



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas

## Reconocimiento Facial

Integrantes

Cerritos Lira Carlos  
Mosqueda García Raúl Isaíd  
Ruiz Puga Ingrid Pamela

Minería de Datos

---

# Índice

<b>1. Antecedentes</b>	<b>3</b>
<b>2. Marco Teórico</b>	<b>3</b>
2.1. Redes Neuronales . . . . .	3
2.1.1. Redes Convoluciones . . . . .	4
2.1.2. FaceNet . . . . .	5
<b>3. Introducción</b>	<b>5</b>
3.1. Reconocimiento Facial . . . . .	5
3.2. Preliminares . . . . .	6
<b>4. Objetivo</b>	<b>7</b>
<b>5. Alcances</b>	<b>7</b>
<b>6. Marco Estratégico</b>	<b>7</b>
<b>7. Alcances</b>	<b>7</b>
<b>8. Definiciones</b>	<b>7</b>
<b>9. ABT</b>	<b>7</b>
<b>10. Modelo</b>	<b>7</b>
<b>11. Resultados</b>	<b>7</b>
<b>12. Conclusiones</b>	<b>7</b>

---

## 1. Antecedentes

Los sistemas de reconocimiento facial se podrían considerar como relativamente modernos. Uno de sus pioneros fue Woodrow Bledsoe, quien junto con su grupo de investigación desarrolló una técnica que fue denominada “reconocimiento facial hombre-máquina”, a mediados de los años 60. Esta técnica requería una clasificación de las fotografías de las caras que eran digitalizadas manualmente. Además, también era necesario indicar las coordenadas donde se encontraban los rasgos faciales dentro de cada imagen, entre los que destacaban los ojos, la nariz y la línea del pelo. Posteriormente, era necesario registrar los datos recogidos junto con el nombre de la persona en una base de datos. Finalmente, con una fotografía de una cara desconocida, el sistema utilizaría un método basado en distancias euclídeas para obtener la imagen en la base de datos que más se aproxime a la fotografía en cuestión. Un acontecimiento importante en el campo del reconocimiento facial se dio en 1988, cuando L. Sirobich y M. Kirby aplicaron la técnica del análisis de componentes principales (PCA) al problema del reconocimiento facial y del que posteriormente surgiría el método de Eigenfaces de la mano de Turk y Pentland en 1991. Aunque hubo unos años en los que el campo del reconocimiento facial no presentó importantes desarrollos, en los años 90 se convirtió en una actividad en crecimiento debido al aumento de la investigación comercial, así como el alza de los clasificadores basados en redes neuronales en tiempo real. Desde entonces, han surgido numerosos tipos de algoritmos, así como variaciones de los primeros, que aún en la actualidad siguen utilizándose.

## 2. Marco Teórico

El problema del reconocimiento facial se basa en una identificación de patrones visuales. Generalmente, un sistema de reconocimiento facial se divide en cuatro fases: detección, preprocesado, extracción de características y comparación y clasificación. La primera fase de detección facial consiste en localizar el rostro en la imagen. A continuación, se realiza un preprocesado a la imagen para alinear y normalizar los rostros, de manera que se obtengan unas características geométricas en común con todas las imágenes que se procesen. Luego, se lleva a cabo una extracción de características faciales para obtener información útil que será usada para distinguir unas caras de otras. Finalmente, el vector de características obtenido será comparado con la base de datos para decidir a quién pertenece el rostro.

