

## RECONOCIMIENTO FACIAL A TRAVÉS DE CÁMARAS DE VIDEO VIGILANCIA

Ing. Gladis Guzmán Guerrero, Dr. José Antonio Montero Valverde, M.T.I. Rafael Hernández Reyna.

*Tecnológico Nacional de México / IT de Acapulco*

México

[gladis\\_guzman@hotmail.com](mailto:gladis_guzman@hotmail.com)

[jamontero1@infinitummail.com](mailto:jamontero1@infinitummail.com)

[rhernan7@yahoo.com.mx](mailto:rhernan7@yahoo.com.mx)

### **Resumen**

*En los últimos años, el desarrollo de programas informáticos para sistemas de seguridad en cámaras de video vigilancia ha tenido un gran avance. El reconocimiento facial es una herramienta importante para la seguridad y control en áreas de interés en particular, es por eso que el objetivo de este proyecto es diseñar una herramienta computacional que sea capaz de hacer reconocimiento facial de manera confiable a través de cámara de Video Vigilancia haciendo uso de un algoritmo de detección de objetos. El algoritmo de Viola Jones es un método de detección de objetos que se usa ampliamente en la detección de caras en imágenes y video. Para realizar la identificación de rostros, se hará uso del algoritmo k-nn (K-Nearest Neighbour).*

**Palabras Clave.** *Detección facial, Reconocimiento Facial, Algoritmo Viola & Jones, Cámaras de Video Vigilancia, imagen Integral, K-Nearest Neighbour.*

FACIAL RECOGNITION THROUGH VIDEO SURVEILLANCE CAMERAS.

### **Abstract.**

In recent years, the development of computer programs for security systems in video surveillance cameras has made great progress. Facial recognition is an important tool for security and control in areas of interest in particular, that is why the objective of this project is to design a computational tool that is capable of reliably making facial recognition through a Video Surveillance camera. making use of an object detection algorithm. Viola Jones' algorithm is an object detection method that is widely used in face detection in images and video. To identify faces, the k-nn (K-Nearest Neighbor) algorithm will be used.

**Keywords.** Facial Detection, Facial Recognition, Viola & Jones Algorithm, Video Surveillance Cameras, Integral image, K-Nearest Neighbor.

### I. ANTECEDENTES

En los últimos años, la Inteligencia Artificial ha tenido grandes avances para la sociedad en general, muchas áreas del conocimiento se han beneficiado de esta disciplina, ya que se ha convertido en una herramienta de ayuda a varios campos de trabajo como, por ejemplo: Diagnóstico médico, comercio de acciones, control robótico, percepción remota, seguridad informática, entre otras.

El uso de Inteligencia Artificial es una posible solución a la seguridad, pues ofrece mayores posibilidades y funcionalidades en video vigilancia. Gracias a las cámaras de seguridad integradas con un software de análisis de video con Inteligencia Artificial, se pueden obtener muchas funciones, no sólo orientadas a la seguridad y vigilancia, también al control de calidad, logística, entre otras.

La inseguridad en nuestro país se ha convertido en un problema. Hoy en día, es una de las principales características de todas las sociedades modernas, ya que vivimos en un mundo en el que la extensión de la violencia se ha desbordado en un clima generalizado de criminalidad. Guerrero es uno de los 5 estados más inseguros y delictivos del país según la revista Forbes [1]. Desafortunadamente esta situación ha pasado a afectar a instituciones educativas, es por eso que se deben tomar medidas de seguridad no solo de personal capacitado sino también haciendo uso de la tecnología inteligente.

