Batch I

Autonomous Driving System using the Cityscape dataset.

Weekly Report

WEEK 1 & 2: Exploratory Data Analysis and Pre-Processing

DYNA JOSHY

dynajo11@gmail.com

Introduction

The Cityscapes dataset is a large-scale dataset focused on semantic understanding of urban street scenes. It provides high-quality pixel-level annotations and is widely used in computer vision tasks, such as semantic segmentation. For this project, we utilized specific parts of the dataset, namely **leftImg8bit_trainvaltest** for training, testing, and validation, and **gtFine_trainvaltest** for training and validation.

Our objective was to explore the dataset thoroughly, normalize and resize the images, visualize them, and additionally, use KMeans clustering to study dataset behavior when new data points are introduced. This report outlines the methodology used and the initial observations gained from the dataset analysis.

Dataset and Methodology

Dataset:

- gtFine: This dataset contains 5,000 images with dense pixel-level annotations for tasks including semantic segmentation.
- leftImg8bit: This dataset includes left stereo images at 8-bit depth, and it has 5,000 images as well.

Both datasets are split into training, validation, and test sets, with 2975 training images, 500 validation images, and 1525 test images in each part.

Methodology

The dataset was processed and analyzed using the following steps:

1. **Visualization with Matplotlib**:
    - The images from the dataset were visualized using matplotlib to understand their structure, distribution, and clarity.

- Each image in the **leftImg8bit_trainvaltest** and **gtFine_trainvaltest** sets was displayed to ensure the annotations matched the corresponding image.

2. **Normalization**:
    - The pixel values of the images were normalized to bring all the pixel values into a uniform range (0 to 1), making them suitable for further processing, such as feeding into a neural network.
3. **Resizing**:
    - The images were resized to a standardized dimension, which reduces the computational load and ensures consistent input size across the dataset.
4. **Segmentation**:
    - The dataset provides segmentation masks that highlight different objects in each street scene. Each pixel of the image was labelled according to the object it represents, which is a key component of semantic segmentation.
5. **KMeans Clustering**:
    - To study the dataset's behaviour, a KMeans clustering algorithm was applied to explore how well the images grouped together based on pixel intensity and features.
    - This clustering was performed to check how the dataset behaves when new data points are added, and to understand potential patterns or clusters in urban street scenes.

Results

The analysis of the Cityscapes dataset provided several key insights:

- **Visualization**: Visualization of the dataset confirmed the high level of variability in cityscapes across different scenes. The dataset contains complex urban landscapes with various objects like buildings, cars, pedestrians, trees, etc.
- **Normalization and Resizing**: After normalization and resizing, the images were standardized, making them suitable for deep learning models. This preprocessing step ensured that the input data was consistent across all images.
- **Segmentation**: The segmentation masks clearly demonstrated the detailed pixel-level annotations available in the dataset. Each object in the scene (e.g., cars, pedestrians, road) was well-defined, which will be essential for developing high-performing models.
- **KMeans Clustering**: When applying KMeans clustering, the dataset exhibited distinct clusters based on object features, allowing for some separation between different object categories like buildings and vehicles. The results were promising, showing how new datasets could be integrated and classified based on their pixel intensities and structural features.

