Secure IoT Systems Monitor Framework using Probabilistic Image Encryption

P.G. Narmadha, K. Pooja, Mrs. K .Muthulakshmi

Department of IT, Panimalar Engineering College, Chennai, India

Abstract— *In recent years, the modeling of human behaviors and patterns of activity for recognition or detection of special events has attracted considerable research interest. Various methods abounding to build intelligent vision systems aimed at understanding the scene and making correct semantic inferences from the observed dynamics of moving targets. Many systems include detection, storage of video information, and human-computer interfaces. Here we present not only an update that expands previous similar surveys but also a emphasis on contextual abnormal detection of human activity , especially in video surveillance applications. The main purpose of this survey is to identify existing methods extensively, and to characterize the literature in a manner that brings to attention key challenges.*

Keywords— *recognition, surveillance, abnormal human behavior.*

I. INTRODUCTION

Wireless Multimedia Surveillance Networks (WMSNs) are part of this IoT-assisted environment , which consists of visual sensors that observe the surrounding environment from multiple overlapping views by continuously capturing images, producing a large amount of visual data with significant redundancy[1-3]. In the surveillance networks research community it is generally understood that the visual data obtained should be processed and that only the useful data should be preserved for future use, such as irregular event identification, case management , data interpretation and video abstraction. The explanation for this is that, due to resources and bandwidth limitations, transmitting all image data across the transmission lines without processing is inefficient. Additionally, the efficient extraction of actionable intelligence from the sheer volume of surveillance data[4] is comparatively difficult and time-consuming for an analyst. Therefore, a mechanism that can collect semantically important visual data autonomously must be exploited by using the processing and transmission capabilities of modern smart visual sensors Such a mechanism can allow the correct view to be intelligently selected from multi-view surveillance data captured from multiple sensors linked through IoT infrastructure. It will allow real-time retrieval of the collected data such that only valid data can be sent to the central database for potential use. Currently there are networks of video cameras available. The amount of data generated by these vision sensor network installed in many settings ranging from protection needs to environmental surveillance easily satisfies big data requirements[5],[6]. The difficulties in analyzing and processing such large video data are apparent whenever an incident occurs which requires foraging through vast video archives to identify interesting events. As a consequence, video summarization, which has gained considerable interest in recent years to automatically retrieve a short and insightful description of these images. Although video summarization has been extensively studied in recent years, many previous methods focused primarily on developing a variety of ways to summarize single-view videos in the form of a key frame sequence or a video skim[7-11]. Another big concern, though, although seldom discussed in this sense, is providing an concise overview from multi-view videos[14],[15]. Multi-view video summarization refers to the question of summarization that attempts to take a series of input images taken from various cameras based on approximately the same fields-of-view (fov) from different perspectives and to generate a video synopsis or main frame sequence that depicts the most important portions of the inputs within a short period (see Fig . 1). In this article, given a range of videos and their

---

1.

shots, we concentrate on creating an unsupervised approach to choosing a subset of shots that make up the overview of multiple views.Such a summary can be very beneficial in many surveillance systems equipped in offices, banks, factories, and crossroads of cities, for obtaining significant information in short time.
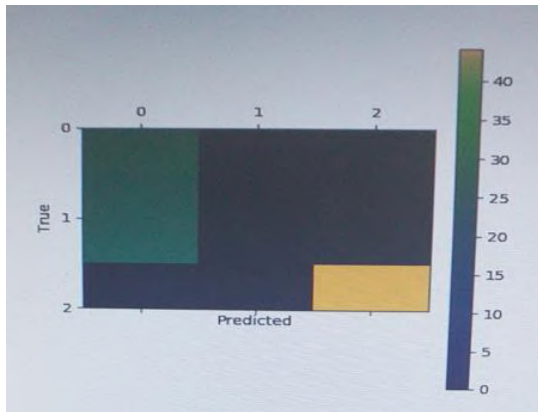

Fig.1: cameras focusing on roughly the same fields-of-view

II. LITERATURE SURVEY

Rameswar Panda et al, in two important ways, represents Multi-view video summary is different from single video summary. First, although the quantity of multi-view data is enormously challenging, a certain structure underlies it. Specifically due to the locations and fields of view of the cameras, there is a large amount of correlations in the data. So to get an informative summary, content correlations as well as discrepancies between different videos need to be properly modeled. Secondly, for the same scenery, these videos are captured with different view angles, and field depth, resulting in a number of unaligned videos. Therefore, variations in lighting, posture, point of view and synchronization problems face a significant challenge in summing up these images. Methods that attempt to derive description from single-view videos do not produce an appropriate set of members when summing up multi-view videos.

