UAV-based Surveillance System: an Anomaly Detection Approach

Amira Chriki
IResCoMath Research Unit
National School of Computer Science
University of Manouba
amirachriki@gmail.com

Haifa Touati
Hatem Bettahar IResCoMath Research Unit
Faculty of Science of Gabes
haifa.touati@cristal.rnu.tn

Hichem Snoussi
Institut Charles Delaunay
(ICD, FRE CNRS 2848)
University of Technology of Troyes
hichem.snoussi@utt.fr

Farouk Kamoun
CRISTAL Laboratory
National School of Computer Science
University of Manouba
farouk.kamoun@ensi.rnu.tn

***Abstract*—Recent advancements in avionics and electronics systems led to the increased use of Unmanned Aerial Vehicles (UAVs) in several military and civilian missions. One of the main advantages that makes UAVs attractive is their ability to reach remote regions that are inaccessible to human operators, i.e. provide new aerial perspective in visual surveillance. Autonomous visual surveillance systems require real time anomalies detection. However, there are many difficulties associated with automatic anomalies detection by an UAV, as there is a lack in the proposed contributions describing abnormal events detection in videos recorded by a drone. In this paper, we propose an anomaly detection approach in a surveillance mission where videos are acquired by an UAV. We combine deep features extracted using a pretrained Convolutional Neural Network (CNN) with an unsupervised classification method, namely One Class Support Vector Machine (OCSVM). The quantitative results obtained on the used dataset show that our proposed method achieves good results in comparison to existing technique with an Area Under Curve (AUC) of** 0.93.

***Index Terms*—UAV, Anomaly detection, Surveillance, Unsupervised Learning, One Class Classification, Convolutional Neural Network.**

I. Introduction

Recently, drones or *U*nmanned *A*erial *V*ehicles (UAVs) equipped with cameras are increasingly used in a wide range of civilian applications of a very varied nature, including: traffic monitoring [1], aerial photography [2], disaster area investigations [3], agricultural [4], fast delivery [5], surveillance [6], human action recognition [7], construction [8], etc. In surveillance context, UAVs equipped with a remote-control abilities, compact size, and large field of view, are often used as mobile cameras to overcome weaknesses of stationary cameras [9][10]. They are usually deemed suitable to reach crowded or disaster-hit regions which are inaccessible to human operators or otherwise challenging to navigate through [11]. UAVs usage can significantly decrease the manual load and allow law enforcement agencies to locate criminals or search for missing persons. Therefore, automatic understanding and treatment of visual data collected from these aerial platforms become highly demanding.

Automated visual surveillance systems have lately attracted increasing interest from the research community of the military and urban security systems. Such systems aim to automatically detect potential suspicious changes or signs of intrusion, and consequently generate a warning to a human operator. These suspicious changes are often referred to as anomalies or outliers [12]. Anomalies are very contextual, for example, running in a bank would be an anomaly, but running at a park would be normal. Furthermore, the definition of anomaly can be vague and usually ambiguous. One may think walking around on a subway station is normal, however, one other may think it should be signaled as an anomaly since it could be suspicious. Without being context specific, anomaly detection mainly consists of identifying changes between frames of the same scene but separated in time. Hence, automated detection of anomalous events in surveillance video can be used to provide more vigilant surveillance, possibly replacing, or assisting, human operators [13]. Moreover, on-board processing of anomaly detection can decrease the network traffics, the bandwidth consumption and the end to end delay by maintaining a minimum communication between UAVs and the *G*round *C*ontrol *S*tation (GCS) [14][15].

In visual surveillance system, primary data is a video which is difficult to model due to noise and its high dimensionality. Anomaly detection consist essentially in applying specific **features extraction techniques** to this video. In some cases, handcrafted features extraction methods such as: *H*istogram of *O*ptical *F*low (HOF), *H*istogram of *O*riented *G*radient (HOG) or *O*ptical *F*low (OF) can become inputs to anomaly detection methods [16-18]. Despite their effectiveness in extracting spatiotemporal features, they are too local to represent effectively complex events and need usually prior knowledge to match the representation with the targeted events. Object trajectories techniques have been also used [19]. These methods are generally not suitable for a crowded scene since they require accurate tracking technique with high computational complexity. Recently, deep learning methods have attracted a lot of interest from the research community. Deep learning based methods have the ability to automatically learn robust representations from raw data without prior knowledge. This

