



Support Vector Machine Based Aerial Image Classification and Object Identification

Prince Daniel T¹, Vibin K², Dr.S.Meena³,

^{1,2} Students, and ³ Faculty
Dept. of Electronics and Instrumentation
Engineering,
St. Joseph's College of Engineering
Chennai, India.
princedaniel3702@gmail.com

Abstract: The acquisition of aerial and satellite imagery data is huge and the gradual increase is being witnessed daily. Analysis of such large quantities of data can be helpful for a number of practical applications. We present an automatic content-based analysis system for aerial imagery in order to classify and detect arbitrary objects or regions in high-resolution images. For that purpose, we proposed a method for automatic object detection based on a support vector machine. A novel two-stage approach for network training is implemented and verified in the tasks of aerial image classification and object detection. Comparative analysis shows SVM's superiority over ConvNeXt-Tiny and ResNet50 in image classification. Uploaded images are matched with trained datasets to classify objects by features, providing accurate predictions. The identified image and predicted class are printed as the result.

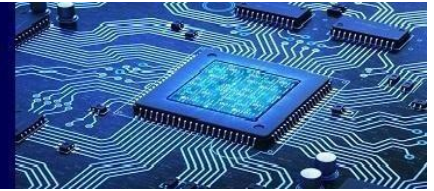
Keywords: Object detection, Feature extraction, Support vector machine, Aerial Images, Convolutional Neural Network, classification.

1. INTRODUCTION:

The recent breakthrough in remote sensing technology made the acquisition and availability of large amount of aerial images grow rapidly. The large amount of aerial images extracted brought up a need for their fast analysis and accurate classification in order to efficiently facilitate their usage in practical civil and military domains. A variety of land use tasks, such as urban planning, surveillance, crop monitoring, flood and fire prevention, etc., benefits on processing of remote sensing imagery. The automatic extraction of valuable information from aerial images became essential which further stimulated the development and improvement of various processing methods with a specific purpose.

In the realm of remote sensing, the process of extracting meaningful information from aerial images is a multidimensional endeavor encompassing several crucial stages. It begins with the meticulous collection of a dataset tailored to the specific task at hand. In this case, the dataset comprises 21 distinct land use classes, each meticulously curated with 100 images, resulting in a rich repository of diverse visual data representative of various environmental contexts.

However, the journey doesn't stop at data collection; rather, it extends into the realm of preprocessing, where the raw images undergo a series of transformations and enhancements to optimize their suitability for subsequent analysis. This preprocessing phase may involve techniques such as normalization, noise reduction, and geometric correction to mitigate



distortions and inconsistencies inherent in aerial imagery.

With a refined dataset in hand, the focus shifts towards model building, a pivotal stage where cutting-edge architectures such as ConvNeXt-Tiny and ResNet50 come into play. Leveraging the power of transfer learning, these models are initialized with weights pretrained on the vast ImageNet dataset, thereby imbuing them with a foundational understanding of visual features essential for image recognition tasks. However, the journey doesn't end here; instead, these pretrained models are customized and fine-tuned to adapt to the nuances of land use classification, ensuring their efficacy in extracting meaningful insights from aerial imagery.

The subsequent phase, training, is characterized by a meticulous orchestration of parameters and hyperparameters to optimize model performance. Techniques such as early stopping and learning rate scheduling are employed to strike a delicate balance between model convergence and generalization, thereby mitigating the risk of overfitting and ensuring the preservation of model integrity.

Trained models are rigorously evaluated and visualized to assess their efficiency. Stakeholders gain insights into the models' capabilities and limitations. Benchmarking against traditional classifiers like SVM provides comparative analysis. This process results in robust, high-performing models for extracting insights from aerial imagery. These models empower stakeholders in civil and military domains to make informed decisions confidently.

