Classification of leakage detections acquired by airborne thermography of district heating networks

Amanda Berg and Jörgen Ahlberg

Amanda Berg and Jörgen Ahlberg, Classification of leakage detections acquired by airborne thermography of district heating networks, 2014, Pattern Recognition in Remote Sensing.
http://dx.doi.org/10.1109/PRRS.2014.6914288

TABLE I
NUMBER OF SAMPLES FOR EACH LAYER AND CLASS

Layer no.	1	2	3	4	5	6
Threshold	0.05%	0.1%	0.5%	1%	3%	5%
Media/Energy	34	39	71	89	99	80
False	71	75	148	237	294	348

The thermal camera is a cooled mid-wave infrared FLIR SC7000 Titanium with a resolution of 640 512 pixels and a field of view of 11°. At an altitude of 800 m, this yields a pixel footprint of 25 25 cm.

In order to minimize the number of false detections, data collection is mainly done during the night or at dawn during spring or autumn. At this time, neither vegetation or snow is blocking the view, the effect from sun heating is minimal and the streets are not covered with cars blocking the view [9].

B. GIS data

The network owner provides pipe location information in the form of vector maps. This information is projected on top of the georectified images creating a rasterized pipe mask for each image. The mask is then used to limit the search for unnaturally high temperatures to areas above the pipes only.

C. Detections

A detection is in this context an area with a certain shape and location pointed out as abnormally warm. That is, it is an extended object, not just a coordinate. In order to extract the detections from the images, we use the anomaly based method described in [1]. Statistics of the ground temperature inside the pipe mask are calculated from all images within one flight and the most deviating pixels above certain thresholds in the high end of the distribution (i.e. the ”warmest pixels”) are marked as detections. The percentage thresholds are; 0.05%, 0.1%, 0.5%, 1%, 3% and 5%, resulting in six different layers of detections.

D. Ground truth data

Acquisition campaigns during the last couple of years have resulted in thousands of thermal images from 17 Scandinavian towns and cities. Three of the most recent acquisitions were selected for this study. The selection was based on the fact that the customers for these flights could provide ground truth, i.e., information about which detections had been investigated further and proven to be real (or false) leakages.

Detections from these cities have been manually labeled as media leakages, energy leakages, or false detections. To do this, interviews with the network owners have been conducted several months after the detection results have been delivered, that is, when the true status of the pipes in many cases have been investigated (by digging). Also, many detections can be pointed out as false by the network owners or by ourselves.

In total, we have 1585 labeled detections, of which 1173 are false. The number of samples for each class and layer can be seen in Table I.

Fig. 1. The proposed approach.

Fig. 2. An example of a building mask (right) generated from the colour raster map OpenStreetMap (left).

III. METHOD

Our approach is illustrated in Figure 1. The green boxes are identical to the scheme by Friman et al., while the blue ones are added or modified. Below we will describe the proposed building segmentation scheme, and then the feature extraction and classification.

A. Removing false detections using building segmentation

A common source for false alarms are detections of objects, e.g., chimneys and atriums, at rooftops with unnaturally high temperatures. These false alarms appear because the pipes sometimes pass beneath buildings. Since we know that real leakages of the district heating network never can appear at rooftops, information on building locations can be used to remove false detections. Friman et al. implemented a building segmentation scheme based on the watershed transform and AdaBoost classification in order to automatically extract building information from the thermal imagery.

The proposed building segmentation method is based on OpenStreetMap[1]. A binary building mask is generated using a simple color segmentation scheme, since buildings in these images have a specific set of colors; an example can be seen in Figure 2. By thresholding the three RGB channels, each pixel is classified as belonging to a building or not. The thermal images as well as the OpenStreetMap-images are stored in GeoTIFF format with world coordinate information related to each pixel, information that is used for image registration.

B. Removing false detections using a classifier

The basic assumption for our approach is that distinguishing features of the different types of detections do exist. Such features could be collected from the imagery or from the detections themselves (shape descriptors). The labeled examples should be studied to find an initial set of discriminating features which is later refined. The extracted features are then used to evaluate a number of classifiers. Below, we describe

[1]http://openstreetmap.org

the selection of features and classifiers to use in the final system.

C. Features

Features were found by studying the labeled samples. The initial selection consisted of 18 scalar features based on (thermal) intensity distribution within the detection, shape and propagation of the detection, and proximity information. Feature selection, described further in section IV-B, was performed using the Mahalanobis distance.

D. Classifiers

Two linear and three nonlinear classifiers were chosen for evaluation. The linear ones are Linear Discriminant Analysis (LDA) [10] and Linear Support Vector Machines [11]. These were mainly included in the experiments to exclude linear methods from further evaluation. Three different nonlinear classifiers were evaluated; the Radial Basis Function Support Vector Machine [11], AdaBoost [12], and Random forest [13]. We used the implementations from [15].