978-1-7281-8086-1/20/$31.00 ©2020 IEEE

major advantage has led to the development of numerous studies for anomaly detection based on deep learning [20][21]. The other fundamental step in the anomaly detection process is **the classification** of the extracted features. It can be divided into two categories: *T*wo-*C*lass and *O*ne-*C*lass *C*lassification. *T*wo-*C*lass methods [22][23] assume that during the learning process, the training data contain labeled instances of both normal and abnormal classes. An event is assumed abnormal if it falls in the abnormal class. However, given their supervised nature, these methods are not suitable to real-world missions. In fact, it is not possible to provide training examples for all abnormal events that can occur in a monitored scene due to their variability.

In *O*ne-*C*lass *C*lassification (OCC) [24][25], all training data have one label. Such methods learn a discriminative boundary around the normal data using a *O*ne-*C*lass *C*lassification algorithm. These methods have achieved good performance on anomaly detection and overcome the limits of the first category, since they use only normal samples during the training phase.

The contribution of this paper is summarized as follows. First, we deal with anomaly detection problem on a dataset containing videos taken by an UAV in a car parking, which is more complex than those commonly used. Second, we propose an approach, combining a pretrained *C*onvolutional *N*eural *N*etwork (CNN), namely GoogLeNet, and unsupervised machine learning technique namely *O*ne-*C*lass *S*upport *V*ector *M*achine (OCSVM), to detect anomalous scenes in the targeted scenario. GoogLeNet is used to extract robust features from UAV's videos, and OCSVM is used for features classification. The rest of this paper is organized as follows. Section 2 presents the related works. The description of our proposed approach and the experimental setup are provided in Section 3 and Section 4. The experimental results are discussed in Section 5. Finally, Section 6 concludes the paper and highlights some open issues.

II. Related work

Anomaly detection in visual surveillance systems is an ongoing project. Taking into account the constraints of mobile-camera (non-stationary background, illumination, camera movement, etc.), the choice of suitable features to extract from videos becomes a critical issue. Therefore, several works are interested in the proposal of a robust anomaly detection approach.

Existing mobile-camera anomaly detection approaches in the literature can be roughly placed into three classes: '*I*gnored-*B*ackground', '*R*eference-*F*rames' and '*D*eep-*L*earning'. In **Ignored-Background** approaches, authors propose to remove the background and focus on extracting the foreground information [18][26][27]. In **Reference-Frame** based methods [28][29], authors usually consider an anomaly-free reference video and compare it to a target video in order to detect anomalous scenes. Their aim is to assign each target frame to a background-reference frame so that their viewpoints are the closest ones. Approaches based on **Deep-Learning** extract automatically hierarchical discriminative features from high dimensional data for a treated task. Various deep-learning techniques have been used for anomaly detection such as: *C*onvolutional-*N*eural-*N*etwork (CNN) [30], *A*uto-*E*ncoders (AE) [31], *L*ong-*S*hort-*T*erm-*M*emory-networks (LSTM) [32], etc.

Some research in the existing literature have used drone as a mobile-camera in surveillance context. To the best of our knowledge, few are the works that have used a publicly available dataset. We find principally the work of [9], which use the dataset proposed by B. Margherita et al in [33]. The main works that used UAVs with camera for a suspicious events detection task are discussed below.

In [34], an anomaly detection and classification system is developed for processing and diagnosing the gas/image/thermal sensing data collected by an UAV in real time or near real time. *P*retrained *R*esidual *C*onvolutional *N*eural *N*etwork is used to perform multitask binary classification wherein the detected changes are classified into abnormal class (such as gas-oil leak, unauthorized human actions, facility failures) and normal class (insignificant/natural) signals.

In [9], authors propose an anomaly detection technique on the mini-drone-video-dataset [33] which consists of surveillance videos taken by an UAV. The proposed anomaly detector is a combination between a *C*onvolutional-*N*eural-*N*etwork (VGG16) and a *R*ecurrent-*N*eural-*N*etwork (LSTM), trained using supervised learning.

