ANOMALY DETECTION IN DRONE VIDEOS FOR PREVENTIVE MAINTENANCE OF POWER LINES

Guillaume FOURRET
LIRMM/ICAR, Univ. Montpellier
CNRS / Drone Geofencing
Montpellier, France
guillaume.fourret@lirmm.fr

Marc CHAUMONT
LIRMM/ICAR, Univ. Montpellier
CNRS, Univ. Nîmes
Montpellier, France
marc.chaumont@lirmm.fr

Gérard SUBSOL
LIRMM/ICAR, Univ. Montpellier
CNRS
Montpellier, France
gerard.subsol@lirmm.fr

Christophe FIORIO
LIRMM/ICAR, Univ. Montpellier, CNRS
Montpellier, France
christophe.fiorio@lirmm.fr

Samuel BRAU
Drone Geofencing
Nîmes, France
samuel.brau@drone-geofencing.com

***Abstract*—Power lines are composed of various components that can deteriorate over time due to use. To detect potential issues with these components and prevent costly network outages, aerial drones are increasingly being utilized. They allow for the rapid inspection of large distances and provide a clear view of the different components and their defects. However, the manual analysis of flight videos by experts is labor-intensive. Moreover, the wide variety of anomalies that can cause network interruptions makes it impractical to develop dedicated automated solutions for each type, particularly due to the limited number of examples available for many anomalies. We thus propose a method to automatically detect anomalies in scenarios where no prior information about the visual appearance of anomalies is available, using drone-acquired videos. The approach relies on the computation of an anomaly score based on a generic feature vector. Results demonstrate that this approach is effective in scenarios without any examples of anomalies and requires very limited computational resources for learning.**

***Index Terms*—Object detection, Anomaly detection, Deep Learning, Aerial drone, Video.**

I. INTRODUCTION

Power outages caused by defects in power line components are extremely costly for electricity providers. These malfunctions can have numerous causes, ranging from surrounding vegetation to component failures. Preventive maintenance is therefore essential for these companies to anticipate such failures. While the inspection of power lines using drones is now common practice, it typically requires visual inspection by technicians and could be automated [1–3]. For instance, the automated analysis of one of the primary components, the insulator, has been studied, particularly in cases involving missing discs [4, 5]. However, these methods are tailored to the analysis of a specific, predefined object and a clearly identified anomaly.

Anomalies, however, can vary significantly in appearance depending on viewing conditions such as distance, lighting, and other factors. Recent methods for more generic anomaly detection leverage classification neural networks trained on large datasets containing objects with known anomalies [6]. By definition, anomalies are rare, making it challenging to collect a dataset large enough for training. This limitation has driven the development of solutions that require no prior information, i.e., no examples of anomalies. These solutions focus on identifying extreme data points (commonly referred to as 'outliers') that deviate from the distribution of the majority of the data.

Many approaches have been proposed in the past for such scenarios (i.e., unsupervised learning). Most of these methods rely on hand-crafted features, including those based on decision trees or random forests (e.g., Isolation Forest), density estimation (e.g., LOF), angular relationships (e.g., ABOD), statistical models (e.g., Gaussian Mixture Models), hypersphere-based methods (e.g., One-Class SVM), clustering techniques (e.g., Spectral Clustering), and autoencoders (where the reconstruction error for anomalies deviates from the learned patterns). For instance, in the case of insulators, a local hand-crafted descriptor is computed for each disc in a chain of insulators as described in [7]. Anomaly detection is then performed in an unsupervised manner using the Local Outlier Factor (LOF) to compare these descriptors.

In this paper, we propose a pipeline designed to assign an anomaly score to all detected objects along a power line. Unlike traditional computer vision methods, which are carefully designed to detect a specific anomaly on an object [8], our approach is not limited to any particular type of anomaly and relies on a generic descriptor that is not specialized for a

We thank the National Association for Research and Technology (ANRT) and Drone Geofencing for their funding. We also thank Enedis for providing the videos.