### II. INTRODUCCIÓN

El rostro humano es un objeto dinámico que tiene un alto grado de variabilidad en su apariencia lo cual hace que su detección sea un problema difícil de tratar en visión por computador. En los inicios de esta tecnología llamada "Reconocimiento facial" se usaba distintos algoritmos muy simples, el cual daba mayor oportunidad a que los errores se produjeran, ya que al ser así el mismo reconocimiento se podría dar para 2 personas diferentes. En la actualidad los algoritmos de detección e identificación han sido exponencialmente mejorados, los errores son mínimos pues la manera en la que se lleva a cabo el

reconocimiento de rostros ha cambiado logrando obtener resultados más confiables. El CETis No. 116 de Acapulco es un centro de estudios tecnológicos de nivel medio superior que ofrece 7 carreras técnicas. El predio del plantel cuenta con una barda perimetral a lo largo de toda su extensión y con servicio de personal de seguridad privada, para la vigilancia de las diversas áreas durante las 24 horas del día, recientemente se implementó un sistema de cámaras de video vigilancia ubicadas en áreas estratégicas, entre ellas se encuentra primordialmente el área de los centros de cómputo, los cuales fueron recientemente remodelados lo que representó una inversión monetaria muy significativa y es de alta importancia para la dirección del plantel mantener el cuidado de esta inversión fortaleciendo el sistema de video vigilancia instalado en el interior de los centros de cómputo haciendo uso de la inteligencia Artificial con un algoritmo de reconocimiento facial que permita identificar a las personas dentro de estas instalaciones.

### III. MARCO TEÓRICO

#### A. Algoritmo de Viola-Jones

El método de detección de objetos de Viola Jones fue propuesto por Paul Viola y Michael Jones en 2001, que fue el primer método que proporciona realmente un alto grado de detección de objetos. Se puede utilizar para detectar objetos en tiempo real, pero se aplica principalmente a la detección de rostros. La tasa de detección de este método es relativamente alta y muy baja con relación a falsos positivos, lo que hace al algoritmo tan robusto y procesa las imágenes rápidamente. Su objetivo principal es la detección de rostros más no el reconocimiento, es decir, distinguir las caras de las no caras, este procedimiento es el primer paso para el reconocimiento facial [2].

El algoritmo se divide en tres etapas como se muestra en la Figura 1, en la primera etapa se realiza una transformación de la imagen generando una nueva llamada imagen integral, en el segundo bloque se realiza la extracción de características usando filtros con base haar1, y por último se usa boosting para la construcción de clasificadores en cascada [3].

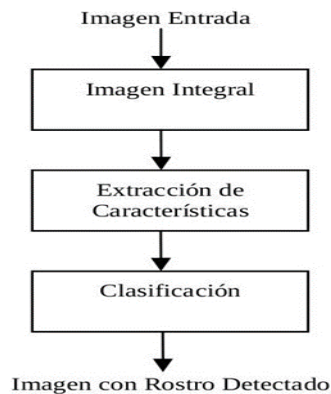


Figura: 1 Diagrama de Bloques [3].

#### B. Imagen Integral

Esta imagen permite extraer de forma rápida características a diferentes escalas ya que no se trabaja directamente con los valores de intensidad si no con una imagen acumulativa que se construye a partir de operaciones básicas [3].

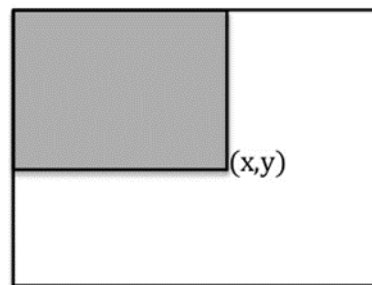


Figura: 2 Imagen Integral [3].

La imagen integral (Figura 2) en la localización  $x, y$ , contiene la suma de los píxeles de la parte superior izquierda de la imagen. La suma de todos los píxeles de cualquier rectángulo en la imagen como se puede calcular a continuación:

$$II(x, y) = \sum_{xj \leq x, yj \leq y} Im(xj, yj)$$

Donde  $II(x, y)$  es la imagen integral e  $Im(x, y)$  es la imagen original.

#### C. Características Haar

En imágenes las características de cada objeto se extraen al aplicar ciertas funciones que permitan la representación y descripción de los objetos de interés de la imagen (patrones). La extracción de características es un paso en el reconocimiento de

patrones en el cuál las medidas u observaciones son procesadas para encontrar atributos que puedan ser usados para asignar los objetos a determinada clase.