In [10], authors present an approach to distinguish anomalous and non-anomalous parts of a blade by using *O*ne *C*lass *S*upport *V*ector *M*achine with *VGG16* features learned from a generic image dataset. They also propose to subsample the frames taken by an UAV, project them to the feature space, and compress them by using *P*rinciple *C*omponent *A*nalysis (PCA) in order to make them learnable.

[9] and [34] use two different supervised learning approaches to classify the detected events. However, it is known that this type of learning approaches is not well suited for anomaly detection problems, as it is difficult, if not impossible, to create a dataset representing all possible anomalous events. In addition, the datasets used in [34] and [10] are not published, which makes the comparison with these works impossible. As already mentioned, [9] is the only work that has used a publicly available dataset for anomaly detection purpose, a comparison with such work is then possible.

III. Proposed model

In this section, we detail the approach we have put forward to automate the abnormal events detection task. We first overview the proposed system's main components and then we detail their principal roles.

A. System overview

The system described here is based on the principle that an abnormal event occurs when the most recent frames of video, captured by the UAV, are significantly different than the frames used in the training phase. Figure 1 depicts the overall

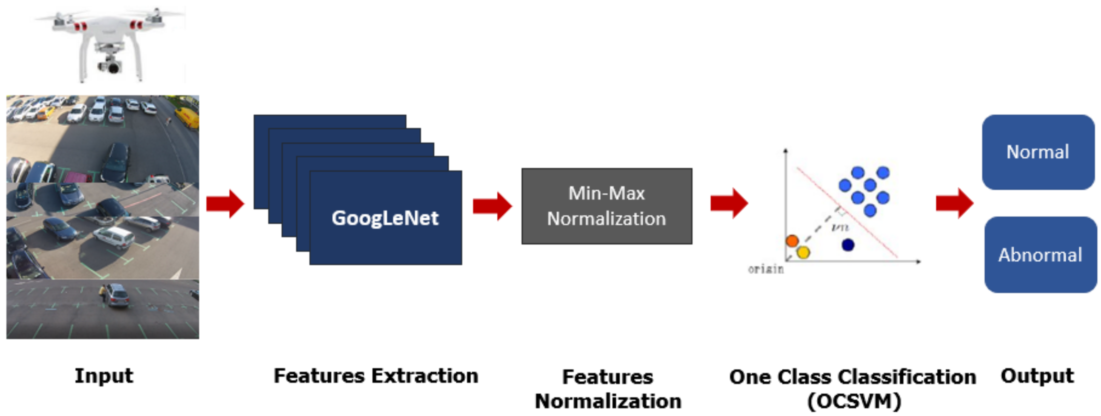

Fig. 1. Proposed model architecture

system architecture along with its principal components. It incorporates different stages starting from data collection using drone, followed by features extraction, features normalization and finally classification of input frames.

B. Feature Extraction

To extract features and learn pattern from input frames, we propose to use a pretrained CNN. The pretrained CNN used in this paper is the popular GoogLeNet [35]. GoogleNet was trained on a large sort of data with different number of classes, such as: ImageNet which contains 1.2 $million$ images with 1000 categories. Considering this fact, GoogLeNet should have learned a robust hierarchy of features, which are rotation, spatial and translation invariant. Therefore, having automatically learned a good hierarchical features representation for over a million images belonging to 1000 different categories, it can behave as a good feature extractor for new images. The main advantage of GoogLeNet is that it trains and runs very fast due to its inception block. It is comparatively smaller than VGG and AlexNet, as it has lower power and memory use (it has a size of only 96 MB). GoogLeNet has also outperformed the state of art classification and detection approaches in the *I*mageNet *L*arge-*S*cale *V*isual *R*ecognition *C*hallenge in 2014 (ILSVRC14).

C. Min-Max normalization

Min-max normalization, known also as feature scaling, is a strategy which linearly rescales a set of data. In most situations, normalization is used to scale the data between 0 and 1. It is defined as:

$$Y_i = [X_i - min(X)]/[max(X) - min(X)] \tag{1}$$

Where min and max are the minimum and maximum values in X, and where X_i is the i^{th} data point of X.

The minimum value in the original set X would be mapped to 0, and the largest value in X would be mapped to 1. Hence, every other value of X would be mapped to a value in the range 0 to 1. SVM algorithms, and in particular OCSVM, are one of the algorithms where feature scaling matters, i.e. normalized data, are easier for OCSVM to fit. Therefore, in this paper, we rescale our feature data using the *min-max* normalization.

