The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

Big Data Enabled Realtime Crowd Surveillance and Threat Detection Using Artificial Intelligence and Deep Learning

Mr. Aquib Hasware[1], Dr Deepali Ujalambkar[2]
[1]Research Scholar, Department of Computer Engineering, AISSMS College of Engineering, Pune.
[2]Assistant Professor, Department of Computer Engineering, AISSMS College of Engineering, Pune.

Abstract
In recent years, security precautions at public events have become increasingly important as a direct response to the growth in the number of disruptive actions. There are several various varieties of closed-circuit televisions that are employed to perform 24-hour surveillance of public areas and the people that live such areas. Every person in a socialised society with a population of 1.6 billion is subjected to imprinted pictures an average of 30 times each day. Since the entire participation of the community is required to safeguard public areas from the most unexpected and lethal of events, it is difficult to discern whether an incident is an exceptional or casual occurrence because continual monitoring of human data makes it difficult to distinguish the difference. Within the scope of this study, we propose a method for identifying potentially threatening actions within footage obtained from closed-circuit television systems. To accomplish this objective, we must first extract individual still frames from the video and then examine the actions of the people seen in those still frames. We have placed a significant amount of reliance on both machine learning and deep learning algorithms to make this a reality. To automate this process, we must first develop a training model that makes use of many photos and a "Convolution Neural Network" that makes use of the Tensor Flow Python package. This model must be created before we can move on to automating the process. Every frame from every video that is provided will be used to train an algorithm that will analyse the film and evaluate whether it contains suspicious content or merely everyday activities. If we conclude that the activity was suspicious, the next phase of the study will concentrate on locating any weapons that may have been concealed on the corpse.
Keywords: Object Detection, Digital Pictures, Convolutional Neural Networks, Security Cameras, and Video Surveillance are some of the topics that will be discussed.

INTRODUCTION:

Patterns of human behaviour in public settings may frequently be used, together with the identification of specific items, to locate potential dangers. Visual information is the major source of such recognition and is used to generate it with the assistance of Image Processing and machine learning techniques. Photos taken by CCTV cameras offer visual evidence that may be used, just like live videos can be rewatched for the sake of future reference. Because of the introduction of an "automation" approach, there has been a significant shift in the way picture analytics are carried out. For instance, studies on crowding and mobility could benefit from the use of regulating systems.

The protection industry and the scientific community are examples of significant areas that make use of video editing technologies. Intelligent algorithms are used to follow the photographs as they move about. While developing a real-time device, it is essential to consider the computing challenges and the time constraints involved. Time-sensitive applications, such as the detection of bank robberies, the tracking of weapons, and the reporting of suspicious activities at railway stations, would benefit the most from a system that employed an algorithm with a relatively low time complexity, used fewer hardware resources, and produced good results.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

This type of system would be most beneficial to a system that employed an algorithm with a relatively low time complexity.

Throughout the course of this research project, we create a Deep Learning based Distributed Intelligent Video Surveillance system that may be implemented at the network's outskirts. We build a multi-layer edge computing architecture for the distributed intelligent video surveillance system as well as a distributed deep learning (dl) training system. the distributed intelligent video surveillance system possesses the capability to offload computational operations from the centre of the network to the perimeter of the network, which has the dual benefit of minimising the enormous network connection overhead while simultaneously offering low-latency and precise video analysis solutions. In this section, we put the proposed distributed intelligent video surveillance system into action and solve some of its issues, including simultaneous training, model synchronisation, and load balancing. Because of this, we provide task-level parallel training techniques in addition to model-level parallel training strategies to hasten the process of video analysis. To our knowledge, this is the first time anybody has described a method for updating model parameters to achieve global DL model synchronisation inside a distributed EC environment. In addition, we provide a method of dynamic data transmission to address the issue of edge nodes having an insufficient amount of processing power relative to the amount of work they are required to do. Video surveillance systems have evolved into a valuable instrument for the protection of both public and private institutions [1-3] these days, video surveillance systems are used to keep a watch on buildings, keep an eye on traffic, and even discover signals of criminal activity. Edge Artificial Intelligence is a technology that combines

Fig 1. Crowd Detection

Source : (https://arxiv.org/pdf/1906.07538.pdf)

artificial intelligence , the internet of things, and edge computing, and it shows considerable potential [4-6]. when used to video surveillance, EAI technology is an innovative and encouraging technique for transferring computational chores from the network's core to its periphery. as a result, this strategy lowers the network's total communication overhead and enables more rapid and precise video analytics.

