Investigation of Machine Learning Algorithm Compared to Fuzzy Logic in Wild Fire Smoke Detection Applications

Azarm Nowzad, Andreas Jock, Klaus Jäckel
IQ wireless GmbH, Berlin, Germany

Azarm Nowzad, Ralf Reulke
Department of computer science, Humboldt-Universität zu Berlin, Berlin, Germany

Abstract

An automatic smoke detection algorithm for real-time detection of wild fires through detection of smoke has been presented in our previous works dealing with forest fires. Implementing a fuzzy logic decision process, the algorithm is based on a chromatic analysis enhanced in the HSV (Hue-Saturation-Value) color mode and dynamic features of smoke. In this work, the quantitative analysis of the mentioned algorithm is evaluated on a real-scene test dataset using a confusion matrix.

Considering the scene complexity of analyzed data, the main concern is, if a well-defined empirical fuzzy membership and rule-based function is sufficient to adapt the system to different environmental conditions, or if additional automatic learning methods can be implemented to reduce the human-assisted training phase.

In this paper we investigate the state-of-the-art of machine learning algorithm used in ground-based image analysis in terrestrial surveillance systems and examine the costs and benefits of their use in the scenario of smoke detection in forest environments.

Keywords: Forest fire detection, Deep machine learning, fuzzy-based smoke detection, CNN in smoke detection

Introduction

Hand-crafted fuzzy-based systems are based on human observations since they need a large quantity of labeled data as training samples. With a proper human-assisted training phase, it is usually possible to obtain a good reliability from the algorithm for a specific location and under a stable condition.

Considering forest images, scene-complexity is a challenging problem. Vegetation type, cloud shadows, changing weather conditions and

shaking cameras are some of the problems in real scene images. The environmental condition of the scene is also correlated to the data site.

The fuzzy-based forest fire detection system is a rule-based function with defined feature extraction blocks. Despite satisfactory results, the system is not able to revise and learn from failures.

To supplement a fuzzy based algorithm on learning complex models, machine learning algorithms can provide techniques such as artificial neural networks (ANN) to approach models for object representation and image understanding. By means of a supervised learning process and backpropagation method, the system is able to build automatic feature extraction blocks and learn how to deal with complex scenarios. The scope of this paper is to explore the application of ANN and compare it to the fuzzy-based smoke detection algorithm on natural scene data, collected from two different forest sites in Germany.

Evaluation of the Fuzzy-based algorithm

In [1], a fuzzy-based smoke detection system is proposed. The scheme of the algorithm is shown in Fig. 1.
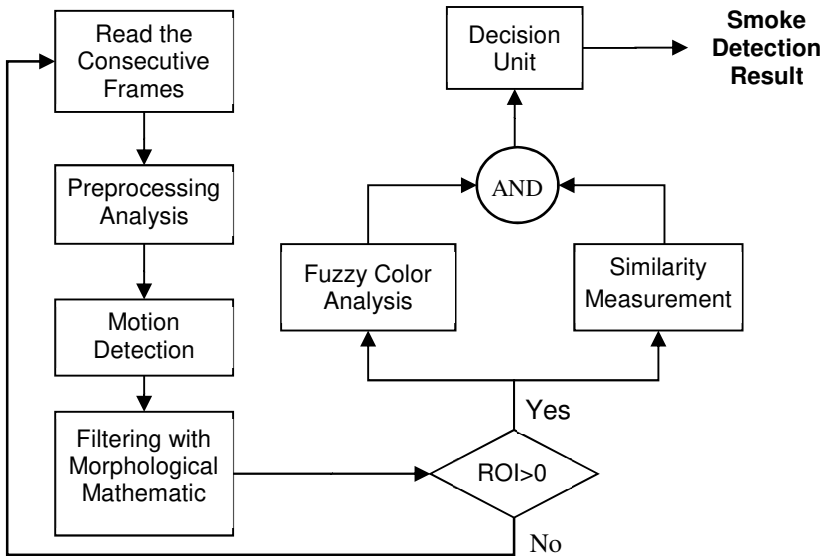

Fig. 1.   Flow chart of the detection scheme [1].

By analyzing the temporal and spatial characteristic of smoke incidences, manual feature extraction blocks define Region Of Interests (ROIs) in the processed data. A rule-base fuzzy logic algorithm then evaluates the ROIs to classify them into smoke and non-smoke events.

A qualitative performance of the proposed algorithm was evaluated on a large number of test data from a tower sited in a German forest. In this

paper, we perform a quantitative analysis of the algorithm using data from two different sites in Germany. To evaluate the classifier, a confusion matrix with four indices is built and a Receiver Operating Characteristic (ROC) plot is derived from this matrix.

Smoke events up to 10 Kilometer away from the camera are considered in the data set. Each image sequence in the data consists of 8 frames taken one second apart.