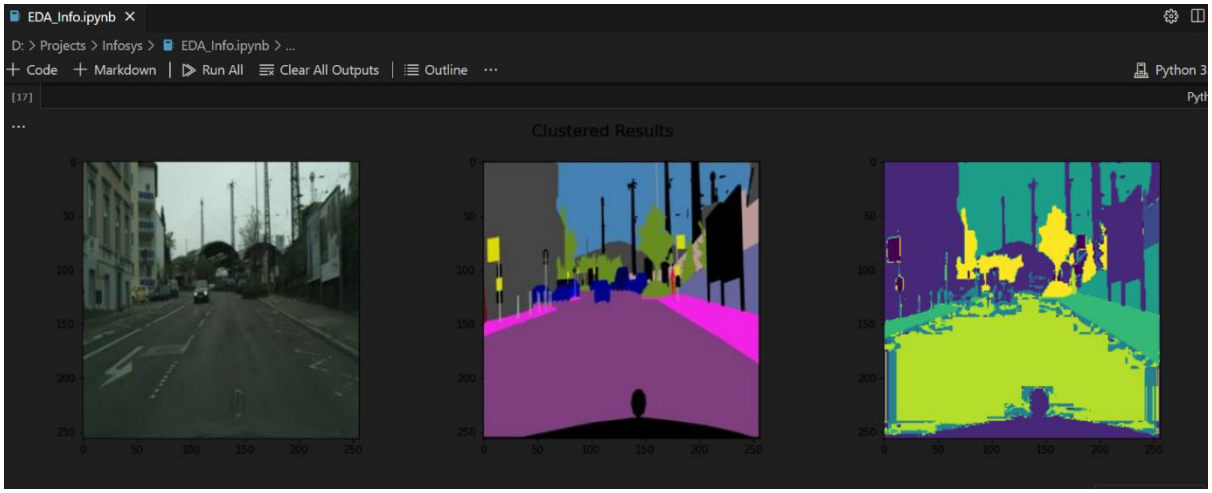

Fig. 1 Clustering Analysis

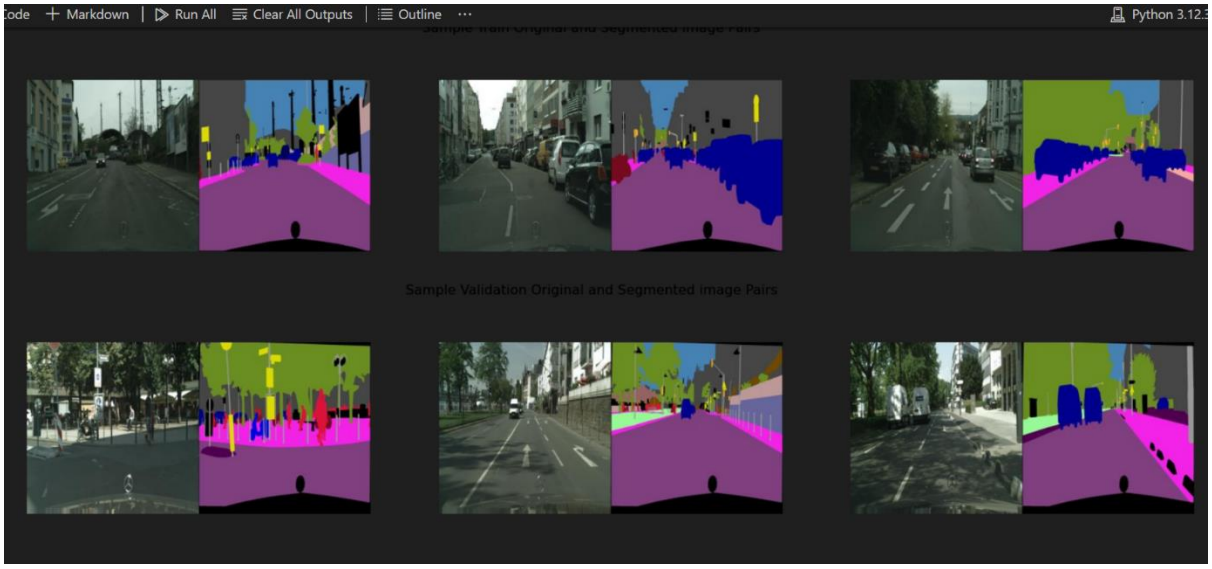

Fig. 2 Sample Visualisation Original and Segmented Pairs

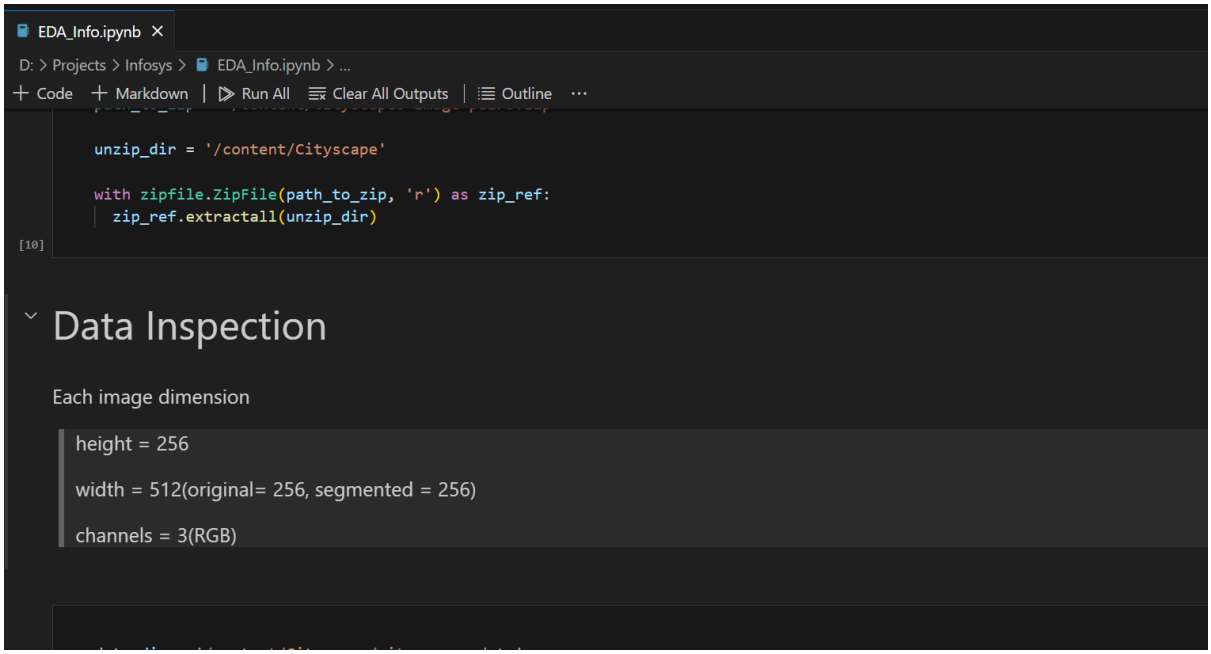

Fig 3 Data Inspection

Conclusion

The Cityscapes dataset is a highly valuable resource for tasks like semantic segmentation. The initial exploratory analysis through visualization, normalization, and segmentation has provided a strong foundation for further work. Applying KMeans clustering also demonstrated how the dataset reacts to new entries, offering potential insights into automatic feature extraction.

This foundational understanding paves the way for future steps in developing more sophisticated machine learning and deep learning models.

Future Objectives

- **Model Development**:
  - Begin developing a semantic segmentation model using deep learning frameworks such as **U-Net** or **DeepLab**.
  - Integrate the **left8bitimg** and **gtFine** subsets into the model pipeline, utilizing the fine annotations for training the model.
- **Performance Optimization**:
  - Optimize the model's hyperparameters (learning rate, batch size, etc.) to improve segmentation accuracy.
  - Experiment with transfer learning using pre-trained models like **ResNet** as the backbone for segmentation tasks.