Unfortunately, this endeavour faces a number of significant challenges, some of which are as follows: (1) how to solve the difficulties of distributed AI model synchronisation in an EC context; (2) how to create a suitable edge computing architecture for the VS system, taking into account large-scale monitor terminals, massive video streams, and massive network

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

communication overhead; and (3) how to maintain workload balance across edge nodes in complex circumstances, such as imbalanced monitor terminal loads.

A handful of the numerous intelligent industrial domains that make use of AI technology include intelligent transportation [2, 7], the Internet of Things [8, 9], smart grids [11, 12], and video surveillance [5]. CNN's and DNNs are two examples of existing AI and deep learning approaches in the field of VS; however, they are typically used for static image analysis rather than image streaming and video analysis [13, 14]. Convolutional Neural Networks (CNN's) and Deep Neural Networks (DNNs) are two examples of existing AI and deep learning approaches in the field of VS. Even though they have a high data transmission cost, high latency, and significant packet loss restrictions, most of the today's distributed VS systems and AI algorithms use traditional centralised or cloud-based solutions [3, 12]. This is the case even though traditional solutions are cloud-based or centralised. Several articles [9, 15, 16] have proposed a variety of distributed AI and Deep Learning (DL) methodologies for use with distributed computing clusters and cloud computing platforms. There are a lot of different ways that distributed AI algorithms and VS systems may be tested in different EC contexts [4, 9, 17]. During this investigation, we build a Distributed Intelligent Video Surveillance system that functions based on a distributed deep learning paradigm and then deploy it in an environment that consists of edge computing.

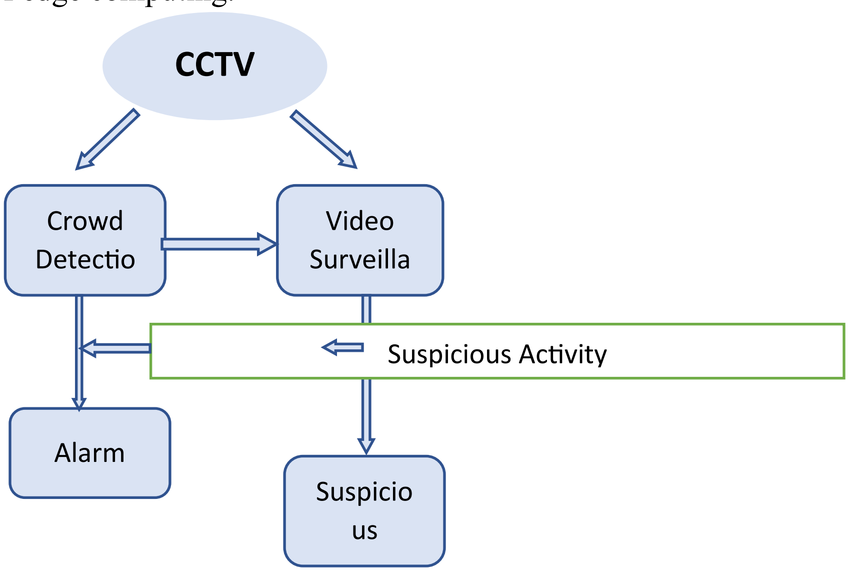


Fig 2. Application Block Diagram

Source: (Self-Created)

The following is a summary of the significant findings that were uncovered by this study:

• To improve the distributed intelligent video surveillance system, we develop a multi-tiered edge computing architecture as well as a distributed deep learning training model. When computational workloads are moved from the network core to the network edges, there is a reduction in the expenses associated with transmission, as well as an increase in the availability of low-latency analytical solutions. In this research, we present concurrent training approaches for the distributed DL model that can be used at both the task-level and the model-level. Each edge node performs numerous jobs of data processing in simultaneously after deploying many

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

DL sub-models, each of which has its own topology. At the edge node level, the CNN model's training operations are further parallelized using distributed computing. It is disclosed that a method for updating model parameters may be used to accomplish model synchronisation of the global DL model on the EC platform while keeping communication costs to a minimum. We present a dynamic data transmission strategy to enhance the workload balance of the Distributed Intelligent Video Surveillance System. Our strategy takes into consideration the unequal connection of monitor terminals as well as the uneven compute power of edge nodes. Video surveillance is a method used to monitor an area, either inside or outside. Managing security risks and keeping tabs on employees are just two of the many uses for video surveillance systems that have been adopted by a wide range of enterprises. Because of the installation of video surveillance equipment, the administrator can keep track of who is doing what at any given moment as well as what is happening at any given time. A comparable capability exists for video surveillance technology to watch and record the activities of any human person. There are a variety of methods that may be implemented while carrying out video surveillance. A library of activity templates that are organised by completion date may be found inside the model templates. Object-based models operate in a comparable manner, in that they base the behaviour of the user on a catalogue of items and the properties of those things. Like how there are a few different scientific