2. METHODOLOGIES:

Block Diagram

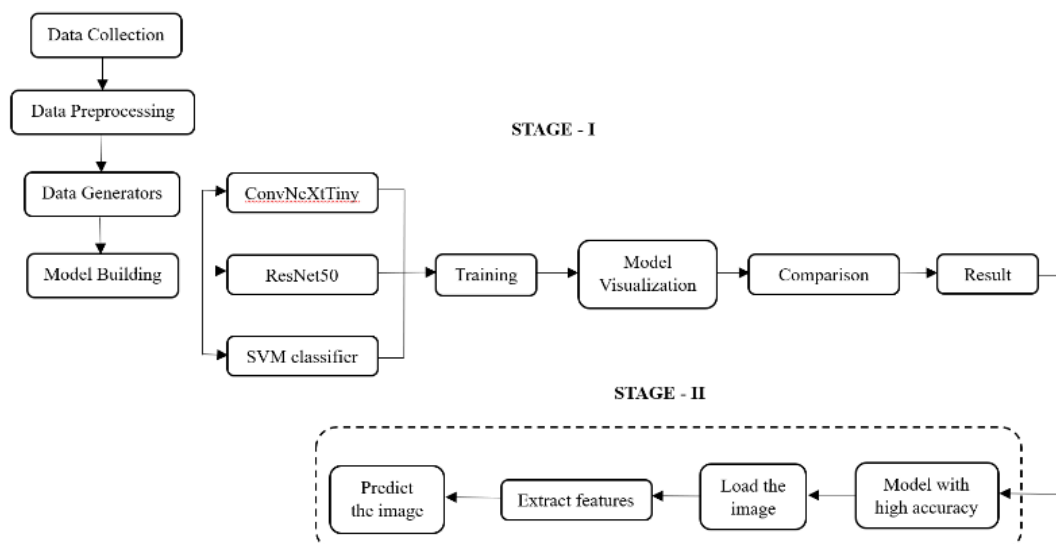


Fig. 1. Block diagram of proposed methodology

2.1 Data Collection

Aerial imagery, sourced from drones, airplanes, or satellites, forms the cornerstone of this



research dataset, comprising 21 land use classes with 100 images each. These high-resolution images, measuring 256x256 pixels, offer intricate details of diverse landscapes, from urban centres to agricultural fields and natural reserves. Carefully curated for research purposes, this dataset captures the rich complexity of land use patterns and dynamics. The dataset's breadth and depth enable researchers to explore a wide range of land use scenarios and phenomena. Through this comprehensive collection effort, a comprehensive understanding of land use dynamics is within reach, facilitating informed decision-making in various domains.

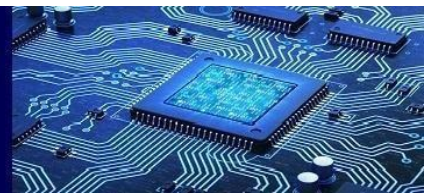
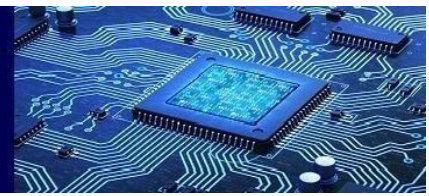


Table. 1. Dataset details

Dataset	
Class Name	Image
agricultural	
airplane	
Baseball diamond	
beach	
buildings	
chaparral	
Dense residential	
forest	
freeway	
Golf course	
Tennis court	
harbour	
intersection	
Medium residential	
Mobile home park	
overpass	
Parking lot	
river	
runway	
Sparse residential	
Storage tanks	

2.2 Data Preprocessing

Data preprocessing is a critical step in preparing image data for analysis, aiming to enhance its quality and suitability for machine learning tasks. In this phase, various techniques are employed to augment the dataset and reduce biases. These include rescaling pixel values, applying random rotations, shifts, shearing, zooming, and horizontal flips. The use of the 'nearest' fill mode ensures smooth transitions and maintains image clarity during transformations. By introducing variability and diversity into the dataset, preprocessing enhances the robustness and generalization capabilities of the



models. Ultimately, these preparatory steps lay the groundwork for accurate and reliable land use classification.

