Tree Crown Classification using Grey-Level Correlation Matrix

Christopher Arif Setiadharma
School of Computer Science and Engineering, Nanyang Technological University

Supervised by Dr Ji-Jon Sit
School of Electrical and Electronics Engineering, Nanyang Technological University

Abstract - Forest inventory-keeping and species identification is crucial to formulate forest management strategies and decision-making in urban planning. However, it requires significant on-the-ground manual identification by professionals as large forests may be vast and inaccessible. Most current studies require the use of expensive and complex sensors such as LIDAR to perform similar work. While other studies have achieved moderate success using computer vision, a large amount of labelled data is required for effective training of machine learning models to achieve robustness. Public datasets for tree species in Singapore forests are unavailable. Therefore, we use digital aerial photogrammetry to collect our own dataset. Each tree crown is then labelled by a professional. Up to this moment, the size of our labelled dataset is still limited.

From spectral bands obtained from drone cameras and using Grey-Level Co-Occurrence Matrix (GLCM) we generate channels for each pixel of the spectral bands. Due to small number of sample per species, we sub-divide the images of the species into smaller images to generate more sets of GLCM-derived features for each species. We use random forest classifier to classify trees based on their GLCM-derived features. Although we find that our classifier is able to recall some species of tree crowns, we find that our method is not yet generalizable across datasets.

Keywords - GLCM, Texture Analysis, Machine Learning, Forestry

1.1. INTRODUCTION

All decision-making, including forestry, requires data. Systems for measuring the extent, quantity, and condition of forest provide the necessary information[1]. Conventionally, tree crown information is collected through field surveys, which are labor-intensive and time-consuming, and limited to a small-scale sample plot [2].

This project focuses on developing a machine learning model that is able to perform tree crown classification in bird-eye view images (Figure 1) taken from Unmanned Aerial Vehicles (UAV, or known as drone) instead of manual identification in Singapore forests. In future, we hope to develop a semi-automatic solution to assist forest restoration work end-to-end in forecasting restoration outcomes, site prioritisation, recommending actions and logistics, and monitoring for adaptive feedback.

To our best knowledge, no open-sourced labelled data of target tree crowns is currently available, as our target tree crowns might be very specific to Singapore. Hence, we are working with our own dataset collected from drone missions at Chestnut Nature Park, Singapore with 18 instances of 15 unique species labelled by forest ecology professionals in our team. The current state-of-the-art for computer vision classification and its downstream tasks revolves around deep learning with focus on Convolutional Neural Networks (CNNs). It requires a large dataset, larger than what we currently have in hand [3]. Even more recently, Vision Transformers is pushing the state-of-the-art by purely using attention mechanism and beating CNNs, albeit only with a very large dataset (>10M images)[4][5].

This paper propose a workflow of classifying tree crowns by first generating texture features using Grey-Level Co-occurrence Matrix (GLCM) for each of the spectral bands from cameras mounted on a drone. After feature extraction using GLCM, we passed the features through a random forest classifier to classify the tree instances. In a concurrent paper, we explore the use of FaceNet. At its best configuration, the random forest classifier achieved 54% overall accuracy, with 8 out of 15 tree species having >0.5 recall.

Although the results find that our model is not yet generalizable across datasets, in this paper we have explored the integrity of GLCM. It is in our best hope that in the future, the current workflow of feature extraction through GLCM will be able to train a robust model as we fly more drone missions and have a larger dataset.

Fig 1: A bird eye view image with labeled tree crowns. We use a bounding box around the label to get image of each tree

2.2. METHODOLOGY

With the data acquired from two different missions, we obtained 2 sets of labeled data, 15 trees with a total of 18 instances. We obtained 8 different spectral bands (Red, Green, Blue, Near Infra-Red, Red-Edge, Wideband Red, Wideband Green, Wideband Blue). For every band, we generate Homogeneity, Contrast, Angular Second Moment, Mean, Variance and Correlation of each of the GLCM pixels.

As we use GLCM for texture analysis and feature extraction of the trees, We will now discuss the GLCM process, followed by a preliminary on random forest classifier.

2.1 UNDERSTANDING THE ROBUSTNESS OF GLCM

GLCM is a classic statistical technique for texture analysis and image classification first introduced by [6]. For each of the spectral bands we collected, we extract spatial features from the single-band images based on the relationship of brightness values to the center pixel with its surrounding defined by a window size. For example, for a red band, each pixel of the image would have a 'gray-scale' value for the red band between 0-255.

A matrix is used to represent the relationship of the brightness values. The matrix was made up of the frequent occurrence of the sequential pair of the pixel values along with a defined direction [7]. From here, GLCM generates a different set of texture information based on gray-scale, window size, and direction.

