Crowd Anomaly Detection Based on Elevator Internet of Things Technology

Chunhua Jia, Wenhai Yi, Yu Wu, Zhuang Li, Shuai Zhu(✉), and Leilei Wu

Shanghai Elevator Media Information Co., Ltd., Shanghai, China
{jiachunhua,yiwenhai,wuyu,lizhuang,zhushuai,wull}@zailingtech.com

Abstract. A work-flow which aims at capturing residents' abnormal activities through the passenger flow of elevator in multi-storey residence buildings is presented in this paper. Firstly, sensors (hall sensor, photoelectric sensor, gyro, accelerometer, barometer, and thermometer) connected with internet are mounted in elevator to collect image and data. Then computer vision algorithms such as instance segmentation, multi-label recognition, embedding and clustering are applied to generalize passenger flow of elevator, i.e. how many people and what kinds of people get in and out of the elevator on each floor. More specifically so-called GraftNet is proposed for fine-grained multi-label recognition task to recognize human attributes (e.g. gender, age, appearance, and occupation). Thirdly, based on the passenger flow data, anomaly detection of unsupervised learning is hierarchically applied to detect abnormal or even illegal activities of the residents. Meanwhile, based on manual reviewed data, Catboost algorithm is implemented for multi-classification task. Experiment shows the work-flow proposed in this paper can detect the anomaly and classify different categories well.

Keywords: IoT · Anomaly · Computer Vision · Machine learning · Big Data · Cloud computing

1 Introduction

In modern city, most people live in apartments of multi-storey buildings, due to the complexity of structure and high density of residents, public safety [1] as well as the management of urban residential community effective is challenged [2]. Thanks to technologies based on Artificial Intelligence (AI) [3] and Big Data and Internet Of Things (IoT), which make it possible to capture or predict direct or potential safety hazard on real-world activities [4,5]. On the other hand, the patterns of behaviors or activities vary from one to one and change continuously [5,6], it is particularly difficult to give a specific definition on residents' activities which would put their own safety in danger. Considering all the aspects above, the elevator could be the most feasible and suitable [7] environment to take operation on because it is legal and reasonable to deploy public surveillance

L. Peñalver and L. Parra (Eds.): Industrial IoT 2020, LNICST 365, pp. 3–17, 2021.
https://doi.org/10.1007/978-3-030-71061-3_1

and people take elevator widely and frequently enough. By real-time collecting, processing and analyzing the video flow in the elevator, important information including the state of the elevator, the states of people can be obtained [8]. Based on IoT, it is not difficult to capture elevator malfunction (e.g. stuck) in real-time through setting several different sensors in elevator [9]. A camera is used to automatically check if any passenger is trapped in elevator box when malfunction happens through Computer Vision (CV) algorithm, e.g. pedestrian detection [10].

Some researchers focus on predictive maintenance based on the detection of abnormal usage of an elevator, taking advantage of sensors information [11]. Common Internet Of Things (IoT) solution can only detect hardware malfunction, however, it is unable to describe the behavior of residents. Video surveillance fits this task better. Through the design of software and hardware, a work-flow is proposed in this paper. Camera and some other sensors including barometers, thermometers, accelerometers and so on mounted in the elevator, which helps to collect real-time data including video, velocity, temperature, pressure and so on. A distribution Powerline Carrier (PLC) Communication Systems [12] is constructed which achieves a stable and high speed transmission and satisfies the demand of data transmission. Once real-time data was collect, computer vision algorithm can be used for objection detection [13,14] and multi-classification model, named GraftNet, to distinguish the attributes and the number of the people, the body area of the people, the distances between people and the moving trajectories of the people in the elevator. Collected data is uploaded and stored in cloud based on Hadoop [15] and the artificial intelligence models are deployed and managed by Kubernetes [16]. The system introduced above is deployed on more than 100000 elevators currently.

The proposed system has the following parts:

- Power line Carrier (PLC) Communication Systems is used to create a smart elevator, which defined as a box equipped with numerous sensors and actuators, e.g. camera, barometers, thermometers, accelerometers. PLC systems are also 24 h a day connected and they have a relative small bandwidth compared to traditional modem communication systems. These properties make PLC systems a good alternative for low bandwidth intensive Internet services such as remote monitoring.
- Data collected from sensors including pictures and time series data are saved in cloud for the flexibility and cost-saving, user trust, privacy and security are also concerns.
- Cloud computing provides the architecture for creating Clouds with market-oriented resource allocation by leveraging technologies such as Virtual Machines (VMs). All the AI models including computer vision, anomaly detection and multi-classification are deployed on cloud.
- More specially, multi-classification model named GraftNet is proposed. GraftNet is a tree-like network that consists of one trunk and several branches and analyses the passenger flow of elevators in residence building.