2.3 Data Generators

Data generators play a crucial role in machine learning pipelines, particularly in the creation of training, validation, and test sets. These generators efficiently load images from directories, resizing them to a specified target size as per the requirements of the model architecture. Images are organized into batches of predefined sizes, optimizing computational resources during model training. Additionally, the generators categorize images based on their directory structure, facilitating accurate labelling and classification. By automating these tasks, data generators streamline the preprocessing pipeline, allowing researchers to focus on model development and optimization. We allocate 80% of the data for training, 10% for testing, and the remaining 10% for validation.

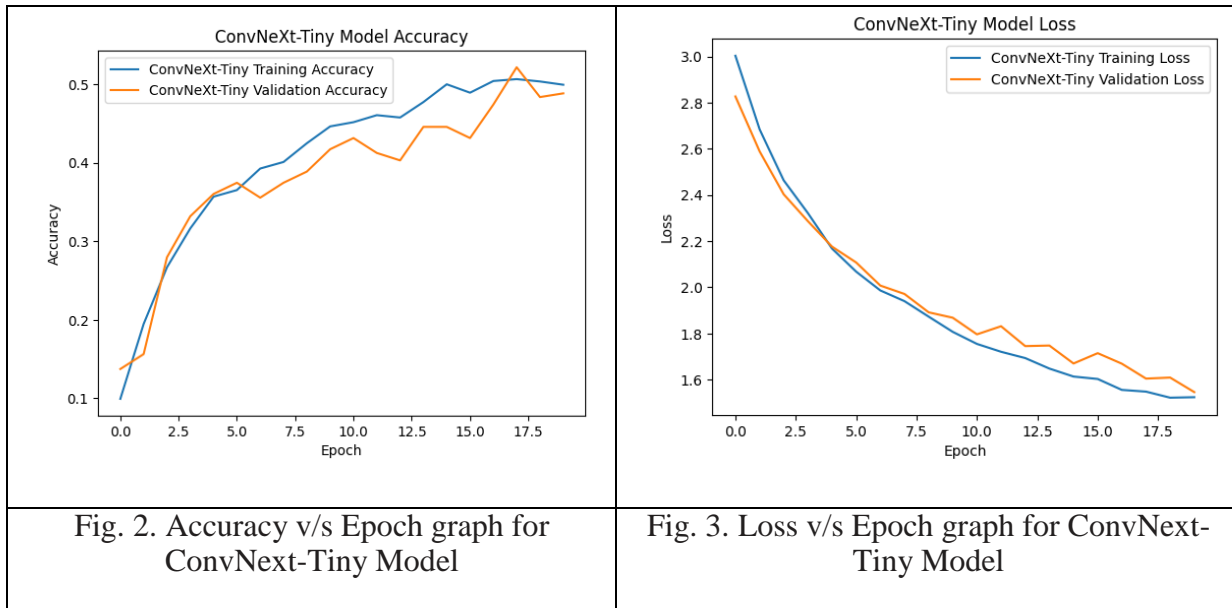
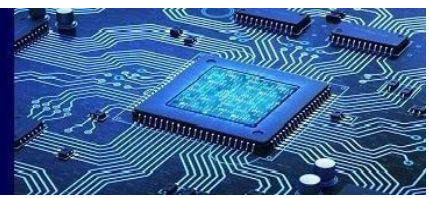
- 80% of the data for training: This subset is used to train the machine learning model.
- 10% of the data for testing: This subset is used to evaluate the model's performance on unseen data.
- 10% of the data for validation: This subset is used to fine-tune model parameters and prevent overfitting.