[6] defined fourteen textural features. However, as mentioned by [7], some of the texture would only provide redundant spatial statistical information, which would only lead to computational overhead. We also plot the different histogram values of the pixels to ensure statistical uniqueness between each tree which is shown in Figure 2. The formulae of the GLCM calculations used would be attached in Appendix A, as discussed by [8]. Eventually, we only used GLCM base, GLCM mean and homogeneity to pass through the final random forest classifier model.

2.2 RANDOM FOREST CLASSIFIER

A random forest consists of many different decision trees (not to be confused with tree crown) that, as much as possible, try to have trees with uncorrelated decisions. Each of the trees would then 'vote' for the final classification result. They take advantage of bagging, where several subsets of the dataset are sampled at random for the different decision trees [9].

Building on this, each of the decision trees have to only use the features from the subset. Hence, this introduces feature randomness, where the random forest is not extremely biased to a feature.

3. FINDINGS

Our GLCM feature extractor is developed in-house. We explored different window sizes and different step sizes when doing the feature extraction with GLCM.

3.1 GLCM DISTRIBUTION

Intuitively, we would expect the GLCM histogram distribution to be well-separated for different species at the respective bands and GLCM features. This is because each tree crown would have different textures that we hope the GLCM can pick up.

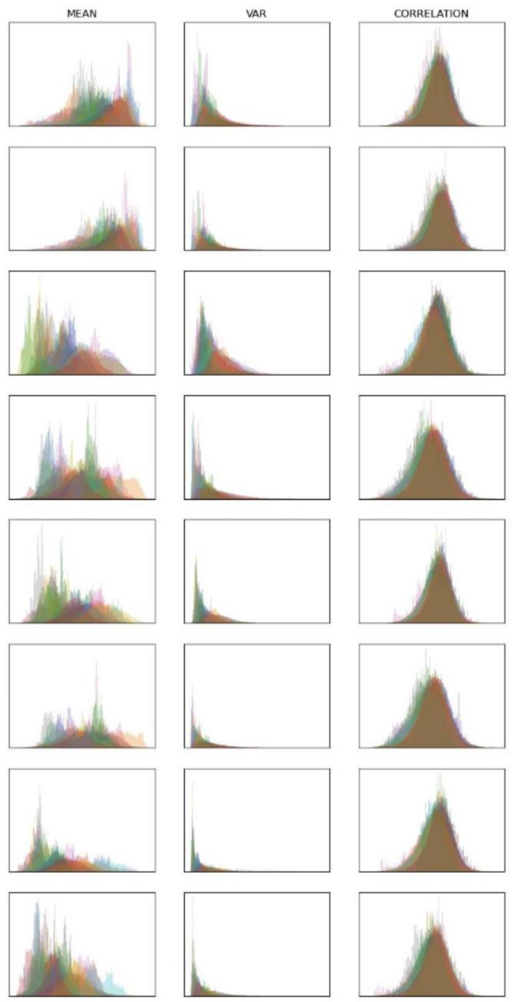


Fig 2: Pixel value histogram of some of our GLCM features. Column 1-3 represents Mean, Var, Correlation. Row 1-8 represent Wideband Red, Wideband Green, Wideband Blue, Red Edge, Blue, Near Infra-Red, Red, Green. Each colour represent the distribution of a tree.

We can see that the GLCM mean values were very nicely separated, while some features like Variance and Correlation are not.

3.2 RANDOM FOREST RESULTS

To multiply the support numbers of the tree, we take in random crops of 20x20 pixels from an original tree crown. These artificial instances can be overlapping each other for a tree instance. Other geometric transformations, such as random horizontal flip/rotation for growing the dataset would not prove to be useful as they would not add statistical significance for the model to learn. Furthermore, more advanced techniques popular for CNNs such as MixUp, CutOut, or ColorJitter would not value-add with our GLCM feature extraction. Indeed, without artificially inflating the support numbers, the model would be very prone to overfitting to a species.

For each of the tree instances, we calculate the Mean, Skew, Variance, Kurtosis for all six GLCM feature, for each of the eight channel. This gives us 4x6x8 = 192 features that we can use for the classifier.

However, in our benchmarking with random forest classifier that achieved the best result, we did not pass in all the features.

The number of random forest estimators and depth is kept at default setting for all our experiments. The number of estimators (decision trees) is 100, and the maximum depth is left for the decision trees to expand.

Species Number	Precision	Recall	Number of instances
1	0.76	0.47	60
2	1.00	0.20	20
3	0.00	0.00	30
4	0.68	0.93	30
5	0.00	0.00	20
6	0.94	0.85	20
7	0.70	0.55	29
8	0.96	0.74	30
9	1.00	0.83	30
10	0.88	0.77	30
11	0.67	0.07	28
12	0.02	0.03	30
13	0.38	0.97	70
14	0.00	0.00	30
15	0.94	1.00	30

Fig 3: Best results table where number of instances is at least 20 and scales with actual image size of tree crown. Name of the actual species is attached in the Appendix.

