Sinkhole Detection by Deep Learning and Data Association

Nam Vu Hoai
Department of Computer Science
Faculty of Information Technology, PTIT
Ha Noi, Viet Nam
namvh@ptit.edu.vn

Nguyen Manh Dung
Deprtment of Information and Communication
Kongju University
Cheonan City, South Korea
Manhdungbk@gmail.com

Soonghwan Ro
Deprtment of Information and Communication
Kongju University
Cheonan City, South Korea
rosh@kongju.ac.kr

Abstract— Accurate tracking of the sinkholes that are appearing frequently now is an important method of protecting human and property damage. Although many sinkhole detection systems have been proposed, it is still far from completely solved especially in-depth area. Furthermore, detection of sinkhole algorithms experienced the problem of unstable result that makes the system difficult to fire a warning in real-time. In this paper, we proposed a method of sinkhole tracking that takes advantage of the recent development of CNN transfer learning. Our system consists of three main parts which are binary segmentation, sinkhole classification, and sinkhole tracking. The experiment results show that the sinkhole can be tracked in real-time on the dataset. These achievements have proven that the proposed system is able to apply to the practical application.

Keywords—sinkhole detection; deep learning; sinkhole tracking; HA algorithm; Otsu algorithm

I. INTRODUCTION

A sinkhole is defined as a depression or hole in the ground caused by some form of the collapse of the surface layer. The human-induced sinkhole has been increasing rapidly with the high number of construction activity in urban areas. This kind of sinkhole is the main source of many disasters which need to be detected to protect loss of human life and infrastructure. Therefore, early warning of the sinkhole is an emergent problem for the government and social organization to improve the quality of human life where sinkhole occurs frequently, especially in the resident area.

Recently, thanks to the availability of large-scale training data and the advance of high-performance GPUs, various deep learning-based methods have been proposed to significantly improve the state-of-the-art of category-specific object detection which is similar to the problem of sinkhole detection. Faster RCNN proposed by S. Ren [2] can achieve high accuracy. However the processing time is not fast enough to be implemented in real-time system. In order to obtain real-time performance, YOLO (You Only Look Once) framework [3] formulates the region proposal as a regression problem. A unique feature of YOLO is that it unifies the separate components of object detection into single convolutional network that simultaneously predicts multiple bounding boxes (regions) and class probabilities for those boxes. The accuracy of YOLO framework is still affected by small object which appear in the processed image. Therefore, it is difficult to apply Faster RCNN and YOLO to the sinkhole detection problem. The first requirement of sinkhole detection problem is real-time performance. In addition, the sinkhole images are captured at different distances, therefore, some sinkholes may appear in small shape.

Unlike other mentioned methods of detecting potential sinkholes, thermal imagery uses the fact that the surface temperature of inside sinkhole area is far different from that of the surrounding area. The paper [4] use a drone mounted with Far-Infrared (FIR) camera to monitor a large area. They observed the optimal time for recording the sinkhole videos by observing the temperature difference over time. From the recorded videos, candidate of sinkhole areas are segmented using adaptive binarization method. Then a convolutional neural network (CNN) [5] and a Boosted Random Forest (BRF) are stacked to classify potential sinkhole into real one. Although the method of detecting sinkhole by FIR camera can give early warning message about the appearing of sinkhole, the result is not stable due to the noise which may be contained in the output image from FIR camera. In addition, the position of sinkhole cannot be tracked to fire correct alarm. In this paper, we introduce automatic sinkhole tracking system to support an early detection and warning of sinkhole using dataset of [4].

The dataset is collected by using FIR camera mounted on drone. Our system consists of three main parts. Firstly, candidate sinkhole areas are extracted from the binary image, then CNN transfer learning is applied to classify sinkhole candidate into real sinkhole. Finally, a data association method is implemented to do sinkhole tracking job. The remaining of this paper is organized as follows. Section 2 describes the proposed system. Section 3 shows some experimental results. Future work and discussion are given in Section 4.

II. PROPOSED METHOD

Our proposed method consists of three main components: candidate sinkhole segmentation, sinkhole classification, and sinkhole tracking, which is described in Fig. 1. The first component in our system is sinkhole segmentation which is a binary segmentation based on the temperature characteristic of sinkhole. This component receives an infrared image and outputs candidate sinkhole location in given video frame. The binary segmentation has been done by the fact that the temperature of sinkhole is much lower than the temperature of surrounding areas. The second component is a sinkhole classifier which uses transfer learning technique to classify the candidate sinkhole into the real sinkhole. In order to deal with the data association problem in real-time sinkhole tracking, we use Hungarian Algorithm [6]. The collected data from the sinkhole classifier are assigned to the tracklets online based on data distribution in the current frame and previous frames. In addition, the problem that the classification stage may output unstable result can be fixed by our developed direction voting technique.

