y estado neutral en una audiencia determinada. Dicha investigación persigue el objetivo de brindar al profesor que dirige el proceso una herramienta que le permita estudiar la actitud de sus alumnos frente a las actividades docentes que se llevan a cabo para así determinar la toma de acciones con vistas a incidir en el aspecto motivacional de los estudiantes. Con el resultado de dicho proyecto el aula tecnológica XAUCE ATcnea cuenta con un módulo para la detección de estados emocionales a través de expresiones faciales en imágenes fijas y frames de video. Para llevar a cabo la investigación propuesta se empleó una serie de técnicas e instrumentos de recolección de datos, específicamente el análisis de fuentes documentales. Para el cumplimiento del objetivo planteado se hace uso de la metodología de desarrollo de software Variación de AUP para la UCI.

Palabras clave: clasificación de emociones; computación afectiva; detección de objetos; extracción de características; K-NN; reconocimiento de expresiones faciales

Abstract

This research is framed in the process of developing a module for recognizing facial expressions capable of recognizing emotional states such as happiness, anger, sadness, disgust, surprise, fear and neutral state in a given audience. This research pursues the objective of providing the teacher who directs the process with a tool that allows him to study the attitude of his students in front of the teaching activities that are carried out in order to determine the taking of actions with a view to influencing the motivational aspect of the students. With the result of this research, the XAUCE ATcnea technological classroom will have a module for the detection of emotional states through facial expressions in still images and video frames. To carry out the proposed research, a series of techniques and data collection instruments were used, specifically the analysis of documentary sources. For the fulfillment of the proposed objective, the methodology of software development is used Variation of AUP for the UCI.

Keywords: classification of emotions; recognition of facial expressions; selection of features; extraction of characteristics; affective computation; K-NN

1. Introducción

“Durante muchos años, la investigación académica en relación con el comportamiento humano ha dejado de lado las emociones, centrándose en el pensamiento racional. Sin embargo, frente a la

contraposición tradicional entre emoción e intelecto, las teorías más recientes indican la influencia de las emociones en los mecanismos del pensamiento racional, ya que influyen en la percepción y en tareas diarias como el aprendizaje, la comunicación o la toma de decisiones”. (Baldasari, 2016)

“La emoción que presenta un individuo determina la acción que es posible que realice. Esto es claro para las emociones más fuertes y las conductas que las caracterizan. Un sujeto enojado no actúa racionalmente, la emoción lo gobierna y lo conduce por un camino que quizás no tomaría si esta emoción no lo dominara. Un individuo triste tampoco es una persona predispuesta a aprender cosas nuevas o a establecer contactos con otros, su emoción lo mantiene en un pozo que le impide generar conexiones con los demás. En resumen, si la emocionalidad define el dominio de acciones posibles de realizar, entonces se puede concluir que la emocionalidad constituye un aspecto de la mayor relevancia para facilitar los aprendizajes en educación. En una primera aproximación, se podría partir de la hipótesis de que las emociones positivas o gratas permitirán la realización de acciones favorables para el aprendizaje. En cambio, emociones negativas o no gratas no lo permitirán”.

“En los últimos años los dispositivos para la detección de emociones han evolucionado notablemente, permitiendo que los ordenadores sean capaces de reconocer el estado emocional de un usuario. A través de cámaras es posible capturar propiedades observables de las emociones, como el color de la piel, los movimientos corporales, gestos, o incluso la detección de expresiones faciales en base al análisis de movimientos musculares o de puntos característicos de la cara”.

En la Universidad de las Ciencias Informáticas (UCI), existe un conjunto de centros de desarrollo de software, entre ellos se encuentra el Centro de Tecnologías para la Formación (FORTES), que cuenta con un equipo de desarrolladores para la creación de aplicaciones que apoyen el aprendizaje de sus usuarios finales. Este centro presenta un software para el aula tecnológica (XAUCE ATcnea), disponible para la formación de estudiantes que sean capaces de interactuar con herramientas que han revolucionado el proceso de enseñanza-aprendizaje convencional, además de proporcionarle al profesor que está dirigiendo el proceso herramientas útiles que facilitan y enriquecen su trabajo. Teniendo en cuenta la utilidad de llevar a cabo un proceso de enseñanza-aprendizaje informatizado, con vistas a realizar un estudio de los estados emocionales presentes en los estudiantes que viven esta experiencia tecnológica. Resulta de interés la detección y registro de las emociones a través las expresiones faciales de los estudiantes en tiempo real, contribuyendo al estudio de las reacciones humanas ante un proceso educacional informatizado y a la determinación de cambios en la clase con vistas a incidir en el aspecto motivacional su audiencia.