Fig. 3 Multiple CCTV Footages

Source (https://www.shutterstock.com/video/search/multiple-cctv)

methodologies from which to pick while conducting behaviour analysis. To categorise the action, the k-means technique would make an estimate of the distance measure based on singular objects that exhibited a known behaviour. How effectively the k-means approach works is highly dependent on the characteristic being analysed as well as the number of samples that are available in each class. The support vector machine has been shown to be beneficial for a wide variety of tasks, including behavioural analysis in the context of video surveillance, which is only one of the many activities. The accuracy of any algorithm is strongly dependent on the characteristics it utilises and the similarity estimation techniques it employs. This is true regardless of the kind of algorithm. Object tracking is the first step in performing behaviour analysis using footage from security cameras. Each still image of the video would have a substantial amount of information and detail. You will need to be able to label the items that may be seen in the video frames if you wish to perform behaviour tracking. This is because the video frames contain a lot of information. When the location of the frame's components has been established, the behaviour may then be analysed. Items can be recognised by the application of saved templates that depict the things that are required for specific procedures.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

With template matching, it might be possible to determine what the items are. In a similar manner, it is extremely important to include contextual aspects while analysing behaviour. Object detection may be made better by separating the characteristics of the foreground first. There have been a lot of different backdrop models employed, but the accuracy of those models is rather bad when it comes to removing background elements. Considering these things to consider, a novel model for multi-variant feature analysis is presented in this article. For behaviour analysis, the method makes use of a wide variety of form features. To complete the categorization procedure, an artificial neural network was utilised. The ANN may be used for behaviour analysis in video surveillance and has a profound impact on many scientific problems.

A technique for identifying potentially malicious behaviour was implemented in this system by analysing footage from closed-circuit television cameras. Several of the modules and data sets that come with the system are use.

CLASSIFIERS BASED ON IMAGE RECOGNITION:

By using image sequences, we can monitor the typical focus areas and filter out the background noise in the images that are currently being displayed (foreground).

1. Recognizing People According to the Behaviours They Exhibit:
   The act of walking, sleeping, and operating a motor vehicle are all examples of common human activities. As a result of the lack of success of scene understanding solutions and the fact that these methods are distinct from one another, the boundary conditions and object tracking rules that work in one domain may not be applicable in another. Cognitive theories have been used to categorise persons by analysing the dynamics of human movement at the level of the complete body as well as the movement of portions of the body. This framework employs a variety of empirical research approaches, each of which is distinct from the others. The moving item and the motion it is a part of my both be identified with the help of the images that were taken by the picture extraction unit and utilised by the image classifiers. These image classifiers need to be able to carry out a broad variety of functions, such as static background extraction, foreground separation, foreground noise reduction, object recognition, posture modelling and identification, and motion detection. However, we brought in experts and employed a complete approach for the ISADF frame function. This strategy considers polygonal shapes on faraway aircraft vehicles and the motion of objects based on certain very low-quality photographs.
2. Instructions for monitoring for suspicious behaviour.
   The subsequent behaviour detection algorithms are extremely simple and can only recognise object models (human activities), not suspicious ones. This is the reason why this is the case. Describe the suspiciousness of an object based on your observations of its motion. It is necessary to provide the system with knowledge in the form of training data so that it can perform the function for which it was designed. Both online and offline data are used to classify the various types of training material. We believe that offline-stored training data is immutable; nevertheless, this cannot be guaranteed. On the other hand, the content for professional development may draw on the experience gained during the early years of employment and be updated to reflect the present. With this approach, we make use of trainable data that is capable of

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

being updated online and that primarily concentrates on the detection of acts that raise ethical concerns.

3. A System for the Identification of Strange Behaviour
For its functions, this mechanism uses a training-data-related classification algorithm. Training methods for human observers are tied to data for both training and testing that has been pre-processed. By comparing video and still images to a database of individuals who have been suspected of wrongdoing, it is possible to identify criminal behaviour. This may be accomplished using one's awareness of their surroundings as well as the signals offered by suspicious video footage.