The rest of the paper is organized as follows. The related work of IoT, computer vision and anomaly detection is briefly discussed in Sect. 2. We explain the infrastructure of internet of things and how to capture and analysis computer vision for passenger flow. The proposed multi-classification model named GraftNet is also described in Sect. 3. Then, based on the collected and analyzed data, hierarchical anomaly detection for abnormal activity is utilized, we discuss this part in Sect. 4. Finally, in Sect. 5, we describe the experiments of our system, conclude our findings and suggest opportunities for future work.

2 Related Work

There were surveys conducted to anomalous events detection which utilize video surveillance. Recognizing human activities of daily living [17], which is an important research issue in building a pervasive and smart environment.

A survey [18] was given visual surveillance of object motion and behaviors, review recent developments and general strategies of all these stages, and proposed a general processing framework including several stages such as environment modeling, motion segmentation, object classification, understanding and description of behaviors, which all belong to computer vision domain. The possible research directions were given, e.g., occlusion handling, a combination of two and three-dimensional tracking. Videos Surveillance system should become more intelligent, crucial, and comprehensive to deal with the situations under which individual safety could be compromised by potential criminal activity. Surveillance systems provide the capability of collecting authentic and purposeful information and forming appropriate decisions to enhance safety [19]. Three generations of contemporary surveillance system and the most recent generation is decomposed into multi-sensor environments, video and audio surveillance, and distributed intelligence and awareness. A novel framework [20] was developed for automatic behavior profiling and online anomaly sampling/detection without any manual labeling of the training data set, which aims to address the problem of modeling video behavior captured in surveillance videos for the applications of online normal behavior recognition and anomaly detection. A real-time computer vision and machine learning system for modeling and recognizing human behaviors in a visual surveillance task were describe [21]. The system was particularly concerned with detecting when interactions between people occur, and with classifying the type of interaction. A novel approach to understand activities from their partial observations monitored through multiple non-overlapping cameras separated by unknown time gaps was propose [22]. Dynamic Probabilistic Networks (DPNs) [23] were exploited for modeling the temporal relationships among a set of different object temporal events in the scene for a coherent and robust scene-level behaviour interpretation.

Human detection and tracking [24] were tasks of computer vision systems for locating and following people in video imagery. Some object detection and object instance segmentation model were very popular at that time. For example, a simple fully-convolutional model for real-time instance segmentation was proposed in Yolact [13] and Mask R-CNN [14]. Some models offer a probabilistic model that allows us to predict the user's next actions and to identify anomalous user behaviours. Predicting the behavior of human participants in strategic settings was an important problem in many domains [4]. Deep learning architecture based on long short-term memory networks (LSTMs) was created [3], which models the inter-activity behaviour. By employing a deep learning approach, Deep Neural Network (DNN) framework was introduced, which has the capability for both modeling social influence and for predicting human behavior [6]. Social Influence Deep Learning (SIDL) [5] is a framework that combines deep learning with network science for modeling social influence and predicting human behavior on real-world activities.

A key goal of surveillance was to detect behaviors that can be considered anomalous [25]. Anomaly detection [26], which is also known as outlier or novelty detection, was a widely studied topic that had been applied to many fields including medical diagnosis, marketing, network intrusion, and to many other applications except for automated surveillance. There were different outlier detection models. A distance-based outlier detection method that finds the top outliers in an unlabeled data set [27]. Proximity-Based methods [28] consider the task of performing anomaly detection in highly noisy multivariate data. Statistical anomaly detection typically focuses on finding individual data point anomalies [29,30]. Neural networks, including Variational Auto-Encoders (VAEs) have shown great potential in the unsupervised learning of data distributions [31]. Outlier detection as a binary-classification issue by sampling potential outliers from a uniform reference distribution was approached [32]. Anomalies are data points that are few and different, as a result of these properties, anomalies are susceptible to a mechanism called isolation. Isolation Forest (iForest) [32] detects anomalies purely based on the concept of isolation without employing any distance or density measure—fundamentally different from all existing methods. An open-source Python toolbox [33] for performing scalable outlier detection on multivariate data contains different kinds of anomaly detection models.

