Phase 1 Report

I. Introduction and Problem Statement

Our project aims to simplify tourists' exploration by creating a model that categorizes photos of various tourist attractions. By identifying landmarks like airports, bus stations, canyons, temples, and markets from images, we aim to streamline travel planning, benefiting the local tourism sector. Since tourist landmarks come in diverse forms and images are taken from various angles, establishing standard classification rules is challenging. To tackle this, we employed decision trees (DT) in Phase 1. DTs are relatively easy to understand but require identifying image features for classification, which can be complex. Our team inspected images to identify common characteristics of each class and used these to preprocess images, such as by forming object edges. More details are provided in Section 2.

In Phase 2, we plan to explore convolutional neural networks (CNNs) **[5]**, a deep learning method well-suited for image classification, to enhance the program's accuracy and robustness.

II. Proposed Methodologies

Our Dataset is called the Places365 **[1]** it consists of 365 different classes and 5000 images for each class. After going through all the classes we selected 5 classes, namely airfield, bus stand, canyon, market, and temple, as per project use as it is comparatively more attractive to tourists. Each class contains 5000 color images of size 512x683 pixels. The total size of our finalized dataset is 1.52 GB and the average size of the image is about 65kB. All of them are outdoor scenes and manmade except the canyon which is natural.

- <u>Supervised Learning</u>

For supervised learning we first wanted to narrow down the feature we wanted to use for decision-making. After enough research and trial and error we decided to use color histograms as the feature set. After extracting the features and storing it in a pickle file, we selected 2 hyperparameters that we saw after trial and error that had the most impact on the classifier. They were max features and max depth. After selecting the most ideal hyperparameter values we trained the classifier on our training dataset.

- <u>Semi-supervised Learning</u>

All the steps in Supervised learning for extracting features and determining the hyperparameters are used for this approach. We used the Train.pkl as a dataset for Semi-supervised learning and divided this set into two parts, The first one is labeled data which consists of 20% of the training set, and the remaining 80% is allocated for unlabeled data. Initially, we trained the Decision tree classifier using the labeled data and identified pseudo labels for unlabeled data. The model is retrained to achieve accurate results using the most confident pseudo-labels.

III. Evaluation of Results

Supervised Learning: The dataset is divided into 3 sets with 60% for training, 20% for validation, and 20% for testing.

Evaluation	Result
Accuracy	.4358
Precision	.4376
Recall	.4358
F1-score	.4340

The trial results are shown in **Figures 1 to 3**.

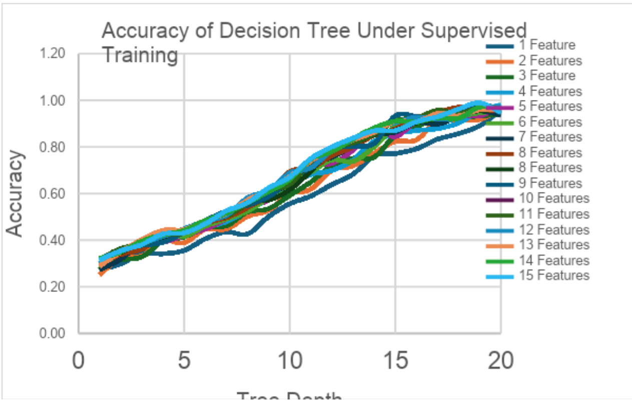

Figure 1 Accuracy of DT under supervised training with different tree depths and number of features

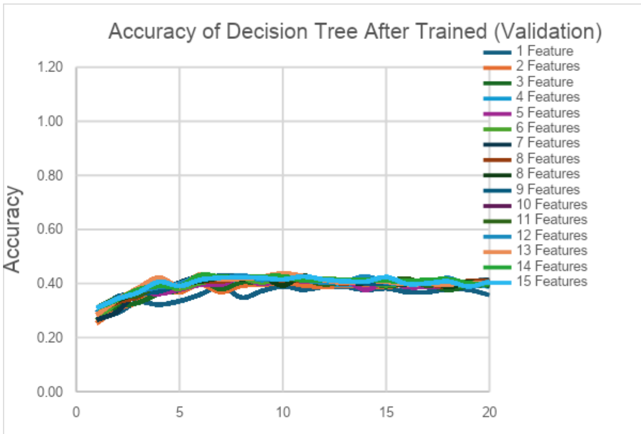

Figure 2 Accuracy of trained DT in validation stage with different tree depth and number of features