METHODOLOGY:

A solution based on optical flow is proposed in this research. The structure of CNN is composed of a total of eight tiers. The training is done with the help of the BVLC caffe framework. First, the parameters are seeded with random values, and then the system is trained using backpropagation, which is based on stochastic gradient descent. Throughout the implementation, the UCSD, UMN, Subway, and U-turn datasets were taken into consideration. Criteria at both the frame-level and the pixel-level are contained in the implementation details provided by UCSD. The focus of the frame level criterion is on the time domain, whereas the pixel level criterion considers both the spatial and temporal domains. Equal Error Rate (EER) and Detection Rate are two useful performance indicators that should be considered together (DR).

FSCB, behaviour detection by feature tracking and image segmentation, is recommended for online real-time crowd behaviour identification in video sequences. These are the actions that must be taken to complete the procedure:

Some of the aspects that may be recognised and filtered over time include activity detection, an activity map, an analysis of the data, and an alert system. In the article , the authors propose a strategy for the analysis of crowds that makes use of deep learning and includes the steps listed below. This strategy is independent of the environment in which the crowds are observed. Abnormality detection in a crowd, crowd attribute recognition, crowd behaviour analysis, crowd tracking, and crowd counting are all possible applications of this technology. To do attribute recognition, the slicing CNN is utilised. Discover the properties of its outside using a two-dimensional CNN model, and then save those discoveries in a cuboid form. It has been discovered that the cuboid has not one, not two, but three different temporal filters in its interior. When the cuboid's characteristics have been recovered, they are added together to form a vector, which is then fed into a classifier. The challenge of accurately predicting and measuring the size of a crowd may be stated as a regression issue. Attribute recognition research of crowds is conducted with the World Wide Web Crowd dataset serving as the basis for the investigation. AUC and AP are two measures that are utilised for the purpose of assessment.

The research paper titled "High Density Crowds in Films" describes various methodologies, including data-driven crowd analysis and density-aware tracking, among others. Data-driven analyses study vast collections of crowd recordings offline to understand crowd motion patterns. Whatever useful pattern you've learned may be ported to other programmes. There

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

are two steps to the process that will solve the problem. Examining the similarities and differences between global crowd scenes and local crowd patches. Figure 2 provides a visual representation of the process's two distinct phases.

In the database of films selected for the purpose of experimental analysis, there are a total of 720 480 titles. The observation of out-of-character actions carried out by members of the crowd is the primary evaluation standard to be applied here. Experiments have demonstrated that data-driven tracking is superior to batch mode tracking in terms of performance. A density-based system for recognising and following persons contains the following components: baseline detection, geometric filtering, and tracking with a detector that is aware of the density of its surroundings.

A review on diagnosing anomalous behaviour in crowd scenarios highlights four primary techniques: the Hidden Markov Model (HMM), the Gaussian mixture model (GMM).

flow, and support vector machine (STT). GMM is augmented with several other ways to improve its ability to catch unexpected behaviours. The general mixture model (GMM), the general mixture model plus the Markov random field (GMM+MRF), the Gaussian Poisson mixture model (GPMM), and the general mixture model plus a support vector machine are all upgraded versions of the original general mixture model (GMM).

Recognizing crowd behaviour with only one shot implies using untrained or sparsely trained recognizers. Attribute-context co-occurrence is the central concept in this method. Attributes that are already known can be used to make predictions about the target behaviour. The technique is comprised of several steps, one of which is the probabilistic zero-shot prediction. In the procedure, the conditional probability of an initial attribute connection that is already established as being valid is figured out. The second step consists of learning context based on visual co-occurrence and learning attribute relatedness from text corpora. Both types of learning are taken from the corpora. Figure 3 depicts the findings in its entirety.

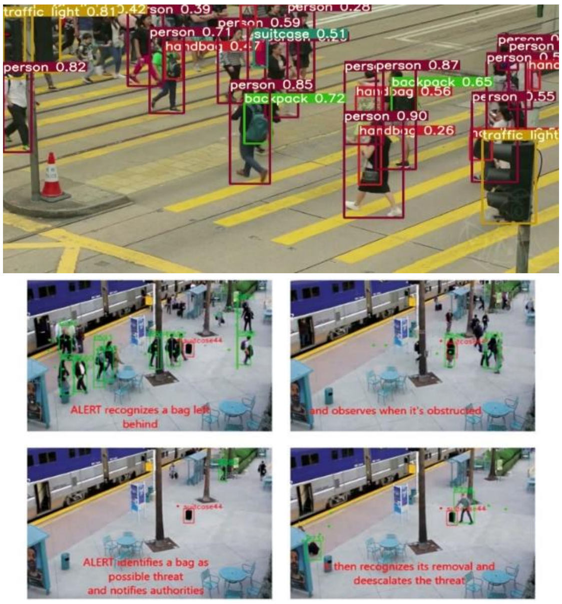


Fig 3.1 Identification of suspicious Objects

Source:

