Caelum Incognitum: A Deep Learning Approach to Classifying Unidentified Flying Objects

Christian Fischer
Hugo Boss

Jonathon Olson
Fresenius

Maya Seifert
SAP

Anis Tajioui
Zurich

Sarah Wonke
SAP

I. Motivation

With the growing presence of drones and other aerial objects, automated image classification has become increasingly important [1]. However, traditional closed-set models fail when facing unknown objects, causing misclassification and reduced reliability [2]. This project addresses these challenges by combining Convolutional Neural Network (CNN)-based classification with an Open-Set Recognition (OSR) mechanism to improve robustness in real-world scenarios.

II. Project Scope

The goal is to build a classifier for Unidentified Flying Objects (UFOs) with a small taxonomy (e.g., bird, drone, aircraft, unknown). The project focuses on single-image classification using RGB images sourced from public datasets, excluding video analysis and real-time tracking. A basic OSR mechanism is included to flag objects outside the known classes.

III. Related Work

Flying object classification has historically relied on CNNs to solve closed-set problems. For instance, classic research utilized CNNs to specifically differentiate between drones and birds in aerial imagery, demonstrating robust performance for pre-defined categories [3]. However, in real-world scenarios, these closed-set models fail to recognize novel objects, often misclassifying them as one of the known classes. To address this, previous work explored transitioning CNN-based architectures into an OSR context by employing the CNN as a mere feature extractor, where the final classification and out-of-distribution rejection are handled by an ensemble method [4]. While the most recent literature points toward highly flexible Open Vocabulary approaches, leveraging large pre-trained vision-language models like CLIP [5] for advanced Open-World detection, this work strategically returns to the established CNN architecture as its core. Our focus is thus on developing a robust, resource-aware OSR system for flying objects based on a reliable CNN backbone.

IV. Data Collection & Preprocessing

Several publicly available Kaggle datasets provide RGB images of labeled aerial objects, which serve as the foundation for training and evaluation. Pre-processing steps include image resizing to ensure consistent input dimensions, normalization to standardize pixel values, data augmentation (flips, rotations) to improve generalization, as well as handling class imbalance [6]. Data will be split into training, validation and test sets [7], with care taken to maintain representative distributions. For OSR, some classes are treated as "unknown" in the test set only, simulating real-world conditions where unseen categories occur during inference. This ensures the model is evaluated both on familiar and truly novel instances.

V. Model / Architecture

The primary model will be a CNN, fine-tuned on the known classes. A simple OSR mechanism will be applied on top of the classifier, using either a confidence- or energy-based threshold to label objects as "unknown" when they fall below the threshold [8], [9]. Additional standard CNN models will be trained as closed-set baselines to quantify the added value of the OSR mechanism.

VI. Model Training

Standard supervised training is applied to the known classes. Hyperparameters such as learning rate, batch size and optimizer are tuned via the validation set. Since the OSR mechanism operates as a post-processing step during inference, it does not require additional training. However, the threshold for unknown detection will be determined using validation data to improve OSR-specific metrics.

VII. Model Evaluation

Evaluation will be conducted across two dimensions: For closed-set classification, metrics such as accuracy, precision, recall, and F1-score will be used to measure performance on known classes [10], [11]. In the open-set setting, OSR-specific metrics including open-set accuracy, F-measure for unknown detection, AUROC, and Youden's index will assess the system's ability to correctly identify previously unseen classes [7].

VIII. Model Deployment & Reproducibility

The trained model will be deployed through a lightweight web interface built with Svelte, enabling users to upload images and receive predictions in real time. This minimal setup guarantees consistent results across different systems while keeping deployment simple and accessible. To ensure reproducibility, the project will follow basic best practices, such as maintaining clear documentation and fixing random seeds where possible.

AI Usage Statement

AI tools were used to assist with phrasing and improving the clarity of selected text sections in this sketch. All AI-generated suggestions were reviewed and edited by the team.

References

[1] Statista, "Prognose zum Volumen des weltweiten Drohnenmarktes in den Jahren 2022 bis 2030." 2025.

[2] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward Open Set Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1757–1772, 2013, doi: 10.1109/TPAMI.2012.256.

[3] F. Mahdavi and R. Rajabi, "Drone detection using convolutional neural networks," in *2020 6th Iranian conference on signal processing and intelligent systems (ICSPIS)*, 2020, pp. 1–5.

[4] L. Balasubramanian, F. Kruber, M. Botsch, and K. Deng, "Open-set recognition based on the combination of deep learning and ensemble method for detecting unknown traffic scenarios," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, 2021, pp. 674–681.

[5] Y. Li *et al.*, "Toward open vocabulary aerial object detection with clip-activated student-teacher learning," in *European Conference on Computer Vision*, 2024, pp. 431–448.

[6] A. Ahmad, A. Wanto, and S. M. Adnan, "Bird and Drone Image Classification Using ResNet CNN: A Deep Learning Approach for Aerial Surveillance," *Bulletin of Computer Science Research*, vol. 5, no. 4, pp. 372–381, 2025.

[7] C. Geng, S.-j. Huang, and S. Chen, "Recent advances in open set recognition: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3614–3631, 2020.

[8] W. Liu, X. Wang, J. Owens, and Y. Li, "Energy-based out-of-distribution detection," *Advances in neural information processing systems*, vol. 33, pp. 21464–21475, 2020.

[9] G. Baye, P. Silva, A. Broggi, N. D. Bastian, L. Fiondella, and G. Kul, "varMax: Towards Confidence-Based Zero-Day Attack Recognition," in *MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM)*, 2024, pp. 863–868.

[10] J. C. Obi, "A comparative study of several classification metrics and their performances on data," *World Journal of Advanced Engineering Technology and Sciences*, vol. 8, no. 1, pp. 308–314, 2023.

[11] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," *arXiv preprint arXiv:2008.05756*, 2020.