The Korean Institute of Communications and Information Sciences

Available online at www.sciencedirect.com

ScienceDirect

ICT Express 2 (2016) 67–70

www.elsevier.com/locate/icte

A deep-learning-based emergency alert system☆

Byungseok Kang*, Hyunseung Choo

Department of Computer Science and Engineering, Sungkyunkwan University, Suwon, 440-746, Republic of Korea

Received 1 March 2016; received in revised form 26 April 2016; accepted 4 May 2016
Available online 21 May 2016

Abstract

Emergency alert systems serve as a critical link in the chain of crisis communication, and they are essential to minimize loss during emergencies. Acts of terrorism and violence, chemical spills, amber alerts, nuclear facility problems, weather-related emergencies, flu pandemics, and other emergencies all require those responsible such as government officials, building managers, and university administrators to be able to quickly and reliably distribute emergency information to the public. This paper presents our design of a deep-learning-based emergency warning system. The proposed system is considered suitable for application in existing infrastructure such as closed-circuit television and other monitoring devices. The experimental results show that in most cases, our system immediately detects emergencies such as car accidents and natural disasters.

Keywords: Emergency alert system; EAS; Deep-learning; Disaster

1. Introduction

The Emergency Alert System (EAS) is a national warning system that was implemented in the United States on November 29, 1997. However, recent years have seen the exposure of a number of drawbacks of the EAS. For example, systems that rely on the use of cellular phone [1,2] or radio broadcast [3] networks are often unable to reach individuals who are located inside of buildings. The interiors of many buildings at universities, research centers, office complexes, manufacturing plants, and other locations often have very poor radio and cellular phone reception because of interference caused by equipment located within the building, or because of a shielding effect created by the building structure itself.

In addition, current emergency systems are not easily able to reach the right people, in the right location, at the right time. Although services relying on Cellular phones services, text messaging services, and e-mail services can target specific individuals, but they will would not be effective for a location-specific emergency because such these services are only able to target individuals people selectively on an individual basis by phone number or e-mail address, regardless of their physical location. Sirens [4] can provide a quick alert, but they may not yield desired results because the sound may not reach all locations, and some individuals in some areas may ignore a siren that provides no specific information about the emergency. People relying solely on cellular telephones would be excluded from the warning. Networks like such as Ethernet and WiFi are prone to failure in times of an emergency due to because potential power outages could shutting down the network or one or more network devices, thereby causing the communication failure with an entire building or geographical area to fail.

Closed-circuit television (CCTV), also known as an emergency or crime monitoring system, involves the use of video cameras to transmit a signal to a specific place, on a limited set of monitors. It differs from broadcast television in that the signal is not publicly transmitted, although it may employ point-to-point (P2P), point-to-multipoint, or mesh wireless links. Lately, CCTV technology has been enhanced with a shift toward Internet-based products and systems [5,6] and other technological developments. For instance, a CCTV at an ATM could be used to capture a user's PIN as it is entered via the keypad, without their knowledge. The devices are small enough not to be noticed, and are positioned such that they

* Corresponding author.
E-mail addresses: byungseok@skku.edu (B. Kang), choo@skku.edu (H. Choo).

Peer review under responsibility of The Korean Institute of Communications Information Sciences.

☆ This paper has been handled by Prof. Dong-Soo Han.

http://dx.doi.org/10.1016/j.icte.2016.05.001
2405-9595/

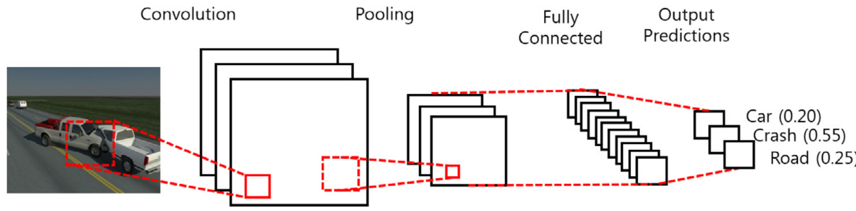

Fig. 1. Deep-learning-based video analyzer.

would enable others to monitor the keypad. Images may be transmitted wirelessly to the criminal.