2.4 Model Building

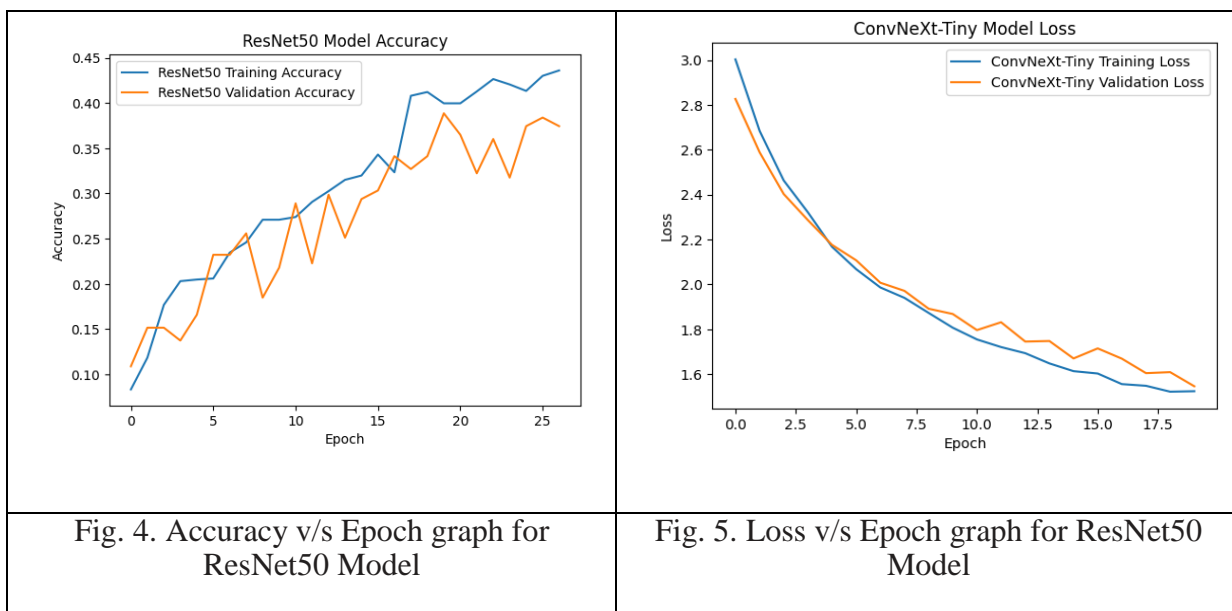
- a) ConvNeXt-Tiny, renowned for its effectiveness in image classification, has demonstrated remarkable performance on tasks such as achieving a top-1 accuracy of 87.8% on the ImageNet dataset (Chen et al., 2023). It is characterized by its lightweight architecture, featuring grouped convolutions and parallel pathways designed for diverse feature extraction. The model places a strong emphasis on minimizing computational cost and memory usage, while still delivering competitive performance in image recognition tasks. Although configurations may vary, ConvNeXt-Tiny prioritizes efficiency without compromising accuracy.
- b) ResNet50: ResNet50 is a deep convolutional neural network (CNN) architecture that was developed by Microsoft Research in 2015. This is a variant of the popular Residual Network architecture, or in short "ResNet". The "50" in the name indicates the number of layers in the network, which is 50 layers deep.
- c) SVM classifier: SVM is a powerful supervised algorithm that works best on smaller datasets but on complex ones. Support Vector Machine, or in short SVM can be used for both regression and classification tasks, but generally, they work best in problems regarding classification. SVMs achieve this by determining the optimal hyperplane that best segregates the data points of different classes in a high-dimensional space.

2.5 Model Training

- a) ConvNeXt-Tiny: The network undergoes iterative optimization to learn the weights of its convolutional layers and other parameters. This process typically involves forward propagation to compute predictions, followed by backward propagation to calculate gradients and update the weights using optimization algorithms such as stochastic gradient descent (SGD) or its variants. Hyperparameters like learning rate and batch size are often tuned to enhance convergence and generalization.



b) ResNet50: First forward propagation computes predictions, followed by backward propagation to calculate gradients and update weights using optimization algorithms like stochastic gradient descent (SGD). Techniques such as batch normalization and residual connections stabilize training, while hyperparameters like learning rate and dropout rate are fine-tuned to enhance performance and prevent overfitting.



c) SVM: The SVM algorithm optimizes a hyperplane to best separate classes by maximizing the margin between support vectors and minimizing classification errors. This involves solving the optimization problem to find the optimal hyperplane coefficients using techniques like quadratic programming or gradient descent. The process also includes tuning hyperparameters such as the regularization parameter and kernel function to improve generalization performance.