2. Materiales y métodos

Se utilizó para la construcción y desarrollo del trabajo como métodos científicos los métodos teóricos: Analítico – sintético con el objetivo de analizar y utilizar la información y la documentación más relevante para el desarrollo y la realización de este módulo. El método histórico – lógico para constatar teóricamente y realizar un estudio de todo lo referente a la identificación de expresiones faciales, la evolución y el nivel de desarrollo que tienen los sistemas. Finalmente, el modelado para representar por medio de diagramas los conceptos asociados a los procesos de identificación de expresiones faciales, obteniendo como resultado un mejor entendimiento de la posible solución a implementar.

Como métodos empíricos se empleó la entrevista para obtener información interactuando con personas que tienen un mayor conocimiento sobre identificación de expresiones faciales y la encuesta para obtener información sobre el nivel de satisfacción del usuario con el producto final.

3. Resultados y discusión

Cuando se habla del término Detección de Objetos se hace referencia al uso de técnicas de Visión por Computador, dichas técnicas están basadas en la detección de objetos en imágenes y video.

El detector de objetos visual, es el sistema que se encarga de determinar si aparecen o no objetos determinados en una imagen. En caso afirmativo, el detector debe proporcionar información sobre ellos. Este tipo de sistemas se componen de dos módulos o fases: en primer lugar, la extracción de ciertas características de la imagen; y, en segundo lugar, la clasificación de las imágenes según algunas de estas características.

En la figura 1 se muestran las etapas esenciales que componen la Secuencia de procesamiento de imágenes que servirá de base a la propuesta realizada en el presente trabajo.



Figura 1. Secuencia de procesamiento de imágenes.

Las imágenes que se procesan pueden venir del espectro visible o de otras partes del espectro como

las Infrarrojas o las de Profundidad y aun así se pueden usar las mismas técnicas.

El sistema detector de objetos siempre tendrá una entrada y una salida, donde la entrada va a ser la imagen y un indicador del objeto que se desea detectar y la salida va a ser una ventana o un conjunto de ventanas con los objetos que se detectaron.

La selección del descriptor que se va a usar para representar el objeto que se busca en cualquier sistema, es una decisión clave para el buen funcionamiento del Sistema Detector, sabiendo que el descriptor está ligado al módulo de extracción de características. Los tipos de descriptores se pueden clasificar:

- Basado en características del píxel: Es un descriptor muy sencillo y muy sensible a pequeños cambios en la imagen, además no toma en cuenta la forma del objeto.
- Basado en toda la imagen: Usa la misma imagen como descriptor y es sensible a pequeños cambios en la imagen.
- Basado en características locales: Un descriptor local es un vector de características que es calculado sobre una pequeña región de interés de la imagen.

3.1 Descriptores de objetos en imágenes

Para abordar cualquier tarea de detección, se necesita una descripción adecuada del objeto a detectar. Para conseguir una descripción fiable se utilizan características que definan por completo al objeto. Estas características se agrupan en un vector típicamente conocido como descriptor de características. Hay que tener en cuenta que diferentes características serán adecuadas para diferentes aplicaciones. La clave está en hacer una elección adecuada de las características: Serán mejores aquellas que acentúen las diferencias entre clases y que eliminen las diferencias dentro de una misma clase (2).

Existe una variedad de descriptores de objetos en imágenes, entre los más utilizados se encuentran: El Descriptor de Patrones Locales Binarios (LBP), el Descriptor de Histograma Orientado a Gradientes (HOG) y el Descriptor Haar, el cual es el seleccionado para el desarrollo del módulo propuesto.

