

FaceTime - Deep learning based face recognition attendance system

En este documento se propone un nuevo sistema de asistencia basado en reconocimiento facial mediante aprendizaje profundo, aprovechando los avances recientes en las redes neuronales convolucionales profundas (CNNs) para tareas de detección y reconocimiento facial. Se detalla el proceso completo de desarrollo de un modelo de reconocimiento facial compuesto por pasos esenciales utilizando técnicas avanzadas actuales: una cascada de CNN para la detección facial y una CNN para generar representaciones faciales. El objetivo principal de esta investigación fue aplicar de manera práctica estos enfoques de aprendizaje profundo de vanguardia en tareas de reconocimiento facial. Dado que las CNNs logran los mejores resultados con conjuntos de datos más grandes, lo cual no es común en entornos de producción, el desafío principal fue aplicar estos métodos en conjuntos de datos más pequeños. Se propone un nuevo enfoque para la ampliación de imágenes en tareas de reconocimiento facial. La precisión general fue del 95.02% en un conjunto de datos reducido de imágenes originales de empleados en un entorno de tiempo real. Se sugiere que el modelo de reconocimiento facial propuesto podría integrarse en otro sistema con o sin algunas modificaciones menores como componente principal o de apoyo para fines de monitoreo y otra de sus debilidades sería la necesidad de que el sistema esté constantemente detectando y monitoreando en busca de los rostros de los empleados, considerando que en 3 meses fueron menos de 300 entradas/salidas por cada uno así que esto no es necesario.

Performance Evaluation and Comparison of Software for Face Recognition, Based on Dlib and Opencv Library

El artículo ofrece una visión general y un análisis de la complejidad temporal de los algoritmos de visión por computadora para el reconocimiento facial. Se centra en comparar dos bibliotecas populares de visión por computadora, OpenCV y dlib, explorando sus características, analizando sus ventajas y desventajas, y determinando en qué situaciones cada una es más adecuada. Se desarrollaron tecnologías de visión por computadora para el reconocimiento facial y se realizó una investigación exhaustiva de estas dos bibliotecas. Se analizan sus características, se evalúan sus pros y contras, y se presentan ejemplos de aplicaciones de reconocimiento facial basadas en histogramas orientados a gradientes para la detección facial, estimación de puntos de referencia para la orientación facial y redes neuronales convolucionales profundas para la comparación con rostros conocidos. El artículo resume el concepto de reconocimiento facial, describiendo la base científica para ello y la construcción de un sistema completo de reconocimiento. Se formulan los principios

básicos de los programas para el reconocimiento facial y se realiza un análisis comparativo de la productividad de ambas bibliotecas en relación con el tiempo de ejecución y el número de iteraciones de los algoritmos aplicados. Se crearon dos aplicaciones simples de reconocimiento facial utilizando estas bibliotecas y se comparó su rendimiento. En la evaluación del modelo, se midió el tiempo que los algoritmos de reconocimiento facial tardaban en identificar a las personas, observándose que tanto con OpenCV como con dlib el tiempo de reconocimiento era menor a 2 segundos, siendo menos de 1 segundo en el caso de OpenCV.

FAREC - CNN based efficient face recognition technique using Dlib

El artículo destaca la creciente atención que ha recibido el reconocimiento facial en las últimas décadas, tanto en la investigación como en los mercados comerciales. Propone una técnica eficiente para un sistema de reconocimiento facial basado en Deep Learning utilizando Redes Neuronales Convolucionales (CNN) con alineación facial Dlib. Describe detalladamente el proceso involucrado en el reconocimiento facial, como la alineación y extracción de características faciales. Además, resalta la importancia de la alineación facial, ya que la precisión y la Tasa de Falsas Aceptaciones (FAR) se observan utilizando la técnica propuesta. El análisis computacional muestra un rendimiento superior a otros enfoques de vanguardia. El estudio se realizó en el conjunto de datos del Face Recognition Grand Challenge (FRGC), logrando una precisión del 96% con una FAR de 0.1.

Se evaluaron la Tasa de Aceptación Verdadera (TAR) y la Tasa de Falsas Aceptaciones (FAR) en el conjunto de datos FRGC para el reconocimiento facial. La TAR se utiliza para medir la aceptación o reconocimiento correcto de rostros y es el porcentaje de veces que un sistema reconoce correctamente la identidad verdadera de una persona. La FAR es la medida de la probabilidad de que el sistema de reconocimiento acepte incorrectamente a una persona no autorizada y es la relación entre el número de aceptaciones falsas y el número total de intentos de reconocimiento.

Face image recognition based on convolutional neural network

El texto presenta el sistema FaceNet, que aborda los desafíos en la implementación eficiente y a gran escala de la verificación y reconocimiento facial, a pesar de los avances recientes en el campo del reconocimiento facial. FaceNet aprende directamente una representación de imágenes faciales en un espacio euclidiano compacto donde las

distancias reflejan la similitud facial. Este espacio permite la implementación sencilla de tareas como reconocimiento, verificación y agrupación facial utilizando incrustaciones de FaceNet como vectores de características. Se emplea una red convolucional profunda para optimizar directamente la incrustación, a diferencia de enfoques previos que usaban una capa intermedia como cuello de botella. El entrenamiento se realiza con tripletes de parches faciales alineados y relacionados/no relacionados, generados mediante un método de minería de tripletes en línea novedoso. La ventaja de este enfoque es una eficiencia representacional mucho mayor, logrando un rendimiento de reconocimiento facial de vanguardia utilizando solo 128 bytes por rostro. En el conjunto de datos Labeled Faces in the Wild (LFW), el sistema alcanza una precisión récord del 99.63%, y en YouTube Faces DB logra un 95.12%. Se reduce la tasa de error en un 30% en comparación con los mejores resultados publicados en ambos conjuntos de datos. La clasificación de precisión obtenida es de $98.87\% \pm 0.15$ con el método de corte fijo y un asombroso $99.63\% \pm 0.09$ con alineación adicional. Este rendimiento supera en más de siete veces el error reportado para DeepFace y en un 30% al estado del arte previo reportado para DeepId2+. Incluso el modelo más pequeño, NN3, alcanza un rendimiento que no es estadísticamente diferente.

