DEVELOPMENT OF MODELS AND ALGORITHMS FOR ANOMALY DETECTION FROM DRONE IMAGES

Abdullayev Jahongir Ilxambay ugli
Independent Researcher of TUIT

Kuchkorov Timurbek Atakhanovich
PhD, Associate Professor
Dean of the Faculty of Computer Engineering of TUIT,

Annotation:
The rapid advancement of unmanned aerial vehicles (UAVs) has enabled large-scale image acquisition for various applications, including environmental monitoring, agriculture, infrastructure inspection, and security. However, the vast amount of data generated necessitates automated anomaly recognition methods to identify irregularities efficiently. This article examines the development of models and algorithms for anomaly recognition from drone images. It reviews state-of-the-art approaches, outlines the theoretical basis of anomaly detection, and proposes methodological frameworks integrating machine learning, deep learning, and computer vision techniques.

Keywords: Drone imagery, anomaly recognition, machine learning, deep learning, convolutional neural networks (CNNs), transformer models, computer vision, UAV data analysis.

Drone technology has revolutionized the way spatial and environmental data are collected. Compared to traditional remote sensing platforms such as satellites, drones offer high-resolution, flexible, and cost-effective imagery. However, the effectiveness of UAV applications depends largely on the ability to automatically detect anomalies within large volumes of data. Anomalies may include damaged infrastructure, disease outbreaks in crops, illegal activities, or environmental hazards.

Introduction to Anomaly Detection in Drone Images: Anomaly detection in drone images, also known as unmanned aerial vehicle (UAV) imagery, is a critical task in computer vision that focuses on identifying unusual patterns, objects, or events within aerial photographs or video frames. These anomalies could include structural defects in bridges or power lines, unauthorized intrusions in restricted areas, sudden environmental shifts like wildfires or flooding, or mechanical failures in agricultural fields. The importance of this technology stems from its wide-ranging applications, such as infrastructure maintenance, search and rescue operations, precision agriculture, environmental monitoring, and rapid disaster response. Drone-captured data is particularly valuable because it provides high-resolution, bird's-eye views over large areas, but it also presents unique challenges. These include inconsistent lighting conditions due to varying altitudes and times of day, motion blur from drone movement, partial occlusions by foliage or terrain, and the presence of multi-scale objects that appear tiny or distorted from afar. Given that drone datasets are often vast and lack comprehensive labels□making supervised learning inefficient□most approaches rely on unsupervised or semi-supervised deep learning techniques. These methods learn representations of "normal" scenes from unlabeled data and then detect deviations, such as through reconstruction errors or statistical outliers, enabling scalable and adaptable systems.

Common Models and Algorithms: The development of models and algorithms for anomaly detection in drone images draws from a rich ecosystem of computer vision techniques, primarily leveraging convolutional neural networks (CNNs) for feature extraction. Unsupervised methods dominate because they require only normal data for training, treating anomalies as outliers in the learned feature space. A foundational approach is the Convolutional Autoencoder (CAE) combined with Support Vector Data Description (SVDD). In this setup, the CAE first processes sequences of drone images to extract spatiotemporal features, capturing both spatial details like textures and edges, as well as temporal dynamics across frames if video is involved. The SVDD then fits a hypersphere around these features in a high-dimensional space, classifying points outside the boundary as anomalies. To handle noisy or complex drone environments, this method incorporates a 0/1 soft-margin loss and optimizes via Bregman Alternating Direction Method of Multipliers (ADMM), making it robust for detecting subtle faults like cracks in solar panels or irregular vegetation patterns. Research has shown this combination achieving accuracies exceeding 95% on adapted UAV fault datasets, highlighting its efficacy for real-world aerial inspections.

For more object-focused detection, an object-centric pipeline integrates YOLOv3 for initial object detection□such as spotting vehicles, humans, or buildings in drone footage□with autoencoders for feature refinement and Support Vector Machines (SVMs) for final classification. Here, YOLOv3 scans the image to localize potential regions of interest, after which semantic segmentation extracts contextual information (e.g., surrounding terrain), optical flow computes motion cues across frames, and appearance features describe visual attributes like color and shape. A few-shot SVM then infers anomalies based on these multi-modal descriptors, excelling in multi-scene scenarios like urban surveillance where anomalies might involve over-speeding vehicles or abandoned objects. On the MUAAD dataset, which features diverse aerial videos, this approach yields an Area Under the Curve (AUC) of 0.712, a notable improvement over baselines lacking contextual integration, demonstrating how fusing modalities enhances localization precision in dynamic drone captures.

More advanced, memory-efficient techniques like PatchCore and PaDiM, available through libraries such as Anomalib, offer state-of-the-art performance without requiring full model retraining. PatchCore builds a memory bank of patch-level features from normal drone images using a pre-trained CNN backbone like ResNet, then scores new patches by comparing them to the nearest neighbors in the bank□low similarity indicates anomalies. PaDiM, on the other hand, employs probabilistic density estimation on multivariate Gaussian distributions of features, flagging low-likelihood regions. Both are particularly suited for edge deployment on drones, with fast inference times, and have demonstrated AUROC scores above 0.95 on industrial defect benchmarks like MVTec AD, which can be readily adapted to drone imagery for tasks like pipeline leak detection.