2413 image sequences are collected from the first site in Halbe (Brandenburg, Germany). 178 sequences are with positive instances. The evaluation of the classification is presented in Table 1.

Table 1. Confusion matrix on smoke classification model "Halbe".

	Smoke (actual)	Non-smoke (actual)
Smoke(predicted)	150	15
Non-smoke(predicted)	28	2220
Overall Accuracy	98,2 %	

The second data set, collected from Kolberg (Brandenburg, Germany), is not very interesting for the existence of smoke incidents. However, the scene characteristic depicts complexity from the point of view of negative examples. 1465 image sequence with 2 relevant smoke regions are analyzed and the classification result is shown in Table 2.

Table 2. Confusion matrix on smoke classification model "Kolberg".

	Smoke (actual)	Non-smoke (actual)
Smoke(predicted)	2	282
Non-smoke(predicted)	0	1181
Overall Accuracy	80,75%	

252 out of 282 False Positive alerts in "Kolberg" are caused by water surface and moving boats.

Strong exposure variations between frame sequences, atmospheric effects, cloud shadow and weather conditions are other causes of false detections. Average overall accuracy of the whole data set is around 84, 5%, which is a remarkable performance with a low number of false negatives alarms (undetected smoke) and a limited number of false positive alarms.

Despite satisfactory results, in order to deal with complex scenarios, an expert operator needs to define manual features for classification purposes. Due to the scene complexity, the variety in vegetation type and tower location, features can be extremely complex to define in some cases. In addition, the idea of back propagation and learning from the positive and negative incidences, would enhance the robustness of the detection algorithm.

Convolutional Neural Network (CNN) and object recognition

The vision algorithm pipeline consists of image pre-processing, regions of interest (ROI) detection, object recognition, and decision making. Object detection is the most challenging part of the pipeline.

Feature extraction and the choice of a proper machine learning algorithm are critical steps in object detection and classification tasks. Computer vision features such as edge detection, blob detection or pixel grouping techniques, such as image segmentation, are applied as information extraction tools of image analysis. In general, choosing the right machine learning algorithm is highly related to the data type and the problem one deals with. Smoke detection applications focus on the localization of an object in the image. The object size and location can vary in each image.

In image classification and object recognition challenges, CNN, inspired by visual cortex, are widely used. The performance of the algorithm in form of Correct Detection Rate (CDR) has achieved outstanding results. As mentioned in [2], a CDR of 99, 77% has been achieved on MNIST handwriting benchmark and, a CDR of 97.47% with the NORB dataset of 3D objects. Also on traffic sign recognition benchmark, CNN outperforms humans by a factor of two.

Four main basic building block of CNN are convolution, nonlinearity function, sub sampling and classification with fully connected network. Fig. 2 shows a simple architecture of CNN.
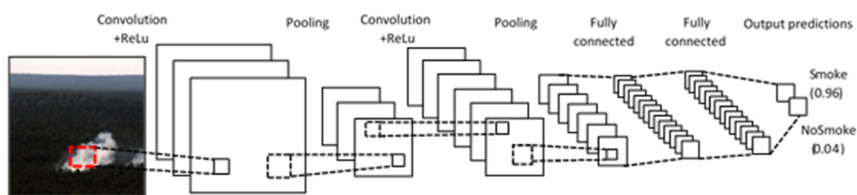

Fig. 2.   A simple architecture of CNN.

The convolutional layers in the early stage try to extract elementary visual features of the image such as edges and corners. Deeper layers are for the higher level and complicated features. Since most of the real-world data are non-linear, an activation function such as Rectified Linear Unit (ReLU) is applied after convolution operation.

Feature maps are then subsampled (pooling layer) to reduce the dimensionality and computational cost. Pooling layer make the network robust against small distortions and transformations in the input.

The output of the *(i,j)*-th entry of the *k*-th feature map in the first convolutional layer is given by

$$(x_{c1})_{i,j}^k = S(\text{ReLU}(\,(W_1^k * \tilde{x})_{i,j} + (b_1^k)_{i,j}\,)), \qquad \text{(Eq. 1)}$$

where W_1^k and b_1^k are the filter and bias values referred to the *k*-th feature map of the first layer. Finally the fully connected layers classify the input based on the high-level features from previous layers.

R-CNN (Regions with CNN) and model architecture

The type of methodology in extracting sub window from the input image varies in different approaches. Region extraction for input patches can be done by sliding a sub-window over the entire image. Each image patch passes through the CNN for further processing. This tedious process is time consuming and complex.

The scene characteristic of the data is a key point to be considered in defining negative regions. Our case study is dealing with natural forest scenes. Extracting patches randomly can lead to poor negative examples. Therefore, a smart search is required to have a variety of diverse examples in the region proposed.

