Forest Fire Monitoring System Based on Aerial Image

Seonghyun Kim, Wonjae Lee, Young-su Park, Hyun-Woo Lee, Yong-Tae Lee
Smart Platform Department
Electronics and Telecommunications Research Institute
Daejeon, Korea
{kim-sh, russell, yspark, hwlee, ytlee}@etri.re.kr

Abstract— **Since natural disaster annually leads to casualties and property damages, developments for ICT-based disaster management techniques are fostering to minimize economic and social losses. For this reason, it is essential to develop a customized response technology for a natural disaster. In this paper, we introduce a smart-eye platform which is developed for disaster recognition and response. In addition, we propose a deep-learning based forest fire monitoring technique, which utilizes images acquired from an unmanned aerial vehicle with an optical sensor. Via training for image set of past forest fires, the proposed deep-learning based forest fire monitoring technique is designed to be able to make human-like judgement for a new input image automatically whether forest fire exists or not. Through simulation results, the algorithm architecture and detection accuracy of the proposed scheme is verified. By applying the proposed automatic disaster recognition technique to decision support system for disaster management, we expect to reduce losses caused by disasters and costs required for disaster monitoring and response.**

Keywords—Forest fire monitoring, deep-learning, decision support technology, disaster management

I. INTRODUCTION

Recently, predictions of natural disasters have become limited, and intensity and scale of natural disasters damage have been increasing due to climate changes. Because natural disaster annually leads to casualties and property damages, economic and social losses are arising as national issues to be solved. For this reason, it is essential to develop a customized response technology for natural disaster in order to reduce costs caused by repeated natural disasters [1][2].

In the case of forest fire induced by thunderstroke, since it is hard to utilize helicopter due to stormy weather, ground staffs should cope with forest fire directly. In addition, human-assisted response and utilization of expensive equipment require costly expenditures [3]. To solve this problem, utilization of unmanned aerial vehicle (UAV) is proposed as a solution [4]. UAV can be effectively utilized owing to low operating costs, improved performance of camera mounted at UAV and reduced weight. To construct disaster management system based on UAV, it is necessary to develop core technology for prediction and analysis of disasters [5].

To provide an accurate forest fire monitoring service, a reliable analysis system is required. There are smoke and flame detection techniques based on conventional computer vision methods. In previous works, wavelets, support vector machine and fuzzy finite automata were utilized for smoke and flame detection [6]-[8]. However, the previous works require a restrictive condition, where videos should be recorded from fixed cameras. This condition is not suitable to UAV-based monitoring system.

In this paper, we propose a deep-learning based forest fire monitoring technique, which utilizes images acquired from UAV with an optical sensor. The main purpose of the proposed scheme is to reduce disaster damages by training the deep model on past disaster information to recognize current disaster situation. In addition, through the proposed scheme, we expect that the secondary damages are prevented, then it leads to decrease of recovery costs for disaster. To this end, it is necessary that the proposed deep-learning based forest fire monitoring technique is designed to be able to make human-like judgement for a new input image automatically. In the following sections, we derive a training methodology for the forest fire monitoring and verify the architecture and detection accuracy of the proposed scheme via simulation results.

II. FOREST FIRE MONITORING SYSTEM

A. *Introduction of smart-eye platform*

Figure 1 represents the procedure of forest fire recognition via the smart-eye platform, which is developed for disaster recognition and response by utilizing an UAV and an image analysis server. As shown in Fig. 1, the UAV captures aerial views over forests or mountains, then the UAV serially transfers the captured images to an image analysis server via a wireless link, where the duration of capturing and transferring per image is three seconds. The image analysis server operates the forest fire monitoring system to recognize automatically whether there is a fire in the captured image or not. The forest fire monitoring system is pre-trained with past forest fire images via deep-learning approach. The procedure including the image capturing, transferring and analyzing is performed in real time in order to provide immediate recognition result to a decision support system for disaster management.

