# KAMETAY TURAR – 2381203 – CSC345B
# Land Cover Detection using Aerial Imagery

## Introduction

The proliferation of digital imagery from aerial and satellite sources has given rise to significant challenges and opportunities in the field of remote sensing, particularly in land cover detection (Wang et al., 2024). Effective analysis of such imagery is crucial for a variety of applications, from urban planning and environmental conservation to agriculture and disaster management. However, the high dimensionality and variability of image data make automated land cover classification a complex problem that demands robust solutions. Recent advancements in artificial intelligence (AI) and deep learning techniques have been leveraged to enhance the interpretation of large datasets and improve the accuracy and efficiency of land cover classification (Gu et al., 2024). These technologies offer innovative methods to manage the complexity of temporal and spatial variations inherent in remote sensing data, addressing the dynamic nature of land cover changes and contributing to more effective monitoring and management strategies (Qquenta et al., 2024).

In this context, this report presents a comparative study of two advanced machine learning models, Convolutional Neural Networks (CNN) and Support Vector Machines (SVM), applied to the UC Merced Land Use dataset. This dataset includes 21 distinct land cover types, each represented by 100 images, providing a comprehensive basis for evaluating model effectiveness in classifying high-dimensional spatial data.

The study explores two scenarios: one with the full complexity of 21 classes and a simplified model focusing on 5 classes. Initial findings reveal that CNN significantly outperforms SVM across both scenarios. Specifically, in the 21-class model, CNN achieves a testing accuracy of 82.3810% compared to SVM's 61.9048% and in the 5-class model, CNN's testing accuracy rises to 98.0000% versus 96.0000% for SVM.

These results underscore the potential of deep learning techniques, particularly CNNs, in handling complex image classifications with higher accuracy and efficiency than traditional methods such as SVM. The insights gained not only contribute to the academic and practical understanding of machine learning in image analysis but also suggest pathways for further improvement in automated land cover detection systems.

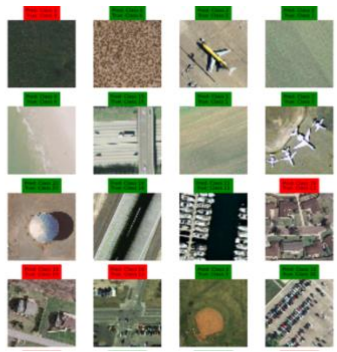

*Figure 1 Image fragment from the Result of Testing Set Visualisation of CNN - 21 classes. Red-Incorrect, Green-Correct*

## Methods

The approach to addressing the problem involved a series of methodical steps beginning with data preparation, followed by model architecture design and training, and culminating in a comprehensive performance evaluation.

In the initial phase, the UC Merced Land Use dataset, consisting of 2100 low-resolution images (61x61 pixels, RGB), was employed. Image data and corresponding labels were first loaded into our environment. The pixel values were normalized to the range [0,1], a necessary step to ensure consistency in input data scale which aids in accelerating the convergence during the training phase. Concurrently, the labels were converted to a one-hot encoded format, which is particularly suitable for multi-class classification, allowing our models to efficiently handle up to 21 different land cover categories. To ensure unbiased model evaluation and training, after normalization, the dataset was shuffled and divided into training, validation, and testing sets with proportions of 80%, 10%, and 10% respectively. This step ensures that the model is trained on a diverse subset of data, fine-tuned on the validation set, and accurately evaluated on the test set, which is crucial for assessing the model's ability to generalize to new data.

**<u>Convolutional Neural Networks (CNN)</u>**

To enhance the robustness of the CNN models and prevent overfitting, data augmentation techniques were utilized. Utilizing the ImageDataGenerator, transformations such as rotations, shifts, shears, zooms, and flips were applied. These augmentations introduce a simulated variability in the training data, mimicking different viewing conditions that might be encountered in real-world scenarios, thus enhancing the model's ability to generalize.

The architecture for both the comprehensive 21-class and the simplified 5-class classification scenarios included multiple convolutional layers paired with ReLU activation functions. These layers effectively capture the spatial and textural patterns in the images and are interspersed with pooling layers which serve to reduce the dimensionality of the feature maps, thereby compressing the information and reducing the computational load. Dropout layers at rates of 0.25 and 0.5 were strategically placed to combat overfitting by randomly omitting subset features during training.

