Smoke Detection Using GMM and Deep Belief Network

Rabeb Kaabi
ENSIT, LR13ES03, SIME
Université de Tunis
Montfleury, Tunisia
rabebkaabi89@gmail.com

Mounir Sayadi
ENSIT, LR13ES03, SIME
Université de Tunis
Montfleury, Tunisia
mounirsayadi@yahoo.fr

Moez Bouchouicha
LIS-CNRS
Université de Toulon-Université Aix-Marseille
Toulon-Marseille, France
moez.bouchouicha@univ-tln.fr

Eric Moreau
Université de Toulon-Université Aix-Marseille
Toulon-Marseille, France
eric.moreau@univ-tln.fr

***Abstract*— The objective of this work is to develop a deep learning model for classification of smoke and no smoke regions in aerial recorded videos. For that, a deep belief network model was selected and implemented. First, frames were extracted from the provided videos. The Gaussian Mixture Model (GMM) was applied as background estimation algorithm. Then, the Deep Belief Network algorithm was applied to detect the smoke for the candidate region. Deep Belief Network was implemented and tested on different datasets. Overall, the obtained results reveal that our implemented model was able to accurately classify smoke and no smoke regions. Through the experiments with input videos obtained from various weather conditions, the proposed algorithms were useful to detect smoke in forests to minimize the damage caused by forest fires onto vegetation, animals and humans.**

Keywords- smoke detection; GMM; Deep Belief Network.

I. INTRODUCTION

Video-based smoke detection systems are composed of two types of methods. The first one relies on static features, such as the color of the smoke. The second method uses dynamic feature like movement, texture, etc.

Our proposed method of smoke detection uses dynamic features based on the Gaussian Mixture Model [1] and the Deep Belief Network (DBN) [2][3] as a classifier to reduce the false alarms and to increase the detection rate of video smoke detection systems.

Several researchers have played a significant role in the development of useful video smoke detection systems. We focused on some of them. Hu et al. [4] extract the motion feature and use the Convolutional Neural Network (CNN) [5] as classifier. Chen et al. [6] extract motion, color and energy features to classify smoke and no smoke regions using Support Vector Machine SVM [7]. Toreyin et al. [8] extract motion, color, energy and texture as features and classify the smoke and no smoke regions using decision trees.

The rest of the paper is organized as follows: the proposed technique intended for smoke detection is presented and detail in Section 2. Experimental results are shown in Section 3. Conclusion and perspectives are presented in the last section.

II. THE PROPOSED METHOD

A. Smoke Motion Detection

First, extracting candidated regions is a crucial step to know the nature of motion (ordinary or chaotic). Labeling motion regions could be done by using three methods: Optical Flow, Background Subtraction and Temporal differencing. The technique intended for background subtraction that we used in simulations is the Gaussian Mixture Model [1]. This method subtracts the background image from the current frame to find regions containing motions. In this approach, the camera is stationary.

Each pixel in the frame is defined by a mixture of K Gaussian distributions. The probability that a pixel represents the intensity is defined by:

$$P(I_t) = \sum_{i=1}^{K} w_i \gamma(I_t, \mu_i, \sigma_i^2) \tag{1}$$

where w_i is the weight.

μ_i is the mean.

σ_i^2 is the covariance for the i th distribution.

γ is a Gaussian probability density function:

$$\gamma(I_t, \mu_i, \sigma_i^2) = \frac{e^{-1/2\sigma^2 (I_t-\mu)^T (I_t-\mu)}}{(2\pi)^{1/2}\sigma} \tag{2}$$

To model the background, Bk is estimated as:

$$B_k = \arg\min(\sum_{i=}^{b} w_i > Th) \tag{3}$$

Th is the minimum fraction of the background model.

 ISBN: 978-1-61208-716-0

B. *Fisher Vector*

Fisher vector [9] represents a dimension reduction technique that can be used for classification as well. It picks a new dimension that gives maximum separation between means of projected classes and minimum variances within each projected class. The Fisher Vector is an image representation obtained by pooling local image features. It is mostly used as an image descriptor in visual classification and improves the classification performance of the representation. Our developed idea is realized as follows:

- Fitting a Gaussian Mixture Model (GMM)
- Saving and loading the fitted GMM
- Computing the Fisher Vectors based on the fitted GMM

C. *Deep Belief Network*

Deep Belief Network is a tool of machine learning that represents a stack of Restricted Boltzmann Machine [10]. After the pre-training process with the RBMs, the network acts like a multi-layer Perceptron [11] using the backpropagation [12] as tool to accomplish the training. The architecture of the used Deep Belief Network is presented in Figure 1.