978-1-5090-5234-9/16/$31.00 ©2016 IEEE

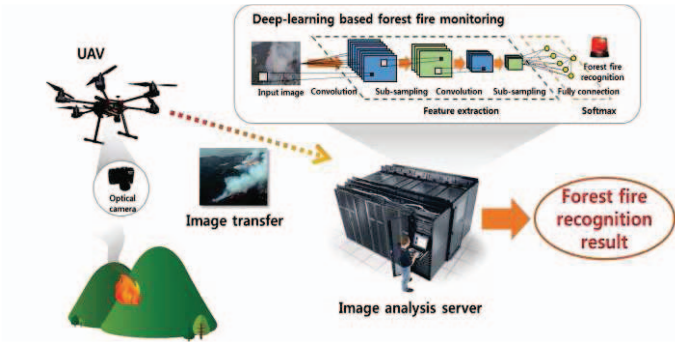


Fig. 1. Procedure of forest fire recognition via the smart-eye platform

B. Forest fire monitoring system based on deep-learning

Figure 2 shows our forest fire monitoring system based on convolutional neural network, which is one of deep-learning approaches [10]. The forest fire monitoring system is a core technique in the image analysis server of the smart-eye platform. When the image analysis server receives an aerial image from the UAV in Fig. 1, the image analysis server performs the forest fire monitoring action to recognize a fire scene automatically.

As shown in Fig. 2, the proposed forest fire monitoring system is constructed by two phases as training phase and testing/application phase. In the training phase, the convolutional neural network is trained by using a dataset which consists of fire image set and non-fire image set. After the training phase, the convolutional neural network is fixed then it is applied to the forest fire monitoring system in order to recognize a fire scene for a new input image.

Concretely, the architecture for the forest fire monitoring system consists of three convolutional layers, three pooling layers, and two fully connected layers. The first convolutional layer has 96 kernels of size $11\times11\times3$ for the $128\times128\times3$ RGB input image. The second and third convolutional layers have 256 kernels of size $4\times4\times96$ and 256 kernels of size $3\times3\times256$, respectively. Each pooling layer in the convolutional neural network summarizes the outputs of the previous layer in the same kernel map. The fully-connected layers include 4096 neurons. For faster saturation of the training phase, a rectified linear unit (ReLU), which is a nonlinear function defined as $f(x) = \max(0, x)$, is applied to the outputs of convolutional and fully-connected layers.

To reduce testing errors and avoid overfitting problem, the dropout technique is adopted, which consists of setting each output of neuron to zero with a probability of 0.5 [11]. In addition, data augmentation, which enlarges the training dataset via transformation for the given dataset, is used to avoid the overfitting problem. Image cropping and vertical/horizontal flipping are used for the data augmentation. Training is based on mini-batch gradient descent with a batch size of 120. Stochastic gradient descent is used to find the optimal network values [12].

C. Data set

To train the convolutional neural network in Fig. 2, the past information for fire and non-fire scenes with adequate label for each scene is required. It means that data collection for the past information is required in advance, and the dataset is classified into subsets with adequate labels where a label is mapped to the corresponding image in a hand-crafted manner [4].