D. Classification

To efficiently use the extracted features and to classify each frame as anomalous or normal, we have opted for OCSVM. OCSVM was extended by Scholkopf et al. [36] to handle training using only positive data. The problem addressed by OCSVM is novelty detection. The basic idea of novelty detection is to identify rare events, i.e. events of which we have very little samples. Therefore, the usual way of training a binary or multiclass classifier will not work in this case. OCSVM algorithm estimates a probability distribution function which makes most of the observed samples more likely than the rest, and a decision function that separates these samples by the largest possible margin. We applied OCSVM using the following steps:

- Train our support vector model with video data composed only of normal scenes.
- The trained model estimates a distribution that encompasses most of the normal samples and infers their properties.
- After the model is properly trained, it will be able to label as “abnormal”, the samples that lie far from the estimated distribution with respect to the inferred normal properties.

IV. EXPERIMENTAL SETUP

This section presents the experimental setup. We detail the used dataset, the implementation, the training process and the evaluation metrics.

A. Mini Drone Video Dataset

In order to evaluate the performance of the proposed approach on the anomaly detection task, experiments were conducted on the Mini-Drone Video Dataset (MDVD) [33]. MDVD was originally created for the design, analysis and evaluation of privacy systems. It is composed of 38 various videos captured in full HD resolution, with a duration of 16 to

24 $seconds$ each, shot with the mini-drone Phantom 2 Vision+ in a car parking. The videos in the dataset can be divided into three categories: normal, suspicious, and abnormal. Categories are almost all defined by the actions of the persons involved in the videos. **Normal scenes** depict several situations such as people walking, conversing, getting in their cars or parking their cars. In **suspicious scenes**, nothing a priori wrong happens but people have a suspicious behavior which would draw the attention of the surveillance staff. **Abnormal behaviors** show people fighting, mis-parking their vehicles, or stealing cars and items. Therefore, the MDVD could be used to evaluate an anomaly detection model.
Videos were recorded with different positions, altitudes and movements of the UAV. In some scenes, the UAV approaches a particular target and remains stationary in others. The dataset also includes videos in which the UAV circles a parking slot to obtain global information. Moreover, the videos were recorded at different times (day and night). Hence, there is a variable luminosity between the videos. Thus, the MDVD is complex both in terms of the variety of contents it includes and the various conditions under which these videos were recorded.
The original dataset was split into two subsets: train and test. The training set includes 15 videos, whereas the test set contains 23 videos. However, in our case, we did not use the dataset as it is. We re-divided the dataset to fit the One-Class Classification problem: the training set contains normal scenes and the test set includes normal and abnormal scenes. Also, we considered suspicious scenes as normal events, as authors in [9] studied both cases "**suspicious scenes as normal events**" and "**suspicious scenes as abnormal events**", and they showed that considering suspicious scenes as normal events gives better results because the model struggles to distinguish suspicious human behaviors from normal ones.

B. Implementation

Publicly available implementation in Matlab [37] of GoogLeNet pretrained on ImageNet was used as feature extractor. In the first, input frames are resized to **(224, 224)** to match the input size of the GoogLeNet network. Next, the pretrained network is taken and its last fully-connected layer is removed. After that, the video's frames were converted to sequences of feature vectors, where those features are the output of the activation function on the last pooling layer ($Pool5-7*7-s1$) of the GoogLeNet network, as shown in Figure 2. Then, the feature vectors are normalized and used to train the OCSVM classifier. LIBSVM library [38] in Matlab was used to perform the one class classification task.

C. Training details

In order to select the optimal hyper-parameters of OCSVM, we used the **GridSearch** method. Indeed, with the standard **Radial Basis Function** (**RBF**) kernel, which is commonly used, two hyper-parameters of OCSVM need to be properly tuned: the RBF kernel width **gamma** (γ) and the regularization coefficient **nu** (ν). Finding the optimal values of those parameters is a very challenging task and can be