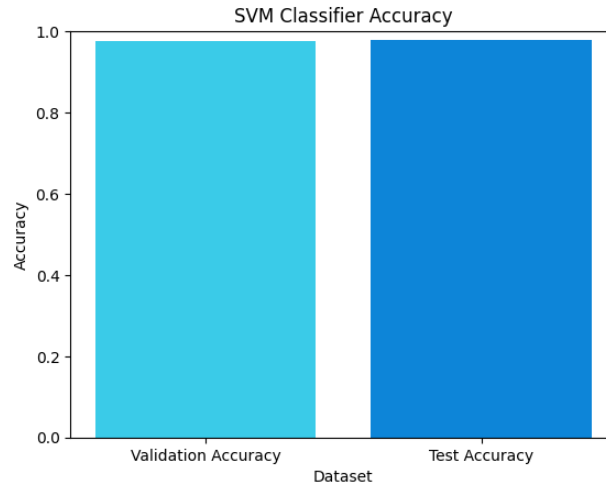
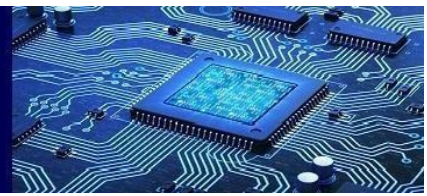


Fig. 6. Bar chart representation of SVM Model Accuracy v/s Epoch

2.6 Model Visualization

In SVM classifier visualization, the decision boundary separating classes in the feature space is paramount. This boundary, depicted in 2D or 3D space for binary tasks, showcases the hyperplane maximizing class separation. For multi-class scenarios, decision boundaries among classes are illustrated. Furthermore, visualization of support vectors, closest to the decision boundary, aids in understanding their impact on classification. Additional techniques like confusion matrices, ROC curves, and precision-recall curves offer insights into the classifier's performance and trade-offs.

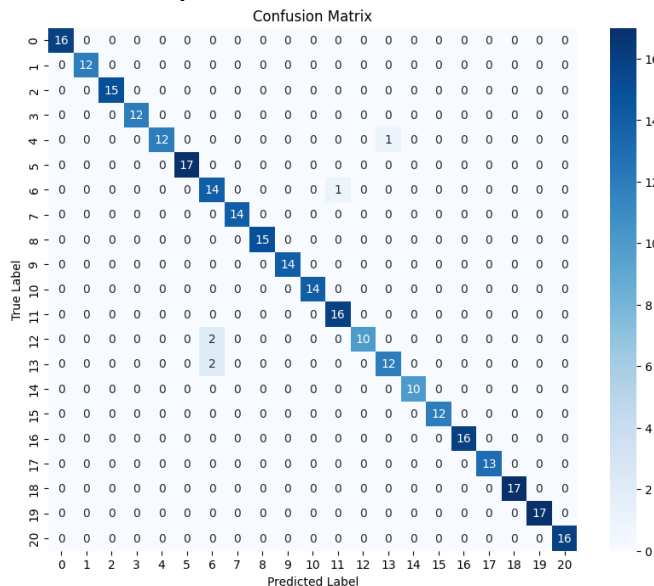


Fig. 7. Confusion matrix for dataset

2.7 Comparison of Accuracies :

In comparison to ConvNeXt-Tiny and ResNet50, the SVM classifier demonstrates robust performance, particularly in scenarios with limited data or where interpretability is crucial. Its supervised learning approach constructs a decision boundary based on training data, yielding competitive accuracy in well-separated class scenarios. While ConvNeXt-Tiny



and ResNet50 excel in learning complex hierarchical features, their performance may be affected by dataset size and computational resources. In situations with constrained resources or smaller datasets, the SVM's simplicity and efficiency can lead to superior or comparable performance. However, the selection of the most suitable model should consider specific task requirements and constraints, weighing factors such as model complexity, interpretability, and computational demands.