3.1.1 Descriptor Haar

Un filtro Haar se puede definir como la diferencia de intensidad de zonas adyacentes de la imagen con diferentes configuraciones que permiten detectar cambios de intensidad en diferentes orientaciones. Los filtros se aplican a múltiples escalas y posiciones en la imagen, de manera que se obtienen un número muy elevado de características (López, Valveny. y

Vanrell, 2019)

Las características de Haar se obtienen como resultado de aplicar por toda la imagen un conjunto de filtros de Haar. (López, Valveny. y Vanrell, 2019)

Los *Filtros Haar* son (Figura 2):

- Una combinación de rectángulos del mismo tamaño adyacentes vertical u horizontalmente.
- Los rectángulos en negro representan zonas con una contribución positiva al filtro.
- Los rectángulos en blanco representan zonas con una contribución negativa al filtro.
- Por tanto, el resultado del filtro va a ser la diferencia en la suma de los valores de los píxeles entre zonas en negro y zonas en blanco.



Figura 2. Filtros de Haar básicos. Fuente: (López, Valveny. y Vanrell, 2019)

Ejemplo de ello se ilustra en la figura 3 donde se muestran los diferentes resultados después de aplicar los filtros básicos de Haar:

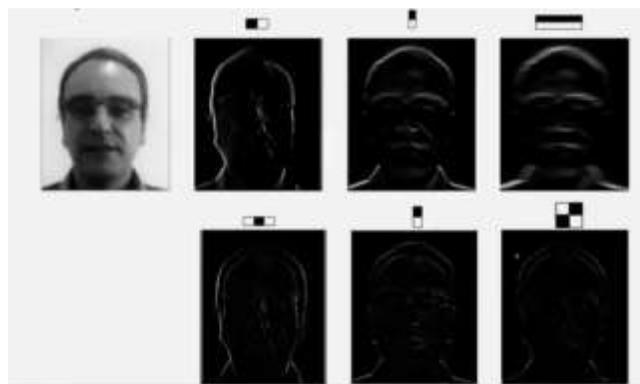


Figura 3. Resultados de aplicar los filtros de Haar básicos. Fuente: (López, Valveny. y Vanrell, 2019)

Filtros de Haar extendidos:

- Algunos son rotaciones de los filtros básicos que permiten detectar características direccionales.
- Algunos son nuevas configuraciones de los filtros para detectar otro tipo de características.

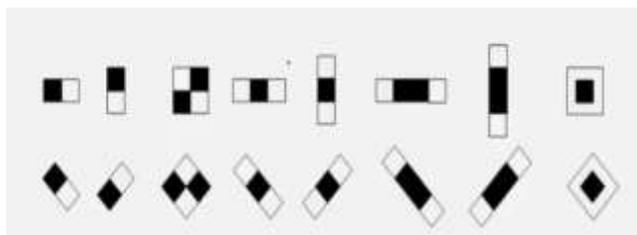


Figura 4. Filtros de Haar extendidos. Fuente: (López, Valveny. y Vanrell, 2019)

Aplicación de los filtros Haar en una imagen

- Cada filtro se aplica a todas las posibles escalas en horizontal y vertical.
- Cada escala se aplica en todas las posibles posiciones de la imagen.
- El resultado de aplicar de cada filtro en cada escala y posición es una característica de Haar.

3.2 Extracción de características del rostro

Un paso fundamental, previo a la aplicación de cualquier técnica de *Machine Learning*, es la selección de las características, o atributos, que se van a utilizar para el entrenamiento y posterior aplicación del sistema de inteligencia. Para el contexto de este trabajo, se encontraron diversas alternativas para la extracción de dichas características. En primer lugar, se puede distinguir entre dos tipos básicos de características: globales y locales (Garrido, 2018).

Las características globales se obtienen teniendo en cuenta todos los píxeles de una imagen. Por ejemplo, Wu & Zhang (2001) aplican PCA como método de extracción de características. Aunque los algoritmos obtenidos con este método, tienen un rendimiento bastante bueno, existe un problema inherente, y es que las características globales son sensibles a las variaciones en la imagen, como la orientación, los cambios en la iluminación o la oclusión parcial (Garrido, 2018).

Las características locales, por el contrario, se ven menos afectadas por estos cambios. Además, la información geométrica y las restricciones en la configuración de las diferentes características locales pueden ser utilizadas de forma tanto implícita como explícita, es decir, no solo importa su valor, si no, también su posición en la imagen (Garrido, 2018).