Métricas Seleccionadas

Seleccionar métricas apropiadas para el desafío implica analizar las evaluaciones y métricas empleadas en los modelos actuales del estado del arte. Tras este análisis, considero que la combinación de la primera y segunda aplicación es la más adecuada para nuestra situación. Sin embargo, debido a la limitación de no poder verificar la asistencia a largo plazo, propongo realizar pruebas de detección y reconocimiento facial a diversas distancias. Se medirá el tiempo de reconocimiento correcto o incorrecto de personas en videos de 30 segundos a distancias variables: 1, 2, 3 y hasta 4 metros.

Resultados

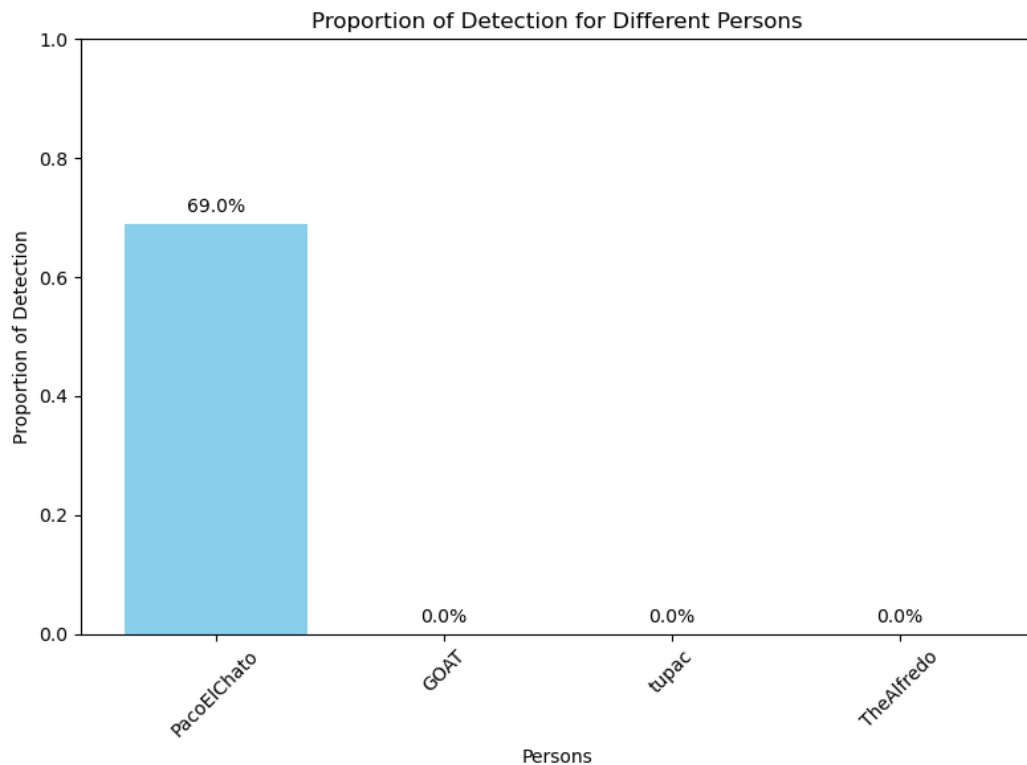
Para los resultados hicimos un video de 18 segundos en el que el alumno se encontraba a distancias variables en condiciones de luz complicadas y girando la cabeza para ocasionar ángulos muertos al algoritmo de face recognition, a pesar de todas estas condiciones adversas obtuvimos un 69.02% de acierto en la clasificación de los frames del video, cabe recalcar que este fue un video y en el caso del reto se planea hacer fotografías esporádicas para verificar la asistencia, algo que es mucho más rápido y fiable, quizá con la indicación de ver a la cámara directamente para maximizar el desempeño del modelo.

Comparaciones

Podemos ver que a comparación de los demás modelos el nuestro resultó un poco por debajo de la media debido a las condiciones del experimento, contra el caso más similar que es el de detectar personas que entran a un cuarto la diferencia es que ahí se les indica que volteen a ver a la cámara, influyendo directamente en la precisión del modelo. Comparando con las otras

soluciones no es tan directo porque las condiciones son muy distintas y lejanas a la realidad, siendo así que nuestro algoritmo en uno de los datasets famosos como Labeled Faces in the Wild benchmark obtuvo un 99.38% de precisión.

Gráficas y conclusiones



Algo que podemos destacar de este gráfico es que nuestro algoritmo solamente detectó correctamente a Paco ~70% de las veces (frames) en los que se detectó un rostro, sin embargo podemos destacar que el algoritmo detectó correctamente a quien se encontraba en el frame un 100% de las veces, es así que podemos darnos cuenta de que nuestro algoritmo funciona muy bien para hacer el face recognition, pero es necesario trabajar o enfocarnos un poco más en detectar caras en ángulos y condiciones de luz complicadas, sin embargo, para un entorno como lo es un aula de clases con condiciones de luz de buena calidad y mejor aún constante, al igual que los rostros de los estudiantes mirando hacia delante la mayor parte del tiempo creemos que nuestra solución cumple satisfactoriamente con solucionar la problemática de pasar asistencia en un salón de clases.