La extracción de características es realizada aplicando a la imagen filtros con bases haar. Estos filtros pueden ser calculados eficientemente sobre la imagen integral, son selectivos en la orientación espacial y frecuencia, y permiten ser modificados en escala y orientación. En la Figura 3 se muestran algunos de los filtros usados para la extracción de características.

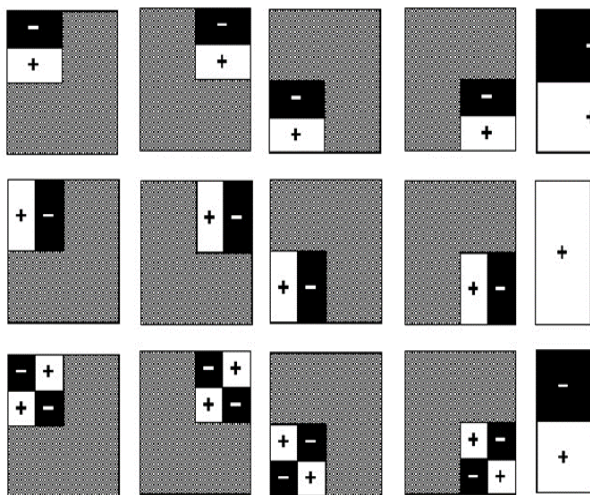


Figura: 3 Filtros Haar rotados, trasladados y con cambios de escala

De esta manera, un conjunto de características pueden ser usadas para codificar los contrastes encontrados en los rostros.

Originalmente se consideran 3 tipos características: (Figura 4)

Característica de dos rectángulos: Su valor se calcula con la diferencia entre la suma de los píxeles dentro de las dos regiones rectangulares.

Característica de tres rectángulos: Su valor se calcula con la suma de los píxeles dentro de las dos regiones rectangulares exteriores y la sustracción de los píxeles de la región interior

Característica de cuatro rectángulos: Su valor se calcula con la diferencia entre las diagonales de los pares de rectángulos.

Rectángulos en negro representan zonas con una contribución positiva al filtro. Rectángulos en blanco representan zonas con una contribución negativa al filtro

200	200	100	100	200	200	100	100
250	250	50	50	250	250	50	50
255	255	255	255	100	100	100	100
255	255	255	255	100	100	100	100
200	200	100	100	200	200	100	100
250	250	50	50	250	250	50	50
255	255	255	255	100	100	200	200
255	255	255	255	100	100	250	250

Figura: 4 Representación de filtros Haar

1) Aplicación de filtros haar

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 cada filtro en cada escala y posición es una característica de haar.

En imágenes de 24x24 píxeles se obtienen 162336 características.

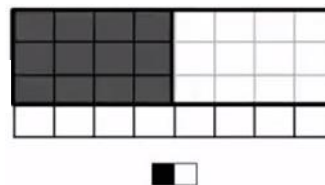


Figura: 5 Aplicación de filtros haar

D. Clasificación del algoritmo de detección

Ésta etapa dentro del algoritmo de detección se encarga de asignar un conjunto de características dado a una clase con la que se encuentra una mayor similitud, de acuerdo a un modelo inducido durante el entrenamiento [4]. En la figura 6 se muestra un esquema de un clasificador en cascada.



Figura: 6 Cascada de clasificadores

#### IV. METODOLOGIA

El reconocimiento facial es una herramienta que nos permite identificar a una persona automáticamente por medio de una imagen digital. Para realizar un reconocimiento facial se deben analizar las características faciales de la persona, las cuales se pueden extraer ya sea de una fotografía o desde un fotograma en una fuente de video, esta se convierte en una plantilla y luego se compara con las imágenes en una base de datos para verificar la identidad de la persona. Se obtiene por medio de un proceso, este proceso se divide en 4 etapas como se muestra en la figura 7, como lo son: Pre-procesamiento, Detección facial, Extracción de Características Y Clasificación en la cual se realiza una consulta en la base de datos para comparar los datos extraídos y finalmente identificar la persona.