Las características "Haar wavelet", se basan en la aplicación de funciones matemáticas para la descomposición de una imagen. Este tipo de características capturan la forma y la estructura de objetos, y dependiendo de la escala a la que se apliquen obtendrán un objeto más o menos nítido, concluyendo que lo importante no era el valor de estos coeficientes, sino su presencia o ausencia (Garrido, 2018).

Otro enfoque que se ha investigado son los filtros Gabor, que transforman la imagen en función de una determinada frecuencia y Orientación que por lo general combinados junto con las características Haar-wavelet mejoran el rendimiento (Garrido, 2018).

En este trabajo, se utilizan el conjunto de características *Facial Network Landmark*, propuestos por (Garrido, 2018). Ya que representan el espacio en su totalidad, lo que resulta complicado cuando se usa un número infinito de muestras. Además, los modelos basados en este tipo de características operan mucho más rápido que los que utilizan las características explicadas anteriormente. Estas características fueron aplicadas para la detección de objetos, sobre todo para el reconocimiento facial, obteniendo muy buenos resultados.

3.3 Clasificación de expresiones faciales

Hay una serie de dimensiones que se deben tener en cuenta para determinar cuál puede ser un algoritmo de clasificación razonable (Pelillo, 2013)

- Número de ejemplos en el conjunto de entrenamiento.
- Dimensionalidad del espacio de atributos o características.
- ¿Se supone que las clases son linealmente separables?
- ¿Los atributos son independientes?
- ¿Cabe esperar que los atributos dependan linealmente de la variable de destino?
- ¿Se sospecha que el sobreajuste (*overfitting*) pueda ser un problema?
- ¿Cuáles son los requerimientos de sistema en términos de velocidad / rendimiento / uso de memoria?

Los elementos listados poseen un alto grado de complejidad para responder. Sin embargo, pueden ser abordados mediante el principio "la navaja de Occam": utilizar el algoritmo menos complicado que puede hacer frente a las necesidades y sólo ir por algo más complicado si es estrictamente necesario (Pelillo, 2013).

3.3.1 Algoritmo K-NN: K vecinos más cercanos (*K Nearest Neighbours*)

El algoritmo K-NN clasifica cada dato nuevo en el grupo que corresponda, y para ello toma en cuenta la clase más frecuente entre sus k vecinos más cercanos. Es decir, calcula la distancia del elemento nuevo a cada uno de los existentes, y ordena dichas

distancias de menor a mayor para seleccionar los primeros k elementos en dicho ordenamiento. La

clase de mayor frecuencia en dicho conjunto es asignada al nuevo elemento. En caso de empates (mayor frecuencia compartida entre clases) se asigna la clase de mayor probabilidad apriori, y si persiste el empate se selecciona aleatoriamente una de dichas clases (Moreno, 2004)

K-NN es un algoritmo de aprendizaje supervisado, es decir, que a partir de un juego de datos inicial su objetivo será el de clasificar correctamente todas las instancias nuevas. El juego de datos típico de este tipo de algoritmos está formado por varios atributos descriptivos y un solo atributo objetivo (también llamado clase) (Moreno, 2004)

En contraste con otros algoritmos de aprendizaje supervisado, K-NN no genera un modelo fruto del aprendizaje con datos de entrenamiento, sino que el aprendizaje sucede en el mismo momento en el que

se prueban los datos de test. A este tipo de algoritmos se les llama *Lazy Learning Methods* (Pelillo, 2013).

3.4 Bases de datos para estudios de emociones faciales

Con el objetivo de poner en práctica los algoritmos para el reconocimiento y clasificación de las emociones se han desarrollado varias bases de datos de imágenes (Alegre, 2013). Estas bases de datos suelen tener características específicas. En algunos casos se pueden encontrar algunas constituidas por personas de ambos sexos o de uno solo; con cambios de iluminación, poses o distintos tipos de expresiones a representar.

En la siguiente tabla se muestran algunos ejemplos de estas bases de datos y se dan a conocer características de las mismas.