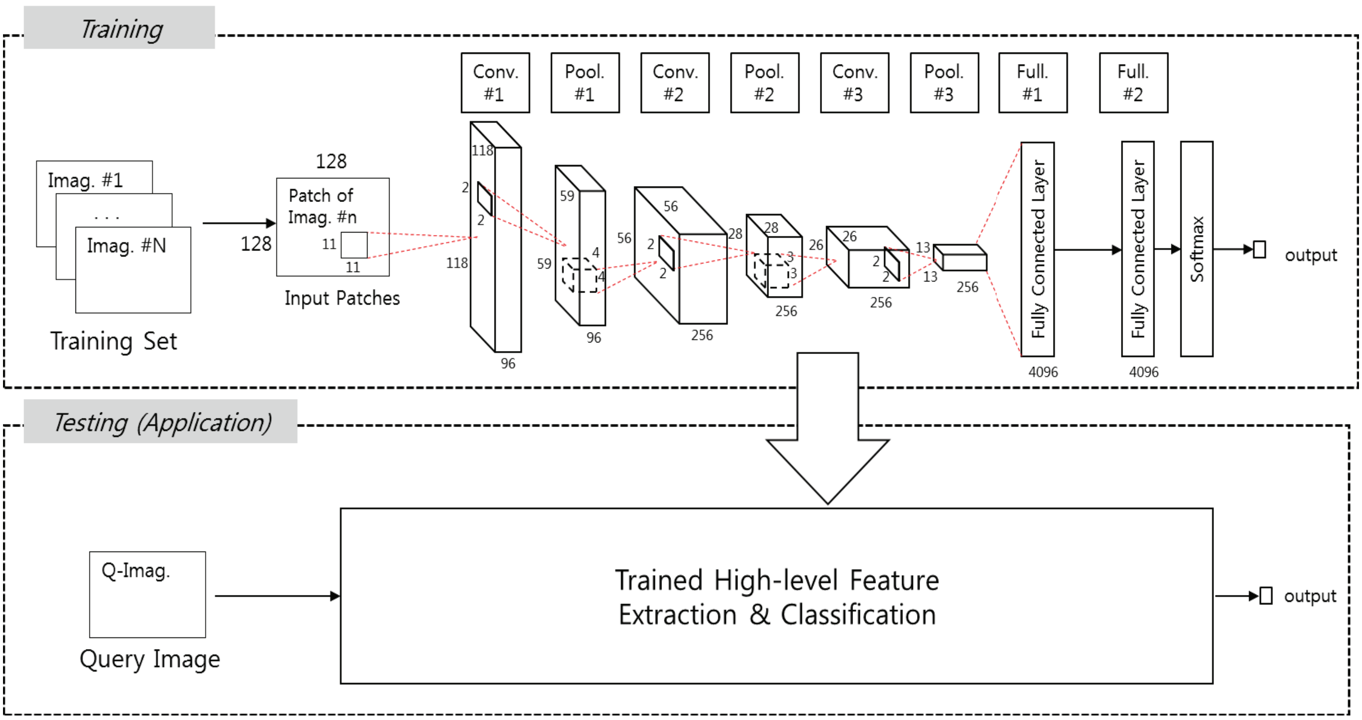
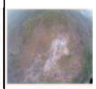


Fig. 2. Forest fire monitoring system based on deep-learning

Table I demonstrates the classified images according to the proposed classification rules, which consist of two types of label-2 and label-6. First, in the classification rule of label-2, images are divided into two categories as fire and non-fire scenes. In the classification rule of label-6, fire scenes in the label-2 are divided into three categories as fire-nighttime, fire-daytime and smoke. In addition, non-fire scenes in the label-2 are divided into three categories as spring-fall, summer and winter. Therefore, the case of label-2 has two classes and the case of label-6 has six classes. By using these classification rules, the convolutional neural network in Fig. 2 is trained iteratively.

TABLE I. DATA SET AND DEFINITION OF LABELS

Label-2	Fire			Non-fire		
Label-6	Fire-nighttime	Fire-daytime	Smoke	Spring-fall	Summer	Winter
Training images						

III. SIMULATION RESULTS

The simulation environment is as follows: CPU i7-5903K, 64GB RAM, GeForce GTX Titan X GPU, and NVIDIA DIGITS 2.0 with Caffe library being used for the implementation of the convolutional neural network [13][14].

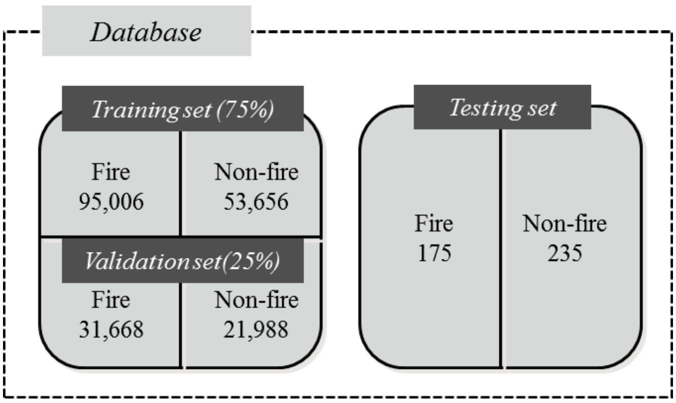


Fig. 3. Database for fire and non-fire images according to training, validation and testing sets.

Figure 3 shows the database of fire and non-fire images used for training, validation and testing sets, which are randomly collected from internet web sites. To train the convolutional neural network of the proposed forest fire monitoring system, 95,006 fire images and 53,656 non-fire images in the training set are used. In addition, 31,668 fire images and 21,988 non-fire images in the validation set are used to verify the trained convolutional neural network, where the validation set is constructed by data augmentation from the training set.

