CSC2005Z

Classification of Crop Trees Images By Clustering On Pixel-Based Features

By Catherin Li

Supervised by Professor Patrick Marais

August 2024

Project Proposal

Table of Contents

Background and Motivation

Crop trees provide invaluable amounts of nutrients and are an important source of food for the population. Tree surveying is performed to ensure production quality and aid the prediction of crop yields. Specific tasks can include counting, locating and assessing the health or growth status of the trees. However, due to the scale of modern commercial farms, traditional manual farm management methods prove to be labour-intensive, time-consuming and susceptible to human errors [11]. Hence, technology is utilised as an alternative approach in tree surveying for more efficient, reliable and timely results. The technologies used to monitor the farms can be split up into three stages: data collection, data processing, and data analysis. This project will focus on data classification during the processing stage.

Classification in this scenario entails placing trees into different categories. It is an invaluable step before starting the analysis stage, because classification enables a nuanced approach in tree management, where different categories of trees can be managed differently. Furthermore, it can clean up the data and prevent complications in the analysis stage by identifying the tree of interest allow other tree species to be filtered out.

However, before object classification, more general data processing tasks must be performed to prepare the data. Common tasks include object detection and object segmentation. There are many deep learning architectures dedicated to object detection that are employed to identify individual trees, which wrap the tree in a bounding box for counting purposes. To take it a step further, some of the object detection architectures also can perform object segmentation for the delineation of trees, to prepare data for classification. However, accurate tree delineation remains a challenge as a result of the tree's irregular shape [3]. This project will not require tree crown polygons to be provided as input for each canopy. Instead, classification will be performed on pixel-based features computed from the pixels of the tree within the bounding box generated from object detection, which is more widely accessible.

Project Description

This project aims to assess which common pixel and image-level features can be used efficiently in different clustering techniques to ensure accurate tree classification, where the dataset would only be labelled with a bounding box. The impact of the different computed features on the chosen clustering techniques will be examined in terms of their ability to categorise different trees. A public dataset consisting of annotated aerial photos of palm oil trees and other species taken by unmanned aerial vehicles (UAV) will serve as the subject of the experiments. The dataset will serve as the basis to develop and test how the pixel-based features interact with the clustering methods. The project hopes to contribute to developing accurate classification methods that make crop tree management simpler, and more resource efficient.

Related Work

The two most common approaches in the classification of remotely sensed images are the object-based approach and the pixel-based approach. Zerrouki and Bouchafra [10] investigated whether pixel or object-based classification is more suitable when using satellite pictures. They found that when the object's number of pixels per region exceeds eight and there is an absence of ideal segmentation, the pixel-based approach is more appropriate. The dataset used in this project is taken significantly closer to land compared to the study, hence the pixel-based is appropriate for this project.

1. Clustering

K-means [14] is a simple centroid-based that creates initial clusters based on the Euclidean distance from randomly selected points. Then, it takes the mean of each cluster and makes it the new centroid. The process of clustering and averaging is repeated until the centroids stabilise. The clustering algorithm is known for its computational efficiency and ease of implementation and has been widely explored to perform pixel classification [10]. Hence, in this project, K-means also acts as a benchmark to compare the performance of the other two clustering techniques against.

DBSCAN [12] is a density-based clustering algorithm that defines clusters by separating areas of high density from areas of low density. DBSCAN is known for its ability to handle noise, outliers, uneven densities and can form clusters of any shape. Deng et al. [19] used DBSCAN to classify their dataset for facial recognition tasks in the preprocessing stage. DBSCAN exhibited superior computational efficiency compared to k-means and hierarchical clustering when the classification was done for noisy datasets. Thus, the clustering technique is investigated for this project as the trees in the dataset were extracted from differing locations that may introduce nose.

In contrast, Spectral clustering is a graph-based technique that first creates a weighted graph based on an affinity matrix derived from the data. Then eigen decomposition is used to find eigenvalues and eigenvectors. This allows the data to be represented in a lower dimensional space, where algorithms such as K-means can be applied to create clusters. It has the potential to handle large-scale and high-dimensional data [2], which will be ideal given the high-dimensional feature descriptors that will be used. However, spectral clustering is mostly explored in fields of image or video segmentation, feature selection and dimensionality reduction [2]. This leaves room for spectral clustering's ability to handle high dimensional data in a pixel-based classification context to be further examined.