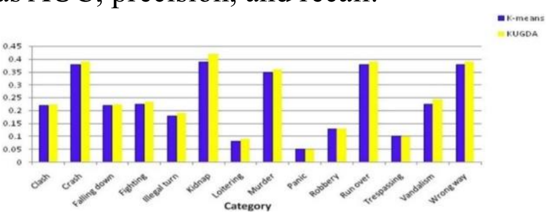
(https://viso.ai/deep-learning/object-detection/)

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

Rejecting Motion Outliers for Efficient Crowd Anomaly Detection is the term of a two-stage technique that was presented in reference to [54]. Retrieve features and identify outliers before moving on. Flow analysis is used to determine which features to extract. The pipeline is comprised of several stages, the first of which is the segmentation of the input video into frames, followed by the segmentation of the frames into super pixels, followed by the extraction of the histogram for each super pixel, followed by the spatial aggregation of the histograms, and finally the concatenation of the combined histograms from consecutive frames to extract the final feature.

A deep network model where features and distance measures are learned simultaneously is a part of Deep Metric Learning for Crowdedness Regression. Studying a precise distance measurement requires the application of metric learning. TensorFlow was the package that was used in the implementation of the model that was suggested. The activate function makes use of the rectified linear unit. During the training phase of the process, the gradient descent algorithm is utilised. Performance may be evaluated based on the mean squared error as well as the mean absolute error. During this analysis, both the world Expo dataset and the Shanghai Tech dataset were consulted.

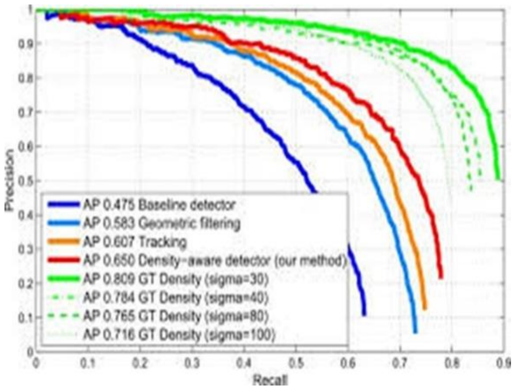
"A Deep Spatiotemporal Perspective for Understanding Crowd Behaviour" is an attempt to shed light on the dynamics of large groups of people by combining a convolution layer with a lengthy short-term memory. The convolution layer is responsible for capturing information about the space, while the LSTM layer oversees limiting the temporal motion dynamics. This method can predict the route that people will take, provide an estimate of where they will end up, and categorise people's actions according to their mobility patterns. This method of path forecasting makes use of two layers of stacked convLSTM of which possesses 128 hidden states. The ConvLSTM algorithm employs a 3x3 kernel that has a stride of 1 and zero padding. The model only makes use of a single convolution layer, and its kernel size is set to 1. A classification of crowd behaviour is achieved through the collaborative efforts of three layers: an average spatial pooling layer, a fully linked layer, and a SoftMax layer. Understanding Crowded Scenes with Deeply Trained Volumetric Slices suggests utilising a deep model in conjunction with a variety of different fusion techniques to decipher complex scenes that contain many people. The architecture is comprised of completely connected layers, as well as convolutional layers and a global sum pooling layer. The design requires the implementation of weight-sharing and slice-fusion technologies to function properly. A deep learning model that can tackle multiple tasks all at once has recently been developed. The results of the survey and some suggestions for moving forward with them. In this section, we will explain the accuracy analysis that was carried out on several the methods making use of a variety of assessment criteria, such as AUC, precision, and recall.



Source : (Python Project)

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

A comparison of the methods that were utilised in the research paper titled "Rejecting Motion Outliers for Efficient Crowd Anomaly Detection" [54] can be found in Figure 4. KUGDA is a classification method that was proposed in the paper titled "Rejecting Motion Outliers for Efficient Crowd Anomaly Detection"



Source : (Python Project)

Figure 5 shows the results of a comparison of the detection performance of several approaches. The results demonstrate that the density-aware detector is superior to the competing approaches.

After conducting an analysis of the current state of available methods, the following issues were discovered. The following are examples of goals that real-world problems aim to achieve: Complications with time, bad weather, real-world dynamics, occult phenomena, and object overlap.