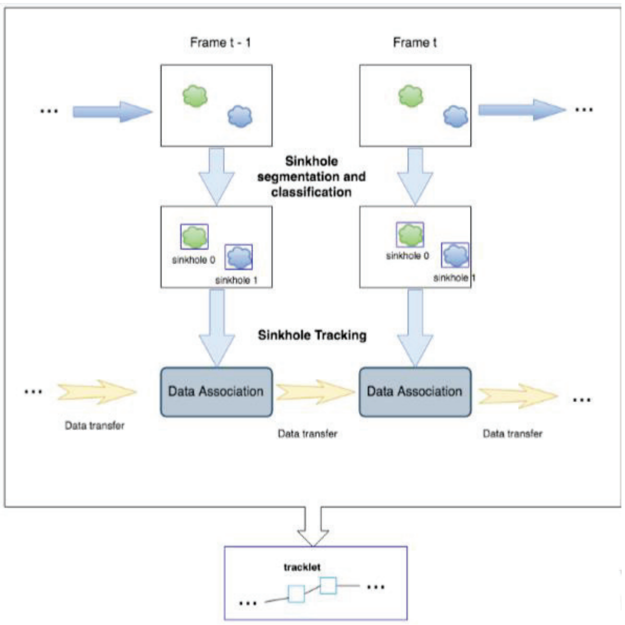

Fig. 1. The flowchart of proposed method

A. Sinkhole Candidate Segmentation

The gray scale image from the infrared camera is the input of this first part. The segmentation algorithm detects the candidate sinkholes by looking for the cold areas in the thermal images. In order to implement this, we apply the dual-thresholding method because the surrounding things may have the grayscale that is much darker than the sinkholes. The process of dual thresholding method is described as follows. I is the input image, and (x, y) is the location of a specific pixel, T_l and T_h are the thresholds.

$$\mathrm{I(x,y)} = \begin{cases} 0 \; if \; I(x,y) \leq T_l \; or \; I(x,y) > T_h \\ 255 \; if \; T_l < I(x,y) \leq T_h \end{cases} \quad (1)$$

The T_l has been found by applying the otsu algorithm [8]. By the experimental result on the dataset, we found that the $T_h = T_l + 20$ give us the best performance.

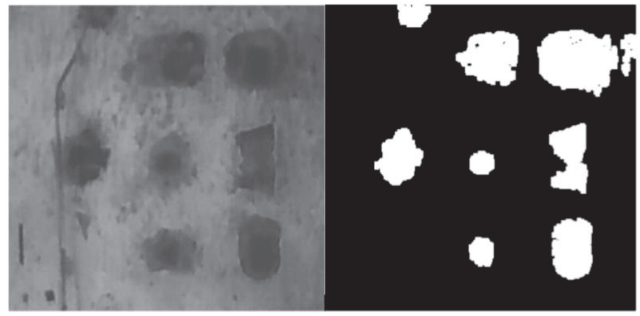
Fig. 2. Binary Segmentation Result

The heuristic filter is implemented to remove noises that appear in the binary image. The final result of sinkhole candidate segmentation is given in Fig. 2.

B. Sinkhole Classification by CNN Transfer Learning

Because of enormous number of parameters, the large dataset and high computation resource are required to train CNN models. That leads to the difficulty when training with lack of dataset. Instead, a pre-trained model can be transferred to work on the categories which not belong to the original dataset. In this paper, the transfer learning approach has been used to train CNN sinkhole classification model.

The ResNet [9] model was trained on the Image Net Dataset [10], which contains millions of images and 1000 object types. In this paper we have chosen ResNet50, which contains 49 convolutional layers, 1 fully connected layer and 1 classification layer that computes for each image its classification score across the 1000 object types. For the purpose of classifying sinkhole regions, we replaced this final layer with a new one which has only two object kinds, sinkhole and non-sinkhole. This layer will be trained from scratch by using back-propagation fine-tune approach with our dataset.

In order to train the CNN for sinkhole image classification in the video frame, we need to prepare one dataset for training and another dataset for evaluation. Our training dataset has 7000 sinkhole images and 7000 non-sinkhole images, while our evaluation dataset has 1000 images for each class. The non-sinkhole images could be regular object in the video frames which have similar surface temperature with real sinkhole like vehicles, humans, or trees; they can also simply be background images. The resulting accuracy of the trained CNN classifier was more than 99%.