Figura: 7 Metodología de Software

##### A. Pre-procesamiento

El Pre-procesamiento de imágenes digitales es el conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar la calidad o facilitar la búsqueda de información. Su objetivo es reducir el entorno que no es de interés para el problema. Fondo, ruido, entre otras. El conjunto de métodos de Pre-procesamiento de imágenes está dividido en tres grupos:

**Algoritmos en el dominio espacial:** Se refiere a métodos que procesan una imagen píxel por píxel, o también tomando en cuenta un conjunto de píxeles vecinos.

**Algoritmos en el dominio de la frecuencia:** Frecuentemente, estos métodos son aplicados sobre los coeficientes resultantes de la Transformada de Fourier de una imagen.

**Algoritmos de extracción de características:** A diferencia de los dos grupos anteriores, los algoritmos de extracción de características están enfocados al análisis de imágenes para la extracción de atributos y regiones de interés, separación de objetos del fondo, detección de bordes o formas,

entre otros. El algoritmo de Viola-Jones utiliza esta técnica [5].

##### B. Detección Facial

La detección facial es una tecnología de visión computarizada que determina el lugar y el tamaño de rostros humanos en imágenes o videos. Es uno de los subtipos de detección de objeto/clase, cuya tarea es encontrar la localización y el tamaño de los objetos en una imagen. Mientras que para el ser humano es trivial esta tarea, una computadora presenta dificultades debido a diversos factores como: variabilidad en la posición del rostro, presencia o ausencia de componentes estructurales (ejemplo: bigote), expresión facial, oclusión (uso de lentes, gorro, entre otras.) y condiciones del ambiente. La detección facial es una etapa importante en un sistema de reconocimiento facial, embargo influye de manera significativa en el estado del proceso, ya sea dado un conjunto de frames o video en tiempo real. Por lo tanto, debe ser capaz de identificar los rostros independientemente de los factores que anteriormente se mencionaron. Los métodos de detección facial se dividen en cuatro categorías (no excluyentes):

**Métodos basados en conocimiento:** Codifican el conocimiento humano mediante distancias y relaciones entre las características humanas (ojos, nariz, labios).

**Métodos basados en características invariantes:** Las características invariantes son aquellas que no se modifican a eventuales cambios de luz, pose o ubicación de la cámara, tales como la ceja, nariz, textura de la piel y línea de pelo.

Este método funciona detectando uno de estos componentes, construyendo un modelo estadístico y con los resultados, verificar la existencia de un rostro.

**Métodos basados en moldes (patrones):** Es la relación entre una imagen de entrada y un patrón o molde previamente definido, cuyo objetivo es capturar características del rostro.

**Métodos basados en apariencia:** Utilizan modelos obtenidos mediante entrenamiento de imágenes, tomando la imagen como un vector de características, es decir, es visto como una variable aleatoria. A diferencia de los métodos basados en moldes, donde el patrón es definido por un "experto", los patrones en este modelo son determinados por el aprendizaje obtenido en el entrenamiento de imágenes.

Existen varios métodos para detectar rostros, sin embargo el algoritmo de Viola-Jones, es el más

eficaz, obteniendo un mayor porcentaje de aciertos respecto a sus pares, en la figura 8 podemos observar una detección facial eficaz y confiable.



Figura: 8 Detección Facial

### C. Extracción de Características

La extracción de características en imágenes y secuencias de imágenes faciales, consiste en extraer información asociada con la activación de los diferentes músculos del rostro, esta tarea puede realizarse en forma global u holística en donde se analiza el rostro como un solo conjunto o localmente en donde se seleccionan regiones de interés del rostro como ojos cejas y boca. Adicionalmente los métodos de extracción de características faciales se pueden clasificar de acuerdo a su enfoque en: características faciales de movimiento o deformaciones del rostro.

Inicialmente, el algoritmo necesita muchas imágenes positivas (imágenes de caras) e imágenes negativas (imágenes sin caras) para entrenar al clasificador. Entonces necesitamos extraer características de él [6]. En imágenes las características de cada objeto se extraen al aplicar ciertas funciones que permitan la representación y descripción de los objetos de interés de la imagen (patrones). La extracción de características es un paso en el reconocimiento de patrones en el cual las medidas u observaciones son procesadas para encontrar atributos que puedan ser usados para asignar los objetos a determinada clase [7]. Inicialmente en esta etapa de extracción de características, el algoritmo necesita imágenes positivas (imágenes de caras) e imágenes negativas (imágenes sin caras) para entrenar al clasificador.