Previous approaches treated each issue independently. There is no one approach that offers a solution to all the goals as separate features. To successfully manage intelligent crowd video analysis in real time, the approach in question needs to be capable of resolving each one of these issues. The use of conventional methods cannot result in a solution that is both economically and temporally viable.

Because high-speed computational resources such as GPUs are now more readily available, it is now feasible to implement deep learning-based solutions for the rapid processing of massive amounts of data. The best results can be obtained by combining the strongest aspects of several different deep learning architectures or models.

CONCLUSION

Discussion centres on various approaches to the intelligent analysis of surveillance footage. Publications that have undergone the process of peer review cover a wide variety of academic disciplines. Tables provided an overview of the methodologies, software, and data sets that were discovered. After providing an overview of the field of video surveillance analysis, the survey then shifts its emphasis to look at crowd analysis. The complexity of crowd analysis can be traced back to the fact that, in real life, crowd sizes are typically not only quite large but also constantly shifting. It is not simple to give a name to each individual thing and the peculiar behaviours it possesses. Several approaches to analysing the dynamics of crowds were covered in this discussion. Compilations of directions have been made for the development of more effective solutions to the problems found with the approaches that are currently being used.

REFERENCES

[1] Kardas K, Cicekli NK. SVAS: *surveillance video analysis system. Expert Syst Appl*. 2017;89:343–61.
[2] Wang Y, Shuai Y, Zhu Y, Zhang J. *An P Jointly learning perceptually heterogeneous features for blind 3D video quality assessment. Neurocomputing.* 2019;332:298–304 (ISSN 0925-2312).
[3] Tzelepis C, Galanopoulos D, Mezaris V, Patras I. *Learning to detect video events from zero or very few video exam- ples*. Image Vis Comput. 2016;53:35–44 (ISSN 0262-8856).
[4] Fakhar B, Kanan HR, Behrad A. *Learning an event-oriented and discriminative dictionary based on an adap- tive label-consistent K-SVD method for event detection in soccer videos*. J Vis Commun Image Represent. 2018;55:489–503 (ISSN 1047-3203).
[5] Luo X, Li H, Cao D, Yu Y, Yang X, Huang T. *Towards efficient and objective work sampling: recognizing workers' activities in site surveillance videos with two-stream convolutional networks*. Autom Constr. 2018;94:360–70 (ISSN 0926-5805).
[6] Wang D, Tang J, Zhu W, Li H, Xin J, He D. *Dairy goat detection based on Faster R-CNN from surveillance video*. Comput Electron Agric. 2018;154:443–9 (ISSN 0168-1699).
[7] Shao L, Cai Z, Liu L, Lu K. *Performance evaluation of deep feature learning for RGB-D image/video classification*. Inf Sci. 2017;385:266–83 (ISSN 0020-0255).
[8] Ahmed SA, Dogra DP, Kar S, Roy PP. *Surveillance scene representation and trajectory abnormality detection using aggregation of multiple concepts*. Expert Syst Appl. 2018;101:43–55 (ISSN 0957-4174).
[9] Arunnehru J, Chamundeeswari G, Prasanna Bharathi S. *Human action recognition using 3D convolutional neural networks with 3D motion cuboids in surveillance videos*. Procedia Comput Sci. 2018;133:471–7 (ISSN 1877-0509).
[10] Guraya FF, Cheikh FA. *Neural networks based visual attention model for surveillance videos.* Neurocomputing. 2015;149(Part C):1348–59 (ISSN 0925-2312).
[11] Pathak AR, Pandey M, Rautaray S. *Application of deep learning for object detection.* Procedia Comput Sci. 2018;132:1706–17 (ISSN 1877-0509).
[12] Ribeiro M, Lazzaretti AE, Lopes HS. *A study of deep convolutional auto-encoders for anomaly detection in videos*. Pattern Recogn Lett. 2018;105:13–22.
[13] Huang W, Ding H, Chen G. *A novel deep multi-channel residual networks-based metric learning method for mov- ing human localization in video surveillance*. Signal Process. 2018;142:104–13 (ISSN 0165-1684).
[14] Tsakanikas V, Dagiuklas T. *Video surveillance systems-current status and future trends.* Comput Electr Eng. In press, corrected proof, Available online 14 November 2017.
[15] Wang Y, Zhang D, Liu Y, Dai B, Lee LH. *Enhancing transportation systems via deep learning: a survey*. Transport Res Part C Emerg Technol. 2018. https://doi.org/10.1016/j.trc.2018.12.004 (ISSN 0968-090X).
[16] Huang H, Xu Y, Huang Y, Yang Q, Zhou Z. *Pedestrian tracking by learning deep features*. J Vis Commun Image Represent. 2018;57:172–5 (ISSN 1047-3203).

