

Predicción de letras del alfabeto en Lenguaje de Señas Español mediante aprendizaje automático

Juan Fernando Cardona Ruiz
Daniela Vásquez Londoño

Fundamentos de Deep Learning
Facultad de Ingeniería
Departamento de Ingeniería de Sistemas



Abril 2024

Contexto de aplicación

El propósito de esta aplicación es predecir la definición correcta asociada a una letra del alfabeto del Lenguaje de Señas Español (ESL), a partir de los gestos realizados con las manos de una persona frente a una cámara. Este enfoque es fundamental para facilitar la comunicación con personas que utilizan el lenguaje de señas y no cuentan con un intérprete disponible.

Objetivo de Machine Learning

El objetivo principal consiste en desarrollar un modelo de aprendizaje automático capaz de reconocer y predecir de manera precisa y eficiente la letra correspondiente al gesto realizado con las manos en el Lenguaje de Señas Español (ESL). Esto implica la utilización de técnicas avanzadas de aprendizaje automático, como el procesamiento de imágenes y redes neuronales, para interpretar los gestos y asignarles una letra específica del alfabeto ESL.

Dataset

El conjunto de datos utilizado en este proyecto consta de imágenes a color que representan gestos realizados con las manos para cada letra del alfabeto del Lenguaje de Señas Español (ESL). Este dataset está organizado en 19 subcarpetas, cada una representando una letra del abecedario ASL, con entre 90 y 110 imágenes JPG en cada una, representando un total de 2017 imágenes.

Métricas de Machine Learning:



1. **Exactitud (Accuracy):** Es la proporción de predicciones correctas respecto al total de predicciones realizadas. En este caso, sería la cantidad de letras del ESL correctamente clasificadas sobre el total de imágenes.
2. **Precisión (Precision):** Mide la proporción de predicciones positivas (letras clasificadas correctamente) entre todas las predicciones y falsos positivos. Es importante para evitar falsas alarmas.
3. **Recuperación (Recall o Sensibilidad):** Representa la proporción de predicciones positivas correctas respecto a todas las instancias verdaderamente positivas. Es relevante para no perder casos positivos.
4. **F1-Score:** Combina precisión y recuperación en una sola métrica. Es útil cuando hay un desequilibrio entre las clases.
5. **Matriz de Confusión:** Proporciona una visión detallada de los aciertos y errores en la clasificación de cada letra.
6. **Curva ROC (Receiver Operating Characteristic):** Evalúa el rendimiento del modelo en función de la tasa de verdaderos positivos frente a la tasa de falsos positivos.

Métricas de Negocio:

1. **Tiempo de Inferencia:** El tiempo que tarda el modelo en predecir una letra a partir de un gesto. Debe ser rápido para aplicaciones en tiempo real.
2. **Precisión en la captura:** Precisión al llevar las predicciones realizadas a un texto como archivo de salida.

Referencias y resultados previos

Para desarrollar este proyecto, se han consultado diversas fuentes y resultados previos relacionados con el reconocimiento de gestos y el aprendizaje automático aplicado al Lenguaje de Señas Español (ESL). Algunas de las referencias relevantes incluyen:

- Spanish Sign Language Alphabet (Static). (2022, 22 julio). Kaggle. <https://www.kaggle.com/datasets/kirlelea/spanish-sign-language-alphabet-static>
- Model Training: Recurso que proporciona información detallada sobre el entrenamiento de modelos utilizando técnicas específicas, como el callback de decaimiento de peso. Minnikeswarrao. (2024, 4 febrero). *Model training*. Kaggle. <https://www.kaggle.com/code/minnikeswarrao/model-training#Weight-Decay-Callback>
- Análisis Exploratorio de Datos (EDA): Recurso que ofrece una visualización animada del conjunto de datos ASLFR, facilitando la comprensión de su distribución y características. Leonidkulyk. (2023, 12 mayo). *[EDA] 🙌ASLFR🙌 - Animated visualization*  . Kaggle. <https://www.kaggle.com/code/leonidkulyk/eda-aslfr-animated-visualization>
- Reconocimiento de deletreo de señas ASL con TensorFlow: Implementación práctica del reconocimiento de deletreo de señas utilizando la biblioteca TensorFlow. Gusthema. (2023, 5 julio). *ASL Fingerspelling Recognition w/ TensorFlow*. Kaggle. <https://www.kaggle.com/code/gusthema/asl-fingerspelling-recognition-w-tensorflow>
- Reducción del tiempo de entrenamiento - Redimensionamiento primero: Estrategia propuesta para reducir el tiempo de entrenamiento del modelo, manteniendo un alto rendimiento. Gpiosenka. (2022, 23 julio). *Reduce Train Time-Resize First F1 score=100%*. Kaggle. <https://www.kaggle.com/code/gpiosenka/reduce-train-time-resize-first-f1-score-100>

Estas referencias y resultados previos servirán como guía y base para el desarrollo y la implementación del modelo de aprendizaje automático destinado a predecir letras en el Lenguaje de Señas Español (ESL).