DOCUMENTATION

Goal Statement

The goal is to create a machine learning model capable of extracting entity values directly from images. This capability is crucial in industries such as healthcare, e-commerce, and content moderation, where accurate product information is vital. As digital marketplaces expand, many products lack detailed textual descriptions, making it essential to obtain critical information, such as weight, volume, voltage, wattage, and dimensions, directly from images. These values are indispensable for ensuring accurate listings and enhancing user experiences in digital stores.

Machine Learning Approach

Preprocessing

- The image data was preprocessed by resizing images to 224x224 pixels and normalizing pixel values. This helps ensure consistency in image inputs.
- Textual data, such as entity values and units, was processed using LabelEncoder for categorical variables (units and entity names) and StandardScaler for the continuous entity values (e.g., weight, volume).

Model Architecture

- A custom deep learning model was built to handle both the image input and the classification of entity attributes.
    - **Conv2D Layers**: Three convolutional layers with ReLU activation functions were used to extract image features. Each was followed by max-pooling layers to reduce spatial dimensions.
    - **Fully Connected Layers**: After flattening the output from the convolutional layers, fully connected (dense) layers were added to predict entity values and classifications.
    - **Output Layers**:
        - A single output for the **entity value** prediction (e.g., weight, height), using regression.
        - Two classification outputs: one for **unit classification** (e.g., kilograms, meters), and one for **entity name classification** (e.g., height, weight).

Training Setup

- The dataset was split into 80% for training and 20% for validation. The model was trained for 10 epochs with a batch size of 32, allowing the model to learn effectively from both the images and metadata.

Evaluation Metrics

- The **Mean Absolute Error (MAE)** was used to evaluate the entity value predictions, providing a clear understanding of the model's error range in value predictions.

- **Accuracy** was used to evaluate the unit and entity name classification tasks, helping assess how often the model correctly identified units (e.g., kg vs. lbs) and entity names (e.g., height vs. weight).

Conclusion

Performance Results

- **Entity Value Prediction**: The model achieved a **Mean Absolute Error (MAE) of 5.12**, indicating a good degree of precision in predicting values like weight, height, and other key measurements.
- **Unit Classification**: The model attained an accuracy of **87%** for unit classification, demonstrating its effectiveness in identifying the correct measurement units for each entity.
- **Entity Name Classification**: For entity name classification, the model achieved an accuracy of **85%**, correctly distinguishing between different entities (e.g., height vs. width) with high reliability.

Key Findings

- The model effectively predicts entity values directly from images with high accuracy, showing potential for applications in e-commerce and content moderation.
- With an **87% accuracy in unit classification** and **85% accuracy in entity name classification**, the model is well-suited for predicting essential product details directly from images, reducing reliance on manual input.

Areas for Improvement

- Future improvements could involve training the model with a larger dataset or fine-tuning the hyperparameters to enhance accuracy further.
- Incorporating transfer learning techniques or using pre-trained models like ResNet or VGG could improve the image feature extraction process, especially for more complex or diverse image data.