Base de Datos	Cant. de personas	Características
Cohn-Kanade AU-Coded Facial Expression	100	Conformada por estudiantes universitarios, donde el 65% son mujeres, 15% son afroamericanos y el 3% son asiáticos o latinos. Incluye un conjunto de 23 proyecciones de distintas AUs o combinaciones de estas.
Natural Visible and Infrared facial Expression (USTC-NVIE)	100	Muestra expresiones espontáneas o posadas con iluminación proveniente de tres direcciones diferentes.
The Hong Kong Polytechnic University NIR	335	Conformada por alrededor de 34 000 imágenes con cambios de la posición y las expresiones.
The Bosphorus Database		Conformada por actores y actrices quienes representan las seis emociones básicas.
Indian Face Database	40	Se representan 4 expresiones. Contiene 11 imágenes diferentes de cada persona, las cuales incluyen hombres y mujeres.
Georgia Tech Face Database	50	Contiene 15 imágenes de cada individuo con variaciones de expresiones, apariencias y condiciones de iluminación; tomadas en diferentes escalas y orientaciones.
The AR Face Database, Purdue University, USA	126	Contiene alrededor de 4000 imágenes de hombres y mujeres, representando distintas expresiones y con diferentes condiciones de iluminación.
Japanese Female Facial Expression (JAFFE) Database	10	Contiene 213 imágenes representando las 6 emociones básicas más la imagen neutral de cada persona. Está conformada solamente por actrices.
Multi-PIE	337	Alrededor de 750 000 imágenes representando distintas expresiones faciales y cambios de iluminación.
PIE Database, CMU	68	Contiene 41 368 imágenes, cada persona con 13 poses diferentes, con presencia de 43 condiciones diferentes de iluminación y representando 4 tipo de expresiones faciales.
The Yale Face Database	15	Contiene 11 imágenes de cada persona, una para cada expresión o configuración en correspondencia con los cambios en la iluminación.
Face Recognition Data, University of Essex, UK	395	Contiene imágenes tanto de hombres como de mujeres de diferentes orígenes raciales y edades.

Tabla 1: Ejemplos de Bases de Datos de emociones faciales.

3.5 Implementación del software

Pasos para el desarrollo del proceso de detección de estados emocionales en el módulo propuesto:

1- *Inicio del procesamiento:*

El proceso de detección es desencadenado por la solicitud del profesor desde la aplicación ATcnea-Teacher-PC, dicha solicitud puede ser para realizar la detección del estado emocional de un estudiante a través de una imagen fija o una captura de video.

1.1-Captura de Imagen fija:

Para la captura de la imagen fija se usan las clases definidas por la librería OpenCV-3.4.0 para el enlace con la *webcam*, dichas librerías facilitan la creación de un objeto capture de tipo *VideoCapture* que permite manipular el comportamiento del dispositivo en todo momento.

1.2- Captura de frame de video:

Para la captura de cada frame de video se explotaron las potencialidades de las librerías Javacv-3.1.4 para el uso del dispositivo *webcam* en captura de video, capacidad extensible incluso a sistemas que usen por sistema operativo alguna distribución de GNU-Linux en donde no sea posible el uso de ficheros .dll para el enlace con los dispositivos de hardware del ordenador.

2.1- Preprocesamiento de la imagen

Para el preprocesamiento de la imagen capturada se aplican diferentes algoritmos con el objetivo de preparar la imagen para lograr el máximo de desempeño durante el proceso de extracción de características. Los algoritmos de preprocesamiento utilizados son: *cvtColor* (transformando la imagen original a escala de grises), *equalizeHist* (ecualizando el histograma de color de la imagen) y *imresize* (transformando las dimensiones de la imagen a 512x512).

2.1.1- Detección de rostros y demás estructuras faciales (ROIs)

Para la detección del rostro y de las regiones de interés ojos, boca y nariz en la imagen capturada se utilizaron el algoritmo Viola-Jones y los descriptores Haar para la construcción de los detectores Haar-CascadeClassifier. Estos detectores permiten definir, mediante las propiedades del modelo de clasificación, el tipo de objeto a detectar en la imagen, el cual será, de forma implícita, rostros.

Este método también ofrece la posibilidad de encontrar en la imagen, los ojos (por separados o juntos), la boca y la nariz. Como resultado de la detección, se obtiene una caja contenedora o BBOX (bounding box), que es una matriz que encierra los objetos detectados. En la imagen de entrada queda señalado entonces un rectángulo con estas medidas. Cada fila de la matriz de salida, BBOX, contiene un vector de cuatro elementos, [x, y, ancho, alto], donde se especifican, en píxeles, las coordenadas de esta caja.