In the data augmentation, images in the training set are transformed, where image transformation includes cropping, shifting and reversing. Then these transformed images are stored in the validation set. It means that images in the validation set are similar to those in the training set due to artificially enlarged data. A ratio of the training set and validation set is 0.75:0.25.

After the training phase is completed and the convolutional neural network is fixed, the finally updated convolutional neural network is applied to the actual forest fire monitoring system. To measure the performance of the forest fire monitoring system, 175 fire images and 235 non-fire images in the testing set are used, where the images in the testing set are randomly collected from online websites and not used in the training phase. The performance of the forest fire monitoring system is presented in terms of filter output, detection accuracy and training time.

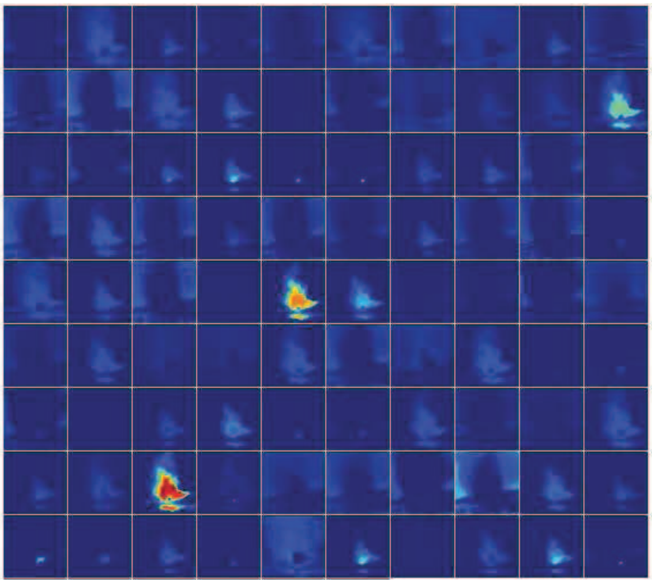
(a) Fire scene

(b) Non-fire scene

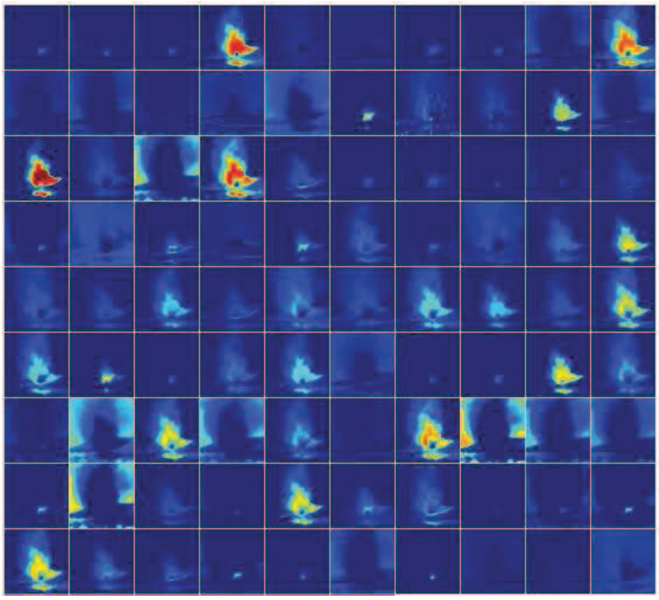
Fig. 4. Test images for fire and non-fire scenes

Figure 4 displays some test images for fire and non-fire scenes collected from internet web sites. As can be seen from Fig. 4 (a), the fire scene contains flame and smoke. In contrast, there is no flame and smoke in Fig. 4 (b). These test images are not used for the training phase but for the testing phase only.

Figure 5 represents the filter outputs of the first convolutional layer for the label-2 and label-6 with a test image for fire scene. As shown in Figs. 5 (a) and (b), each filter output has the activated responses for the flame and smoke in Fig. 4 (a). In comparison with Figs. 5 (a) and (b), the filter output of the label-6 has more activated responses for the components related to fire scene than that of the label-2.