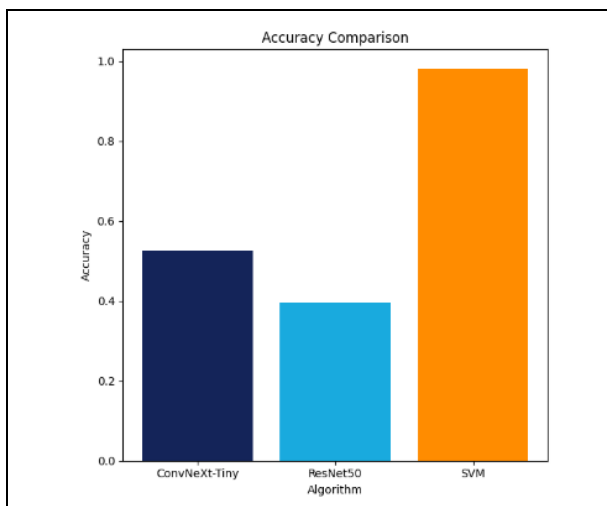


Fig. 8. Bar chart representation of algorithm models v/s accuracy

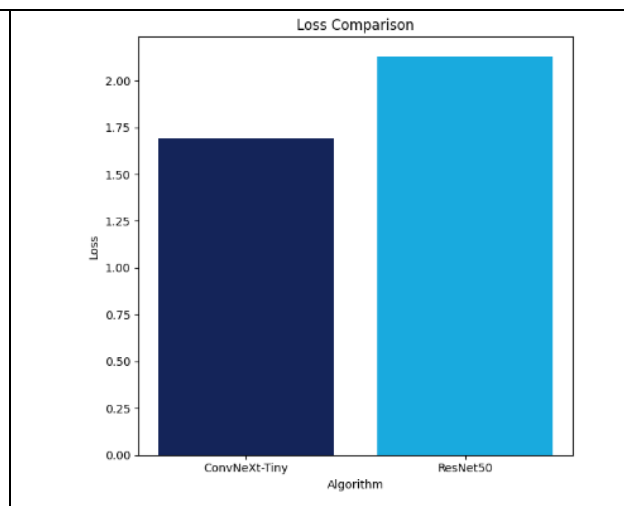


Fig. 9. Bar chart representation of algorithm models v/s loss

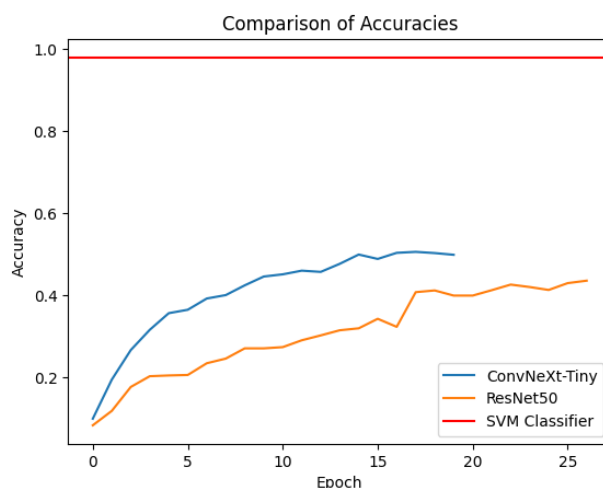
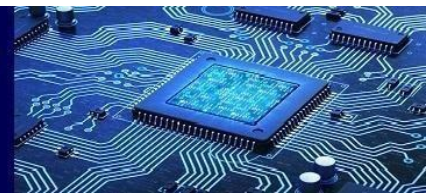


Fig. 10. Accuracy v/s Epoch graph for Model

3. RESULTS:

Table.2. Data Implementation Results For Algorithm Models

Algorithm	Accuracy	Loss
-----------	----------	------



ConvNeXt-Tiny	0.526504	1.689482
ResNet50	0.395474	2.127327
SVM	0.980000	NaN

The comparison among the algorithms, ConvNeXt-Tiny, ResNet50 and Support Vector Machine proved that the SVM algorithm performs better than the others under image classification and object identification. The image uploaded in the algorithm is compared with the trained data set in the respective classes to classify the object with respect to its features. After this process, the uploaded image is identified and the predicted class is printed as the result.