A. A. Steffi et al. proposed a modified algorithm for encryption and decryption of images using Lorenz and Baker map. Among that, the authors presented encryption process comprising of two stages: confusion and diffusion. In both stages, the pixel positions and values are changed based on one of two chaotic systems (Lorenz and Baker). To improve security of the algorithm, separate keys are used for generating the chaotic sequence. However, for decryption

stage, reverse operation is performed to obtained original image.

X. Zhang et al. proposed an chaos-based image encryption scheme based on large permutation with chaotic sequence [12]. The image encryption scheme proposed in this paper consists of multiple rounds of permutation and diffusion. The permutation process is used to permute all the pixels. After that, the diffusion process modified the pixel value. The pseudo number is generated by logistic map. For the decryption algorithm the only difference comes out in the inverse of iteration. Their test results and analysis have demonstrate that this scheme is much faster than the other works suggested by Fridrich et al. [16], G. Chen et al. [17] and G. Zhang et al. [18], because the proposed chaotic image encryption is well suitable for real-time transmission.

III. PROPOSED SYSTEM

The main purpose of this survey is to identify existing methods extensively, and to characterize the literature in a manner that brings attention to key challenges. The block diagram shown in Fig, for this proposed procedure. 2.
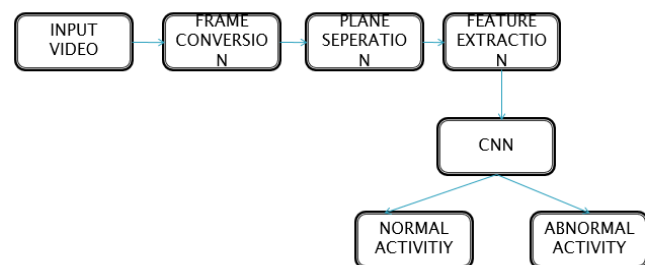

Fig.2: Proposed system block diagram

A. INPUT VIDEO

For multi-view surveillance videos recorded in industrial environments, their processing capacities may be used to evaluate the video stream to identify keyframes and then delete obsolete and redundant visual data, thereby reducing the requirements for bandwidth. In addition, keyframe protection can be assured by applying the Gaussian blurring theory, taking into account the computing capacities, memory and transmission constraints.

B.FRAME CONVERSION

With each frame taken by the visual camera, the integral image is computed, then background bootstrapping is done which is important with eliminating background motion and

correct measurement of salient motion. Salt motion can be measured in neighboring frames by calculating changes in image block values. Even small background motion is robust, as it uses temporal gradients based on the background model and the integral image for salient motion. This can be ascertained from Fig. 3, where the salient motion recognition is demonstrated by our scheme using a few frames from a reference film of an illegal border crossing.
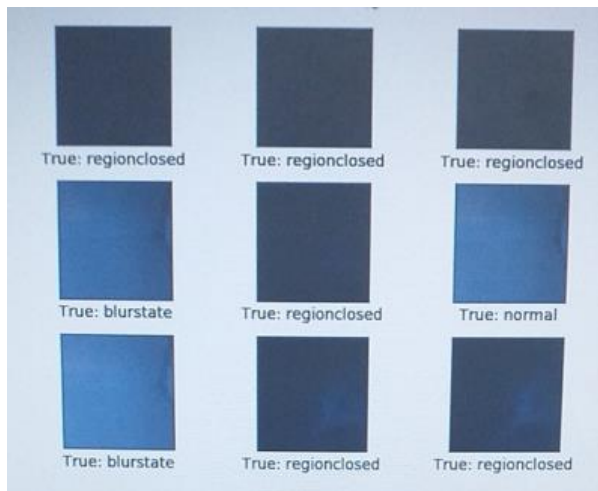