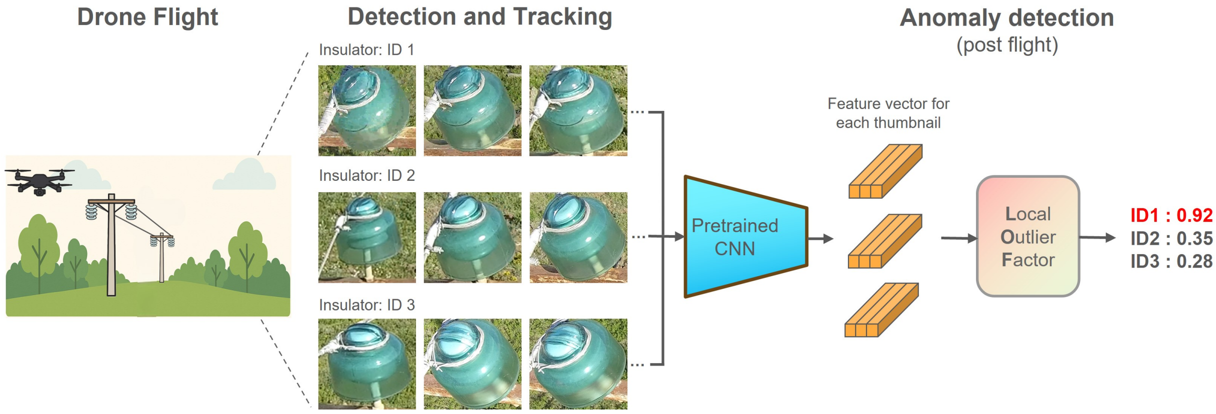

Fig. 1. Overview of our anomaly detection pipeline. Anomalies in power line components are identified as deviations from others of the same class detected during the drone's flight, assuming most are defect-free. This approach eliminates the need for reference or training defect examples, making it particularly suitable given the inherent challenges of data collection in practice.

specific object. During the drone flight, most, if not all, of the observed objects (e.g., insulators, steel pole caps, tops of wooden poles, etc.) are expected to be in normal condition. Most objects detected along the power line are assumed to be in good condition, with anomalies clearly standing out. An anomalous object is expected to exhibit a significantly different score, facilitating its detection. Our general idea is to compare all objects of the same class after the flight to identify those that deviate from the norm. This approach eliminates the challenging task of collecting data for various types of anomalies, which is difficult to obtain for training a discriminator network. Moreover, our anomaly detection method does not require defect-free examples to learn a distribution of normal objects, unlike most autoencoder-based approaches [9].

The overall process for detecting anomalies in drone videos is illustrated in Figure 1, and begins with the identification of each object of interest. This is achieved using an object detection network to extract sub-images (thumbnails) of the objects seen during the flight. Subsequently, a convolutional neural network (CNN) is used to extract a feature vector from each thumbnails, effectively representing the object. Finally, a score is computed by comparing these feature vectors, enabling the identification of potential outliers.

II. State-of-the-art

A. Multiple Object Tracking

Multiple Object Tracking (MOT) algorithms aim to identify and track multiple objects without any prior knowledge of their number. Object detection algorithms output a collection of rectangular bounding boxes, defined by their coordinates, height, and width, surrounding the thumbnail of the object. In addition, MOT algorithms also assign a target ID to each bounding box, enabling the tracking of each individual object throughout its appearance in the video [10].

Most MOT algorithms follow the following steps: first, detect objects in the current frame; second, optionally extract features describing each object and use a motion predictor to estimate their positions in the next frame; third, compute a similarity or distance score between pairs of detections using these features; and fourth, associate the bounding boxes from the current frame to those in the next frame [10].

The first step is typically performed using an object detection algorithm. In our case, we used YOLOv5 [11], a one-stage detector from the YOLO family [12]. While more recent detectors are available (e.g. DETR [13]), YOLOv5 was sufficient for demonstrating feasibility of our anomality detection pipeline. Additionally, it is a well-known, accurate, lightweight, and fast detector, making it suitable for deployment on embedded systems aboard drones, especially when optimized with inference engines (e.g., TensorRT) and/or precision quantization [14, 15].

The second step can leverage features provided by the object detector. Finally, the third and fourth steps, which involve associating bounding boxes, can be accomplished using trackers [16–18], deep learning approaches (including end-to-end methods) such as [19], or graph-based methods [20].

B. Feature extraction and anomaly score computation

1) A generic feature vector:

Once the thumbnails of each class for each frame have been extracted using the MOT process, it is necessary to compute a generic and discriminative feature descriptor to both compare vectors along consecutive frames and to calculate an anomaly score. A feature vector for a thumbnail should encapsulate the object appearance. We utilize the last convolutional layer of a VGG16 neural network [21], pre-trained on ImageNet. While ResNet [22] and EfficientNet [23] were also evaluated as feature extractors, VGG16 was selected as it experimentally yielded the best results for anomaly discrimination. VGG16 produces a feature vector capable of capturing subtle appearance variations within objects of the same class.