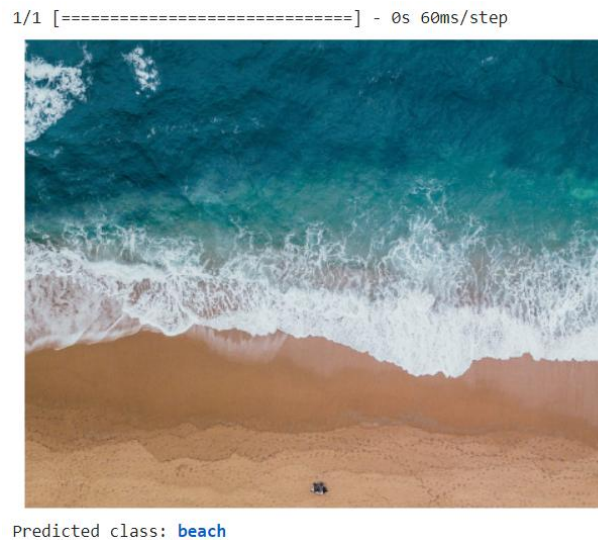
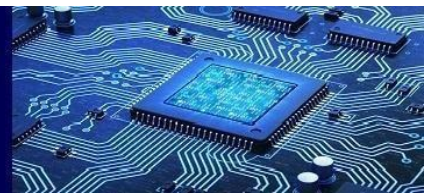


Fig. 11. Input image and predicted class

4. CONCLUSION:

The investigation into SVMs for aerial image analysis revealed that the proposed approach surpassed both feature-based methods and other network-based solutions. Subsequently, a network-based method was implemented for automatic content-based object detection on high-resolution aerial images, showcasing the accuracy of the proposed method across various types of targeting objects. Through several examples, the effectiveness of the approach was verified, underscoring its versatility and robustness. Moreover, the superior performance of SVMs in aerial image analysis underscores their suitability for tasks requiring precision and accuracy in object detection. By leveraging the strengths of SVMs, the proposed method achieves notable advancements in automated content-based object detection, offering valuable insights for applications in surveillance, urban planning, and environmental monitoring. The robustness of the approach is evidenced by its consistent performance across diverse targeting objects, indicating its potential for real-world deployment. Overall, this research highlights the efficacy of SVMs as a key component in



advancing the capabilities of object detection algorithms for aerial imagery analysis.

5. References:

- [1] I. Golightly, J. Dewi. "Visual control of an unmanned aerial vehicle for power line inspection," International Conference Advanced Robotics(ICAR05), Jul. 2005, pp. 228-295, doi:10.1109/ICAR.2005.1507426.
- [2] T.Li, F.Lijin, and W.Hongguang. "Development of an inspection robot control system for 500KV extra-high voltage power transmission lines," SICE 2004 Annual Conference, Aug. 2004, vol.2, pp.1819-1824.
- [3] T.Mingwen, et al. "Application of unmanned aerial vehicle in inspecting transmission lines," Electric Power, Mar. 2013,vol.3, pp. 35-38.
- [4] Z.Jingjing, et al. "High speed automatic power line detection and tracking for a UAV-based inspection," International Conference on Industrial Control and Electronics Engineering (ICICEE2012), Aug. 2012, pp. 266-269, doi:10.1109/ICICEE.2012.77.
- [5] C.Weiran, et al. "High voltage transmission line detection for uav based routing inspection." ASME International Conference on Advanced Intelligent Mechatronics (AIM2013), Aug.2013,pp.554-558,doi:10.1109/AIM.2013.6584150.
- [6] L.Zhengrong, et al. "Knowledge-based power line detection for UAV surveillance and inspection systems." 23rd International Conference Image and Vision Computing New Zealand(IVCNZ2008), Nov.2008,pp.1-6, doi: 10.1109/IVCNZ.2008.4762118.
- [7] Z.Lipo, et al. "Research on real-time detection and recognition algorithm of high-voltage transmission line for inspection with unmanned aerial vehicle." Journal of Chinese Computer Systems,vol 33(4), Apr.2012, pp. 882-886.