Our aim is to improve the performance of smoke detection. An efficient method is presented using deep belief network. The algorithm is separated into two major phases. The first phase is to segment the candidate regions in video sequences. The Fisher criterion is applied to maximize the ratio of the separation of the two classes with respect to their dispersions. The Fisher criterion [9] is similar to Principle components analysis but it focuses on maximizing the separability among known categories. Second, the normalized candidate areas are identified by the novel structure based on deep belief network. Finally, the corresponding alarm is given by the identification results. The algorithm schematic diagram is shown in Figure 2.

III. EXPERIMENTAL RESULTS

A. *Database*

The following simulations were done on a PC Processor Intel(R) Core (TM) i5-3337U CPU @ 1.80GHz, 1801 MHz, RAM 4GB. Some of the used frames extracted from videos in our database are presented in Figure 3.

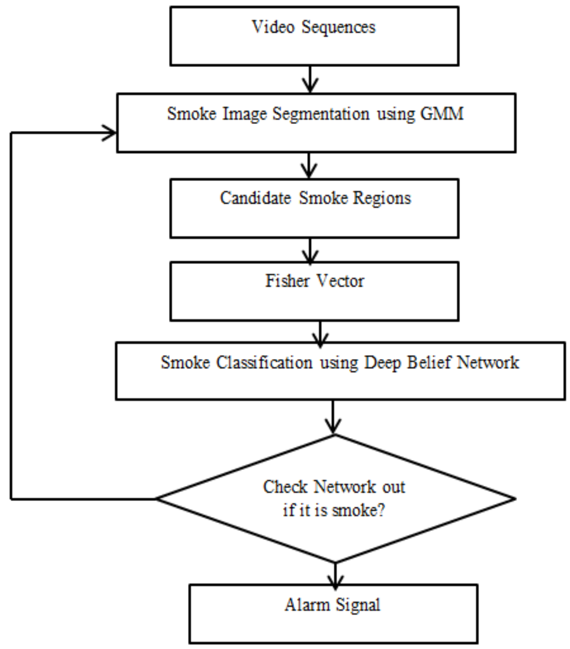


Figure 2. Algorithm Shematic Diagram

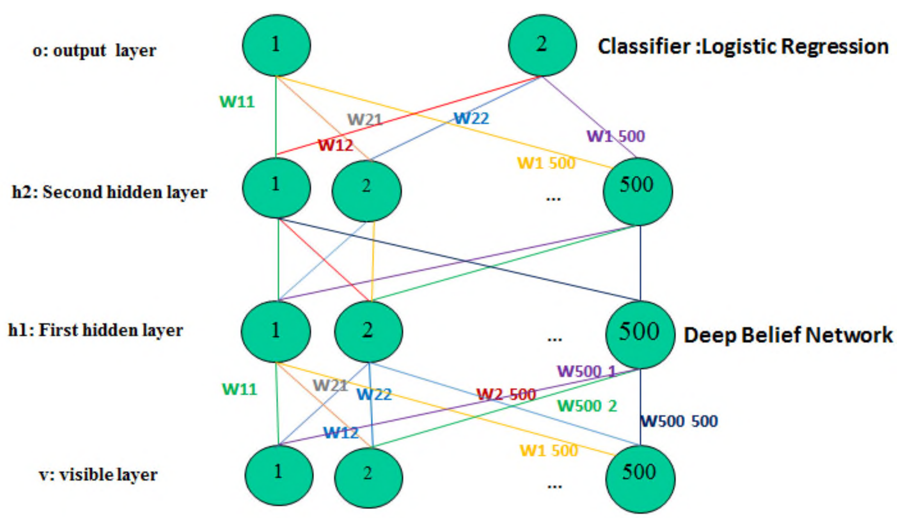


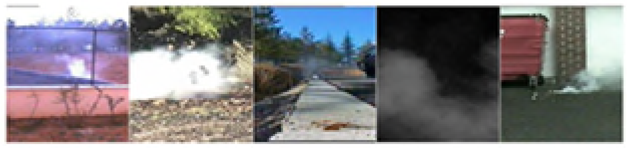
Figure 1. Deep Belief Network Architecture

 ISBN: 978-1-61208-716-0

B. Discussion

First of all, we have extracted frames from smoke-based videos. The size of each frame of the videos is set to 320 × 240 pixels. The main idea of our work is to extract the candidate regions related to the smoke movement which will inserted in a vector of characteristics.