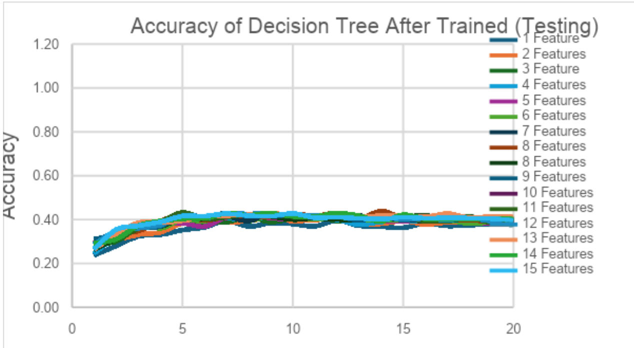

Figure 3 Accuracy of DT tree in the testing stage with different tree depths and number of features

<u>The results showed the following characteristics:-</u>

- In general, the accuracy of the decision tree was increased with increasing in number of features. The increase is quite linearly to a max. of 15 features. The max. training accuracy could achieve over 0.98
- The number of features has only a minor impact on the training accuracy. Thus, the number of features can be minimized to simplify the computational effort. Anyway, it is proposed to have the DT with a minimum number of features, say 5, in order to minimize the data variations.
- For the validation stage, the max. accuracy only achieved to about 0.4 which was considered low. Furthermore the accuracy reaches to this max. value if the tree depth was increased to about 6. Further increase of the tree depth could not increase the accuracy. Similar to the training stage, the accuracy of the DT would not be significantly affected by the number of features.
- In testing, the performance of the trained DT is comparable to that of the validation. The max. accuracy is about 0.4 which was considered low.

Semi-Supervised Learning: The Dataset is divided into 20 % labeled data, and 80% unlabeled data.

	precision	recall	f1-score	support
airfield	0.52	0.40	0.45	600
bus stand	0.30	0.38	0.34	600
canyon	0.33	0.34	0.33	600
market	0.40	0.38	0.39	600
temple	0.30	0.29	0.29	600
accuracy			0.36	3000
macro avg	0.37	0.36	0.36	3000
weighted avg	0.37	0.36	0.36	3000

Figure 4 Precision, recall, f1-score, support

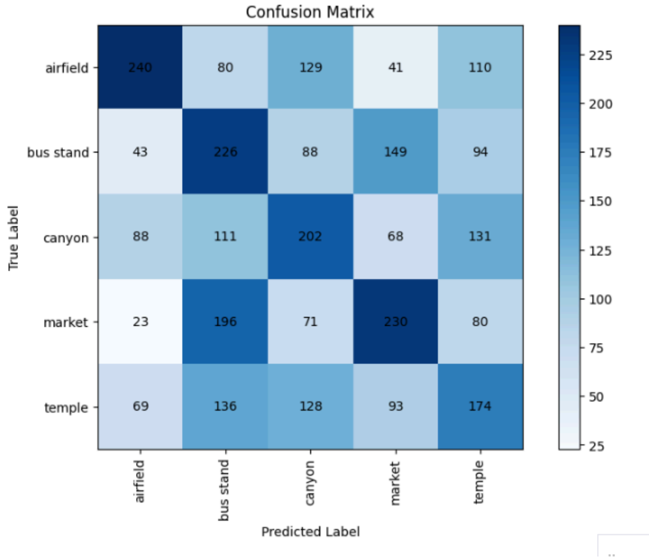

Figure 5 Confusion Matrix

IV. Future Improvements

At this stage, DT is being used for image classification. Based on our results, DT might not be the best method for this purpose. In the first attempt, a color histogram was used to investigate the color distributions of the images. With the color distribution, the images were classified according to the color distribution patterns. However, the accuracy of this method could only achieve approximately 40%, which is relatively low. It implies that DT, as a conventional machine learning technique, has its own limitations in image classification. Thus, it triggered the project's Phase 2 to use another method, namely convolutional neural network (CNN), which is a kind of deep learning, for image classification. In many literature and reference books, CNN is particularly effective for this purpose.

References

[1] Dataset - http://places2.csail.mit.edu/download-private.html

[2] B. Jijo and A. Abdulazeez, Classification Based on Decision Tree Algorithm for Machine Learning, Journal of Applied Science and Technology Trends, Vol.02, No. 01, pp. 20-28 (2021)

[3] V. Costa and C. Pedreira, Recent Advances in Decision Trees: an Updated Survey, Artificial Intelligence Review, 56:4765-4800, (2023)

[4] B. Zhou et al, Object Detectors Emerge in Deep Scene CNNs, Conference paper at ICLR 2015, pp. 1-10, (2015)

[5] https://franky07724-57962.medium.com/semi-supervised-learning-and-image-classification-3c3d3f57cb52