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

[17] Yuan Y, Zhao Y, Wang Q. *Action recognition using spatial-optical data organization and sequential learning frame- work. Neurocomputing.* 2018;315:221–33 (ISSN 0925-2312).

[18] Perez M, Avila S, Moreira D, Moraes D, Testoni V, Valle E, Goldenstein S, Rocha A. *Video pornography detec- tion through deep learning techniques and motion information. Neurocomputing*. 2017;230:279–93 (ISSN 0925-2312).

[19] Pang S, del Coz JJ, Yu Z, Luaces O, Díez J. *Deep learning to frame objects for visual target tracking*. Eng Appl Artif Intell. 2017;65:406–20 (ISSN 0952-1976).

[20] Wei X, Du J, Liang M, Ye L. *Boosting deep attribute learning via support vector regression for fast moving crowd counting*. Pattern Recogn Lett. 2017. https://doi.org/10.1016/j.patrec.2017.12.002.

[21] Xu M, Fang H, Lv P, Cui L, Zhang S, Zhou B. D-stc: *deep learning with spatio-temporal constraints for train drivers' detection from videos*. Pattern Recogn Lett. 2017. https://doi.org/10.1016/j.patrec.2017.09.040 (ISSN 0167-8655).

[22] Hassan MM, Uddin MZ, Mohamed A, Almogren A. *A robust human activity recognition system using smartphone sensors and deep learning.* Future Gener Comput Syst. 2018;81:307–13 (ISSN 0167-739X).

[23] Wu G, Lu W, Gao G, Zhao C, Liu J. *Regional deep learning model for visual tracking. Neurocomputing.* 2016;175:310–23 (ISSN 0925-2312).

[24] Nasir M, Muhammad K, Lloret J, Sangaiah AK, Sajjad M. *Fog computing enabled cost-effective distributed summa- rization of surveillance videos for smart cities.* J Parallel Comput. 2018. https://doi.org/10.1016/j.jpdc.2018.11.004 (ISSN 0743-7315).

[25] Najva N, Bijoy KE. *SIFT and tensor-based object detection and classification in videos using deep neural networks*. Procedia Comput Sci. 2016;93:351–8 (ISSN 1877-0509).

[26] Yu Z, Li T, Yu N, Pan Y, Chen H, Liu B. *Reconstruction of hidden representation for Robust feature extraction*. ACM Trans Intell Syst Technol. 2019;10(2):18.

[27] Mammadli R, Wolf F, Jannesari A. *The art of getting deep neural networks in shape.* ACM Trans Archit Code Optim. 2019;15:62.

[28] Zhou T, Tucker R, Flynn J, Fyffe G, Snavely N. *Stereo magnification: learning view synthesis using multiplane images.* ACM Trans Graph. 2018;37:65

[29] Fan Z, Song X, Xia T, Jiang R, Shibasaki R, Sakuramachi R. *Online Deep Ensemble Learning for Predicting Citywide Human Mobility.* Proc ACM Interact Mob Wearable Ubiquitous Technol. 2018;2:105.

[30] Hanocka R, Fish N, Wang Z, Giryes R, Fleishman S, Cohen-Or D. ALIGNet: *partial-shape agnostic alignment via unsupervised learning.* ACM Trans Graph. 2018;38:1.

[31] Xu M, Qian F, Mei Q, Huang K, Liu X. DeepType: *on-device deep learning for input*.

[32] Potok TE, Schuman C, Young S, Patton R, Spedalieri F, Liu J, Yao KT, Rose G, Chakma G. *A study of complex deep learning networks on high-performance, neuromorphic, and quantum computers*. J Emerg Technol Comput Syst. 2018;14:19.

[33] Pouyanfar S, Sadiq S, Yan Y, Tian H, Tao Y, Reyes MP, Shyu ML, Chen SC, Iyengar SS. *A survey on deep learning: algorithms, techniques, and applications*. ACM Comput Surv. 2018;51:92.

[34] Tian Y, Lee GH, He H, Hsu CY, Katabi D. *RF-based fall monitoring using convolutional neural networks*. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2018;2:137.