To capture activities with potential public safety hazard, anomaly detection of unsupervised learning was the choice from algorithm perspective because the definition of safety hazard on people's activity was ambiguous and such activities were relatively rare. By real-time collecting, precessing and analyzing the video flow in the elevator and using background subtraction algorithm and human tracking algorithm, some important information including the state of the door of the elevator, the states of people can be obtained [8].

The previous work tried to analysis the video with the help of computer vision algorithms (see, [3,4,24]) which inspired our job. A work-flow is proposed based on the results of computer vision as well as some other sensors. Combined the different information of sensors, unsupervised learning (for example, Isolation Forest) and multi-classification, are given to make the conclusion which is valuable for the management of property management company. Our system can detect the anomaly directly and based on the passager flow data, we can easily distinguish the normal residence, group rental, apartment, office, decoration, dormitory and so on.

3 Structure of Our System

In this section, the architecture of the proposed work flow is given, the related computer vision algorithms will be discussed as well. In the first part, the IoT system will be described. The second part discusses how to calculate the number of person in/out the elevator. The proposed GraphNet for fine-grained multi-label classification recognizes different attributes of elevator passengers. With the help of PLC networks, connecting is much more reliable and efficient.

3.1 Infrastructure of Internet of Things (IoT) for Data Collection

Camera and several kinds of sensors (Hall sensor, photoelectric sensor, gyro, accelerometer, barometer, thermometer, etc.) are mounted in or on the elevator as the deployment of our system. Our work flow is a solution which is easy to promote with low cost and does not require access to the elevator's ECU, which means our solution could be deployed without influence of the elevators. Once the sensors is properly setup, we are able to monitor the trace of elevator running. Images captured from the camera help to collect passenger data. One snapshot is taken once the elevator moves from one floor to another. Alongside with the image, time stamp and floor numbers of start/end are also collected. Images with extra information are uploaded and stored on cloud for further analysis. The mechanism of snapshooting is based on one simple assumption: people get in or out of the elevator only when it stops at certain floor, and thus we can generalize how many and what kinds of people get in and out by comparing images successively captured, as shown in Fig. 1.

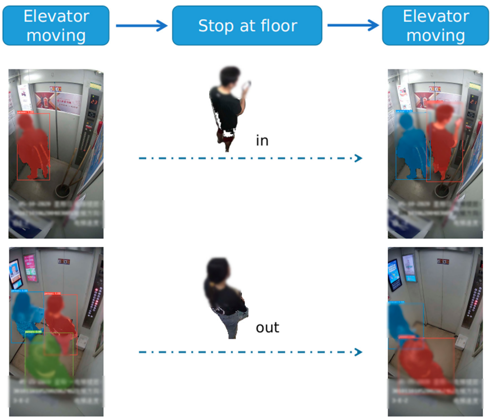

Fig. 1. Comparison of two images successively taken.

As shown in Fig. 1, by comparing the stage change of two images snapshooted serially. Conclusion a person in or out special floor can be easily make.

3.2 Instance Segmentation, Embedding and Clustering

Considering occlusion is inevitable when there are several passengers in the elevator at the same time, for feature extraction on each individual passenger, instance segmentation is needed. Different from object detection, instance segmentation can accurately segment all objects at pixel level and minimize the impact of occlusion and background. It could be considered as a pre-process similar to attention mechanism, so that other CV models could focus on human target itself completely.

YOLACT [13], a representative one-stage method which was proposed to speed up instance segmentation, is utilized to segment target person from background and other non-targets as shown in Fig. 2.

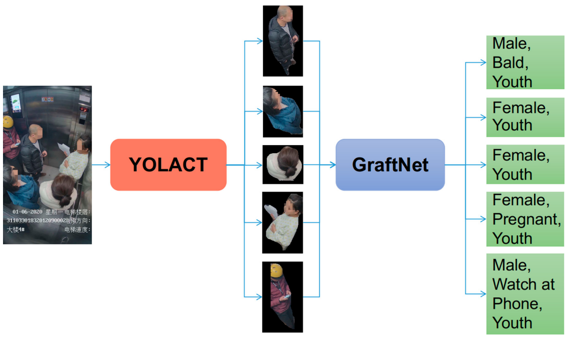

Fig. 2. Process of YOLACT + GraftNet.

As mentioned above, the information of people getting in or out of each floor could be generated by comparing two successively captured images. We assume the variance of passengers' overall appearances is greater than that of their faces, so FaceNet [34] designed for face embedding and clustering should be sufficient to identically represent the appearance of each segmented individual. As a result, FaceNet is re-trained with our own dataset of segmented passengers and utilized as feature extractor.