2. Pixel-based features

Previous developments in image recognition utilised global features derived from the overall picture. This contrasts with local descriptors like SIFT, Daisy and ORB where feature extraction is only focused on the image's local regions that is centred around some key point features. Global features include colour histograms, colour moments, edge histograms, texture correlations and co-occurrence matrixes [8]. Some of these features are commonly used in tasks such as medical image analysis [18], object detection [11], and facial recognition [8]. Hence, my project will contribute to the application of texture and spatial descriptors in the crop tree management domain.

Grey-scaled histograms, HOG, LBP and for and GLCM are selected as the features used to assess the algorithm's ability to create clusters.

Colour histograms are known for their effectiveness in discerning colour textures in ideal conditions [1]. However, considering the "curse of dimensionality" in clustering and the high dimensional vectors the feature could generate, a grey-scaled version is utilised instead to discriminate grey-scaled textures.

HOG [17] describes the edges and gradient structure of an image. The resulting descriptor is a three-dimensional histogram made of gradient locations and orientations, which are calculated from the changing pixel intensity of local image patches.

LBP reflects the local texture in an image by assigning 0 to adjacent pixels if the grey level is less than that of the central pixel and 1 otherwise. It is an implementation of the bag-of-visual-words model, which maps local features to their distribution probability to describe images.

GLCM [13] analyses spatial relationships of an image to capture texture. It computes how likely the intensity of pairs of pixels, separated at a given distance, will occur. These probabilities are then stored in a two-dimensional NxN matrix, where N is the number of grey levels.

3. Principle Component Analysis (PCA)

The feature extractors discussed above will result in feature vectors of high dimensionality and PCA is often applied to reduce the dimensionality. PCA achieves this by employing a vector space transform that only keeps the most significant values in the original value and returns a reduced vector called Principle Components [6]. Due to the spectral clustering's built-in dimension reduction quality, PCA likely has the least impact on its clustering abilities. However, DBSCAN and K-means suffer from the curse of high dimensionality as they use distances between points to determine clusters. Hasan and Abdulazeez [9] reviewed 15 papers on PCA that were published in the three years before the review was published. They have found papers using PCA to reduce the dimensionality of GLCM and LBP [16] and that PCA can improve the performance of the machine learning algorithms if the dataset utilised has high dimensionality [15]. However, there seems to be a lack of research in PCA application to high dimensional pixel-based features to clustering algorithms for classification.

Research Questions

The project will be attempting to answer the following questions:

2. To What extent does PCA reduce the accuracy of feature-based classification for our chosen clustering methods and improve efficiency?
3. Which combination of features and clustering algorithms yield the best performance?

Methodology

1. Dataset Description

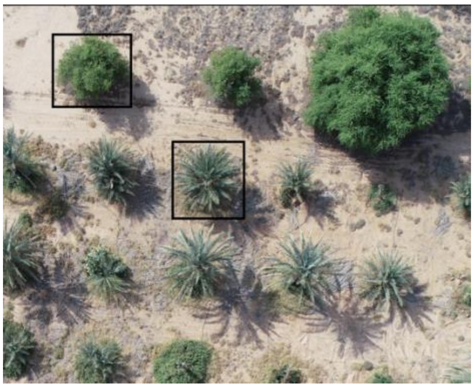


Figure 1 Palm Tree and Other Tree from Kharj Region with examples of the bounding box

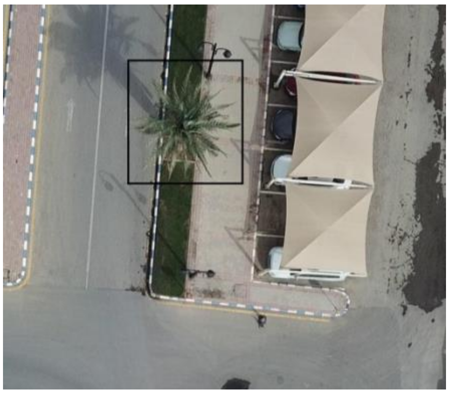


Figure 2 Palm Tree from Prince Sultan University Campus with examples of the bounding box