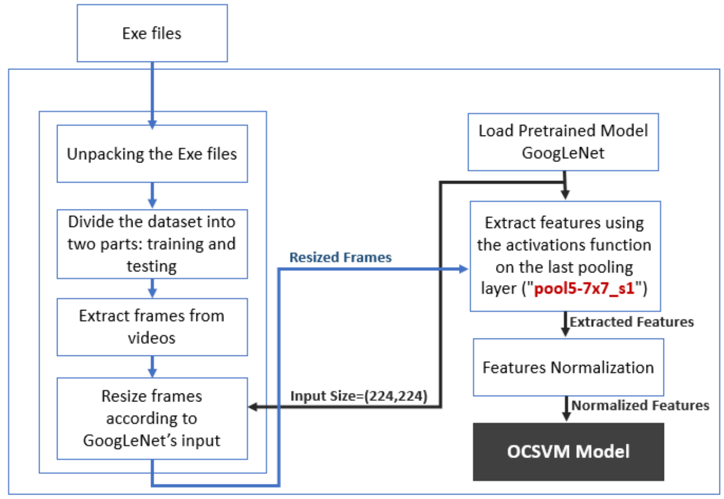

Fig. 2. Features extraction Steps

treated as a search problem [39]. As shown in Table 1, the possible and optimal values of **nu** (ν) and **gamma** (γ) for our model are, respectively, 0.1 and 0.4.

TABLE I
OPTIMAL VALUES OF THE HYPER-PARAMETERS OF OUR MODEL

Possible values of ν and γ	GridSearch optimal values
Nu (ν) : [0.01 : 0.5]	0.1
Gamma (γ) : [2^{-15} : 2^{3}]	0.4

A GridSearch algorithm must be guided by some performance metric, typically measured by cross validation on the training set or evaluation on a held-out validation set. In this paper, the **K-Fold Cross-Validation [40]** is used, with $k = 5$.

D. Evaluation metrics

Quantitative evaluation of the performance of the proposed approach is done using the well-known *confusion matrix*, the *R*eceiver *O*perating *C*haracteristic (ROC) and the *A*rea *U*nder *C*urve (AUC), known also as *A*rea *U*nder the *R*OC *C*urve.
The confusion matrix is a table with 4 different combinations of actual and predicted values. It is often used to measure the performance of a machine learning classification problem on a set of test samples for which the true values are known. It is extremely useful for measuring most performance metrics: Recall, Precision, Specificity, Accuracy and most importantly AUC-ROC Curve.
The ROC curve is a probability graph representing the performance of a classification model for all classification thresholds. This curve plots *T*rue *P*ositive *R*ate (TPR) against the *F*alse *P*ositive *R*ate (FPR). The higher TPR, the lower FPR, in other words, the sharper the ROC curve, the better the detection performance.
The AUC is a metric for classification approaches. According to [41], it is the second most popular metric, after the accuracy. It is frequently used to evaluate the performance of OCC

models and methods for novelty detection [6]. An excellent model has an AUC near to 1, which means it has good performance. A model with AUC = 0.5, is a model with no class separation capacity.

V. Experimental results

Once our proposed model is trained, we can evaluate the performance of our approach by feeding in testing data (data from normal and abnormal scenes) and checking whether it is capable to detect anomalous events while keeping the rate of false alarm low. We used a test dataset containing 5854 frames from both normal and abnormal scenes.

The confusion matrix of our model is given in Table 2. It can be interpreted as follows:

- Our model predicts all normal frames as normal, therefore, there is no false alarm. This is can be seen in line (1), *T*rue *P*ositive (TP) = 2857 and *F*alse *N*egative (FN) = 0.
- In line (2), the proposed model predicts 376 frames as normal but they are actually not, i.e. we have 376 non-detection.

TABLE II
Confusion Matrix of our Model

		Predicted Values	
		Normal (1)	Abnormal (-1)
Actual Values	Normal (1)	TP = 2857	FN = 0
	Abnormal (-1)	FP = 376	TN = 2621

By analyzing the frames which are object to a non-detection, we noticed that these are the frames where anomaly has not yet occurred. As all frames of an abnormal video are labeled as abnormal, but which do not necessarily contain an abnormal scene. For example, in a car stealing scene, we can see a man walking normally at the beginning of the video, then he tries to steal the car. The first frames are regarded as abnormal when they are not really so. Therefore, the proposed model does not differentiate them from normal frames. Based on the hypothesis that an anomalous event cannot appear or disappear abruptly, in real case, we assume that a video is considered as abnormal if at least 40% of the frames are abnormal, by analogy with Pixel-Level detection in a frame [42]. Thus, this non-detection does not have a great impact on the final precision of our model.