As shown in Fig. 3, there are M passengers in image $p1$ captured right before the elevator stops at certain floor, and N passengers in image $p2$ captured right after the elevator leaves that floor. All those segmented passengers are fed to FaceNet and corresponding feature vectors are returned. Then we build an association matrix D with order $M * N$. The element d_{ij} represents the Euclidean distance between the feature vectors of passenger i in $p1$ and passenger j in $p2$. Minimal value searching on each row (or column) with a threshold t is carried out on D to get the best match for each passenger (no match found if the minimal value is greater than t). The passengers in *p2* without any match from *p1* are considered as "get out of that floor" and those in *p1* without match from *p2* are considered as "get into the floor". Then attributes recognition are only applied on above two kinds of passengers for each floor.

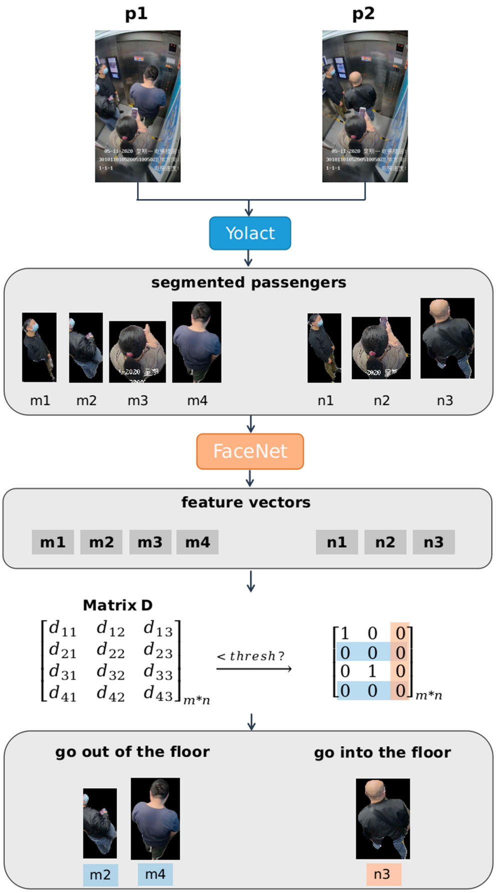

Fig. 3. Segmentation, embedding and clustering to generalize passenger flow of each floor.

3.3 GraftNet: A Solution for Fine-Grained Multi-label Classification

To analyse the passenger flow of elevator in residence building, data with more descriptive information is needed, e.g. gender, age, occupation, appearance, etc. GraftNet is proposed, which id a solution for fine-grained multi-label recognition task, i.e. to recognize different attributes of elevator passengers.

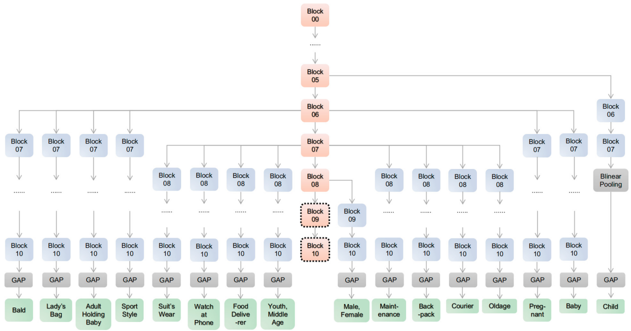

Fig. 4. Architecture of GraftNet (based on Inception V3).

GraftNet is a tree-like network that consists of one trunk and several branches corresponding to attributes, as shown in Fig. 4. The trunk is used to extract low-level features such as shapes and textures, which can be commonly represented with generic features. And branches is mainly used to generate high-level features and thus can be customized for different attributes. For GraftNet, we propose a two-step training procedure. Instead of to annotate all samples for all attributes overall, samples are collected and annotated for one single attribute separately, so that we get the sub-datasets corresponding to attributes. In first step, InceptionV3 is pre-trained on the collection of all sub-datasets by using a dynamic data flow graph, as shown in Fig. 4. The 11 blocks with pre-trained weights could be considered as the trunk of GraftNet. The second step is to separately fine-tune and graft branches onto the trunk for each attribute. By training trunk and branches in a two-step way, GraftNet could save time and labor for both annotation and training. Sub-datasets of different attributes could be maintained separately and incrementally, i.e. new attributes or samples could be added without any re-work on the existing data set. Besides, training task of one-branch-for-one-attribute (the iteration of samples re-collecting and model fine-tuning) is more manageable in practice for a team-work. So to speak, the very basic consideration of the design of GraftNet is that the requirement of recognizing a new attribute could come at any time and we don't want any re-work because of that (Fig. 5).