The research will be conducted on the palm tree dataset, which has a total of 349 annotated images with 258 taken from the Kharj region (which exhibits a rural background) and 91 taken from Prince Sultan University campus (which exhibits an urban background). As a result, the two sets of images have different densities and aspects of the trees depending on the region. The dataset contains 13071 instances of trees that were manually labelled using Labelbox and categorised either as palm or tree (for tree species that aren't palm trees). This information is summarized in CSV and .xml files. The palm trees appear jagged around the edges and rough in texture due to their leaves. Its colour is often a muted, cool-toned green. In contrast, other tree species in the dataset appear more vibrant and warmer in colour. The texture appears wispier and more spotty in comparison to the palms.

2. Proposed Approach

2.1 Data Pre-processing and Environment

Python offers an abundance of libraries for machine learning and image processing. Thus, the research will be conducted using a Python environment.

All the images in the dataset will first be converted to greyscale images. For each instance of a tree in each image of the dataset, the region within the bounding box will be extracted and resized to get the same dimensions for the feature extraction. Then, the tree instances will be normalised to ensure a consistent image intensity by using either OpenCV or Scikit images.

2.2 Feature extraction

This project will use image descriptors found in the Scikit image library. These descriptors include grey-scaled histograms, histogram of oriented gradients (HOG), local binary patterns (LBP) and grey-level co-occurrence matrices (GLCM).

The palm trees in the project's dataset appear to have an intense contrast between their light and dark parts, while the other tree species' colours appear less variable. Grey-scaled histograms capture the intensity distribution of images, which is appropriate for the project.

A palm tree object detection study [5] was able to achieve above 94% accuracy when detecting palms using HOG in combination with a support vector classifier [17]. The dataset they worked with also had other tree species mixed with palms. They took it a step further in certain test sets where they even included overlapping palm trees. The authors attributed part of the success of the detection study to HOG's ability to capture the unique texture of palm trees. Hence, HOG could be especially helpful in this project as the project's dataset has similar characteristic to the study.

LBP is widely used today due to its simplistic concept and implementation, low computational demand as well as its ability to discriminate against different textures[7]. As a result, it will be studied in this context.

The GLCM is chosen for its versatility, conciseness and efficiency in texture analysis. It set the precedent that sparked the creation of many other grey-scaled analysis methods [7] and thus is valuable to investigate.

In the project, the following will be applied to each instance of a tree:

- Each descriptor will be calculated and stored in a feature vector.
- Histogram equalizers will be applied, likely through OpenCV, on the grey-scaled histogram.
- The vectors will be normalised for PCA to be applied
- Each vector will be labelled with a unique identifier, made up of the name of the image and the coordinates of the bounding box that contains the tree.

2.3 PCA Transformation

The derived features naturally result in high dimensional vectors.

Thus, python's implementation of PCA will be applied to lower the dimensional space in the feature vector while preserving key information.

First, a training set of 200 tree instances that is representative of the data will be chosen from the dataset. The specific qualities of the four different types are displayed in the table below. The PCA will be trained four separate times, each time for a different feature extractor on the corresponding feature vectors.

No. of instances	Palm	Other Tree
Rural	50	50
Urban	50	50

Then, a scree plot will be plotted of cumulative explained variance (y) versus the number of principle components(x). The point on the x-axis where an "elbow" appears (reflecting diminishing returns) is the amount of principle components I will use.

2.4 Clustering

The DBSCAN, K-means and spectral clustering will be performed through scikit learn's implementation. The experiment will be split into two categories: clustering on vectors produced before applying PCA and after applying PCA. Each clustering algorithm (three in total) will cluster on each type of feature vector (five in total). Then, the data points in the resulting cluster will be assigned a class label that is determined by the majority label in that cluster. Clustering will be performed 24 times in total and the run times of each experiment will be timed. For each experiment, the parameters of the clustering algorithm will need to be fine-tuned before the result is recorded.

3. Evaluation Framework

The evaluation mechanisms are available through scikit learn and graphed through Matplotlib. The metrics listed below will measure the ability of clustering algorithms to classify palm trees against other trees and the ability of each feature extractor to assist in the classification process.