Each thumbnail of a given class is represented by a feature vector denoted as $\mathbf{X}_i \in \mathbb{R}^d$, where d is the dimension of the vector, $i \in \{1, .., N\}$ is the index of the thumbnails and, N is the total number of thumbnails. From the flight sequence, we thus have a set of feature vectors $\{\mathbf{X}_i\}_{i=1}^{i=N}$. Some of these vectors may represent objects exhibiting anomalies.

2) Local Outlier Factor (LOF):

Detected object of the same class acquired during the same fly over a power line should present roughly the same appearance. A defective object could be then considered as an outlier. We thus use the Local Outlier Factor (LOF) [24] as a measure for evaluating the local deviation of a given feature vector with respect all the others.

The LOF is based on the concept of 'local reachability density', denoted 'lrd', with respect to the k-nearest neighbors:

$$lrd_k(\mathbf{X}_i) = 1/\left(\frac{\sum_{\mathbf{X}_j \in N_k(\mathbf{X}_i)} d_k(\mathbf{X}_i, \mathbf{X}_j)}{|N_k(\mathbf{X}_i)|}\right), \quad (1)$$

where $d_k(\mathbf{X}_i, \mathbf{X}_j)$ represents the reachability distance, defined as the maximum between the Euclidean distance between $\mathbf{X}_i$ and $\mathbf{X}_j$, and the Euclidean distance between $\mathbf{X}_j$ and its k-th nearest neighbor. $N_k(\mathbf{X}_i)$ represents the set of the k-nearest neighbors of $\mathbf{X}_i$. Based on this local reachability density, we compute the Local Outlier Factor of a vector, which is the ratio between the local reachability density of $\mathbf{X}_i$ and that of its neighbors $N_k(\mathbf{X}_i)$:

$$LOF_k(\mathbf{X}_i) = \frac{\sum_{\mathbf{X}_j \in N_k(\mathbf{X}_i)} lrd_k(\mathbf{X}_j)}{|N_k(\mathbf{X}_i)| \times lrd_k(\mathbf{X}_i)}. \quad (2)$$

This allows us to assign an anomaly score based on the local density of a vector. The more isolated a vector is (fewer neighbors and/or distant neighbors), the higher its LOF score will be compared to other vectors.

III. Experimental setup

The used image detector was YOLOv5m6 model [11], which we trained on 143 images of low-voltage "insulators" using an Nvidia RTX3070 Laptop GPU. The 4,096×2,160 resolution images were extracted from two videos captured during real preventive maintenance drone missions. The hyperparameters and weights used are those from the pretraining on MSCOCO conducted by Ultralytics[1]. We set a batch size of 4, an image resolution of 1,280×720, and the training lasted for 85 epochs.

We used tracking algorithms like SORT [18] and StrongSORT [25]. An ID is assigned to each detected object in an image sequence, based on motion estimation using a Kalman filter and texture similarity through the correlation of feature vectors computed on the thumbnails. This approach allowed us to enhance insulator detection by lowering the confidence threshold at which YOLO detections are considered valid. Potential false positives were then filtered by considering only tracked objects that appear in more than 15 consecutive frames.

Finally, we used VGG16 to extract the feature vectors of dimension $d = 4\,096$, for each thumbnail and then calculated the LOF for each one.

[1]https ://github.com/ultralytics/yolov5

IV. Results and Discussions

The first step consists in running an object detection in the MOT pipeline. Figure 2 shows the detection of insulators on a test image.

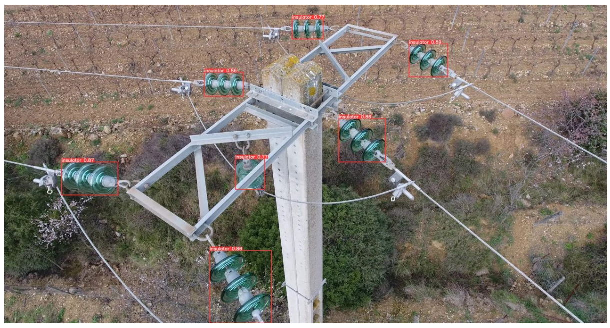

Fig. 2. Example of insulator detection by YOLOv5m6.