Since GraftNet is deployed on cloud but not embedded devices in our system, we focus more on extendability rather than efficiency. From some perspectives, our work is quite like an inverse process compared to network pruning or weights compression [35]. Rather than to reduce the redundancy of neural networks for a fixed task, what we do is to leverage the over-parameterization and maximum its usage to recognize new attributes with a few extra branches.

3.4 Power Lines Communication (PLC)

In circumstances of elevator, it may be too expensive and not convenient for a cable or telephone company to install the infrastructure necessary with

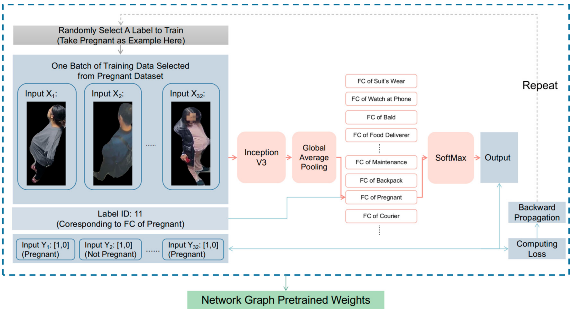

Fig. 5. GraftNet: process of training with dynamic data flow graph.

broadband Internet access. One potential solution for this problem is power line communication, since it can use existing infrastructure to deliver broadband Internet access. Meanwhile, the wireless routers have drawbacks such as signal interference from walls, floors, and other objects. Power line communication (PLC) is a technology that can allow data to be transmitted by an existing electricity infrastructure and can be used the existing wiring in a home or elevator. These systems typically involve an adapter that allows a modem to be connected to a power outlet via an Ethernet cable. Additional adapters can then be plugged into outlets around the elevator, allowing computers or other devices access to the network. PLC technology has the following advantages: low implementation cost, large reach, low running cost and indoor high speed. These advantages lead to more implementations of PLC networks in various industries.

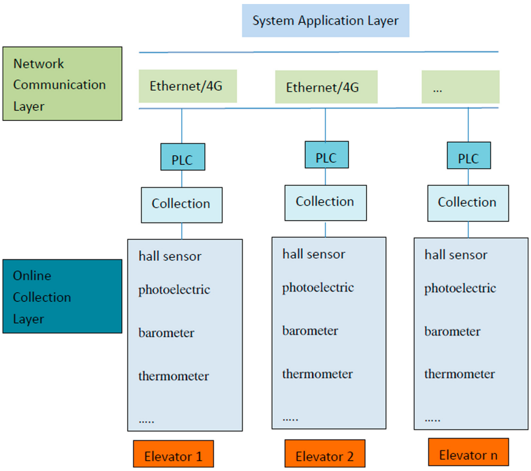

Fig. 6. PLC for elevator system.

As show in Fig. 6, sensor data are uploaded and communicated with the help of PLC technology. After collection, PLC networks link to application layer which contains the work station via Ethernet or 4G/5G.

4 Hierarchical Anomaly Detection for Abnormal Activity

The architecture of our system is described in Sect. 2. Once data is collected, computer vision algorithm, supervised and unsupervised machine learning model can be used for the anomalies. In this section, the details of our model deployed on cloud will discussed.

4.1 Isolation Forest: Anomaly Detection of Unsupervised Learning

Isolation forest (IForest) is an unsupervised learning algorithm dealing with anomaly detection problem. Instead of building a model of normal instance like common techniques use, the principle of Isolation Forest is isolating anomalous points which means the tendency of anomalous instances can be easier to be split from the rest samples in a data set, compared with normal points, because anomalous data points has two quantitative properties: different and fewer than normal data points.

Anomaly detection with Isolation Forest is a process composed of two main stages. In the first stage, a training data set is used to build isolation trees (iTrees). In the second stage, each instance in test set is passed through the iTrees build in the previous stage, and a proper 'anomaly score' is assigned to the instance according to the average path length of iTrees. Once all the instances in the test set have been generated an anomaly score, it is possible to mark the point whose score is greater than a predefined threshold as anomaly.

As one of the most famous anomaly detection algorithms, IForest has outstanding advantages: It has a low linear time complexity and a small memory requirement. IForest can be used in huge amounts of data sets because of its basic approach of random forest. The independence of Isolation trees ensures this model can be employed on large scale distributed system to accelerate Computing Platform. At the same time, large number of Isolation trees makes algorithm more stable.