C. Sinkhole Tracking by HA Algorithm

After having sinkhole detected and recognized from the CNN classifier, the problem of tracking is to assign sinkhole to respective tracklets which are the trajectory of objects in consecutive frames. In order to solve this problem, we have implemented Hungarian Algorithm (HA) for data association.
Suppose that we have N detected sinkhole in a video frame, the problem is how to identify the tracklet that each detected sinkhole belongs to. The detected sinkhole is assigned bases on

the score that calculated by using correlation between image of detected sinkhole and image of sinkhole in the tracklet. It is obvious that the correlation score is high, then the probability the sinkhole belongs to the respective tracklet is high. The HA method works only on the correlation score matrix which includes all distances between detected sinkholes and tracklets. There are several steps in HA method, which include matrix subtraction, matrix addition and zero element finding. The HA method is implemented to find optimal assignment solution in this paper because of its speed and simplicity to apply to the real time application.

The dataset that we collected is the video captured from drone. Therefore, the sinkhole and other objects are stationary, only the camera was moving, thus the video was not stable. Sometimes, the segmentation and classification step are failed to detect the real sinkhole in the image. In that case, we will use the previous sinkhole location in the previous frame with a small translating distance to place into the current frame. The translating distance will be calculated based on the optical flow algorithm. This translation correction has been done by assuming that all sinkholes in the video frame should be moving in the same direction. This approach will help the system to overcome the problem of missing sinkhole in the detection and recognition process. If in the video frame no sinkhole is detected, then the tracklet will reset, and when waiting for the sinkhole to appear again.

III. EXPERIMENTAL RESULTS

In order to evaluate the proposed sinkhole tracking system, we have been used the collected videos captured from the drone. The input videos are segmented into training set and testing set. To evaluate the performance, we estimated the average detection precision (AP) and average recall (AR) by following the equation (6), (7).

$$AP = \frac{TP}{TP+FP} \tag{6}$$

$$AR = \frac{TP}{TP+FN} \tag{7}$$

Where TP is the number of true positives, FN is the number of false negatives, and FP is the number of false positives in the dataset. Based on the overlapping threshold we can find when a detected sinkhole is FP or TP. a FP is achieved by testing all detected sinkholes of the test data that overlapping by less than the overlapping threshold with a ground-truth sinkhole, and a TP is achieved by testing all detected sinkhole in the test data that overlapping by more than the overlapping threshold with a ground-truth sinkhole.

TABLE I. ACCURACY EVALUATION

overlapping threshold	AR	AP
0.3	93.2%	94.1%
0.4	92.0%	92.4%
0.5	**88.7%**	**89.0%**
0.6	86.6%	87.4%

The AR and AP of proposed method is given in the table 1. It is obvious that when the overlapping threshold is increased, the AP and AR score of proposed method are decreased. In general, the overlapping threshold = 0.5 is commonly used in the literature to evaluate performance of tracking method.

IV. CONCLUSION

In this paper, we proposed an approach, which is the combination of CNN object classifier and HA data association to efficiently do real-time sinkhole tracking with the hardware configuration. Experimental results show that the proposed method is achieved 89% and 88.7% for AP and AR respectively at overlapping threshold of 0.5. Our proposed method has experienced some weaknesses that when the input video is not stable, the segmentation step is failed to detect to potential sinkhole candidate. In addition, the classifier performance is not good in some cases when the input potential candidates are missing from the training set. In spite of some weaknesses, our proposed method shows promising result which is suitable for practical applications such as surveillance system using infrared camera. Our future plan should overcome challenge of lacking dataset.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2018R1D1A1B07049146).

Sinkhole video used in this paper for learning and testing had been provided by Prof. Byong Chul Ko from Keimyung University.

REFERENCES

[1] T.L. Dobecki, S.B. Upchurch, Geophysical applications to detect sinkholes and ground subsidence, Lead. Edge (2006) 336?341.

[2] S. Ren, K. He, R. Girshick, and J. Sun, ?Faster R-CNN: Towards realtime object detection with region proposal networks,? in Proc. Advances in Neural Information Processing Systems, 2015, pp. 91?99.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ?You only look once: Unified,real-timeobjectdetection,?inProc.IEEEConf.ComputerVision and Pattern Recognition, 2016, pp. 779?788.

[4] Lee, Eun Ju, Sang Young Shin, Byoung Chul Ko, and Chunho Chang. "Early sinkhole detection using a drone-based thermal camera and image processing." Infrared Physics and Technology 78 (2016): 223-232.

[5] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), 541-551.

[6] MunkresJ(1957)Algorithms for assignment and transportation problems. J Soc Ind Appl Math 5(1):32-38

[7] Nguyen Manh Dung and Soonghwan Ro. 2018. Algorithm for Fire Detection using a Camera Surveillance System. In Proceedings of the 2018 International Conference on Image and Graphics Processing (ICIGP 2018). ACM, New York, NY, USA, 38-42

[8] Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." IEEE transactions on systems, man, and cybernetics 9.1 (1979): 62-66.

[9] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[10] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009.