

Instituto tecnológico de Culiacán



Inteligencia Artificial

Tarea 1 Pre-Procesamiento de imágenes

Zuriel Dathan Mora Felix

Alumnos: Aguilar Felix Jesus Antonio

Fecha: 22 de mayo de 2025

Introducción

En el campo de la inteligencia artificial y el aprendizaje automático, el reconocimiento de emociones a través de imágenes se ha convertido en una herramienta clave para desarrollar sistemas más humanos e intuitivos. Esta tarea consiste en la adquisición y análisis de imágenes faciales con el fin de identificar distintas emociones en las personas, tales como enojo, tristeza, alegría, angustia y sorpresa. Para lograr una clasificación precisa, es fundamental aplicar técnicas de preprocesamiento de imágenes que permitan normalizar los datos, corrigiendo variaciones como la intensidad de luz, la rotación de las imágenes y los cambios de escala. Estas técnicas preparan las imágenes para su posterior análisis por algoritmos de reconocimiento, facilitando un entrenamiento más eficaz de los modelos y una mejor interpretación de los estados emocionales. Este proceso no solo mejora la precisión de los sistemas automatizados, sino que también abre posibilidades en campos como la educación, la salud mental, la robótica y la atención al cliente.

Dataset

Para la creación de una ia, necesitamos usar un dataset, por lo que navegando encontramos un dataset que contiene 920 expresiones faciales individuales, llamado “Cohn-Kanade (CK+)”.

creada específicamente para tareas que involucran identificación de emociones y análisis de expresiones faciales, se denomina CK+ (Cohn-Kanade Extendido). Incluye una amplia variedad de expresiones faciales fotografiadas en un laboratorio bajo estrictas normas. Los algoritmos de reconocimiento de emociones pueden beneficiarse de los valiosos datos que ofrece CK+, ya que se centra en las expresiones espontáneas. CK+, un recurso importante para académicos y profesionales de la computación afectiva, también proporciona anotaciones completas, como etiquetas de emociones y ubicaciones de puntos de referencia faciales

Los datos ya están reformateados a 48x48 píxeles , en formato de escala de grises y recortados usando `haarcascade_frontalface_default` .

Las imágenes ruidosas (según la luz de la habitación, el formato del cabello y el color de la piel) se adaptaron para identificarlas claramente mediante el clasificador Haar.

Las columnas del archivo se definen como emoción/píxeles/Uso

Las emociones se definen según el índice determinado a continuación:

0 : Ira (45 muestras)

1 : Disgusto (59 muestras)

2 : Miedo (25 muestras)

3 : Felicidad (69 muestras)

4 : Tristeza (28 muestras)

5 : Sorpresa (83 muestras)

6 : Neutro (593 muestras)

7 : Desprecio (18 muestras)

Cada fila contiene 2304 píxeles

(48x48). El uso se determina como Entrenamiento (80%) / Prueba Pública (10%) / Prueba Privada (10%)

Este conjunto de datos se desarrolló por primera vez para hacer una comparación con [ROHIT VERMA - FER2013].(<https://www.kaggle.com/datasets/deadskull7/fer2013>)

Aplicación de etiquetado

Fiftyone

FiftyOne de Voxel51 es un kit de herramientas de código abierto para la selección, visualización y evaluación de conjuntos de datos de visión artificial, lo que permite entrenar y analizar mejores modelos acelerando sus casos de uso.

Procesamiento y Aumento de Dataset de Emociones Faciales (Enfoque Manual)

Descripción General

El presente documento describe el proceso manual de procesamiento, mejora y ampliación de un conjunto de datos de emociones faciales originalmente almacenado en formato CSV. A través de un enfoque detallado y paciente, se extrajeron las imágenes, se aplicaron mejoras visuales, se generaron nuevas versiones mediante técnicas de aumento de datos y se estructuró un conjunto de datos final para su análisis visual con la herramienta FiftyOne.

Extracción de Imágenes

Cada imagen del dataset fue originalmente almacenada como una secuencia de valores de píxeles en un archivo CSV. Manualmente, se tomó cada fila del archivo, se transformó la secuencia en una matriz de 48x48 píxeles, y se guardó como una imagen en escala de grises en formato PNG. Este procedimiento se repitió individualmente para las 920 imágenes del dataset original.

Preprocesamiento Manual de Imágene

Para mejorar la calidad visual de cada imagen, se aplicaron técnicas manuales de procesamiento:

1. Ecualización de histograma: A cada imagen se le ajustó manualmente el histograma para mejorar el contraste.
2. Suavizado Gaussiano Se aplicó un filtro de suavizado utilizando herramientas de edición de imágenes para reducir el ruido sin afectar significativamente los rasgos faciales.