As previously mentioned, we lowered the confidence threshold at which YOLO detections are considered valid to reduce false negatives. Potential false positives were then filtered by retaining only tracked objects that appear in more than 15 consecutive frames. Figure 3 compares the final detection results over an entire video sequence under three conditions: without tracking at a higher confidence threshold, without tracking at a lower confidence threshold, and after applying MOT processing and filtering. In this example, the drone flew over three poles containing 3, 7, and 3 insulators, respectively. The results show that the number of detected insulators is stable and consistent with the ground truth.

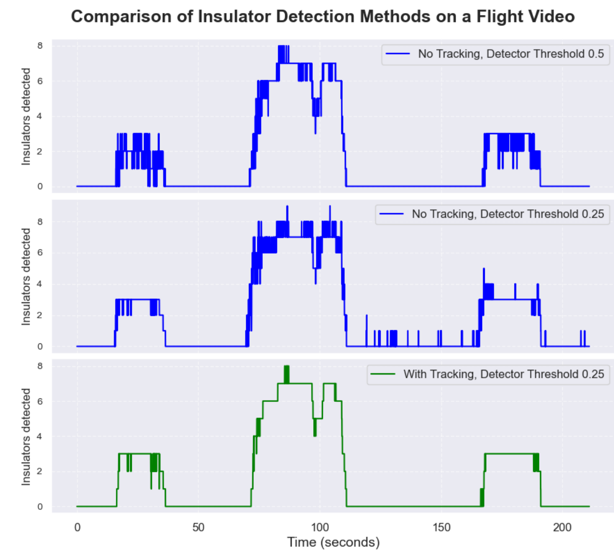

Fig. 3. Number of insulators detected as a function of the time in seconds. The drone flew over three poles, containing 3, 7, and 3 insulators, respectively. It can be observed that the number of detected insulators is stable and consistent with the ground truth.

Note that not all insulators are always detected, partly due to occlusions. However, this does not pose a significant issue for anomaly detection, as it is preferable to focus on visually representative instances within the tracklet (i.e., thumbnails displaying the complete object) to determine the presence of anomalies.

We applied the entire processing pipeline (detection + MOT + LOF) to a drone video provided by Enedis[2], the company responsible for managing and developing 95% of the electricity distribution network in France. In this video, one insulator exhibits a crack. Figure 4 shows the LOF scores of all the thumbnails of objects detected as insulator, sorted in descending order. The anomaly scores corresponding to the images of the defective insulator are highlighted by the red points. It can be observed that the scores for the thumbnails of this insulator are among the most distinct and stand out clearly. Note that the negative LOF values in the graph result from the sklearn implementation, which outputs the negative of the LOF value.

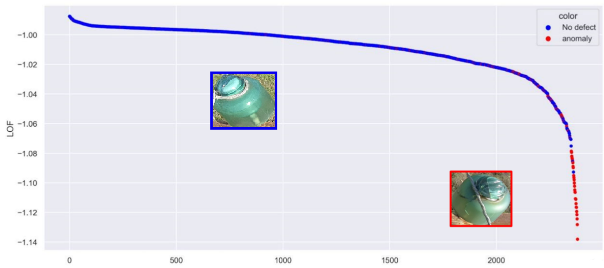

Fig. 4. LOF scores (k=400) associated with the thumbnails of insulators detected by YOLO during a flight. The thumbnails are sorted in descending order of their anomaly scores. Cracked insulators exhibit LOF scores that are distinct from those of normal insulators.

We also evaluate the performance of our VGG16 feature extractor coupled with LOF computation using a set of thumbnails of 'steel electric pole caps' and 'wooden poles' manually extracted from a video. Some of the wooden poles were damaged by termites, and some of the steel poles had their caps disappeared. It can be observed in Figures 5 and 6, that these defective objects also exhibit LOF scores that stand out from those of the defect-free objects.

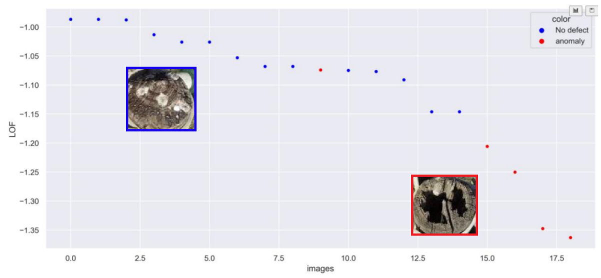

Fig. 5. LOF scores (k=2) associated with the thumbnails of wooden poles. The thumbnails are sorted in descending order of their anomaly scores. The poles damaged by termites, having a different texture, generally receive distinct anomaly scores.