For each nonlinear classifier, the choice of parameters was evaluated and chosen iteratively starting at the default value and moving in the direction of steepest descent of false positive rate until convergence. The chosen parameters and the default values can be seen below (chosen/default value).

- RBF-SVM: σ = 140 / 1
- Random forest: No. of decision trees: 120 / 50
- AdaBoost: No. of weak classifiers: 192 / 100

Apart from these parameters, the weak classifier used in AdaBoost was the decision stump and the number of random features used for node splitting in Random forest was one.

E. Evaluation and selection methodology

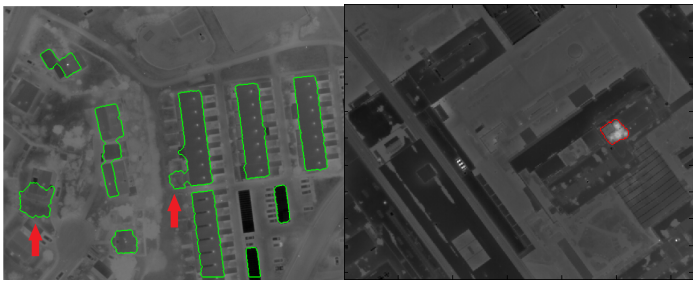
There are three main types of detections; media, energy and false detections. Since both media and energy detections are interesting for the network owner, we chose to combine the media and energy samples into one class, hereafter called true detections. Furthermore, incorrect classification of one of these classes as the other one is not as critical as incorrect classification between media/false and energy/false.

As mentioned before, the cost for classifying a true detection as a false one is much higher than classifying a false detection as a true one. Therefore, the true positive rate was set to a specific value, namely 99%, and the error measurement used for the evaluation was the false positive rate.

IV. Experimental results

A. Building segmentation

Two building segmentation schemes were evaluated. The scheme proposed by Friman et al. [1] using the thermal imagery, and the scheme using OpenStreetMap. It was observed that both schemes sometimes result in errors, but the errors were of different character. The method in [1] sometimes suffer from non-building areas being classified as buildings, which could potentially make true detections (i.e., real leakages) be discarded. In contrast, the OpenStreetMap-based scheme sometimes suffer from missing buildings, which might result in false detections classified as true detections. As mentioned, the cost for the latter is much lower, thus we selected the OpenStreetMap-based scheme. Examples of the performance of both methods can be seen in Figure 3.



(a) (b)

Fig. 3. (a) shows an example of the performance of the building segmentation method proposed by Friman et al. [1]. The green boundaries indicate the areas that have been classified as buildings. The arrows show the problem areas. Unacceptable classification of ground as buildings which could lead to missed real leakages. (b) shows instead an example of a successfully classified false detection, indicated with red boundaries, when the OpenStreetMap-based scheme is used.

Using the OpenStreetMap-based scheme described, removing all detections which lie 100% on top of a building, 19% of the false detections in the data set could be removed without removing any of the energy or media samples.

B. Feature selection

Feature selection was done by calculating the Mahalanobis distance between class means for each feature. As the Mahalanobis distance increases and approaches infinity, the probability of error decreases. Therefore, the Mahalanobis distance is a common choice for measuring the ”goodness” of a feature [14].

The final set of features are the eight features for which the ground truth samples had the largest average Mahalanobis distance. The features are listed and described in Table II.

C. Classifier selection

Each classifier was evaluated using 10-fold cross-validation and the 8 features described above. The classifiers were trained and tested on each layer individually.

The averaged results over all layers are shown, for each individual classifier, in Figure 4. Apparently, the Random forest classifier outperforms the others. Combining the results from the building based rejection and the classification, the weighted (with respect to number of false samples in each layer) averaged false positive rate across layers is 58%.

D. Classification and detection layers

As mentioned, the detection method results in several layers of detections, where the first layer contains the ”worst” detections only (the ones with the most anomalous temperatures). In the previous sections, training and classification were done separately for each layer. We have investigated whether these

TABLE II
FEATURES USED FOR CLASSIFICATION

Feature	Description
Median intensity	Median intensity within the detection.
Standard deviation	Standard deviation of the intensity within the detection.
Coverage	Ratio of detection area inside heat pipe mask.
Elongatedness	$\frac{\text{area}}{4d^2}$, where d is the number of erosions needed to make the detection disappear.
Concentricity	Measurement of how central the maximum intensity value is within the detection.
Connected components	Number of other detections which lie within a certain radius from the detection.
Border average	Mean intensity within an area around the detection.
Distance to building	Distance from maximum intensity value to the wall of the closest building.