The ROC curve and the AUC of our model are given respectively in Figure 3 and Table 3. As we can see in Figure 3, the ROC curve hugs the upper left corner of the plot, which means that the proposed model has a good separation ability between normal and abnormal scenes. Also, Table 3 gives a comparison between our anomaly detection approach and the method proposed in [9]. We can obviously remark that our proposition outperforms the model presented in [9] by achieving an AUC value of 0.93.

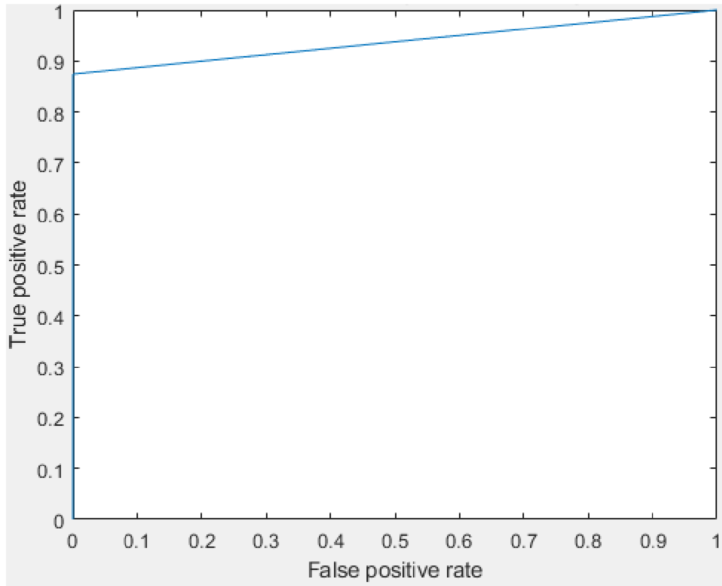

Fig. 3. ROC curve of our model

TABLE III
AUC comparison with Henrio et al. [9]

Approaches	**AUC**
Our solution	0.93
Henrio et al. [9]	0.72

VI. Conclusion and future work

In this paper, we introduced an approach, based on coupling deep learning Pretrained CNN (GoogLeNet) and unsupervised machine learning technique (OCSVM), to detect anomaly events in UAV-based surveillance systems. GoogLeNet was used to extract robust features from input videos, then, the feature vectors were normalized using the *min-max* normalization and used to train the OCSVM classifier. Differently to other state-of-the art approaches that attempt to solve anomaly detection problem, we used a more complex and challenging dataset, containing videos taken by an UAV in a car parking. The advantage of our approach is its semi-supervised nature, i.e. the unique ingredient needed is videos containing normal scenes. The quantitative results obtained on the MDVD dataset show that our proposed method achieves good results in comparison to existing technique. The limitation of the proposed method is the non-detection of some abnormal frames due to their resemblance to normal ones. In the future, research can be conducted on the effect of source dataset choice. Indeed, in order to obtain the exact precision of the proposed method, we aim to rearrange the used dataset, so that the abnormal videos contain only abnormal scenes. Moreover, other kinds of anomalous scenes and a large number of tests will be conducted to further validate the proposed approach. Furthermore, an in-depth study of the energy consumption that is needed to implement the proposed framework will be carried out.

References

[1] KE, Ruimin, LI, Zhibin, TANG, Jinjun, et al. Real-time traffic flow parameter estimation from UAV video based on ensemble classifier and