4.2 Feature Generation and Hierarchical Anomaly Detection

For the generalized passenger flow we mentioned in Sect. 3, we treat them as two parts hierarchically. The first is the result of instance segmentation, embedding and clustering, which is called flow-count data here, i.e. how many people get in/out of elevator. And the other is the result of human attributes recognition, which is called attributes data, i.e. what kinds of people get in/out of elevator. The hierarchy is defined here because the flow-count data is actually byproduct of our system in commercial use but attributes data is not. Since our IoT system is deployed in more than 100000 elevators which transport tens of millions people daily, to additionally recognize attributes for all these passengers is computationally not affordable for us. Besides, the original requirement from our customer at the very beginning is to find out over crowded residence (e.g. more than 20 illegal migrants live in one small apartment), it's reasonable that we put more consideration on flow-count data. Therefore we decide to perform anomaly detection hierarchically, first on flow-count data then on attributes data.

Flow-count data is used for the first round of anomaly detection. For the weekdays in the last 15 days, we calculate the mean $m1$ and standard deviation $s1$ of passenger flow per floor per elevator, $m2$ and $s2$ of the flow per elevator, and $m3$ and $s3$ of the flow of all elevators in the same residential estate. Considering citizens are highly civilized and socialized, $m4-m6$ and $s4-s6$ are calculated for the weekends in the last 15 days, the same as $m1-m3$ and $s1-s3$ on weekdays. Including the floor number, we get feature vectors of length *13* per floor per elevator.

The contamination parameter of isolation forest during training procedure prescribes the proportion of outliers in the data set. We set it as 0.2 to output as many records as possible for the next second round anomaly detection.

Attributes recognition is only performed on the output of the first round anomaly detection to reduce computation. Besides the $l2$ values of mean and standard deviation generated in first round (floor number excluded), the attribute features are adopted by calculating the mean of attribute recognition result per floor per elevator. In detail, for passengers who get in/out elevator of certain floor, attributes recognition with GraftNet is performed to get feature vectors (22 classification results and 22 corresponding scores), then mean of these attribute feature vectors are calculated. The head count and distribution of time of appearance in 24-h are also included.

The contamination parameter of isolation forest is set as 0.01 for the second round detection. There are around 1 million records for each floor from 100000 elevators. In our experiment, after two rounds of anomaly detection, we finally get 643 outliers output.

4.3 Manual Review and Analysis

One fact that we have to admit is that sometimes the outliers obtained from anomaly detection of unsupervised learning might not be activities with public safety hazard. The anomaly of such outliers could be caused by malfunction of IoT system (e.g. malfunction of sensors, camera, network), or misentries in our database. For example, most of the elevators in our system are from residence buildings, and very few are from non-residence buildings, such as office building, hospital, school, and shopping mall. The activities of people in such non-residence buildings are obviously different from residence buildings. Our abnormal activity capture is supposed to perform on residence buildings only. However, some non-residence buildings are misentried as residence ones in our system, which could probably make them captured as anomaly. Similarly malfunction of IoT system could also cause exception data which might be captured.

Fig. 7. Interface of inspection tool.

To verify the outliers and keep improving the result, we build inspection tools to review and analyze the output manually. Corresponding to each output, the image-pairs successively captured will be reviewed, together with the statistical data such as distribution of attributes and time of appearance in 24 h, as shown in Fig. 7. All data exceptions caused by malfunction/misentries and abnormal activities clearly without any public safety hazard (e.g. running company in home office, decorators getting in/out, etc.) will be logged and excluded from next round of anomaly detection. Meantime confirmed records with suspicions of safety hazard will be reported to our customer.

Out of the 643 records output in our experiment we randomly pick 412 and review them one by one with the inspection tool. The result is shown in Table 1.

Table 1. Results after reviewed

Review result	Safety hazard	Comments	Number
Positive	Probably yes/unsure	Something different is there. Need to be checked by property manager	289
Positive	No	Caused by malfunction of sensors	13
Positive	No	Apartment under decoration	32
Positive	No	Dormitory/hotel	27
Positive	No	Shopping mall/entertainment venue	2
Positive	No	Office building	40
Positive	Yes	Catering service running in apartment	3
Positive	Yes	Over crowded residence	6

4.4 Anomaly Multi-classification

Based on manual review and analysis data, supervised machine learning can be used to distinguish the anomaly. A data set containing 15440 samples (floors) is built whose features come from 1286 lifts, and affiliated to 13 categories including normal residence, group rental, apartment, office, decoration, dormitory, system abnormality, shopping mall, entertainment place, unable to judge, hotel, restaurant and educational institution. The ratio of different categories, which is unbalanced, can be seen in Fig. 8. In details, the normal residence takes up 57.2% while educational institution holds only 0.19%. Meanwhile, as show in Subsect. 4.3, the feature contains both numerical and category type[1]. Catboost [36] is used for multi-classification task under this situation.