The metrics used are precision and recall, where Precision = True Positive / (True Positive + False Positive) and Recall = True Positive / (True Positive + False Negative), while Accuracy is the Total Correct Prediction / Total Prediction Made.

To ensure predictions are done on unseen dataset, the random forest classifier result shown is trained on features from our first data collection mission (December 2020) and tested on the second mission (May 2021). It achieved 54% accuracy.

When we swap the train and test dataset around (train on May 2021 and test on December 2020), the accuracy dropped to 48% as seen below.

Species Number	Precision	Recall	Number of instances
1	0.56	0.95	60
2	0	0	20
3	0.23	0.23	30
4	0.07	0.10	20
5	0	0	30
6	1.00	0.15	20
7	0.95	0.66	29
8	0.88	0.93	30
9	0.60	0.21	20
10	0.53	0.87	30
11	0.28	0.40	20
12	0.21	0.23	30
13	0.40	0.43	70
14	0	0	20
15	0.90	0.90	30

Fig 4: same configuration with train and test datasets swapped

The spectral bands used were Red, Green, Blue, Red-Edge and Near Infra-Red. We extract their base, mean, homogeneity GLCM features and calculate the mean, variance, skew, kurtosis for all the pixels in one instance of a tree crown. There were a total of 5 x 3 x 4 = 60 features for the classifier, out of the possible 192. Sometimes, less is more. Window size used was 15x15 with step size 4 and nbins=128. The confusion matrices for the two predictions are attached in the appendix.

3.3 RESULT COMPARISON

To understand the effect of using different GLCM-derived features on its classification accuracy, we conduct a result comparison of using different combinations of features. We compare the best result of using 60 features to using all possible 192 features and also comparing to using 20 GLCM-mean derived features.

Total # features	Spectral bands	GLCM features	Accuracy
4 x 8 x 6 = 192	R, G, B, RE, NIR, WR, WG, WB	all 6 features	45%
4 x 5 x 1 = 20	R, G, B, RE, NIR	mean	42%
4 (Mean, Var, Kurtosis, Skew) x 5 x 3 = 60	R, G, B, RE, NIR	base, mean, homogeneity	54%

Fig 5: The different classification accuracy with varying number of features used.

Most recently, we have also added new tree species and instances from most recent labelling by the ecology team. However, we did not compare the results of using more species for the multi-class classification in this paper.

4 CONCLUSION.

In conclusion, we have seen how GLCM has the potential to be used to classify tree crowns from a multispectral camera in Chestnut National Park due to its statistical significance. Despite the results yet to be adequate, it is a good starting point as we build up the dataset of the unique trees. Moving forward, 1) more dataset can be collected and even published to open-source to encourage others to find ways to push the state-of-the-art of forest inventory-keeping and management through the use of machine learning. As the dataset grows, we can 2) further confirm the robustness of GLCM to classify tree crowns and 3) iterate our in-house GLCM algorithm. Furthermore, with deep learning being far superior for computer vision application, we can also explore 4) how to pivot towards using CNNs. Other classification models such as SVM can also be used.

ACKNOWLEDGMENT

I would like to acknowledge the funding support from Nanyang Technological University – URECA Undergraduate Research Programme for this research project.

I would also like to acknowledge my supervisor Dr Ji-Jon Sit for letting us explore our tasks independently. Many thanks also goes to the team, Benjamin, Aaron, and especially John who has always been the cornerstone of the GLCM algorithm development and doing results verification.

I would like to acknowledge the NUS ecology team for identifying the relevant tree species with their domain expertise, and lastly NPArks for providing us with access to their site data.

REFERENCES

[1] Penman, Jim et al. (eds.). (2003). Definitions and methodological options to inventory emissions from direct human-induced degradation of forests and devegetation of other vegetation types. 32 p. Intergovernmental Panel on Climate Change. IPCC-XXI/Doc. 8. Kanagawa, Japan: IPCC Secretariat.

[2] Yang, Mou et al. (2022). Detecting and mapping tree crowns based on convolutional neural network and Google Earth images,International Journal of Applied Earth Observation and Geoinformation,Volume 108 ,102764,ISSN 1569-8432,

[3] Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11976-11986).

[4] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*

[5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

[6] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621.

[7] Iqbal, N., Mumtaz, R., Shafi, U., & Zaidi, S. M. H. (2021). Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*, *7*, e536.

[8] Chang J., Integrity of GLCM (2021) Unpublished.