optical flow. IEEE Transactions on Intelligent Transportation Systems, 2018, vol. 20, no 1, p. 54-64.
[2] FLORES, Donate A., SAITO, Carlos, PAREDES, Juan A., et al. Aerial photography for 3d reconstruction in the peruvian highlands through a fixed-wing uav system. In : 2017 IEEE International Conference on Mechatronics (ICM). IEEE, 2017. p. 388-392.
[3] KIMA, S. S., YOOB, S. H., PARKA, J. S., et al. Rapid Disaster Mapping through Data Integration from UAVs and Multi-sensors Mounted on Investigation Platforms of NDMI, Korea. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018, vol. 42, no 3/W4.
[4] TOKEKAR, Pratap, VANDER HOOK, Joshua, MULLA, David, et al. Sensor planning for a symbiotic UAV and UGV system for precision agriculture. IEEE Transactions on Robotics, 2016, vol. 32, no 6, p. 1498-1511.
[5] BHATTARAI, Saleen, POUDEL, Kushal, BHATTA, Nishant, et al. Modeling and Development of Baseline Guidance Navigation and Control System for Medical Delivery UAV. In : 2018 AIAA Information Systems-AIAA Infotech@ Aerospace. 2018. p. 0508.
[6] MOTLAGH, Naser Hossein, BAGAA, Miloud, et TALEB, Tarik. UAV-based IoT platform: A crowd surveillance use case. IEEE Communications Magazine, 2017, vol. 55, no 2, p. 128-134.
[7] SULTANI, Waqas et SHAH, Mubarak. Human Action Recognition in Drone Videos using a Few Aerial Training Examples. arXiv preprint arXiv:1910.10027, 2019.
[8] DOMINICI, Donatella, ALICANDRO, Maria, et ZOLLINI, Sara. The smart management of" construction site" in a post-seismic scenario using UAV photogrammetry. In : EGU General Assembly Conference Abstracts. 2018. p. 15620.
[9] HENRIO, Jordan et NAKASHIMA, Tomoharu. Anomaly Detection in Videos Recorded by Drones in a Surveillance Context. In : 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2018. p. 2503-2508.
[10] WANG, Yinan, YOSHIHASHI, Ryota, KAWAKAMI, Rei, et al. Unsupervised anomaly detection with compact deep features for wind turbine blade images taken by a drone. IPSJ Transactions on Computer Vision and Applications, 2019, vol. 11, no 1, p. 1-7.
[11] CHRIKI, Amira, TOUATI, Haifa, SNOUSSI, Hichem, et KAMOUN Farouk. FANET: Communication, mobility models and security issues. Computer Networks, 2019, vol. 163, p. 106877.
[12] CHANDOLA, Varun, BANERJEE, Arindam, et KUMAR, Vipin. Anomaly detection: A survey. ACM computing surveys (CSUR), 2009, vol. 41, no 3, p. 1-58.
[13] HAERING, Niels, VENETIANER, Péter L., et LIPTON, Alan. The evolution of video surveillance: an overview. Machine Vision and Applications, 2008, vol. 19, no 5-6, p. 279-290.
[14] CHRIKI, Amira, TOUATI, Haifa, SNOUSSI, Hichem, et KAMOUN Farouk. UAV-GCS Centralized Data-Oriented Communication Architecture for Crowd Surveillance Applications. In : 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC). IEEE, 2019. p. 2064-2069.
[15] CHRIKI, Amira, TOUATI, Haifa, SNOUSSI, Hichem, et KAMOUN Farouk. Centralized Cognitive Radio Based Frequency Allocation for UAVs Communication. In : 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC). IEEE, 2019. p. 1674-1679.
[16] CAI, Yixin, MA, Li, et LIU, Gang. A Night-time Anomaly Detection System of Hog Activities Based on Passive Infrared Detector. Applied Engineering in Agriculture, 2019, vol. 35, no 4, p. 481-493.
[17] MISHRA, Soumya Ranjan, MISHRA, Tusar Kanti, SARKAR, Anirban, et al. Detection of Anomalies in Human Action Using Optical Flow and Gradient Tensor. In : Smart Intelligent Computing and Applications. Springer, Singapore, 2020. p. 561-570.
[18] WANG, Siqi, ZHU, En, YIN, Jianping, et al. Anomaly detection in crowded scenes by SL-HOF descriptor and foreground classification. In : 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE, 2016. p. 3398-3403.
[19] FAHN, Chin-Shyurng, LING, Jer, YEH, Ming-Yuan, et al. Abnormal Maritime Activity Detection in Satellite Image Sequences Using Trajectory Features. International Journal of Future Computer and Communication, 2019, vol. 8, no 1.
[20] NAWARATNE, Rashmika, ALAHAKOON, Damminda, DE SILVA, Daswin, et al. Spatiotemporal Anomaly Detection Using Deep Learning for Real-Time Video Surveillance. IEEE Transactions on Industrial Informatics, 2019, vol. 16, no 1, p. 393-402.