The field of artificial intelligence (AI) is in a period of unprecedented improvement. Deep-learning technology [7–9] is behind most of the recent breakthroughs in object recognition, natural language processing, and speech recognition. This paper introduces an emergency alert system based on the use of deep-learning technology. The main advantage of the proposed system is that it does not require additional devices or infrastructure. We adapt a deep-learning-based real-time video analyzing module inside the CCTV device. Our system uses computer simulation to immediately detect accidents and natural disasters. The remaining part of this paper is organized as follows. We detail our proposed emergency alert system in Section 2. Experiments and measurement results are provided in Section 3. Finally, we conclude the paper in Section 4.

2. Proposed emergency alert system

2.1. General feature

Convolutional neural networks [10,11] are hierarchical machine-learning models capable of learning a complex representation of images using vast amounts of data. They are inspired by the human visual system and learn multiple layers of transformations, which extract a progressively more sophisticated representation of the input.

Our system uses heuristic/knowledge-based machine-learning technology, in which the process to generate descriptors starts with an early step to discover all the labels included in our problem domain (knowledge pool). This is done by processing all the available video images with the analyzer API, and then removing the duplicated label values. The proposed analyzer uses basic Deep Neural Network (DNN) [12] architectures for object detection and parsing to generate compositional models. This step resembles a codebook or dictionary generation process. Each input image is then encoded as a list of probability values whose length is determined by the size of the dictionary. The ith value in this descriptor corresponds to the probability returned by the analyzer for the ith label, or zero otherwise. Fig. 1 shows the various steps of the real-time video analyzing process. Our system has the following benefits:

- Scalable: DNNs [13,12] scale to billions of parameters giving them the capacity to learn highly complex concepts and thousands of categories. Modern hardware and an abundance of data enable our system to train larger and more powerful networks.
- Fast: the trained model stores its knowledge compactly in learned parameters, making it easy to deploy in any environment. There is no need to store any additional data to make predictions for new inputs. This means that we could easily use them in embedded devices such as CCTV to provide responses in milliseconds.
- Flexible: unlike traditional computer vision approaches [14] our system learns to extract discriminative features from the input using the provided training data, instead of using hand-engineered feature extractors such as SIFT [15] and LBP [16]. This makes these features easy to adapt to problems in any domain.

2.2. Emergency alerting

Our EAS has the ability to aggregate emergency information from CCTV devices. Below are the basic steps involved in alerting emergency services such as the fire/police station to an emergency situation. 1. CCTV stores the captured video data in their own storage device. 2. The analyzer module periodically monitors the stored video data based on the deep-learning technology. 3. When the system detects disaster, it sends an emergency alert message directly to the fire/police station. 4. Fire department/police station receives the emergency alert message accompanied by other related information such as location, type of disaster, analyzed results, and image data (optional). In step 4, to avoid network congestion, the system optionally sends the image files.

3. Experimental results

We evaluated our system by developing a simulator and use two different types of disasters for experimentation purposes. The first is a fire and the other is a car accident. We performed the simulation by using the following system: Intel Core i7-4810MQ (4 Cores -8 Threads), Nvidia Quadro K4100M (4 GB) GPU, and 16 GB RAM. The use of a GPU to train deep neural networks enabled us to increase the runtime, ranging from 3x faster to 15x faster [17,18]. Our simulator is programmed with Python. The simulator randomly generates an emergency situation based on the Poisson process [19]. We simulated the two emergency situations more than 100 times at each measuring point while progressively increasing the level of strength (weak, normal, strong) and the lambda value ($\lambda = 0.1, 0.3, 0.6$). Values of $\lambda = 0.1$ and 0.6 mean an event (disaster) occurred on average every 10 and 5 s, respectively. In Figs. 2 and 3, lines indicated with various colors have different meanings. Each line was plotted using data from our database.