[35] Roy P, Song SL, Krishnamoorthy S, Vishnu A, Sengupta D, Liu X. NUMA-Caffe: *NUMA-aware deep learning neural networks*. ACM Trans Archit Code Optim. 2018;15:24.

[36] Lovering C, Lu A, Nguyen C, Nguyen H, Hurley D, Agu E. *Fact or fiction*. Proc ACM Hum-Comput Interact. 2018;2:111.

[37] Ben-Hamu H, Maron H, Kezurer I, Avineri G, Lipman Y. *Multi-chart generative surface modeling*. ACM Trans Graph. 2018;37:215

[38] Ge W, Gong B, Yu Y. *Image super-resolution via deterministic-stochastic synthesis and local statistical rectification*. ACM Trans Graph. 2018;37:260

[39] Hedman P, Philip J, Price T, Frahm JM, Drettakis G, Brostow G. *Deep blending for free-viewpoint image-based rendering*. ACM Trans Graph. 2018;37:257

[40] Sundararajan K, Woodard DL. *Deep learning for biometrics: a survey.* ACM Comput Surv. 2018;51:65.

[41] Kim H, Kim T, Kim J, Kim JJ. *Deep neural network optimized to resistive memory with nonlinear current–voltage characteristics*. J Emerg Technol Comput Syst. 2018;14:15.

[42] Wang C, Yang H, Bartz C, Meinel C. *Image captioning with deep bidirectional LSTMs and multi-task learning*. ACM Trans Multimedia Comput Commun Appl. 2018;14:40.

[43] Yao S, Zhao Y, Shao H, Zhang A, Zhang C, Li S, Abdelzaher T. RDeepSense: *Reliable Deep Mobile Computing Mod- els with Uncertainty Estimations.* Proc ACM Interact Mob Wearable Ubiquitous Technol. 2018;1:173.

[44] Liu D, Cui W, Jin K, Guo Y, Qu H. DeepTracker: *visualizing the training process of convolutional neural networks*. ACM Trans Intell Syst Technol. 2018;10:6.

[45] Yi L, Huang H, Liu D, Kalogerakis E, Su H, Guibas L. *Deep part induction from articulated object pairs*. ACM Trans Graph. 2018. https://doi.org/10.1145/3272127.3275027.

[46] Zhao N, Cao Y, Lau RW. *What characterizes personalities of graphic designs?* ACM Trans Graph. 2018;37:116.

[47] Tan J, Wan X, Liu H, Xiao J. QuoteRec: *Toward quote recommendation for writing.* ACM Trans Inf Syst. 2018;36:34.

[48] Qu Y, Fang B, Zhang W, Tang R, Niu M, Guo H, Yu Y, He X. *Product-based neural networks for user response predic- tion over multi-field categorical data.* ACM Trans Inf Syst. 2018;37:5.

[49] Yin K, Huang H, Cohen-Or D, Zhang H. P2P-NET: *Bidirectional point displacement net for shape transforms*. ACM Trans Graph. 2018;37:152.

[50] Yao S, Zhao Y, Shao H, Zhang C, Zhang A, Hu S, Liu D, Liu S, Su L, Abdelzaher T. SenseGAN: *Enabling deep learning for internet of things with a semi-supervised framework*. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2018;2:144.

[51] Saito S, Hu L, Ma C, Ibayashi H, Luo L, Li H. *3D hair synthesis using volumetric variational autoencoders*. ACM Trans Graph. 2018. https://doi.org/10.1145/3272127.3275019.

[52] Chen A, Wu M, Zhang Y, Li N, Lu J, Gao S, Yu J. *Deep surface light fields.* Proc ACM Comput Graph Interact Tech. 2018;1:14.

[53] Chu W, Xue H, Yao C, Cai D. *Sparse coding guided spatiotemporal feature learning for abnormal event detection in large videos.* IEEE Trans Multimedia. 2019;21(1):246–55.

[54] Khan MUK, Park H, Kyung C. *Rejecting motion outliers for efficient crowd anomaly detection*. IEEE Trans Inf Foren- sics Secur. 2019;14(2):541–56.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2003 - 2015

[55] Tao D, Guo Y, Yu B, Pang J, Yu Z. *Deep multi-view feature learning for person re-identification.* IEEE Trans Circuits Syst Video Technol. 2018;28(10):2657–66.

[56] Zhang D, Wu W, Cheng H, Zhang R, Dong Z, Cai Z. *Image-to-video person re-identification with temporally memorized similarity learning.* IEEE Trans Circuits Syst Video Technol. 2018;28(10):2622–32.