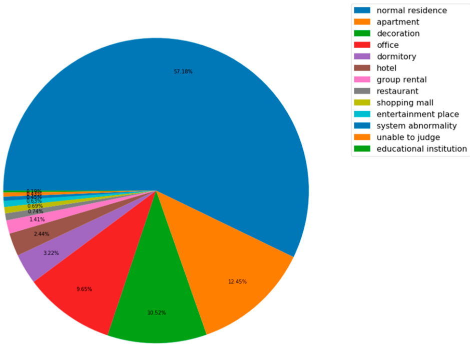

Fig. 8. Data distribution of different classifications

The data set is split-ted into training and test set, which account 80% and 20% respectively. Catboost model achieves 85.6% accuracy and 84.7% recall for the test stage. A plot of the ROC curve is also created in Fig. 9.

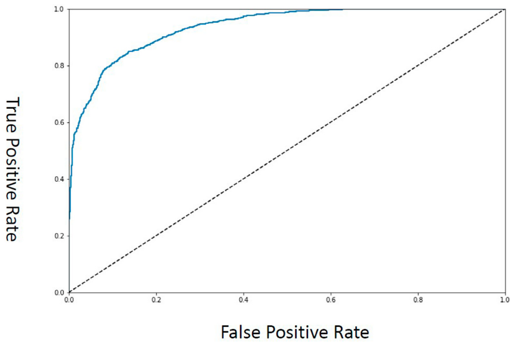

Fig. 9. ROC curve

[1] https://pan.baidu.com/s/10Cty8nJcpbp9B0_oGXjPww, extraction code: llqi.

5 Conclusion

As for now, a work flow is proposed which helps to detect the anomalies of floor. The proposed model is different from traditional ones which make the conclusion merely based on the results of computer vision with the help of deep machine learning. Our work flow is comprehensive system which has been deployed more than 100 thousands elevators in different places of China. By collecting pictures and some other data of sensors, supervised and unsupervised model are combined to make the discussion.

References

1. Hanapi, N.L., Sh Ahmad, S., Ibrahim, N., Razak, A.A., Ali, N.M.: Suitability of escape route design for elderly residents of public multi-storey residential building. Pertanika J. Soc. Sci. Humanit. **25(s)**(2017), 251–258 (2017)
2. Yan-Bin, W.: Creating effective mechanism of management of urban residential community. J. Yuxi Teachers Coll. 40–43 (2002)
3. Aitor, A., Gorka, A.: Predicting human behaviour with recurrent neural networks. Appl. Sci. **8**(2), 305 (2018)
4. Hartford, J., Wright, J.R., Leyton-Brown, K.: Deep learning for predicting human strategic behavior. In: Proceedings of the 30th International Conference on Neural Information Processing Systems, NIPS 2016, pp. 2432–2440. Curran Associates Inc., Red Hook (2016)
5. Luceri, L., Braun, T., Giordano, S.: Analyzing and inferring human real-life behavior through online social networks with social influence deep learning. Appl. Netw. Sci. **4**(1), 1–25 (2019). https://doi.org/10.1007/s41109-019-0134-3
6. Phan, N., Dou, D., Wang, H., Kil, D., Piniewski, B.: Ontology-based deep learning for human behavior prediction with explanations in health social networks. Inf. Sci. **384**, 298–313 (2017)
7. Zihan, M., Shaoyi, H., Zhanbin, Z., Shuang, X.: Elevator safety monitoring system based on internet of things. Int. J. Online Eng. (iJOE) **14**(08), 121 (2018)
8. Yi-Ping, T., Hai-Feng, L.U.: Intelligent anti-violence surveillance system in elevator based on computer vision. J. Zhejiang Univ. Technol. **6**, 591–597 (2009)
9. Gui-Xiong, L., Hai-Bing, Z., Ruo-Quan, H.E., Yu-Hui, K.E.: Design of elevator real-time energy efficiency recorder and system. China Measurement & Test (2012)
10. Dollar, P.: Pedestrian detection: a benchmark. In: Proceedings/CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 304–311 (2009)
11. Zhou, Y., Wang, K., Liu, H.: An elevator monitoring system based on the internet of things. Procedia Comput. Sci. **131**, 541–544 (2018)
12. Li, J., Luo, C., Xu, Z.: A novel timing synchronization metric for low-voltage OFDM powerline communication system. In: Power & Energy Engineering Conference (2012)
13. Bolya, D., Zhou, C., Xiao, F., Yong, J.L.: YOLACT: real-time instance segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2019) (2020)
14. Kaiming, H., Georgia, G., Piotr, D., Ross, G.: Mask R-CNN. IEEE Trans. Pattern Anal. Mach. Intell. **PP**, 1 (2017)