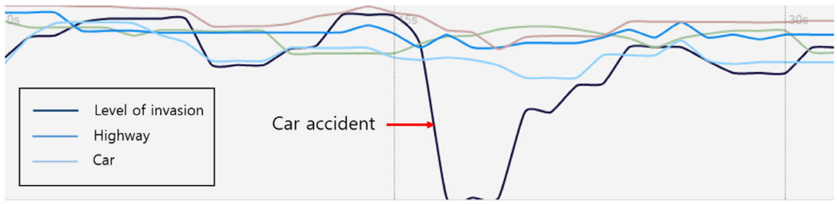

Fig. 2. Monitoring result for car accident. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

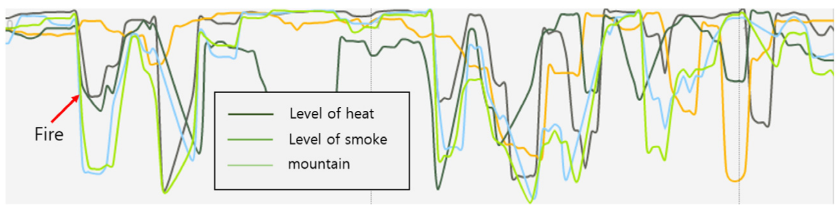

Fig. 3. Monitoring result for fire. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Simulation parameters and values.

Parameter	Value (s)
Type of event	Fire, car accident
Event frequency (λ)	01, 0.3, 0.6
Strength of event	Weak, normal, strong
Event duration time (s)	1.5
Simulation time (min)	5
Number of simulations	100 times at each point

Fig. 2 shows that the analyzer normally detects a car accident within 600 ms. In this figure, the dark-blue line means a car invaded the yellow line and collided with a car traveling in the opposite direction. Disasters that were generated by the system disappeared after 1.5 s. Fig. 3 shows the monitoring results for a fire situation with each curve showing the level of the unexpected situation. In this case, three or four factors (lines) indicate the fire situation.

Tables 1 and 2 present the experimental parameters and results for a fire and car accident, respectively. In Table 2, the term avg. detection time refers to the time that passed after the event occurred, whereas accuracy refers to the probability of detecting events. The results show that the proposed system can detect fire situations around 99% of the time within 400 ms, which means that our system can easily detect fire situations. However, in the case of a car accident, the detection accuracy of our system is below 96% because it is quite difficult and requires time to recognize an "accident" or "collision". Therefore, our system has the ability to detect fire outbreaks more rapidly than situations involving car accidents. This indicates that our system would need to accumulate a certain amount of knowledge by using a heuristic process to enable it to detect car accidents more accurately.

Table 2
Experimental results.

Event	Strength	Frequency	Avg. Detection Time(ms)	Accuracy (%)
Fire	Weak	0.1	400	96.24
		0.3	410	96.89
		0.6	390	96.23
	Normal	0.1	290	98.84
		0.3	320	98.30
		0.6	340	97.79
	Strong	0.1	260	99.45
		0.3	280	99.34
		0.6	270	99.42
Car accident	Weak	0.1	400	91.56
		0.3	410	89.82
		0.6	390	89.39
	Normal	0.1	290	92.10
		0.3	320	91.78
		0.6	340	92.57
	Strong	0.1	260	94.56
		0.3	280	95.32
		0.6	270	96.19

4. Conclusion

In this paper, we propose an emergency alert system, designed to be suitable for natural disaster detection. The proposed system uses deep-learning technology to detect and analyze disasters. We carried out experimental measurements to assess the performance of our proposed system while increasing the disaster strength and event frequency. The evaluation showed that the average detection time and accuracy for situations involving "fire" demonstrated a higher detection rate than for those involving "car accidents". These experimental results could be applied in practice by adapting our EAS system to real CCTV or other monitoring devices. Although our computer simulation only generated the above-mentioned two types of emergency events, in real environments, there exist many different emergency situations, not only disasters but also various types of criminal situations. Therefore, in future, we plan to extend our EAS to detect criminal activities and clarify the advantages of the proposed deep-learning-based detection over conventional surveillance systems in terms of accuracy and detection delay.