Fig.3: Salient motion detection

C. PLANE SEPARATION

In the given sample video, there is significant motion clutter due to the true and prediction that continuously change the background pattern, like normal and blur state, thus making the salient motion detection more challenging. Despite these challenges, this approach detects the salient motion correctly, as shown in Fig. 4.
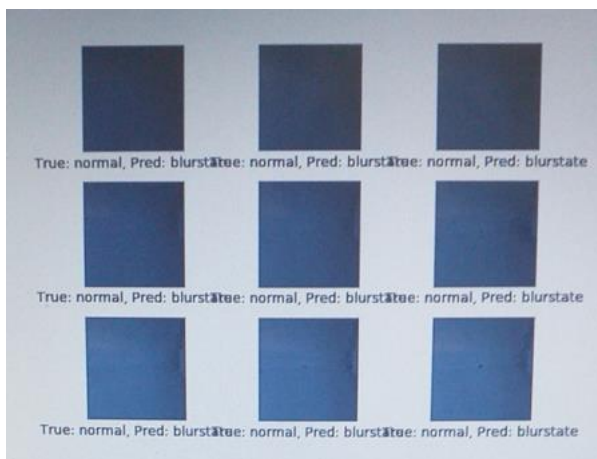

Fig.4: Plane separation of salient motion detection

D.FEATURE EXTRACTION

Feature Extraction expects to decrease the number of highlights in a dataset by making new features from the current ones. This new diminished arrangement of the feature should then have the option to outline the vast majority of the data contained in the first arrangement of the feature. Right now, outlined adaptation of the first features can be made from a blend of the first set.

IV.   DESIGN METHODOLOGY

A.CONVOLUTIONALNEURAL NETWORKS

A CNN contains is a supervised learning algorithm, for training Multi-Layer Perceptions. It is a general, hierarchical feature extractor that will map input image pixel intensities into a feature vector. This will be classified by several fully connected layers in the next step. All adjustable parameters are optimized by minimizing the misclassification by reducing the error over the training set. Each convolutional layer performs a 2D convolution of its input maps with a filter of different size 3 x 3, 5 x 5, 7 x 7. The subsequent activations of the output maps are given by the total of the past convolutional responses which are gone through a nonlinear activation function. Max pooling layer will perform the dimensionality reduction. The output of a thin layer is given by the most extreme activation over non-covering rectangular areas. Max-pooling makes location invariance and down-samples the image along every direction over the bigger neighborhood. Filter size of convolutional and max-pooling layers are selected in such a way that a fully connected layer can combine the output into a one-dimensional vector. The last layer will always be a fully connected layer which contains one output unit for all classes. Here rectification linear unit is used as the activation function.

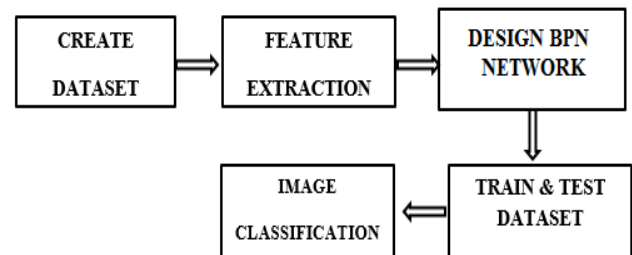### B.IMAGE CLASSIFICATION BASED ON CNN

Fig.5.1: Image classification Using CNN

Object detection is the way toward discovering occasions of certifiable items, for example, faces, structures, and bike in pictures or recordings. Along these lines, the working procedure of picture order dependent on the CNN appears on Figure 4.1. Object detection calculations commonly use separated highlights and learning calculations to perceive the occurrences of an item class. It is regularly utilized in applications, for example, picture recovery, security and propelled driver help frameworks.
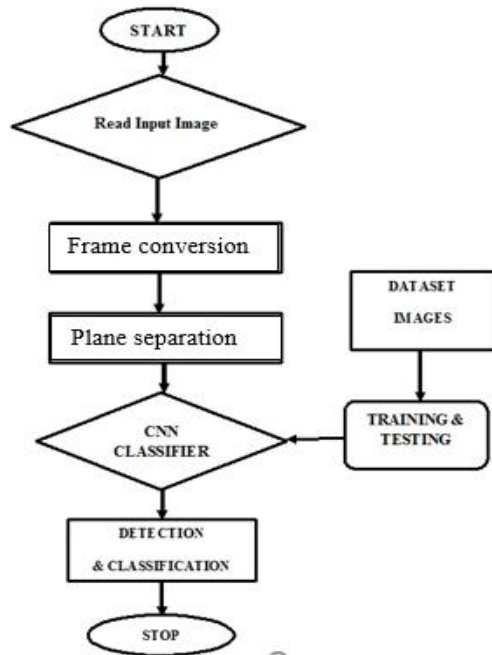
Fig.5.2 Flow Chart of of Detection