The use of criterion of Fisher allowed us to keep the most important values in this feature vector to make the classification then by the deep belief network. As we mentioned before, we used the GMM [1] as a technique to detect the motion of smoke that is considered a chaotic movement. The results are shown in the Figure 4.



(a)

(b)

(c)

Figure 3. (a) Frames containing smoke extracted from our Database without noise, (b) Frames containing smoke with noise (fog, moving people, etc.), (c) Frames without smoke

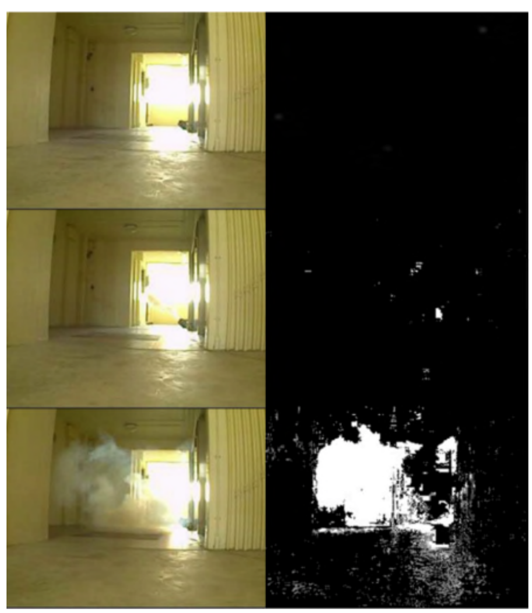


Figure 4. Background Substraction using GMM [1]

 ISBN: 978-1-61208-716-0

At this stage, we have extracted frames from smoke-based videos. The size of each frame of the videos is set to 320 × 240 pixels. The fixed parameters of the deep belief network are defined as:
Number of hidden layers=2
Number of epochs=400
Learning rate = 0.6

After fixing the features that we classified using deep belief network, we changed the classifier to SVM [7]. The comparison between the proposed method and the method using SVM is based on these criteria:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1score = 2.\frac{Precision.Recall}{Precision+Recall} \quad (7)$$

where TP: True Positive smoke Frames
TN: True Negative Frames
FP: False Positive Frames
FN: False Negative Frames

In this section, the frames containing noise such as moving people and fog give us an idea about the robustness of the proposed method. We notice a slight modification in the different criterion of comparison. These results are exhibited in the Table I.

This table shows that the developed method presents in the absence of the noise the best value of accuracy and Recall compared to the classification with SVM [7]. The presence of the noise affects slightly the accuracy and the precision. These values decrease slowly with noise. Moreover, the application of our proposed method helps us to find the smoke and the no smoke regions, as shown in Figure 5.

TABLE I. COMPARISON BETWWEN THE PROPOSED METHOD USING DBN AND THE METHOD USING SVM CLASSIFIER FOR SMOKE DETECTION

Condition	Classifier	Accuracy	F1 score	Precision	Recall
-Noise	SVM	93.43	0.952	1	0.912
	DBN	94.86	0.962	1	0.928
+Noise	DBN	92.57	0.936	0.88	1
	SVM	91.54	0.92	0.86	1