[21] CHALAPATHY, Raghavendra et CHAWLA, Sanjay. Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407, 2019.
[22] KAWACHI, Yuta, KOIZUMI, Yuma, MURATA, Shin, et al. A two-class hyper-spherical autoencoder for supervised anomaly detection. In : ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019. p. 3047-3051.
[23] CARREÑO, Ander, INZA, Iñaki, et LOZANO, Jose A. Analyzing rare event, anomaly, novelty and outlier detection terms under the supervised classification framework. Artificial Intelligence Review, 2019, p. 1-20.
[24] WANG, Tian, QIAO, Meina, ZHU, Aichun, et al. Abnormal event detection via covariance matrix for optical flow-based feature. Multimedia Tools and Applications, 2018, vol. 77, no 13, p. 17375-17395.
[25] XU, Dan, YAN, Yan, RICCI, Elisa, et al. Detecting anomalous events in videos by learning deep representations of appearance and motion. Computer Vision and Image Understanding, 2017, vol. 156, p. 117-127.
[26] FAROOQ, Muhammad Umer, KHAN, Najeed Ahmed, et ALI, Mir Shabbar. Unsupervised video surveillance for anomaly detection of street traffic. International Journal of Advanced Computer Science and Applications (IJACSA), 2017.
[27] CHAN, Yi-Tung, WANG, Shuenn-Jyi, et TSAI, Chung-Hsien. Extracting foreground ensemble features to detect abnormal crowd behavior in intelligent video-surveillance systems. Journal of Electronic Imaging, 2017, vol. 26, no 5, p. 051402.
[28] THOMAZ, Lucas A., JARDIM, Eric, DA SILVA, Allan F., et al. Anomaly detection in moving-camera video sequences using principal subspace analysis. IEEE Transactions on Circuits and Systems I: Regular Papers, 2017, vol. 65, no 3, p. 1003-1015.
[29] Diego, Ferran, Georgios Evangelidis, and Joan Serrat. "Night-time outdoor surveillance with mobile cameras." 2012.
[30] ZHOU, Shifu, SHEN, Wei, ZENG, Dan, et al. Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes. Signal Processing: Image Communication, 2016, vol. 47, p. 358-368.
[31] RIBEIRO, Manassés, LAZZARETTI, André Eugênio, et LOPES, Heitor Silvério. A study of deep convolutional auto-encoders for anomaly detection in videos. Pattern Recognition Letters, 2018, vol. 105, p. 13-22.
[32] MEDEL, Jefferson Ryan et SAVAKIS, Andreas. Anomaly detection in video using predictive convolutional long short-term memory networks. arXiv preprint arXiv:1612.00390, 2016.
[33] BONETTO, Margherita, KORSHUNOV, Pavel, RAMPONI, Giovanni, et al. Privacy in mini-drone based video surveillance. In : 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). IEEE, 2015. p. 1-6.
[34] ZHAI, Xiang, LIU, Kui, NASH, William, et al. Smart Autopilot Drone System for Surface Surveillance and Anomaly Detection via Customizable Deep Neural Network. In : International Petroleum Technology Conference. International Petroleum Technology Conference, 2020.
[35] SZEGEDY, Christian, LIU, Wei, JIA, Yangqing, et al. Going deeper with convolutions. In : Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 1-9.
[36] SCHÖLKOPF, Bernhard, WILLIAMSON, Robert C., SMOLA, Alex J., et al. Support vector method for novelty detection. In : Advances in neural information processing systems. 2000. p. 582-588.
[37] Web page.https://ch.mathworks.com/help/deeplearning/ref/googlenet.html
[38] CHANG, Chih-Chung et LIN, Chih-Jen. LIBSVM: A library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2011, vol. 2, no 3, p. 1-27.
[39] CHICCO, Davide. Ten quick tips for machine learning in computational biology. BioData mining, 2017, vol. 10, no 1, p. 35.
[40] STONE, Mervyn. Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society: Series B (Methodological), 1974, vol. 36, no 2, p. 111-133.
[41] BRADLEY, Andrew P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern recognition, 1997, vol. 30, no 7, p. 1145-1159.
[42] BOUINDOUR, Samir, HU, Ronghua, et SNOUSSI, Hichem. Enhanced Convolutional Neural Network for Abnormal Event Detection in Video Streams. In : 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE). IEEE, 2019. p. 172-178.