C. SOFTWARE SYSTEM

A.   Python 2.5 / 3.5

Python is a scripting language of high quality, interpreted, dynamic and object-oriented. Python is designed to be easily readable Python is object-oriented − Python supports object-oriented programming style or method encapsulating code inside objects. Meanwhile, it released Python 3.0 in 2008. Python 3 does not fit backwards with Python 2. Due to its growing popularity as a scientific programming language and the free availability of many state-of-the-art image processing tools within its ecosystem, Python is an excellent choice for these types of image processing tasks.

B.MACHINELEARNINGLIBRARIES   AND FRAMEWORKS

Many popular ML frameworks and libraries already offer the possibility to use GPU accelerators to speed up learning process with supported interfaces Some of them also allow the use optimised libraries such as CUDA (cuDNN), and OpenCL to improve the performance even further. The main feature of many-core accelerators is that their massively parallel architecture allows them to speedup computations that involve matrix-based operations. The software development in the ML/DL direction community is highly dynamic and has various layers of abstraction as depicted.

A.   OpenCV-Python

OpenCV (Open Source Computer Vision Library) is one of the most widely used computer-vision libraries. OpenCV-Python forms the OpenCV Python API. Not only is OpenCV-Python fast, since the background consists of code written in C / C++, it is also easy to code and deploy (because of the Python wrapper in the foreground). This makes computationally intensive computer vision programs a great choice.

4.5.4 DEEP   LEARNING   LIBRARIESAND FRAMEWORKS

Deep learning (DL) is an artificial intelligence branch that allows computers the ability to learn without express programming. Here machine is trained to identify objects of various kinds. Object image is given as an input to the machine, and the processor tells if it is the same object or not. Until the DL era, the apps were selected and designed manually, and then a classifier followed. The revolutionary part of ML is that features are mostly learned automatically by using Convolutionary Neural Network (CNN) from the training data. CNN's use renders a classifier efficient in image recognition process. Deep learning is a branch of machine learning that has some of the best results in these fields. In order to facilitate the implementation of those approaches, a set of software frameworks have been developed and are currently available. TensorFlow is the second generation of Google artificial intelligence learning system library that has got much attention and affirmation in the field of the machine learning in all over the world. TensorFlow is written with a Python API over a C/C++ engine, this makes it run faster. It is available on Linux, Mac, Windows OS and embedded platforms like Android OS and Raspberry Pi. It provides good accuracy with better detection

speed. Its detection speed can improve by efficient algorithm.

A. Tensorflow

TensorFlow is an open-source numerical computing software library using data flow graphs TensorFlow was created and maintained by the Google Brain team within the Machine Intelligence research organization for ML and DL at Google. This is officially available under the open-source license Apache 2.0. TensorFlow is designed for the distributed testing and inference on a wide scale. Nodes in the graph represent mathematical operations, while the edges of the graph represent the shared multidimensional data arrays (tensors) between them. The distributed architecture of TensorFlow involves centralized master and worker facilities of Kernel implementations. These include 200 standard operations written in C++ including mathematical operations, array manipulation, control flow, and state management operations. TensorFlow is designed for use in research, production and manufacturing systems. It can run hundreds of nodes on single CPU systems, GPUs, handheld devices, and distributed large scale networks.

V. RESULTS AND DICUSSION

In this, the project presents experimental results and discuss the suitability of the best performing representation and model over the others. The architecture our model is based on classification of CNN with normal and abnormal activities using CCTV surveillance. In the figure 5.1 represents the feature extraction of the sample input images.

Fig.5.3: Results of Normal Stages Classification

The training validation uses for conduct experiments to have fair validation of the performance of the proposed approach. In Figure 5.3 presents the results of the experiments for the classification of 'Normal' vs. 'Abnormal Images' using proposed CNN and the existing method used for comparison for both the datasets.