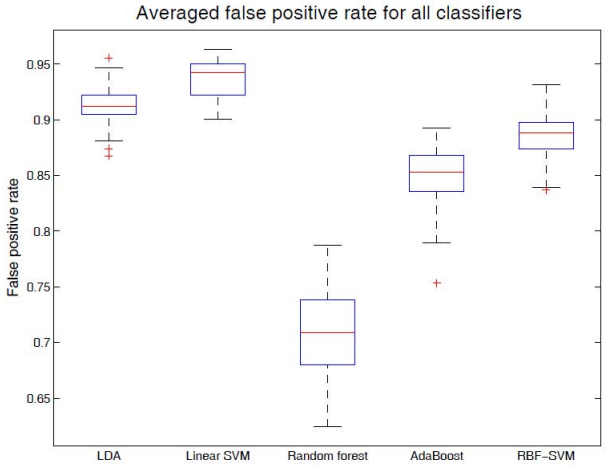


Fig. 4. Averaged false positive rate for all evaluated classifiers and thresholds. The red line shows the median, the blue box the 25th and 75th percentile and the whiskers mark the most extreme samples not considered as outliers. Outliers are plotted as red crosses.

classifiers can be combined, since (i) if a detection is present in several layers, adding information from the other layers might improve robustness, and, (ii) the differences in classification output from the different layers indicate the robustness of the classifiers.

1) Voting: Since detections corresponding to the same area often are present in several layers, the possibility of improving the result by voting between layers was investigated. However, in all the available data, no occasion of dissimilar labeling by different threshold classifiers was observed. Thus, the classifier is apparently robust to changing temperature thresholds.

2) Layer invariant classification: For the evaluation of layer invariant classification, all the 1585 detections from the six layers were combined to form one large data set.

Since the Random forest classifier have shown to achieve the best general performance for all percentage thresholds, it was chosen to be the classifier also for the threshold invariant classification. Combining the results from the building based rejection and the classification, the false positive rate was 42%. Thus, to our surprise, this combined classifier gives better performance (on average) than a set of specialized classifiers. However, the results for the set of specialized classifiers is somewhat unreliable for layer 1 and 2 due to the small number of samples.

V. CONCLUSION

We have presented an improvement to a previously published method for finding leakages in district heating networks by classifying its resulting detections using trained classifiers. We have evaluated various features and five different classifiers. The classifier that generally produced the smallest false positive rate while maintaining a true positive rate of 99% is a Random forest classifier with 120 trees, an average tree depth of 10 and splitting at nodes based on one randomly selected feature. The classification of different layers of detections have been investigated, and the results indicate that the classification is consistent over layers.

We have also shown that a pre-processing step for building extraction using publicly available GIS-data is preferable compared to the previously used method, not due its superior performance in general, but due to the different type of errors produced by the two methods.

In the end, we were able to achieve a false positive rate of 42%, that is, we can discard 58% of the detections given by the previous system, thus significantly enhancing the usability.

REFERENCES

[1] O. Friman et al., "Methods for large-scale monitoring of district heating systems using airborne thermography," *IEEE Trans. Geoscience and Remote Sensing*, 52(8):5175-5182, 2014.

[2] A. Poredos and A. Kitanonovski, "District heating and cooling for efficient energy supply," in *Proc. Int. Conf. Electrical and Control Engineering*, 2011, pp. 5238–5241.

[3] M. Olsson, "Long-term thermal performance of polyurethane-insulated district heating pipes," Ph.D. thesis, Chalmers Univ. of Techn., 2001.

[4] M. Fröling, "Environmental and thermal performance of district heating pipes," Ph.D. thesis, Chalmers Univ. of Techn., 2002.

[5] S.-A. Ljungberg, "Aerial thermography - a tool for detecting heat losses and defective insulation in building attics and distric heating networks," in *Proc. SPIE Thermosense IX*, 1987, pp. 257–265.

[6] S. R. J. Axelsson, "Thermal modeling for the estimation of energy losses from municipal heating networks using infrared thermography," *IEEE Trans. Geoscience and Remote Sensing*, 26(5):686–692, 1988.

[7] B. Bøhm and M. Borgström, "A comparison of different methods for in-situ determination of heat losses from district heating pipes," Dept. of Energy Engineering, Technical Univ. of Denmark, 1996.

[8] H. Zinko et al., "Quantitative heat loss determination by means of infrared thermography – the TX model," Int. Energy Agency, 1996.

[9] S. Sjökvist at al., "Kvantifiering av värmeläckage genom flygburen IR-teknik — en förstudie," Fjärrsyn, Rep. 2012:17, 2012.

[10] R. O. Duda et al., *Pattern Classification*, John Wiley & Sons, 2001.

[11] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*, 2nd ed. Springer, 2008.

[12] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[13] L. Breiman, "Random forests," *Machine learning*, 45(1):5–32, 2001.

[14] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(2):153–158, 1997.

[15] R. P. W. Duin et al., "PR-Tools4.1: A Matlab Toolbox for Pattern Recognition," *Delft University of Technology*, 2007, http://www.prtools.org.