The results will either discourage or motivate the usage of pixel-based features for clustering when classification is needed. Either way, it will hopefully positively contribute to the research into better-automated surveying methods for crop trees.

3.1 Overall Accuracy

The overall accuracy of the results, which is the most used empirical measure [10], obtained from the clustering is defined as the percentage of tree instances being correctly labelled. The higher the percentage, the better and more accurate the results. Because the dataset provides ground truth labels, confusion matrices will be computed for each of the 24 experiments, and accuracy per experiment will be obtained as follows:

$$\text{Accuracy} = \frac{truePositive + trueNegative}{truePositive + trueNegative + falsePositive + falseNegative}$$

The overall accuracy per clustering algorithm will be obtained by taking the average of the 8 experiments' accuracy that will use that clustering algorithm.

The overall accuracy per feature extractor will be obtained by taking the average of the 6 experiments' accuracy that will use that feature extractor.

To obtain the overall accuracy before and after PCA is applied, the average of the 12 experiments that use feature vectors without PCA will be calculated. The same goes for the experiments that use feature vectors reduced by PCA.

3.2 Normalized Mutual Information (NMI)

For the clustering algorithms, NMI will be used to measure the extent to which ground truth labels and the labels obtained through clustering agree, while accounting for precision and recall [4]. It offers a balanced evaluation of the quality of clustering by giving a score ranging from 0 to 1, where 0 is no mutual information and 1 means a perfect agreement of true and assigned labels. The NMI formula is as follows:

$$\mathrm{NMI}(T, C) = \frac{2 \times \mathrm{MI}(T, C)}{H(T) + H(C)}$$

Where T is the set of ground truth labels, C is the set of the labels obtained from clustering, MI (T, C) represents the mutual information between T and C, while T and C's entropies are represented via H(T) and H(C).

NMI will be averaged the same way as overall accuracy, where the average will be obtained in terms of feature extractors, clustering algorithms and PCA application.

3.3 Speed-up graph

To measure the difference in computational efficiency of applying the PCA, a speed-up versus type of feature extractor graphs will be plotted to compare the effect of PCA. Speed-up versus Clustering algorithm will not be plotted as the number of principle components kept will be different for different feature extractors. The higher the value of the speedup, the more computationally efficient the PCA made the resulting feature vectors.

Anticipated outcomes

The expected findings of this project include the relative performance of the clustering algorithms against each other and the relative ability of the feature extractors to describe an image for classification purposes. Additionally, PCA's effect on each feature will be evaluated in terms of accuracy, speed-up obtained, precision and recall. However, an overall degradation in precision, accuracy and recall may be expected after PCA, because of the information lost. Care will be taken to investigate the explanations behind the results.

The project will also involve producing Python environments to perform preprocessing on the data, feature extraction, PCA and clustering. The main challenge here will be what parameters to use for the clustering algorithm and feature extractors.

The project will be a success if 1) The report compares the clustering algorithms' and the feature extractor's ability to classify tree instances. Possible explanations for the results are investigated and a conclusion is made on whether pixel-based clustering is appropriate for classification tasks in crop tree management. 2) PCA's accuracy and efficiency is discussed based on the results obtained.

Work details

Ethical Issues

The dataset was downloaded from the internet and has an unknown licence. However, the authors requested users of the dataset to give credit to with a specific reference and that is what I will do.

The images in the dataset do not include people and human data. It will not infringe on anyone's right to privacy in that regard.

The approaches in methodologies are not original and instead found through research and in the libraries of Python. Thus, credit will be given where it is due to prevent plagiarism.

The accuracy of the experimental set-up (especially concerning code) will be checked to mitigate errors in the analysis and misleading readers.

This research intends to enhance agricultural practices, and is not done for commercial purposes. This limits the potential for misuse of the data occurring. Furthermore, the results of the experiment will be used in food production and to aid decision-making around food security.