2.2- Preprocesamiento del frame de video

Para el preprocesamiento de cada frame de video se aplican los siguientes algoritmos: *CV_BGR2GRAY* (transformando la imagen original a escala de grises) y *imresize* (transformando las dimensiones de la imagen a 512x512). Estas dimensiones son selec-

cionadas debido a que una dimensión menor le daría mucho peso a un píxel, el cual pudiera perderse en el proceso. No se seleccionan dimensiones mayores para ahorrar costo computacional.

3- Segmentación del rostro en la imagen

La segmentación del rostro en la imagen se realiza para evitar durante el proceso de extracción de características información redundante o que pudiese añadir ruido. Para realizar la segmentación en la imagen fija o en frame de video se utiliza la función *imcrop*.

3.1- Extracción de los puntos característicos del rostro en imagen fija

Para la extracción de los puntos característicos en imagen fija se utiliza el algoritmo *KazemiFacemark* este algoritmo es capaz de cargar un modelo de datos usando la función *loadModel* ("face_landmark_model.dat"), modelo previamente entrenado para la detección de las unidades de acción (AUs) en el rostro.

Para la utilización de la información contenida en el modelo de datos es necesario convertirla de vectores a puntos usando la función *Point2fVector*. Para el cálculo de las características expresadas como las distancias entre determinados puntos faciales se utiliza la función distancia correspondiente.

3.2- Extracción de los puntos característicos del rostro en frame de video

Para la extracción de los puntos característicos en frame de video se utiliza el algoritmo *FNlandmark* este algoritmo carga un modelo de datos usando la función *loadFLandmarkModel* ("flandmarkModelFile.dat"), este modelo fue entrenado con el mínimo de características necesarias para la extracción, en función de minimizar el coste computacional del cargado en memoria entregando altos tiempos de respuesta, haciendo posible su uso en video capturas.

Para el cálculo de las características expresadas como las distancias entre determinados puntos faciales se utiliza la función distancia correspondiente.

4- Clasificación de las características, etiquetado y salva de datos

Para la clasificación de las características se utiliza el algoritmo K-Vecinos más cercanos (K-NN). Para su implementación este algoritmo se sirve de dos listas de arreglos *List<double[]> testfeatures* unas con los datos de pruebas (datos recopilados en el proceso anterior en forma de características expresadas como las distancias entre determinados puntos faciales) y *List<double[]> trainfeatures* la lista de datos de entrenamiento donde se encuentran las características asociadas a las modelos que integran la base de

datos de entrenamiento Jaffe.

Para el etiquetado el algoritmo utiliza la función de distancia euclidiana *euclideanDistance* y de forma iterativa se hallan las distancias entre los valores de la lista de prueba y los datos de entrenamiento almacenados en la base de datos Jaffe comparando los resultados arrojados para escoger las k-clases más cercanas.

La clase seleccionada para etiquetar la nueva instancia es la de mayor frecuencia entre los k-prospectos. La selección de la etiqueta está acompañada por un resultado de precisión calculado como: $(Class /$

$knn_value) * 100) \%$.

Finalmente, el resultado de la clasificación y la precisión serán almacenados de forma persistente en una base de datos para futuros análisis.

5- Fin del procesamiento

Una vez concluido el procesamiento se muestra el resultado de la detección del estado emocional del estudiante en forma de emoción en la parte inferior izquierda del cuadro del estudiante en la vista de la clase ATcnea-Teacher-PC.