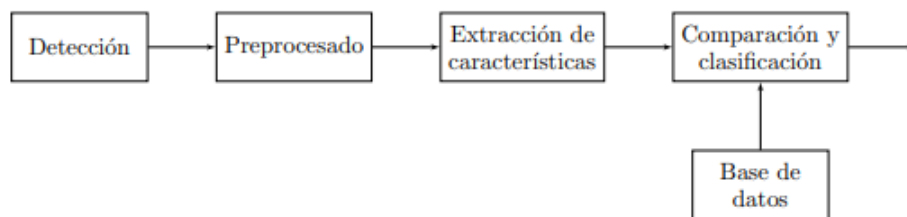


Figura 1: Proceso para reconocimiento

### 2.1. Redes Neuronales

Una red neuronal es un modelo simplificado que emula el modo en que el cerebro humano procesa la información: Funciona simultaneando un número elevado de unidades de procesamiento interconectadas que parecen versiones abstractas de neuronas.

Las unidades de procesamiento se organizan en capas. Hay tres partes normalmente en una red neuronal : una capa de entrada, con unidades que representan los campos de entrada; una o varias capas ocultas; y una capa de salida, con una unidad o unidades que representa el campo o los campos de destino. Las unidades se conectan con fuerzas de conexión variables (o ponderaciones). Los datos de entrada se presentan en la primera capa, y los valores se propagan desde cada neurona hasta cada neurona de la capa siguiente. al final, se envía un resultado desde la capa de salida.

---

Al principio, todas las ponderaciones son aleatorias y las respuestas que resultan de la red son, posiblemente, disparatadas. La red aprende a través del entrenamiento. Continuamente se presentan a la red ejemplos para los que se conoce el resultado, y las respuestas que proporciona se comparan con los resultados conocidos. La información procedente de esta comparación se pasa hacia atrás a través de la red, cambiando las ponderaciones gradualmente. A medida que progresa el entrenamiento, la red se va haciendo cada vez más precisa en la replicación de resultados conocidos. Una vez entrenada, la red se puede aplicar a casos futuros en los que se desconoce el resultado.

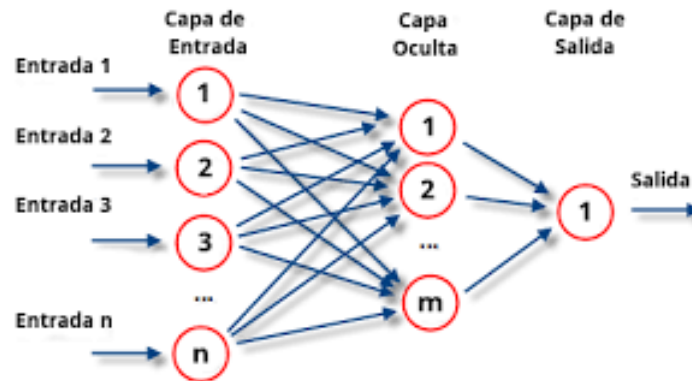


Figura 2: Arquitectura de red neuronal

### 2.1.1. Redes Convoluciones

Las Redes neuronales convolucionales son un tipo de redes neuronales artificiales donde las “neuronas” corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria de un cerebro biológico. Este tipo de red es una variación de un perceptrón multicapa, sin embargo, debido a que su aplicación es realizada en matrices bidimensionales, son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones.

Las redes neuronales convolucionales consisten en múltiples capas de filtros convolucionales de una o más dimensiones. Después de cada capa, por lo general se añade una función para realizar un mapeo causal no-lineal.

Como cualquier red empleada para clasificación, al principio estas redes tienen una fase de extracción de características, compuesta de neuronas convolucionales, luego hay una reducción por muestreo y al final tendremos neuronas de perceptrón mas sencillas para realizar la clasificación final sobre las características extraídas.

La fase de extracción de características se asemeja al proceso estimulante en las células de la corteza visual. Esta fase se compone de capas alternas de neuronas convolucionales y neuronas de reducción de muestreo. Según progresan los datos a lo largo de esta fase, se disminuye su dimensionalidad, siendo las neuronas en capas lejanas mucho menos sensibles a perturbaciones en los datos de entrada, pero al mismo tiempo siendo estas activadas por características cada vez más complejas.