The CNN models were trained over 100 epochs for the 21-class model for the 5-class model, using batch sizes of 32 and 8 respectively, with the Adam optimizer, known for its adaptive learning rate capabilities, making it well-suited for handling the non-uniformity of image data. Smaller number of batch size were chosen because it helps ensure that each batch will contain a diverse set of samples. The categorical crossentropy loss function was employed to gauge the discrepancy between the predicted and actual labels, providing a quantitative measure for model adjustments during the iterative training process. The use of data generated by the augmentation pipeline for training allowed for a dynamic adjustment of model parameters in response to the evolving complexity of feature recognition as training progressed.

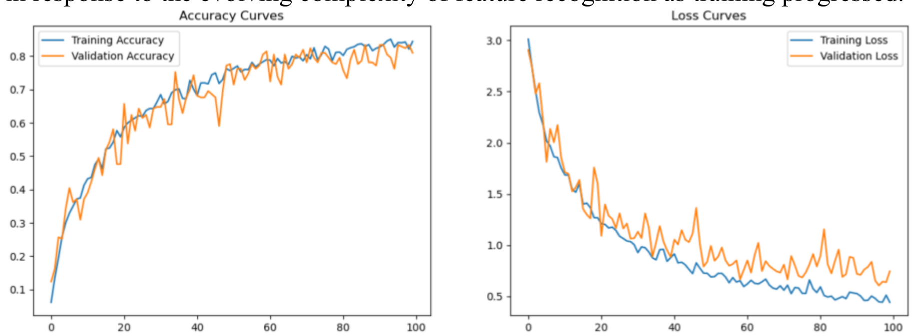


*Figure 2 CNN 21 class model Training, Validation Accuracy and Loss Curves*

**<u>Support Vector Machine (SVM)</u>**

In parallel to the CNN approach, Support Vector Machine (SVM) models were implemented for both the 21-class and the simplified 5-class classification scenarios, each tailored to the specific complexity and characteristics of the dataset. The SVM models, renowned for their performance in high-dimensional spaces, were particularly adjusted to employ the Histogram of Oriented Gradients (HOG) for feature extraction.

The HOG feature extractor was used to transform each RGB image into a grayscale format, simplifying the input while retaining crucial structural information in the form of gradient orientations. This transformation was pivotal in capturing edge and texture information, which are essential for distinguishing between different types of land cover. To manage computational resources effectively and handle the high dimensionality of the data, the images were processed in batches of 100, with HOG features extracted sequentially to create a comprehensive feature matrix for the entire dataset.

For the 21-class SVM model, following feature extraction, the data underwent rigorous preprocessing steps including standardization and dimensionality reduction. Features were standardized by removing the mean and scaling to unit variance, ensuring that the SVM model was not biased by the scale of the input features. Principal Component Analysis (PCA) was then applied to reduce the dimensionality of the data, retaining components that explained a substantial amount of the variance. This step was crucial to enhance the training speed and effectiveness of the SVM classifier, simplifying the model while potentially improving performance.

Conversely, for the simplified 5-class SVM model, while the same initial steps of feature extraction and standardization were applied, PCA was not employed as it was observed to degrade the model's performance. Instead, the simpler feature set was directly used for classification, which consisted of images randomly selected from five classes chosen to represent a broad range of land cover types.

The SVM classifier, equipped with a radial basis function (RBF) kernel and a regularization parameter of C=10, was then trained on these prepared datasets. The choice of the RBF kernel was motivated by its ability to handle non-linear relationships in the data, which is often required for complex image classification tasks. The training process focused on optimizing the decision boundaries based on the HOG features, aiming to achieve maximum margin separation between the different classes

For both CNN and SVM, performance evaluation was multifaceted, focusing on accuracy as the primary metric to quantify effectiveness across training, validation, and test datasets. Additionally, loss metrics offered insights into the model's learning efficiency and the effectiveness of the backpropagation in minimizing error. Detailed classification reports and confusion matrices were generated post-training to evaluate precision, recall, and F1-scores, providing a granular view of model performance across individual classes.

The decision to employ a combination of CNN and SVM models aimed to provide a balanced view of the capabilities of modern deep learning techniques against more traditional machine learning methods in processing high-dimensional image data. The comprehensive evaluation across varied scenarios and classes highlighted the scalability and adaptability of the proposed methods, offering valuable insights into their application in real-world land cover detection tasks.