Fig.5.3: Results of Abnormal Stages Classification

VI. CONCLUSION

A significant amount of redundant video data is generated thanks to recent advances in IoT-assisted surveillance networks in industrial environments. Its transmission, analysis and management are difficult and demanding, requiring prioritization of the image. For this job, an effective video description approach is first used to retrieve the informative frames from the video surveillance data and can be used to identify suspicious activities. As the derived keyframes are essential for further research, their privacy and protection during transmission is of utmost importance. Hence, we proposed a quick probabilistic keyframe prior to transmission, taking into account the memory and processor requirements of restricted devices, which enhance its suitability for industrial IoT systems.

REFERENCES

[1] R. Hamza, K. Muhammad, Z. Lv, and F. Titouna, "Secure video summarization framework for personalized wireless capsule endoscopy," Pervasive and Mobile Computing, 2017.

[2] K. Muhammad, M. Sajjad, M. Y. Lee, and S. W. Baik, "Efficient visual attention driven framework for key frames extraction from hysteroscopy videos," Biomedical Signal Processing and Control, vol. 33, pp. 161-168, 2017.

[3] K. Muhammad, M. Sajjad, and S. W. Baik, "Dual-level security based cyclic18 steganographic method and its application for secure transmission of keyframes during wireless capsule endoscopy," Journal of medical systems, vol. 40, pp. 1-16, 2016.

[4] J. Wu, B. Cheng, M. Wang, and J. Chen, "Energy-Aware Concurrent Multipath Transfer for Real-Time Video Streaming over Heterogeneous Wireless Networks," IEEE Transactions on Circuits and Systems for Video Technology, 2017.

[5] R. Panda and A. R. Chowdhury, "Multi-View Surveillance Video Summarization via Joint Embedding and Sparse Optimization," IEEE Transactions on Multimedia, 2017.

[6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001, pp. I-I.

[7] Z. Hua and Y. Zhou, "Image encryption using 2d logistic-adjusted-sine map," Information Sciences, vol. 339, pp. 237-253, 2016.

[8] M. Machkour, A. Saaidi, and M. Benmaati, "A Novel Image Encryption Algorithm Based on the Two-Dimensional Logistic Map and the Latin Square Image Cipher," 3D Research, vol. 6, pp. 1-18, 2015.

[9] X. Wu, D. Wang, J. Kurths, and H. Kan, "A novel lossless color image encryption scheme using 2D DWT and 6D hyperchaotic system," Information Sciences, vol. 349, pp. 137-153, 2016.

[10] A. Belazi, A. A. A. El-Latif, and S. Belghith, "A novel image encryption scheme based on substitution-permutation network and chaos," Signal Processing, vol. 128, pp. 155-170, 2016.

[11] X. Wei, L. Guo, Q. Zhang, J. Zhang, and S. Lian, "A novel color image encryption algorithm based on DNA sequence operation and hyper-chaotic system," Journal of Systems and Software, vol. 85, pp. 290-299, 2012.

[12] Zhang, X., Shao, L., Zhao, Z., Liang, Z. (2014). An image encryption scheme based on constructing large permutation with chaotic sequence. Comput. Electr. Eng., vol. 40, no. 3, pp. 931 941.

[13] Y. Zhou, Z. Hua, C.-M. Pun, and C. P. Chen, "Cascade chaotic system with applications," IEEE transactions on cybernetics, vol. 45, pp. 2001-2012, 2015.

[14] X. Huang, "Image encryption algorithm using chaotic Chebyshev generator," Nonlinear Dynamics, vol. 67, pp. 2411-2417, 2012.

[15] L. Xu, Z. Li, J. Li, and W. Hua, "A novel bit-level image encryption algorithm based on chaotic maps," Optics and Lasers in Engineering, vol. 78, pp. 17-25, 2016.

[16] Fridrich, J. (1998). Symmetric Ciphers Based on Two-Dimensional Chaotic Maps. Int.J.Bifurc. Chaos, vol. 08, no. 06, pp. 1259 1284.

[17] Chen, G., Mao, Y., Chui, C.K. (2004). A symmetric image encryption scheme based on 3D chaotic cat maps. Chaos Solitons Fractals, vol. 21, no. 3, pp. 749/761.

[18] Zhang, G., Liu, Q. (2011). A novel image encryption method based on total shuffling scheme. Opt. Commun., vol. 284, no. 12, pp. 2775 2780