(a) Label-2

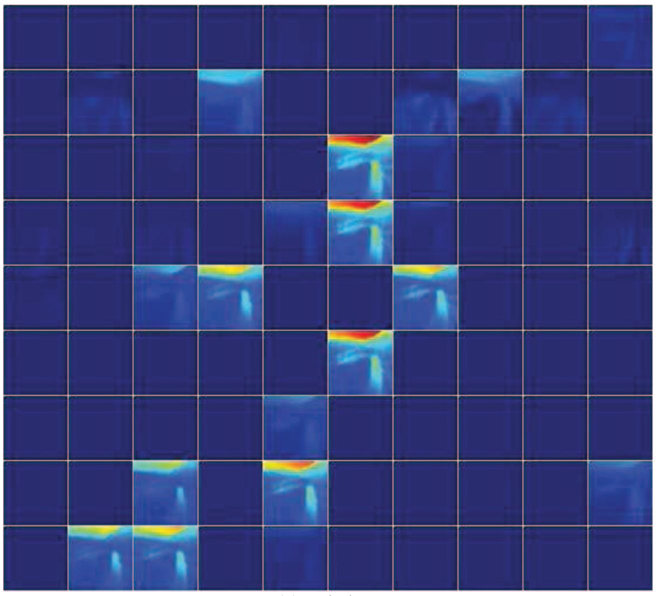


(b) Label-6

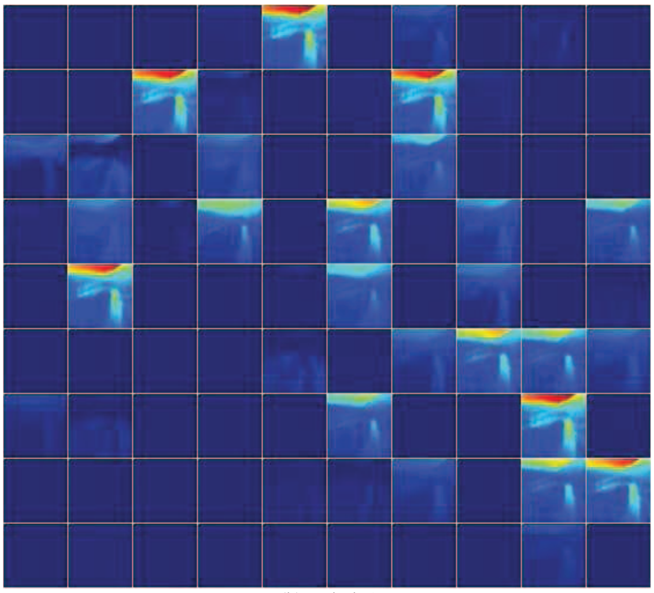
Fig. 5. Filter outputs of the first convolutional layer for the label-2 and label-6 with a test image for fire scene.

Figure 6 demonstrates the filter outputs of the first convolutional layer for the label-2 and label-6 with a test image for non-fire scene. As shown in Figs. 6 (a) and (b), each filter output has the activated responses for the sky and stream in Fig. 4 (b). In comparison with Figs. 6 (a) and (b), the filter output of the label-6 has more activated responses for the components related to non-fire scene than that of the label-2.

As observed in the filter output comparisons of Figs. 5 and 6, it can be known that labels with detailed classification lead to meaningful outputs with more activated response. The difference of the filter outputs between the label-2 and label-6 affects the performance in terms of detection accuracy.



(a) Label-2



(b) Label-6

Fig. 6. Filter outputs of the first convolutional layer for the label-2 and label-6 with a test image for non-fire scene.

Table II presents the comparison of learning time and accuracy between the label-2 and label-6, where the accuracy is defined as *(accuracy)=(number of accurate detections)/(total number of test images) ×100 (%)*. As observed in Table II, the label-6 exhibits better accuracy than that of the label-2, but the learning time of the label-6 is longer than that of the label-2. The reason is that the detailed classification of the label-6 leads to more specific features for the fire and non-fire scenes, and the specific features affects to improvement of the detection accuracy. However, the detailed classification of the label-6 leads to larger learning time due to the increase of the weight variables among the neurons in the convolutional neural network.