For all the results, the thumbnails of outliers show a higher LOF score compared to the other thumbnails. This result is interesting as it highlights the potential defective objects.

[2]The video is different from these used for the YOLOv5m6 training.

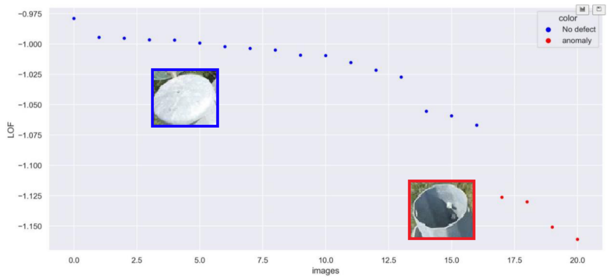

Fig. 6. LOF scores (k=2) associated with the thumbnails of steel poles. The thumbnails are sorted in descending order of their anomaly scores. The absence of a cap on certain poles significantly affects their visual appearance, causing their anomaly scores to stand out prominently.

Setting the LOF threshold value is given to the drone operator or the maintenance technician. Figure 7 illustrates a graphical interface where each object is represented by an horizontal bar on the flight timeline. The color red signifies that our method suspects an anomaly. Clicking on a bar opens a window with a representative thumbnail of the corresponding tracklet for further examination.

Fig. 7. Graphical User Interface. The x-axis represents the flight video timeline. Each detected object is depicted as a colored bar, with the color indicating its normality (green = normal, red = anomalous). The user can click on a bar to view the representative image of the corresponding tracklet for further analysis.

V. Conclusion

In this paper, we propose a generic method for visual defect analysis. A detector (YOLO) combined with a tracking approach, trained on a few hundred examples, allows for the extraction of all objects of interest from a video sequence. Then, by using a generic feature extractor (VGG16) and a geometric measure of outliers (LOF), we are able to detect defective objects. In the case of a drone flying above a power line, we can detect insulators exhibiting cracks, steel poles with missing caps, and wooden poles damaged by termites.

Future extensions could include the use of Few-Shot Learning for detection [26], the use of other feature extractors such as DINO [27] or semi-supervised anomaly detection [28], an analysis of the sensitivity of the hyperparameter k in the computation of LOF, and finally, active/incremental learning to incorporate expert feedback and continuously improve our model [29]. Additionally, further research could focus on automatically preventing the computation of LOF for detected objects that are partially occluded, as their altered extracted features inherently increase the likelihood of being incorrectly flagged as anomalous by our pipeline.

References

[1] V. N. Nguyen, R. Jenssen, and D. Roverso, "Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 107–120, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061517324444

[2] X. Liu, X. Miao, H. Jiang, and J. Chen, "Data analysis in visual power line inspection: An in-depth review of deep learning for component detection and fault diagnosis," *Annu. Rev. Control.*, vol. 50, pp. 253–277, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:225138270

[3] D. Sadykova, D. Pernebayeva, M. Bagheri, and A. P. James, "In-yolo: Real-time detection of outdoor high voltage insulators using uav imaging," *IEEE Transactions on Power Delivery*, vol. 35, pp. 1599–1601, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:208843642

[4] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, pp. 1486–1498, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:126126016

[5] X. Li, H. Su, and G. Liu, "Insulator defect recognition based on global detection and local segmentation," *IEEE Access*, vol. 8, pp. 59 934–59 946, 2020.

[6] V. N. Nguyen, R. Jenssen, and D. Roverso, "Intelligent monitoring and inspection of power line components powered by uavs and deep learning," *IEEE Power and Energy Technology Systems Journal*, vol. 6, no. 1, pp. 11–21, 2019.

[7] M. Oberweger, A. Wendel, and H. Bischof, "Visual recognition and fault detection for power line insulators," in *Computer Vision Winter Workshop*, 2014.

[8] X. Zhang, J. An, and F. Chen, "A method of insulator fault detection from airborne images," in *2010 Second WRI Global Congress on Intelligent Systems*, vol. 2, 2010, pp. 200–203.

[9] A. Bauer, S. Nakajima, and K.-R. Müller, "Self-supervised training with autoencoders for visual anomaly detection," *arXiv preprint arXiv:2206.11723*, 2022.