### D. Clasificación

Esta etapa dentro del algoritmo de detección se encarga de asignar un conjunto de características dado a una clase con la que se encuentra una mayor

similitud, de acuerdo a un modelo inducido durante el entrenamiento.

#### 1) Algoritmo K Nearest Neighbours.

Es un método que busca en las observaciones más cercanas a la que se está tratando de predecir y clasifica el punto de interés basado en la mayoría de datos que le rodean, es un algoritmo:

*Supervisado:* esto quiere decir que tenemos etiquetado nuestro conjunto de datos de entrenamiento, con la clase o resultado esperado dada una fila de datos.

*Basado en Instancia:* Esto quiere decir que nuestro algoritmo no aprende explícitamente un modelo (como por ejemplo en Regresión Logística o árboles de decisión). En cambio memoriza las instancias de entrenamiento que son usadas como base de conocimiento para la fase de predicción.

El algoritmo K Nearest Neighbours será utilizado en este proyecto para la identificación de rostros, este algoritmo clasifica cada dato nuevo en el grupo que corresponda, según tenga k vecinos más cerca de un grupo o de otro. Es decir, calcula la distancia del elemento nuevo a cada uno de los existentes, y ordena dichas distancias de menor a mayor para ir seleccionando el grupo al que pertenecer como se muestra en la figura 9. Este grupo será, por tanto, el de mayor frecuencia con menores distancias.

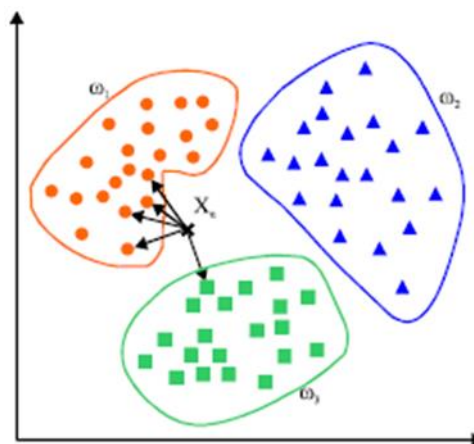


Figura: 9 Clasificador KNN

## V. RESULTADO ESPERADOS

Mediante la implementación de este algoritmo de reconocimiento facial se busca identificar positivamente al 80% de los sujetos que sean detectados de forma eficaz a través de las cámaras de video vigilancia, se pretende que cada persona que esté en posición frontal con el rostro visible sea reconocida por el sistema de reconocimiento como

se observa en la figura 10. Estos sistemas han sido de gran utilidad a lo largo de mucho tiempo, el propósito de este proyecto es robustecer esta tecnología con la capacidad de identificar positivamente con un alto grado de certeza a los sujetos dentro del campo de visión de las cámaras para una gestión mejor de la seguridad dentro del área a supervisar.



Figura: 10 Reconocimiento facial

## VI. REFERENCIAS

- [1] Ameth, E. (2015). Los municipios más inseguros de México. *FORBES*.
- [2] Guevara, M. L., Echeverry Correa, J. D., & Ardila Urueña, W. (2008). Detección de caras en imágenes digitales usando clasificadores en cascada. *Revista Scientia et Technica*.
- [3] VIOLA, P., & JONES., M. J. (2003). Robust Real-Time Face Detection. *International Journal of Computer Vision*.
- C.P. Papageorgiou, M. O. (2018). A general framework for object detection. *IEEE*.
- Ysiquio, I. A., Gil, D. M., Starostenko, D. O., & Rodríguez, D. D. (2009). *Modelado de Sistemas de control de un robot manipulador basado en procesamiento digital de Imágenes*.
- Barroso Heredia, C., & Díaz de María, F. (2014). Implementación del algoritmo de detección de caras de Viola y Jones sobre una FPGA.
- [7] Correa, M. S., & Chichizola, F. (2015). Diseño de Sistemas de Reconocimiento de rostros.