TABLE II. COMPARISON OF LEARNING TIME AND ACCURACY

	Learning time	Accuracy
Label-2	5h15m	81.46%
Label-6	6h20m	88.54%

To provide more details for the accuracies of the label-2 and label-6, a confusion matrix is calculated in Tables III and IV, which includes true positive, false negative, false positive and true negative. In the confusion matrix, the true positive indicates that the forest fire monitoring system has the output of alarm for a fire input image, and the true negative specifies that the forest fire monitoring system has the output of non-alarm for a non-fire input image, i.e., the true positive and true negative represent the adequate results of the proposed system. In contrast, the false negative and false positive represent that the proposed system has wrong output for an input image.

TABLE III. CONFUSION MATRIX OF THE LABEL-2

Label-2	Fire	Non-fire
Alarm	True positive 77.71%	False positive 15.74%
Non-alarm	False negative 22.29%	True negative 84.26%

TABLE IV. CONFUSION MATRIX OF THE LABEL-6

Label-6	Fire	Non-fire
Alarm	True positive 89.71%	False positive 12.34%
Non-alarm	False negative 10.29%	True negative 87.66%

Tables III and IV represent the confusion matrices of the label-2and label-6. In comparison with Tables III and IV, it can be observed that the label-6 outperforms the label-2, where the true positive is about 22.00% higher and the true negative is about 3.40% better. As shown in the comparison results, in particular, the true positive of the label-6 is significantly improved. Therefore, owing to the detail features and filter outputs based on the increased number of classes, the label-6 has more reliable detection outputs than the label-2.

IV. CONCLUSION

In this paper, we proposed the forest fire monitoring system based on the deep-learning and present the performance of the proposed scheme. In the simulation results, we showed that the detailed classification of the label leads to more specific features for the fire and non-fire scenes and the improved detection accuracy. Through the automatic recognition property of the proposed scheme, the decrease of disaster damages and reduced costs for monitoring and response can be acquired. Currently, the proposed forest fire monitoring system is designed with an optical sensor, but it will be improved by using an infrared thermal camera.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (R0190-16-2034, Development of Unmanned Aerial Vehicle Sensor-based Smart Eye Technology for Local Disaster Monitoring and Situational Response)

REFERENCES

[1] P. Rawat, M. Haddad, and E. Altman, “Towards efficient disaster management: 5G and Device to Device communication,” 2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), pp. 79-87, Nov. 2015.

[2] S. Rasouli, O. C. Granmo, and J. Radianti, “A methodology for fire data analysis based on pattern recognition towards the disaster management,” 2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), pp. 130-137, Nov. 2015.

[3] Luis Merino et al., "An Unmanned Aircraft System for Automatic Forest Fire Monitoring and Measurement" J. Intell. Robot. Syst., vol. 65, no. 1, pp 533-548, Jan. 2012.

[4] I. Colomina and P. Molina, “Unmanned aerial systems for photogrammetry and remote sensing: a review,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 92, pp.79-97, 2014.

[5] A. J. Santoso, F. K. S. Dewi, and T. A. P. Sidhi, “Natural disaster detection using wavelet and artificial neural network,” 2015 Science and Information Conference (SAI), pp. 761-764, Jul. 2015.

[6] J. Gubbi, S. Marusica, and M. Palaniswamia, “Smoke detection in video using wavelets and support vector machines,” Fire Safety J. 44(8), 1110-1115, 2009.

[7] S. J. Ham, B. C. Ko, and J. Y. Nam, “Vision based forest smoke detection using analyzing of temporal patterns of smoke and their probability models,” Image Processing: Machine Vision Applications IV, 7877:1-6, 2011.

[8] B. C. Ko, S. J. Ham, and J. Y. Nam, "Modeling and formalization of fuzzy finite automata for detection of irregular fire flames," IEEE Trans. Circuits Syst. Video Technol., vol. 21, no. 12, pp. 1903-1912, 2011.

[9] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436-444, May 2015.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems, pp. 1097–1105, Dec. 2012.

[11] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R.R. Salakhutdinov, “Improving neural networks by preventing co-adaptation of feature detectors.” arXiv preprint arXiv:1207.0580, 2012.

[12] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel. “Backpropagation applied to handwritten zip code recognition,” Neural Comp., 1(4):541-551, 1989.

[13] Y. Jia, “Caffe: An open source convolutional architecture for fast feature embedding,” http://caffe.berkeleyvision.org, 2013.

[14] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” arXiv:1408.5093, 2014.