Risks

Risks	Severity (0-10)	Likelihood	Impact	Mitigation
Spectral Clustering's performance deteriorating due to double reduction in the feature vector.	5	High	The features that went through PCA may experience a great reduction in accuracy.	Explain in the report why this is (i.e. goes through two eigen decompositions)
DBSCAN and spectral clustering's sensitivity to the parameters chosen resulting in inaccurate clusters	7	High	DBSCAN won't form the correct amount of clusters and spectral clustering might produce very inaccurate results.	I set 20 days aside just for the experiments, which is roughly half the 2nd term.
Not finishing the experiments	9	Medium	The report will be incomplete and have a poor basis to be written on. Results may be invalid or inconclusive.	Conduct the experiments in such a way that I will at least have two clustering algorithms and two features to compare against in the end. I will also be performing the experiments on multiple iShango lab computers to experiment with the different clustering algorithms in parallel.

None of the clustering algorithms produce satisfactory results	6	low	The report will not go as planned in the proposal.	More research will be done on why this is the case and summarised in the report as the results of this project.

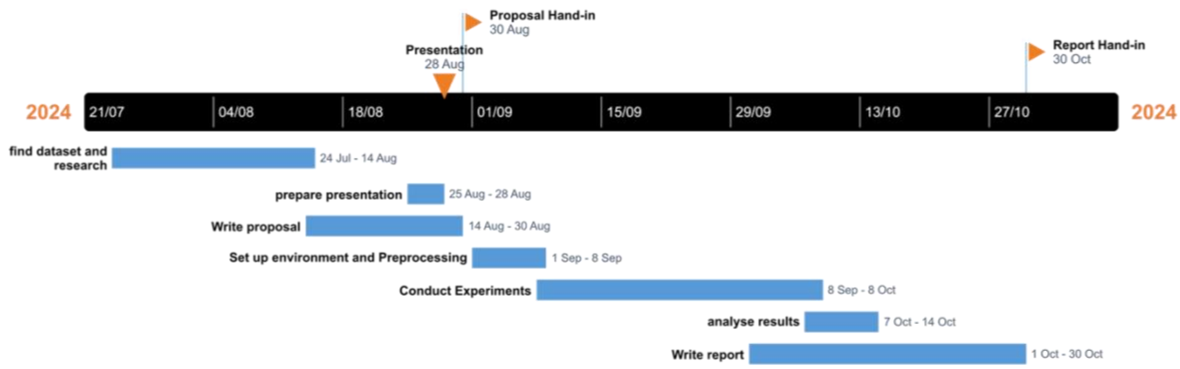
Resources

The experiments will be performed using Python in a notebook such as Jupyter or Google Collab to track progress and view the local results of the code.

The iShango labs will be accessed to conduct the experiment on multiple computers.

The implementation of the clustering algorithms and PCA will be taken from the Scikit learn library. The implementations of the feature extractors will be taken from the Scikit image library. Other Python libraries that I will use in the preprocessing stage include CV2, numpy, matplotlib, and pandas.

Timeline and Milestones



Deliverables	Dates
Proposal Presentation	2024/08/28
Proposal	2024/08/30

Data preprocessing code and “cleaned-up” tree instances	2024/09/08
Feature extractor code and feature vectors (normalised)	2024/09/15
PCA, Feature Extractor and Clustering code and results	2024/10/8
Report	2024/10/30

Bibliography

[1] T. Mäenpää and M. Pietikäinen, "Classification with Color and texture: Jointly or Separately?," *Pattern Recognition*, vol. 37, no. 8, pp. 1629–1640, Aug. 2004, doi: https://doi.org/10.1016/j.patcog.2003.11.011.

[2] Y. Zhao, Y. Yuan, F. Nie, and Q. Wang, "Spectral Clustering Based on Iterative Optimization for large-scale and high-dimensional Data," *Neurocomputing*, vol. 318, pp. 227–235, Nov. 2018, doi: https://doi.org/10.1016/j.neucom.2018.08.059.

[3] Asli Ozdarici-Ok and Ali Özgün Ok, "Using Remote Sensing to Identify Individual Tree Species in orchards: a Review," *Scientia Horticulturae*, vol. 321, no. 112333, pp. 112333–112333, Nov. 2023, doi: https://doi.org/10.1016/j.scienta.2023.112333.