Las Redes neuronales Convolucionales, CNN aprenden a reconocer una diversidad de objetos dentro de imágenes, pero para ello necesitan “entrenarse” de previo con una cantidad importante de “muestras” y a su vez, poder generalizarlo, a esto es lo que se le conoce como el proceso de “aprendizaje de un algoritmo”. Nuestra red va a poder reconocer por ejemplo un cierto tipo de célula porque ya la ha “visto” anteriormente muchas veces, pero no solo buscará células semejantes sino que podrá inferir imágenes que no conozca pero que relaciona y en donde podrían existir similitudes, y esta es la parte inteligente del reconocimiento.

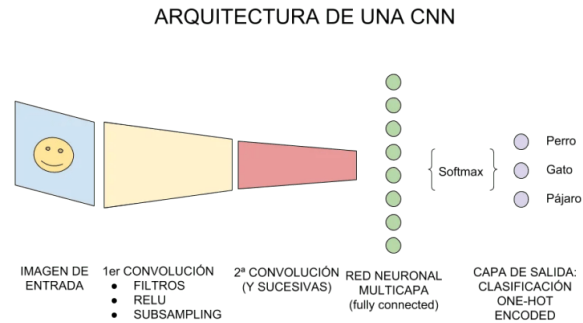


Figura 3: Arquitectura de red convolucional

Por un lado, tenemos la propia imagen a la que queremos aplicar la convolución, y otra matriz más pequeña llamada kernel. El resultado de la convolución es otra matriz de las mismas dimensiones que la imagen original. Para calcular la matriz de convolución vamos desplazando el kernel por cada uno de los elementos de la matriz principal y hacemos la suma de los productos de cada uno de los elementos de ambas matrices.

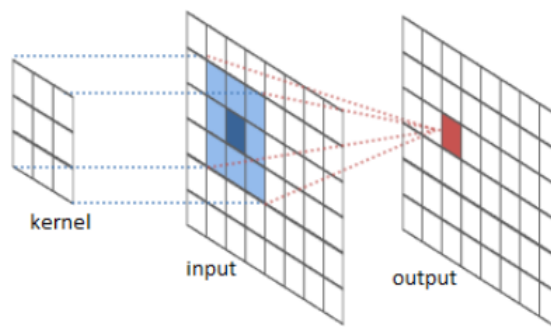


Figura 4: Kernel de a red convolucional

### 2.1.2. FaceNet

Ds

## 3. Introducción

### 3.1. Reconocimiento Facial

El reconocimiento facial se refiere a la tecnología capaz de relacionar una cara humana obtenida de una imagen digital o un vídeo en una base de datos de caras.



Figura 5: Sistema de reconocimiento facial en el metro de Osaka, Japón

A pesar de que la exactitud de este tipo de sistemas es menor que la de reconocimiento por iris o huella digital, estos han sido adoptados ampliamente, debido a su proceso, ya que este no requiere de contacto, esto ha llevado que sean mundialmente empleados por empresas y gobiernos.

### 3.2. Preliminares

Para lograr los objetivos deseados con el proyecto, se requiere un conjunto de datos correspondientes a los datos de entrenamiento que serán representados por las imágenes preprocesadas para el modelo, el conjunto de prueba correspondiente a una partición de las imágenes recolectadas de los integrantes de la generación, además de la arquitectura del modelo de aprendizaje.

La estructura del modelo de aprendizaje estará inspirada en el modelo de FaceNet, cambiando parámetros y modelos de la arquitectura, con la finalidad de encontrar una mejora en la predicción de las etiquetas (nombres de una persona dada su imagen). Habiendo entrenado el modelo, se utilizará la base de datos realizada para hacer la validación del modelo y las pruebas de rendimiento de este.

- 
4. Objetivo
  5. Alcances
  6. Marco Estratégico
  7. Alcances
  8. Definiciones
  9. ABT
  10. Modelo
  11. Resultados
  12. Conclusiones
- Referencias