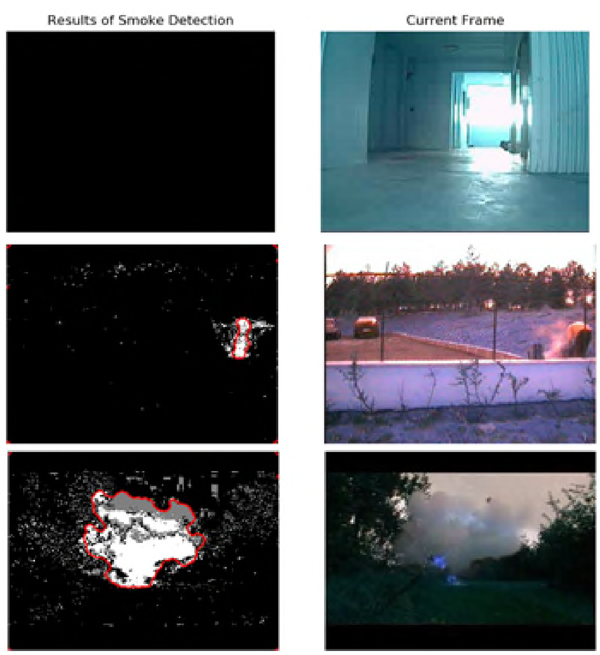


Figure 5. Smoke Detection Results

IV. Conclusion and Perspectives

In this paper, we proposed a novel approach for smoke detection using Deep Belief Network. We extracted the candidate regions by using GMM [1].

For a better classification, we used the Fisher vector which gives us maximum separation between the means of different classes and keeps just the most important values. Then, the extracted feature vector was fed into deep belief network to calculate many criteria such as accuracy, F1 score, Precision and Recall. The robustness of this method is tested by adding noise. Finally, to evaluate the noise influence, we tested our method on noisy data. These promising results provide clear evidence about the capability of such a network in recognizing smokes in our recorded frames easily. Train the developed method with a significant amount of data will undoubtedly provide more promising results and would help in accelerating the next generation of surveillance and real-time monitoring systems. As perspectives to our work, the future step is to combine the deep belief network to FasterR-CNN to classify and localize simultaneously smoke and no smoke regions.

Acknowledgment

This work was supported by the Cooperation Project: PHC Utique 41755XB (CMCU 19G1126) and the CARTT-IUT: University of Toulon.

References

[1] J. Yang, X. Yuan, X. Liao, P. Liul, G. Sapiro, and L. Cari, “Video Compressive Sensing Using Gaussian Mixture Models,” IEEE Transactions on Image Processing, 2014, pp. 4863-4878.

[2] Y. LeCun, Y. Bengio, and G .Hinton, “Deep Learning,” International Journal of Science Nature, 2015, pp. 436 - 444.

[3] G. Hinton, S. Osindero, and Y. Whye, “A fast learning algorithm for deep belief nets,” Neural Computation, Toronto, 2006, pp.1527-1554.

[4] Y. Hu and X. Lu, “Real-time video fire smoke detection by utilizing spatial-temporal ConvNet features,” Multimedia Tools and Applications, 2018, pp. 29283 – 29301.

[5] Z. Yin, B. Wan, F. Yuan, X. Xia, and J. Shi, “A Deep Normalization and Convolutional Neural Network for Image Smoke Detection,” IEEE Transactions on Digital Object Identifier, 2017, pp. 18429-18438.

[6] J. Chen, Y. He, and J. Wang, “Multi-feature fusion based fast video flame detection”, Build. Environ. 45 (May 2010) pp. 1113–1122.

[7] C. Hsu and C. Jen Lin, “A Comparison of Methods for Multiclass Support Vector Machines”, IEEE Transactions on neural networks, Taiwan 2002 , pp. 1–26.

[8] U. Toreyin, Y. Dedeoglu, and E. Cetin, “Contour Based Smoke Detection in Video Using Wavelets”, 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, 2006, pp. 1-5.

[9] S. Mikat, G. fitscht, J. Weston, B. Scholkopft, and K. Mullert, “Fisher discriminant analysis with kernels”, Neural Networks for Signal Processing IX, 1999, pp. 41 - 48.

[10] N. Zhangab, S. Dingab, J. Zhangab, and Y. Xuec, “An overview on Restricted Boltzmann Machines”, Neurocomputing, 2018, pp. 1186-1199.

[11] A. Khotanzad and C. Chung, “Application of Multi-Layer Perceptron Neural Networks to Vision Problems”, Neural Computing and Applications, London, 1998, 9, pp. 249 - 259.

[12] C. Chang and R. Chao, “Application of back-propagation networks in debris flow prediction”, Engineering Geology, 2006, pp. 270–280.

 ISBN: 978-1-61208-716-0