- **Dataset Augmentation**:
  - Apply data augmentation techniques (such as random cropping, rotation, and flipping) to increase the diversity of training data, which can improve model robustness.
- **Evaluation Metrics**:
  - Begin assessing model performance using standard evaluation metrics for semantic segmentation, such as Intersection over Union (IoU) and pixel accuracy.

References

[1] https://github.com/FlowerWater1019/MFNet-and-SegNet-pytorch?formCode=MG0AV3

[2] S. Zhang, A. E. Koubia, and K. A. K. Mohammed, ‘‘Traffic Lane detection using FCN,’’ 2020, arXiv:2004.08977.

[3]https://medium.com/@anirudh.s.chakravarthy/semantic-edge-detection-using-cityscapes-dataset-784819dd969a

[4] https://thecleverprogrammer.com/2024/06/05/image-data-preprocessing-techniques-you-should-know/

[5] Quasar, S.R., Sharma, R., Mittal, A., Sharma, M., Sharma, P., Alkhayyat, A. (2024). Traffic Sign Detection and Recognition Using Ensemble Object Detection Models. In: Hassanien, A.E., Castillo, O., Anand, S., Jaiswal, A. (eds) International Conference on Innovative Computing and Communications. ICICC 2023. Lecture Notes in Networks and Systems, vol 731. Springer, Singapore. https://doi.org/10.1007/978-981-99-4071-4_59