[9] Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

APPENDIX

DATASET AND ITS LABELS

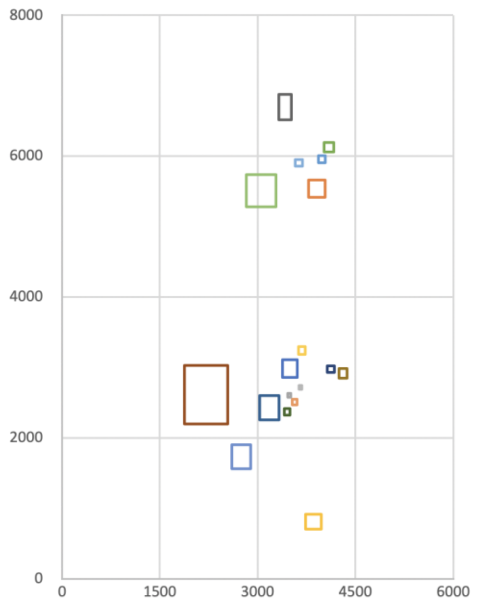


Fig 6: The coordinates of the bounding boxes in the bird-eye view image.

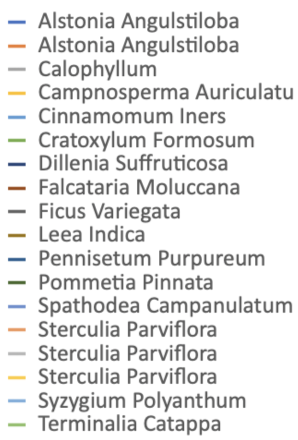


Fig 7: The key of the bounding boxes colour in Fig 4. The species name is in alphabetical order and it also follow the numbering for Fig 3 (Alstonia is 1, Calophyllum is 2, and so on)

GLCM EQUATIONS

$$\text{Contrast} = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$$

$$\text{ASM} = \sum_{i,j=0}^{N-1} P_{i,j}^2$$

$$\text{GLCM Mean}_i = \mu_i = \sum_{i,j=0}^{N-1} P_{i,j} \times i$$

$$\text{GLCM Mean}_j = \mu_j = \sum_{i,j=0}^{N-1} P_{i,j} \times j$$

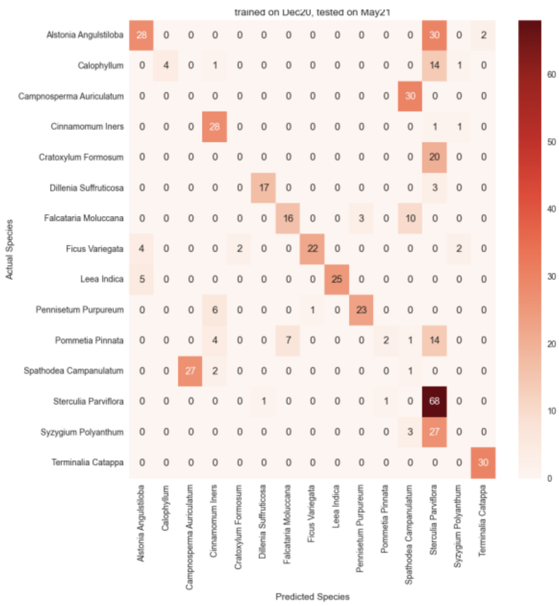
$$\text{GLCM Variance}_i = \sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j}(i-\mu_i)^2$$

$$\text{GLCM Variance}_j = \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j}(j-\mu_j)^2$$

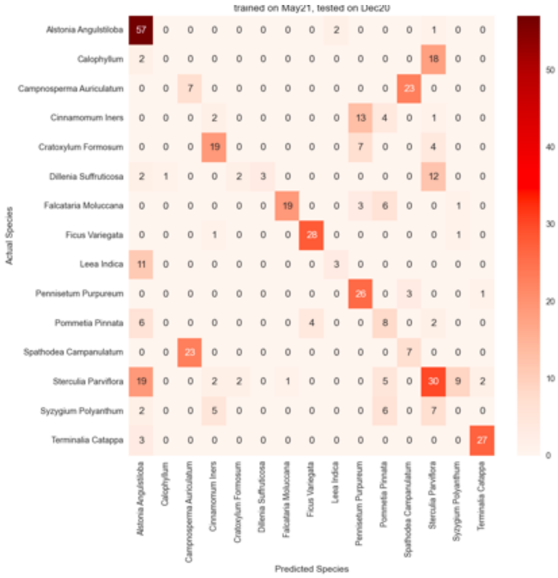
$$\text{Correlation} = \sum_{i,j=0}^{N-1} P_{i,j} \times \frac{(i-\mu_i) \times (j-\mu_j)}{\sqrt{\sigma_i^2 \times \sigma_j^2}}$$

$$Homogeneity = \frac{\sum p(i,j)}{1+|i-j|}$$

CONFUSION MATRIX OF BEST RESULTS



Trained on Dec, Tested on May with 60 features



Trained on May, Tested on Dec with 60 features