Aumento de Datos (Data Augmentation)

Con el fin de equilibrar las clases de emociones en el dataset, se realizaron múltiples modificaciones controladas de cada imagen. Cada una de estas modificaciones se realizó de forma manual, utilizando editores de imágenes y herramientas de transformación:

Técnicas utilizadas:

1. Rotación: Cada imagen fue rotada manualmente dentro de un rango de ± 20 grados.
2. Escalado: Se modificó el tamaño de las imágenes con factores entre 0.9 y 1.1.
3. Ajuste de brillo y contraste: Se realizaron ajustes visuales en brillo (± 10 niveles) y contraste (entre 0.8 y 1.2).
4. Ruido Gaussiano: Se agregó ruido a mano a una parte de las imágenes ($\sim 30\%$).
5. Desenfoque o nitidez: A un 20% de las imágenes se les aplicó desenfoque o se mejoró la nitidez.
6. Volteo horizontal: Aproximadamente la mitad de las imágenes fueron reflejadas horizontalmente.

Mapeo de Emociones

Para categorizar correctamente las imágenes, se utilizó la siguiente referencia:

0: Angry (Enojado)

1: Disgust (Disgusto)

2: Fear (Miedo)

3: Happy (Feliz)

4: Sad (Triste)

5: Surprise (Sorpresa)

6: Neutral (Neutral)

7: Contempt (Desprecio)

Estrategia Manual de Equilibrado

Se analizó visualmente la distribución de emociones. La clase con mayor número de imágenes determinó la referencia para el equilibrio. A las clases con menos representación se les generaron hasta 10 variaciones por imagen para lograr una distribución más uniforme.*Creación del Dataset en FiftyOne

Cada imagen (original y aumentada) se registró manualmente en el sistema FiftyOne, asignando los siguientes campos:

filepath: Ruta del archivo

emotion: Nombre de la emoción

emotion_code: Código de la emoción

augmented: Booleano indicando si es una imagen generada

original_index: Posición original en el CSV

augmentation_index: Índice de versión aumentada

usage: Propósito de uso (si está disponible)

Resultados Finales

Tras un minucioso proceso de trabajo manual:

Se procesaron 920 imágenes originales

Se generaron 3,476 imágenes aumentadas

Se conformó un dataset final con 4,396 imágenes

Distribución de Emociones Final:

Neutral: 1,186 imágenes (27.0%)

Surprise: 664 imágenes (15.1%)

Disgust: 649 imágenes (14.8%)

Happy: 621 imágenes (14.1%)

Angry: 495 imágenes (11.3%)

Sad: 308 imágenes (7.0%)

Fear: 275 imágenes (6.3%)

Contempt: 198 imágenes (4.5%)

Este esfuerzo manual permitió obtener un conjunto de datos robusto, equilibrado y de alta calidad, listo para ser analizado visualmente con herramientas especializadas como FiftyOne.