A weak classifier can be used as a pre selection step to perform an exhaustive search to extract all possible candidates and reduce computational costs. R_CNN uses the method of external region proposal in its pipeline. The predicted regions are compared with labeled ground truth images to find out if the Intersection over Union (IOU) is bigger than a predefined threshold.

Dealing with image sequence in our dataset, temporal features are extracted and evaluated beside the spatial information. Based on the data characteristic, a big number of false positives are generated due to the moving objects. The potential movement in the image is then extracted as a preprocessing step to the CNN.

The temporal motion information is collected into a single image using Motion History image (MHI), as a temporal motion volume. The history of temporal changes are kept at each pixel location. Fig. 3 shows the MHI process in a four frame image sequence.

Applying a fast region proposal to CNN, reduce the number of ROIs to be processed. The advantage of using temporal information can lead to a better performance of CNN.
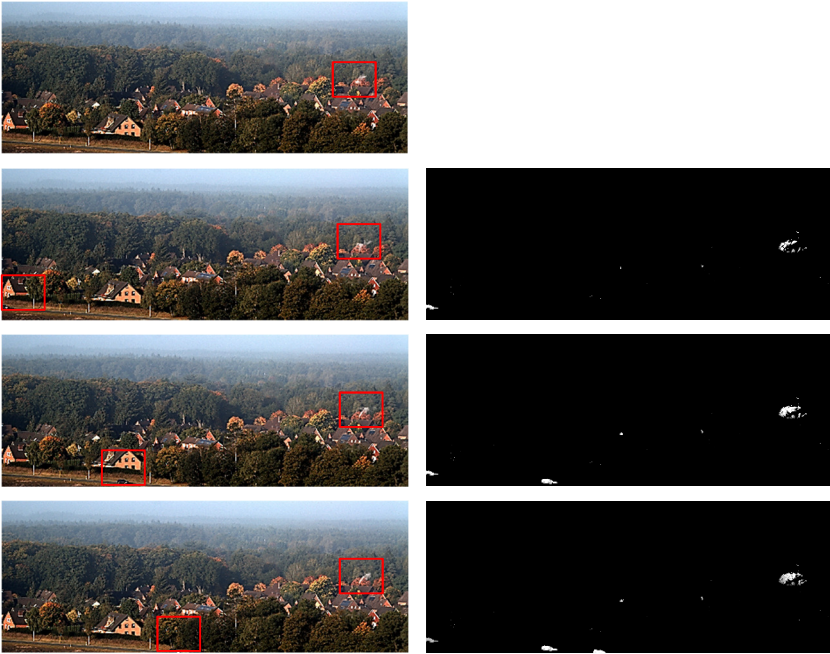
Fig. 3. MHI process, left column four image sequence, right column MHI results.

Next steps

The hyper parameters and the network architecture has been adjusted and need to be finalized for the analysis over the test data. The labeled ground truth data is generated over the same dataset as the fuzzy logic algorithm for the purpose of comparison.

The finalized outcome needs to be evaluated and demonstrated in a ROC plot with fuzzy algorithm final results.

References

[1] A. Nowzad, A. Jock, U. Krane, K. Jäckel, H. Vogel, *Development of an Automatic Smoke Detection Algorithm using Color Images and a Fuzzy Logic Approach for Real-Time Forest Fire Detection Applications*, AUBE '14, Duisburg, Germany, October 2009.

[2] S. Hijazi, R. Kumar, C. Rowen, *Using Convolutional Neural Networks for Image Recognition*, www.cadence.com.

[3] S. Pierre, Y. LeCun. *Traffic Sign Recognition with Multi Scale Networks*. Courant Institute of Mathematical Sciences, New York University. 2011.

[4] S.P. Hohberg, *Wildfire Smoke Detection using Convolutional Neural Networks*, Master thesis, Freie Universität Berlin Fachbereich Mathematik und Informatik Berlin, Germany, September 2015.

[5] Ujjwalkarn, *An Intuitive Explanation of Convolutional Neural Networks*, the data science blog, retrieved on : June 2017, retrieved from: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/, August 2016.

[6] K. Sinhal, *Object Detection using Deep Learning for advanced users*, retrieved on June 2017, retrieved from: https://medium.com/ilenze-com/object-detection-using-deep-learning-for-advanced-users-part-1-183bbbb08b19, iLenze blog, September 2016.

[7] A. Mousavi, R.G. Braniuk, *Learning to invert: Signal Recovery Via Deep Convolutional Networks*, Department of Electrical and Computer Engineering Rice University Houston, Texas, January 2017.

[8] S. Frizzi, R. Kaabi, M. Bouchouicha, J.M. Ginoux, E. Moreau, F. Fnaiech, Convolutional Neural Network for Video Fire and Smoke Detection, IECON 2016 - 42nd Annual Conference of the IEEE, October 2016.