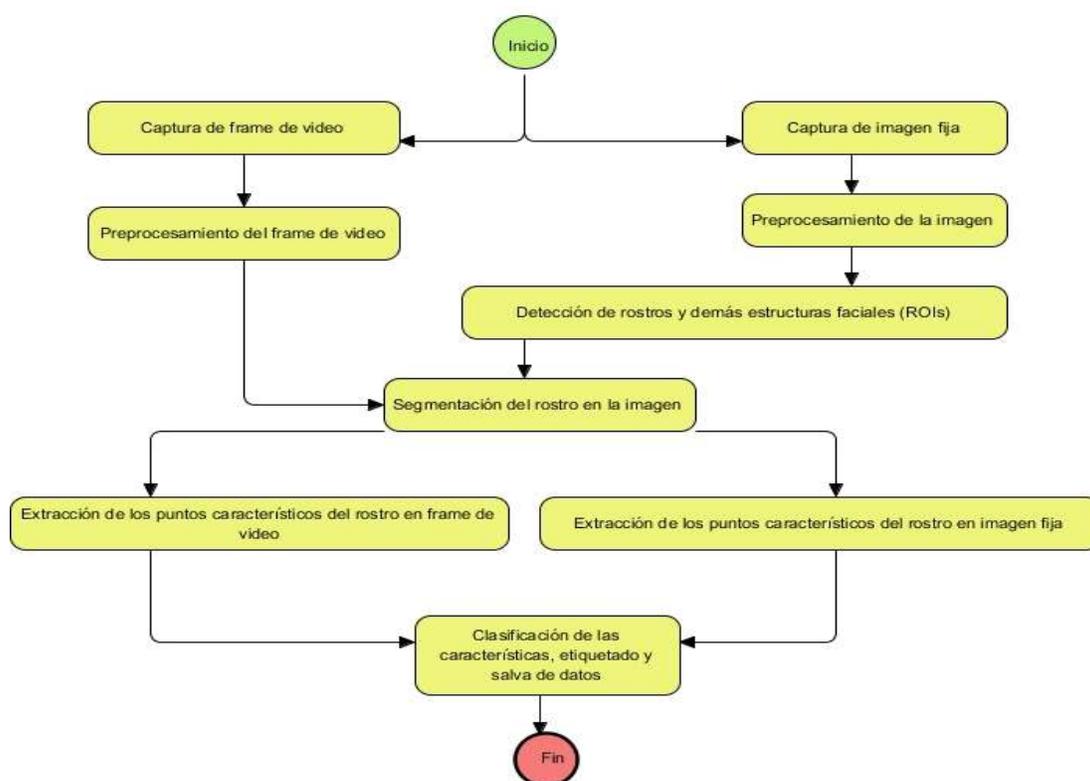


Figura 5: Proceso de detección de estados emocionales

4. Conclusiones

Con la investigación y el desarrollo del “Módulo para la detección de estados emocionales a través de expresiones faciales en el aula tecnológica XAUCE ATcnea” se concluye que:

- Al definir el marco teórico conceptual de la investigación se evidenció la no existencia de un módulo integrable al software XAUCE ATcnea que cumpliera con los requisitos especificados por el cliente.

- El diseño de la propuesta de solución permitió generar los artefactos acordes a la metodología de desarrollo de software AUP-UCI.
- La implementación del sistema a través de las herramientas y lenguajes seleccionados permitió obtener una aplicación capaz de usar una imagen o una captura de video para detectar el estado emocional de los estudiantes que se encuentren haciendo uso de la aplicación informática ATcnea-Estudiante-PC.
- La etapa de prueba del “Módulo para la detección



de estados emocionales a través de expresiones faciales en el aula tecnológica XAUCE ATcnea” permitió la detección de no conformidades y su posterior corrección, además, demostraron que el módulo constituye una solución funcional.

5. Referencias bibliográficas

- Alegre, E. (2013) *Bases de datos de imágenes: arquitectura de los sistemas de recuperación de imágenes basados en contenido*.
- Baldasari, S. (2016) *Computación afectiva: tecnología y emociones para mejorar la experiencia de usuario*. Zaragoza- España: Bit & Byte, Vol. 2.
- Blanco Garrido, C. (s/f) *Extracción de características para algoritmos de aprendizaje*
- Jain, A.K. y Li, S.Z. (2005) *Handbook of face recognition*.
- Londhe, R. y Pawar, V. (2012) *Facial expression recognition based on affine moment invariants*. S.I.: International Journal of 11. Computer Science Issues (ijcsi), vol. 9.
- López, A; Valveny, E. y Vanrell, M. (2019) *Detección de objetos*. [curso en línea] <https://es.coursera.org/learn/deteccion-objetos>.
- Moreno Seco, F. (2004) *Clasificadores eficaces basados en algoritmos rápidos de búsqueda del vecino más cercano*.
- Pelillo, M. (2013) *Avances in computer vision and pattern recognition*.

Fecha de recepción: 6 de octubre de 2019

Fecha de aceptación: 19 de diciembre de 2019