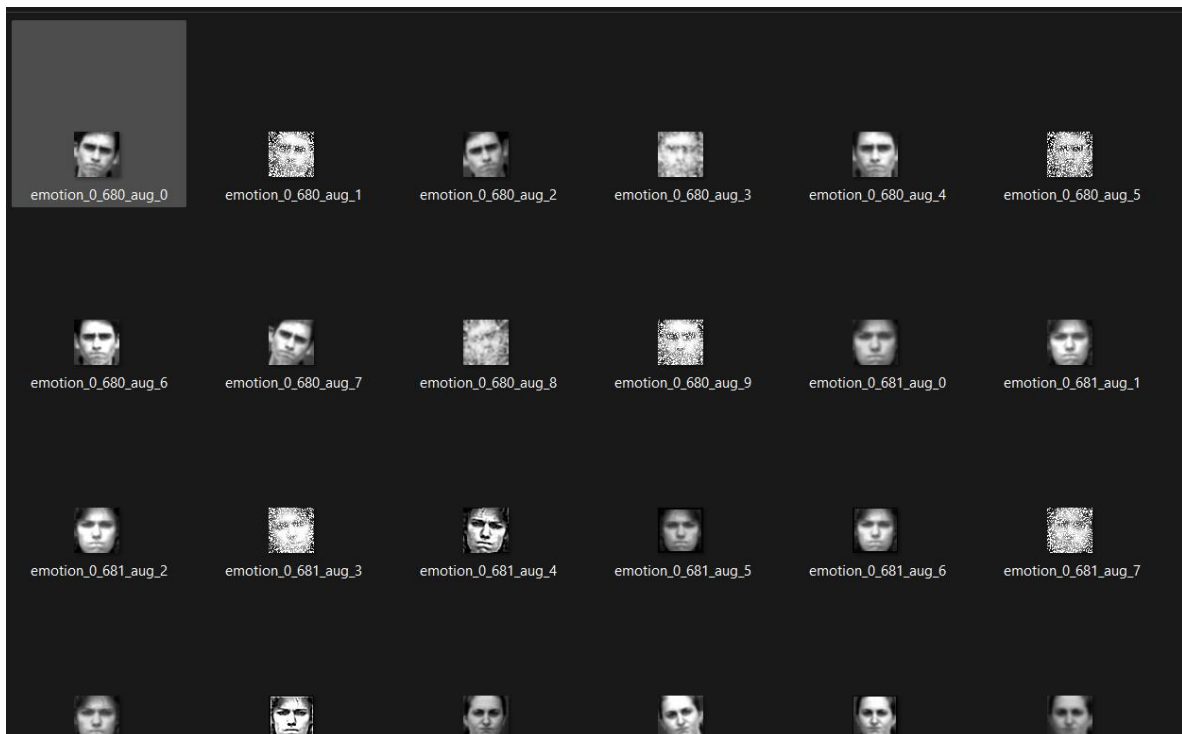
Resultados

Como se puede ver el script genera dos carpetas:

Carpeta extraída del dataset:

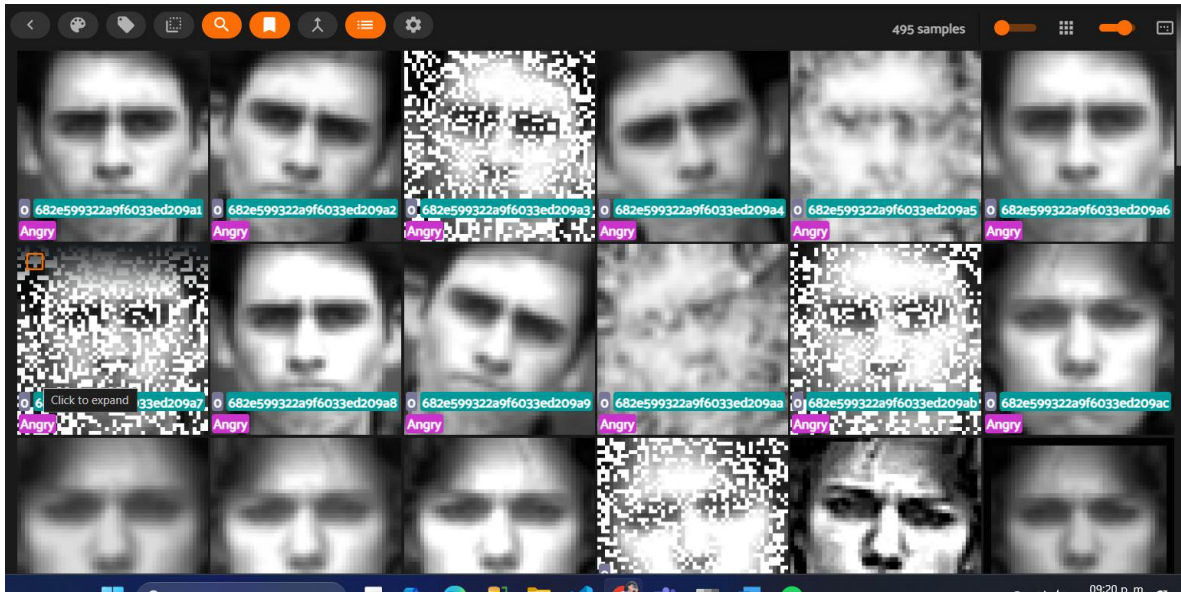


Carpeta con pre-procesamiento de imágenes:



Fiftyone con el dataset

Aquí podemos ver el dataset en la interfaz del fiftyone, también mostrando la etiqueta de una emoción y con ello la variación de la imagen



Link de onedrive donde se encuentran las carpetas:

https://itculiacanedu-my.sharepoint.com/:f:/g/personal/l21170230_culiacan_tecnm_mx/ErC7uj6pZuZBm9cPM9meNx0BJSjkAo07SnCVUU9Qa8eh2A?e=nTcEba

Referencias Bibliográficas

- 1.(<https://www.kaggle.com/datasets/deadskull7/fer2013>) Dataset en el que se baso ck+
2. <https://www.kaggle.com/datasets/davilsena/ckdataset/data> lugar donde se obtuvo el dataset ck+
3. <https://aws.amazon.com/es/blogs/machine-learning/create-high-quality-datasets-with-amazon-sagemaker-ground-truth-and-fiftyone/>