## Results

This section delves deeply into the performance metrics and accuracy levels recorded during the experiments, offering an analytical perspective on why these results manifest and their implications for remote sensing applications.

*Table 1 Performance Metrics*

| Metric | CNN 21 | SVM 21 | CNN 5 | SVM 5 |
|---|---|---|---|---|
| Precision | 0.87 | 0.65 | 0.98 | 0.97 |
| Recall | 0.82 | 0.62 | 0.98 | 0.96 |
| F1-Score | 0.83 | 0.62 | 0.98 | 0.96 |

*Table 2 Accuracy Level*

| Model/Classes | Accuracy (Train) | Accuracy (Testing) | Accuracy (Validation) |
|---|---|---|---|
| CNN - 21 | 90.7143% | **82.3810%** | 80.9524% |
| SVM - 21 | 100.0000% | **61.9048%** | 55.2381% |
| CNN - 5 | 99.5000% | 98.0000% | 96.0000% |
| SVM - 5 | 100.0000% | 96.0000% | 90.0000% |

The performance of the CNN models in both the 21 and 5 class scenarios demonstrated a consistently high level of precision, recall, and F1-score, as seen in the performance metrics (0.87, 0.82, and 0.83), and test accuracy level of 82.38%. Notably, the CNN models excelled in handling the complex spatial hierarchies and high-dimensional data typical of aerial imagery. The strength of CNNs in feature extraction from image data—thanks to their deep layers that capture an extensive range of features from basic to complex—plays a crucial role in their superior performance. This capability enables CNNs to effectively recognize patterns and variations in land cover, which are crucial for accurate classification.

In contrast, the SVM models showed relatively lower performance (P: 0.65, R:0.62, F1:0.62) in the 21-class scenario but improved results in the 5-class scenario (P,R,F1: 0.98). The drop in performance in the 21-class scenario can be attributed to SVM's linear nature and its handling of high-dimensional data. Although SVMs are powerful classifiers known for their effectiveness in binary classification tasks, their performance can diminish as the complexity and number of classes increase. This is particularly evident from the perfect training accuracy observed, suggesting an overfitting scenario where the SVM models memorize the training data but fail to generalize well on unseen data. This overfitting is corroborated by the lower testing and validation accuracies.

In the simplified 5-class scenario, both models showed improved results, with the CNN model slightly outperforming the SVM. The reduction in class complexity likely alleviated some of the challenges faced by the SVM model, allowing it to better categorize the distinct features of the fewer classes. However, the CNN's ability to leverage its convolutional layers for robust feature extraction still provided it with an edge over the SVM, as evidenced by the slightly higher metrics across the board. The accuracy levels across training, testing, and validation phases further underline these points. The CNN model maintained closer performance levels across these phases, indicating a better generalization capability. On the other hand, the disparity between the training and validation/testing accuracies for the SVM model highlighted its struggles with overfitting, especially in the 21-class scenario.

CNN's confusion matrices displayed fewer off-diagonal elements, indicative of fewer misclassifications across classes. This precision underscores the CNN's ability to manage the inter-class variability effectively—a critical factor in the multi-class setting of land cover classification. These results not only reinforce the robustness of CNNs in handling complex and high-dimensional datasets but also highlight the need for careful consideration when applying SVMs to similar tasks. The analysis suggests that while SVMs can be tuned to perform better in less complex scenarios, CNNs generally offer more reliability and higher accuracy for comprehensive image classification tasks in remote sensing.