Acknowledgments

This work was supported in part by the Ministry of Science, ICT and Future Planning, South Korea, Institute for Information and Communications Technology Promotion through the G-ITRC Program under Grant IITP-2015-R6812-15-0001 and in part by the National Research Foundation of Korea within the Ministry of Education, Science and Technology through the Priority Research Centers Program under Grant 2010-0020210.

References

[1] V.H. Soragavi, S. Balundagi, S. Mulla, S. Kumbalavati, Droid alert system using 3g network.
[2] L. Hariprasath, R. Dhivya, S. Adithya, C. Kazhipattur, Emergency alert system using android, IJREAT Int. J. Res. Eng. Adv. Technol. 1 (1).
[3] M.-S. Baek, Y.-H. Lee, G. Kim, S.-R. Park, Y.-T. Lee, Development of T-DMB emergency broadcasting system and trial service with the legacy receivers, IEEE Trans. Consum. Electron. 59 (1) (2013) 38–44.
[4] K. Richard Ridderinkhof, B.U. Forstmann, S.A. Wylie, B. Burle, W.P. van den Wildenberg, Neurocognitive mechanisms of action control: resisting the call of the sirens, Wiley Interdiscip. Rev. Cogn. Sci. 2 (2) (2011) 174–192.
[5] H. Kruegle, CCTV Surveillance: Video Practices and Technology, Butterworth-Heinemann, 2011.
[6] D. Choi, N. Baik, CCTV monitoring technology of nighttime low-light situations at the ITS Center in Korea, in: 20th ITS World Congress, 2013.
[7] L. Deng, G. Hinton, B. Kingsbury, New types of deep neural network learning for speech recognition and related applications: An overview, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2013, pp. 8599–8603.
[8] J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, Q.V. Le, A.Y. Ng, On optimization methods for deep learning, in: Proceedings of the 28th International Conference on Machine Learning (ICML-11), 2011, pp. 265–272.
[9] I. Sutskever, J. Martens, G. Dahl, G. Hinton, On the importance of initialization and momentum in deep learning, in: Proceedings of the 30th International Conference on Machine Learning, ICML-13, 2013, pp. 1139–1147.
[10] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[11] S. Ji, W. Xu, M. Yang, K. Yu, 3d convolutional neural networks for human action recognition, IEEE Trans. Pattern Anal. Mach. Intell. 35 (1) (2013) 221–231.
[12] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, et al., Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups, IEEE Signal Process. Mag. 29 (6) (2012) 82–97.
[13] K. Veselỳ, A. Ghoshal, L. Burget, D. Povey, Sequence-discriminative training of deep neural networks., in: INTERSPEECH, 2013, pp. 2345–2349.
[14] Y. Shirai, Three-dimensional Computer Vision, Springer Science & Business Media, 2012.
[15] M. Niemeijer, M.K. Garvin, K. Lee, B. van Ginneken, M.D. Abràmoff, M. Sonka, Registration of 3D spectral OCT volumes using 3D SIFT feature point matching, in: SPIE Medical Imaging, International Society for Optics and Photonics, 2009, pp. 72591I–72591I.
[16] J. Shelton, G. Dozier, K. Bryant, J. Adams, K. Popplewell, T. Abegaz, K. Purrington, D.L. Woodard, K. Ricanek, Genetic based LBP feature extraction and selection for facial recognition, in: Proceedings of the 49th Annual Southeast Regional Conference, ACM, 2011, pp. 197–200.
[17] A. Kerr, G. Diamos, S. Yalamanchili, Modeling GPU-CPU workloads and systems, in: Proceedings of the 3rd Workshop on General-Purpose Computation on Graphics Processing Units, ACM, 2010, pp. 31–42.
[18] M. Arora, S. Nath, S. Mazumdar, S.B. Baden, D.M. Tullsen, Redefining the role of the CPU in the era of CPU–GPU integration, IEEE Micro. 32 (6) (2012) 4–16.
[19] W. Fischer, K. Meier-Hellstern, The Markov-modulated Poisson process (MMPP) cookbook, Perform. Eval. 18 (2) (1993) 149–171.