15. White, T.: Hadoop: The Definitive Guide. O'reilly Media Inc., Sebastopol (2012). 215(11): 1–4
16. Acuña, P.: Kubernetes (2016)
17. Duong, T.V.: Activity recognition and abnormality detection with the switching hidden semi-Markov models. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer (2005)
18. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. IEEE Trans. Syst. Man Cybern. **34**(3), 334–352 (1986)
19. Räty, T.D.: Survey on contemporary remote surveillance systems for public safety. IEEE Trans. Syst. Man Cybern. Part C **40**(5), 493–515 (2010)
20. Xiang, T., Gong, S.: Video behavior profiling for anomaly detection. IEEE Trans. Pattern Anal. Mach. Intell. **30**(5), 893–908 (2008)
21. Oliver, N.M., Rosario, B., Pentland, A.P.: A Bayesian computer vision system for modeling human interactions. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 255–272 (2000)
22. Loy, C.C., Gong, X.S.: Time-delayed correlation analysis for multi-camera activity understanding. Int. J. Comput. Vision **90**, 106–129 (2010)
23. Gong, S., Xiang, T.: Recognition of group activities using dynamic probabilistic networks. In: Proceedings of the Ninth IEEE International Conference on Computer Vision (2003)
24. Davis, J.W., Sharma, V., Tyagi, A., Keck, M.: Human detection and tracking. In: Li, S.Z., Jain, A. (eds.) Encyclopedia of Biometrics, pp. 882–887. Springer, Boston (2009). https://doi.org/10.1007/978-0-387-73003-5_35
25. Sodemann, A.A., Ross, M.P., Borghetti, B.J.: A review of anomaly detection in automated surveillance. IEEE Trans. Syst. Man Cybern. Part C **42**(6), 1257–1272 (2012)
26. Kalinichenko, L.A., Shanin, I., Taraban, I.: Methods for anomaly detection: a survey. In: Proceedings of the 16th Russian Conference on Digital Libraries RCDL 2014, CEUR Workshop Proceedings (CEUR-WS.org), vol. 1297, pp. 20–25. ceur-ws.org (2014)
27. Angiulli, F., Basta, S., Pizzuti, C.: Distance-based detection and prediction of outliers. IEEE Trans. Knowl. Data Eng. **18**(2), 145–160 (2005)
28. Idé, T., Lozano, A.C., Abe, N., Liu, Y.: Proximity-based anomaly detection using sparse structure learning. In: Proceedings of the SIAM International Conference on Data Mining, SDM, 30 April 2009–2 May 2009, Sparks, Nevada, USA, p. 2009 (2009)
29. Akouemo, H.N., Povinelli, R.J.: Probabilistic anomaly detection in natural gas time series data. Int. J. Forecast. **32**(3), 948–956 (2016)
30. Xiong, L., Póczos, B., Schneider, J.G., Connolly, A., Vanderplas, J.: Hierarchical probabilistic models for group anomaly detection. J. Mach. Learn. Res. **15**, 789–797 (2011)
31. Niu, Z., Yu, K., Wu, X.: LSTM-based VAE-GAN for time-series anomaly detection. Sensors **20**(13), 3738 (2020)
32. Liu, F.T., Ting, K.M., Zhou, Z.H.: Isolation-based anomaly detection. Acm Trans. Knowl. Discovery Data **6**(1), 1–39 (2012)
33. Zhao, Y., Nasrullah, Z., Li, Z.: PyOD: a python toolbox for scalable outlier detection. J. Mach. Learn. Res. **20**(96), 1–7 (2019)
34. Schroff, F., Kalenichenko, D., Philbin, J.: FaceNet: a unified embedding for face recognition and clustering (2015)
35. Song, H., Mao, H., Dally, W.J.: Deep compression: compressing deep neural networks with pruning, trained quantization and Huffman coding. In: ICLR (2016)
36. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A.: CatBoost: unbiased boosting with categorical features (2017)