[10] G. Ciaparrone, F. Luque Sánchez, S. Tabik, L. Troiano, R. Tagliaferri, and F. Herrera, "Deep learning in video multi-object tracking: A survey," *Neurocomputing*, vol. 381, pp. 61–88, 2020.

[11] G. Jocher, A. Chaurasia, A. Stoken, and al., "ultralytics/yolov5: v6.1 - tensorrt, tensorflow edge tpu and openvino export and inference (v6.1)." [Online]. Available: https://zenodo.org/records/6222936

[12] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, 2023. [Online]. Available: https://www.mdpi.com/2504-4990/5/4/83

[13] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I*. Berlin, Heidelberg: Springer-Verlag, 2020, p. 213–229. [Online]. Available: https://doi.org/10.1007/978-3-030-58452-8_13

[14] M. A. Farooq, W. Shariff, and P. Corcoran, "Evaluation of thermal imaging on embedded gpu platforms for application in vehicular assistance systems," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 2, pp. 1130–1144, 2023.

[15] C. Fite, M. Yedroudj, M. Chaumont, P. Serres-Noïtaky, F. Comby, F. Forget, G. Subsol, L. Mannocci, M. Capello, M. Tolotti, and L. Dagorn, "Some considerations to design an autonomous buoy system to enumerate pelagic sharks at Fish Aggregating Devices," LIRMM (UM, CNRS), Tech. Rep. 1, Aug. 2024. [Online]. Available: https://hal-lirmm.ccsd.cnrs.fr/lirmm-04895187

[16] E. Bochinski, V. Eiselein, and T. Sikora, "High-speed tracking-by-detection without using image information," *14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, 2017.

[17] E. Bochinski, T. Senst, and T. Sikora, "Extending iou based multi-object tracking by visual information," in *15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, pp. 1–6.

[18] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 3464–3468.

[19] Z. Wang, L. Zheng, Y. Liu, Y. Li, and S. Wang, "Towards real-time multi-object tracking," in *Computer Vision – ECCV 2020*, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds. Cham: Springer International Publishing, 2020, pp. 107–122.

[20] P. Chu, J. Wang, Q. You, H. Ling, and Z. Liu, "Transmot: Spatial-temporal graph transformer for multiple object tracking," *ArXiv*, vol. abs/2104.00194, 2021.

[21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition." in *ICLR*, Y. Bengio and Y. LeCun, Eds., 2015. [Online]. Available: http://dblp.uni-trier.de/db/conf/iclr/iclr2015.html#SimonyanZ14a

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[23] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 09–15 Jun 2019, pp. 6105–6114. [Online]. Available: https://proceedings.mlr.press/v97/tan19a.html

[24] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: identifying density-based local outliers," *SIGMOD Rec.*, vol. 29, no. 2, p. 93–104, May 2000. [Online]. Available: https://doi.org/10.1145/335191.335388

[25] Y. Du, Z. Zhao, Y. Song, Y. Zhao, F. Su, T. Gong, and H. Meng, "Strongsort: Make deepsort great again," *IEEE Transactions on Multimedia*, 2023.

[26] A.-K. N. Vu, N.-D. Nguyen, K.-D. Nguyen, V.-T. Nguyen, T. D. Ngo, T.-T. Do, and T. V. Nguyen, "Few-shot object detection via baby learning," *Image and Vision Computing*, vol. 120, p. 104398, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885622000270

[27] M. Oquab, T. Darcet, T. Moutakanni, H. V. Vo, M. Szafraniec, V. Khalidov, P. Fernandez, D. HAZIZA, F. Massa, A. El-Nouby, M. Assran, N. Ballas, W. Galuba, R. Howes, P.-Y. Huang, S.-W. Li, I. Misra, M. Rabbat, V. Sharma, G. Synnaeve, H. Xu, H. Jegou, J. Mairal, P. Labatut, A. Joulin, and P. Bojanowski, "DINOv2: Learning robust visual features without supervision," *Transactions on Machine Learning Research*, 2024. [Online]. Available: https://openreview.net/forum?id=a68SUt6zFt

[28] L. Jézéquel, N.-S. Vu, J. Beaudet, and A. Histace, "Anomaly detection via learnable pretext task," in *2022 26th International Conference on Pattern Recognition (ICPR)*, 2022, pp. 1178–1185.

[29] D. Zha, K.-H. Lai, M. Wan, and X. B. Hu, "Meta-aad: Active anomaly detection with deep reinforcement learning," *2020 IEEE International Conference on Data Mining (ICDM)*, pp. 771–780, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:221738930