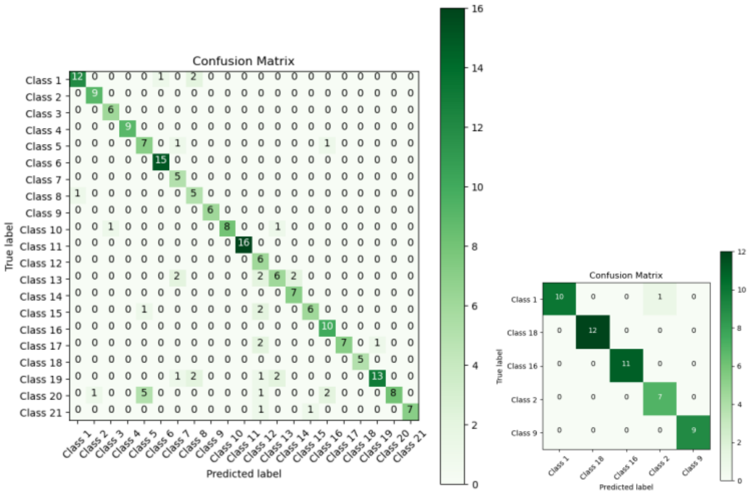


*Figure 3 Confusion Matrix. CNN-21 and CNN-5*

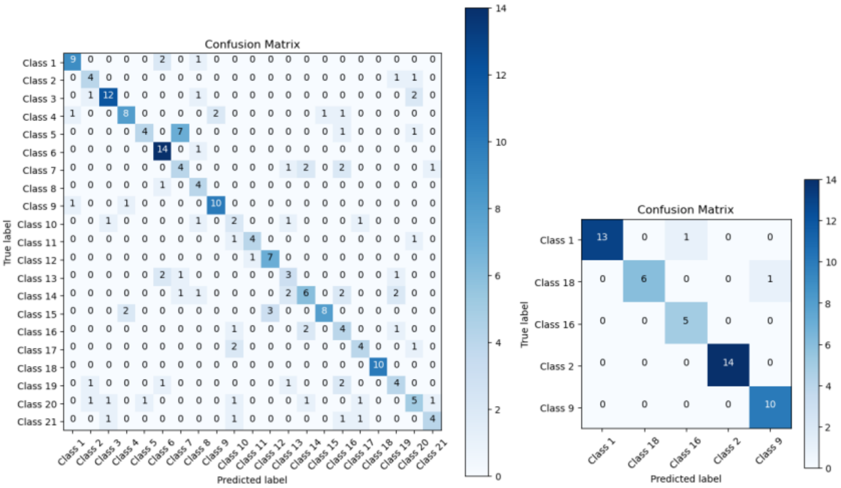


*Figure 4 Confusion Matrix. SVM – 21 and SVM -5*

## Conclusion

There was the evaluation of the effectiveness of Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) in the classification of land cover from satellite imagery, offering an in-depth exploration of each model's capabilities across two different classification scenarios (21, and 5). The methodology involved comprehensive data preparation, model training, and evaluation, leveraging precision, recall, F1-score, and accuracy metrics to compare performance and building confusion matrixes.

The findings reveal that CNNs exhibit superior generalization capabilities in handling both 21-class and simplified 5-class scenarios, achieving high accuracy and robust performance metrics. In contrast, SVMs, while showing commendable results in simpler scenarios, struggled with overfitting and generalization in the more complex 21-class setting. These results underscore the challenges inherent in applying SVMs to high-dimensional and diverse datasets without significant tuning and constraint.

Critically, the study highlights the need for models that can balance complexity and performance without sacrificing generalizability. Future work could explore hybrid models that combine the robust feature extraction capabilities of CNNs with the classification finesse of SVMs, potentially overcoming the limitations observed in standalone applications. Additionally, integrating techniques such as transfer learning and semi-supervised models could further enhance the models' ability to learn from limited labeled data and improve their applicability to broader remote sensing tasks. Further research into more sophisticated feature engineering and advanced data augmentation techniques could also provide the models with a deeper and more abstract understanding of the input data, reducing the model's reliance on the quantity of data and enhancing its ability to generalize from fewer examples.

In conclusion, while both Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) demonstrate significant capabilities in land cover classification, their suitability varies depending on the complexity of the task at hand. CNNs, in particular, stand out due to their ability to effectively manage high-dimensional data and capture complex patterns through their deep learning architecture, making them especially powerful for tasks involving detailed and varied imagery. However, it is not perfect as it requires much more memory and training time compared to SVM.

## Reference:

Gu Z, Zeng M. The Use of Artificial Intelligence and Satellite Remote Sensing in Land Cover Change Detection: Review and Perspectives. *Sustainability*. 2024; 16(1):274. https://doi.org/10.3390/su16010274

Qquenta J, Rau P, Bourrel L, Frappart F, Lavado-Casimiro W. Assessment of Bottom-Up Satellite Precipitation Products on River Streamflow Estimations in the Peruvian Pacific Drainage. *Remote Sensing*. 2024; 16(1):11. https://doi.org/10.3390/rs16010011

Wang L, Zhang M, Gao X, Shi W. Advances and Challenges in Deep Learning-Based Change Detection for Remote Sensing Images: A Review through Various Learning Paradigms. *Remote Sensing*. 2024; 16(5):804. https://doi.org/10.3390/rs16050804