[4] H. Yin, A. Aryani, S. Petrie, A. Nambissan, A. Astudillo, and S. Cao, "A Rapid Review of Clustering Algorithms," *arXiv.org*, Jan. 14, 2024. https://arxiv.org/abs/2401.07389 (accessed Feb. 18, 2024).

[5] Y. Wang, X. Zhu, and B. Wu, "Automatic Detection of Individual Oil Palm Trees from UAV Images Using HOG Features and an SVM Classifier," *International Journal of Remote Sensing*, vol. 40, no. 19, pp. 7356–7370, Sep. 2018, doi: https://doi.org/10.1080/01431161.2018.1513669.

[6] S. Ranjan, D. R. Nayak, K. S. Kumar, R. Dash, and B. Majhi, "Hyperspectral Image classification: A K-means Clustering Based Approach," in *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Jan. 2017, pp. 1–7. doi: https://doi.org/10.1109/icaccs.2017.8014707.

[7] F. Bianconi, A. Fernández, F. Smeraldi, and G. Pascoletti, "Colour and Texture Descriptors for Visual Recognition: A Historical Overview," *Journal of Imaging*, vol. 7, no. 11, p. 245, Nov. 2021, doi: https://doi.org/10.3390/jimaging7110245.

[8] Chao Zhu, Charles-Edmond Bichot, and Liming Chen, "Visual Object Recognition Using DAISY Descriptor," in *IEEE International Conference on Multimedia and Expo*, Jul. 2011, pp. 1–6. doi: https://doi.org/10.1109/icme.2011.6011957.

[9] B. M. Salih Hasan and A. M. Abdulazeez, "A Review of Principal Component Analysis Algorithm for Dimensionality Reduction," *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, Apr. 2021, doi: https://doi.org/10.30880/jscdm.2021.02.01.003.

[10] N. Zerrouki and D. Bouchaffra, "Pixel-based or Object-based: Which Approach Is More Appropriate for Remote Sensing Image classification?," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, San Diego, CA, USA, Oct. 2014, pp. 864–869. doi: https://doi.org/10.1109/SMC.2014.6974020.

[11] Hela Jemaa, Wassim Bouachir, B. Leblon, A. LaRocque, Ata Haddadi, and Nizar Bouguila, "UAV-Based Computer Vision System for Orchard Apple Tree Detection and Health Assessment," *Remote Sensing*, vol. 15, no. 14, pp. 3558–3558, Jul. 2023, doi: https://doi.org/10.3390/rs15143558.

[12] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, *A density-based Algorithm for Discovering Clusters in Large Spatial Databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI Press, 1996, pp. 226–231. Available: https://cdn.aaai.org/KDD/1996/KDD96-037.pdf

[13] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: https://doi.org/10.1109/tsmc.1973.4309314.

[14] J. Macqueen, "Some Methods for Classification and Analysis of Multivariate Observations," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, 1967, Available: https://digitalassets.lib.berkeley.edu/math/ucb/text/math_s5_v1_article-17.pdf

[15] G. T. Reddy et al., "Analysis of Dimensionality Reduction Techniques on Big Data," IEEE Access, vol. 8, no. 5, pp. 54776–54788, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2980942.

[16] Vishlavath Saraswathi and D. Gupta, "Classification of Brain Tumor using PCA-RF in MR Neurological Images," in *2019 11th International Conference on Communication Systems & Networks (COMSNETS)*, Jan. 2019, pp. 440–443. doi: https://doi.org/10.1109/comsnets.2019.8711010.

[17] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, pp. 886–893. doi: https://doi.org/10.1109/cvpr.2005.177.

[18] F. Agüera, F. J. Aguilar, and M. A. Aguilar, "Using texture analysis to improve per-pixel classification of very high resolution images for mapping plastic greenhouses," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 6, pp. 635–646, Nov. 2008, doi: https://doi.org/10.1016/j.isprsjprs.2008.03.003.

[19]

N. Deng, Z. Xu, X. Li, C. Gao, and X. Wang, "Deep Learning and Face Recognition: Face Recognition Approach Based on the DS-CDCN Algorithm," *Applied sciences*, vol. 14, no. 13, pp. 5739–5739, Jul. 2024, doi: https://doi.org/10.3390/app14135739.