Generative models, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs) variants like AnoGAN or GANomaly, provide another pillar. These reconstruct input images from a latent space learned solely on normal data; anomalies are pinpointed by elevated reconstruction errors, often measured via mean squared error (MSE) or perceptual losses. For drone applications, adaptations incorporate 1D or 2D convolutions to handle swarm footage or wide-area scans, with Bayesian optimization tuning hyperparameters like latent dimension size. In binary classification of real flight data, these have proven effective for isolating anomalous channels, such as erratic drone paths signaling hardware issues, though they require careful regularization to avoid overfitting to aerial-specific artifacts like shadows.

Key Datasets for Development: To benchmark and train these models, several public datasets tailored to drone imagery are indispensable. The InsPLAD dataset stands out for industrial applications, comprising over 10,000 high-resolution images captured by drones inspecting power line assets. It includes nearly 29,000 annotated instances across 17 classes of components, such as insulators and conductors, with anomalies categorized into six defect types like corrosion, fractures, or bird nests. This binary labeling of normal versus anomalous scenes makes it ideal for unsupervised training, enabling models to learn fine-grained aerial defects under varying weather conditions.

The MUAAD dataset, or Manipal UAV Anomalous Activity Dataset, shifts focus to behavioral anomalies in multi-scene aerial videos, totaling 60 clips from nine urban and rural locations at 1024x720 resolution. It annotates frame-level events across seven anomaly patterns, including traffic violations like no-parking infractions or over-speeding, providing a rich ground for temporal models that process drone video streams in real-time.

For broader scalability, the A2Seek benchmark offers a massive collection exceeding 2.4 million frames, with over 398,000 anomalous examples spanning diverse aerial events like crowd disturbances or vehicle malfunctions in expansive scenes. This dataset emphasizes large-scale challenges, such as varying drone altitudes and viewpoints, making it a go-to for evaluating generalization.

Smaller but specialized sets like UIT-ADrone target search-and-rescue missions, featuring color images of environmental and urban anomalies, such as debris fields or lost hikers, derived from simulated drone flights. Complementing these, repositories like ADRepository curate over a dozen image datasets adaptable from autonomous driving contexts, while video-focused compilations on GitHub, such as awesome-video-anomaly-detection, aggregate resources for sequential drone data.

Steps for Developing Models and Algorithms: Developing these systems follows a structured pipeline, starting with data collection and preprocessing. Begin by gathering drone imagery using commercial platforms like DJI models at 4K resolution, ensuring diverse captures across altitudes, angles, and conditions. Preprocess by applying augmentations□rotations, flips, brightness adjustments□to simulate variability; normalize pixel values via Z-score; and resize to standard inputs like 256x256 pixels. Partition the data with 80% normal samples for training, reserving mixed sets for validation and testing to mimic real deployments.

Next, select and train the model using frameworks like PyTorch or TensorFlow. For unsupervised paradigms, minimize objectives like reconstruction loss on normal data alone, employing CNN backbones for efficiency. Hyperparameter tuning via grid search or Bayesian methods optimizes elements such as learning rates (often around 1e-3) and batch sizes (typically 32), with early stopping to prevent overfitting.

Conclusion

The development of models and algorithms for anomaly recognition from drone images represents a critical step toward advancing the practical applications of UAV technology in agriculture, environmental monitoring, infrastructure management, and security. The review of existing approaches highlights a clear transition from traditional image-processing methods to AI-driven techniques, particularly deep learning and transformer-based models. Experimental results confirm that CNNs remain highly effective for feature extraction, while transformer architectures and hybrid frameworks improve robustness and contextual understanding.

Despite notable progress, challenges such as limited labeled datasets, environmental variability, and the need for real-time onboard processing remain unresolved. Addressing these issues requires the integration of self-supervised learning, lightweight neural networks, and multi-sensor data fusion. Establishing standardized UAV anomaly detection benchmarks will further accelerate advancements.

References

1. Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. Proceedings of the International Conference on Engineering and Technology (ICET), 1□6. IEEE. https://doi.org/10.1109/ICEngTechnol.2017.8308186
2. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/2010.11929
3. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
4. Kang, J., Körner, M., Wang, Y., Taubenböck, H., & Zhu, X. X. (2018). Building instance classification using street-view textural and structural features. IEEE Transactions on Geoscience and Remote Sensing, 56(3), 1561□1576. https://doi.org/10.1109/TGRS.2017.2765838
5. Li, Y., Yang, X., Sun, X., & Luo, J. (2020). A survey on deep learning in UAV-based remote sensing. Remote Sensing, 12(17), 2826. https://doi.org/10.3390/rs12172826
6. Mohammadi, V., Salehi, B., Mahdianpari, M., Quackenbush, L., Brisco, B., & Homayouni, S. (2020). A meta-analysis on UAV remote sensing for precision agriculture: Current trends and future directions. Remote Sensing, 12(21), 3555. https://doi.org/10.3390/rs12213555
7. Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. ACM Computing Surveys (CSUR), 54(2), 1□38. https://doi.org/10.1145/3439950
8. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767. https://arxiv.org/abs/1804.02767
9. Zhang, C., Li, W., & Travis, D. (2018). Gaps in the development of UAV remote sensing applications: A review. Remote Sensing, 10(9), 1379. https://doi.org/10.3390/rs10091379
10. Zhao, Z., Zheng, P., Xu, S., & Wu, X. (2019). Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 30(11), 3212□3232. https://doi.org/10.1109/TNNLS.2018.2876865.