International Journal of Advanced Smart Convergence Vol.10 No.1 82-88 (2021)
http://dx.doi.org/10.7236/IJASC.2021.10.1.82

IJASC 21-1-8

Automatic Detection of Work Distraction with Deep Learning Technique for Remote Management of Telecommuting

Wan Yeon Lee

Professor, Dept. of Computer Science, Dongduk Women's University, Seoul 02748, South Korea
Email: wanlee@dongduk.ac.kr

Abstract

In this paper, we propose an automatic detection scheme of work distraction for remote management of telecommuting. The proposed scheme periodically captures two consequent computer screens and generates the difference image of these two captured images. The scheme applies the difference image to our deep learning model and makes a decision of abnormal patterns in the difference image. Our deep learning model is designed with the transfer learning technique of VGG16 deep learning. When the scheme detects an abnormal pattern in the difference image, it hides all texts in the difference images to protect disclosure of privacy-related information. Evaluation shows that the proposed scheme provides about 96% detection accuracy.

Keywords: *Automatic detection, Work distraction, Telecommuting, Remote management, Deep learning*

1. Introduction

Covid-19 pandemic makes the telecommuting of workers popular. In the telecommuting environment, the remote management of workers during on-duty hours becomes more important. Real-time monitoring of workers is not preferred due to privacy infringement. In this paper, we proposed an automatic management scheme that detects the work distraction without privacy infringement. The proposed scheme operates separately and independently in the local computer of remote workers and analyzes the computer screen with the deep learning technique.

Recently the image analysis scheme with the deep learning technique has been progressed very quickly. In 2012, a research team of Toronto university introduced an evolutional automatic image analysis scheme, called ImageNet, designed with the deep learning technique [1]. This image recognition scheme dramatically increases the accuracy of image classification. In 2015, another image analysis scheme, called ResNet, based on the deep learning technique shows 96.43% accuracy, which is better than 94.90% accuracy of humans [2]. While the image analysis scheme with the deep learning technique has been intensively focused, the video anomaly analysis scheme with the deep learning technique has been rarely studied. In this paper, we investigate a video anomaly analysis scheme that detects abnormal behaviors of remote workers. Some recent studies

Manuscript Received: January. 13, 2021 / Revised: January. 17, 2021 / Accepted: January. 21, 2021
Corresponding Author: Wan Yeon Lee, email: wanlee@dongduk.ac.kr
Tel: +82-2-940-4685, Fax: +82-2-940-4170
Professor, Dept. of Computer Science, Dongduk Women's University, Seoul 02748, South Korea

[3, 4, 5, 6] dealt with the video anomaly detection issue. Nakazawa and Kulkarni proposed a neural network scheme with the deep convolutional encoder-decoder that automatically finds the wafer defects on the image of semiconductors [3]. Morais et al. proposed a machine learning scheme that analyzes the skeleton of humans in videos and detects abnormal behaviors [4]. Zhu and Newsam proposed a temporal augmented neural network that gives more weight on important areas for improved detection of abnormal behaviors such as stealing money boxes [5]. Birada et al. proposed a time-aware machine learning scheme that detects abnormal moving actions of automobiles with consideration of traffic light signals on road videos [6]. All of these studies did not consider the automatic detection of abnormal patterns on the computer screen. To the best of our knowledge this study is the first approach for the automatic detection of abnormal patterns on the computer screen of workers.

The scheme proposed in this paper periodically captures two consequent computer screens and generates the difference image of these two captured images. The scheme applies the difference image to our deep learning model and makes a decision of abnormal patterns in the difference image. Our deep learning model is designed with the transfer learning technique [7]. When the worker performs normal jobs on the computer screen, the difference image contains a few of regular spontaneous changes, whereas the difference image contains a lot of irregular spontaneous changes when a worker plays an online game or watches a movie on the computer screen. If the scheme detects an abnormal pattern in the difference image, it hides all texts in the difference images to protect disclosure of privacy-related information and sends the difference image to the human resource manager.

To evaluate the proposed scheme, we implement the scheme with tensorflow and keras libraries on python language. Recorded videos of playing an online game are used as train and validation dataset of *abnormal* working patterns. Recorded videos of displaying a powerpoint file are used as train and validation dataset of *normal* working patterns. Evaluation shows that the proposed scheme shows about 98% accuracy on the train and validation dataset and about 96% accuracy on the test dataset.

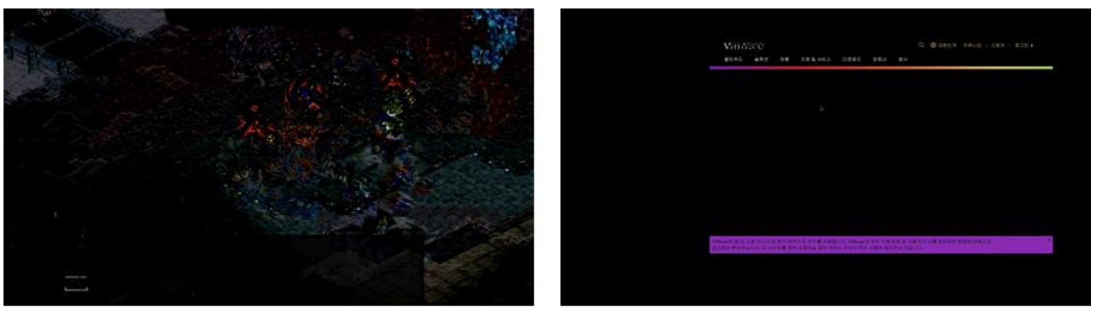
The rest of this paper is organized as follows; Section 2 explains the system model considered in this article. Section 3 describes the proposed scheme in detail. Section 4 shows evaluation results and Section 5 provides concluding remarks.

2. System Model

When establishing an elaborate deep learning model, it is essential to construct a big data set priorly against the normal working pattern and the abnormal working pattern. Unfortunately there is no big data platform collected for the normal working pattern or the abnormal working pattern to the best of our knowledge. So, in order to compensate insufficient data set, we adopt the transfer learning technique that is the remodeling of learning in a new task through the transfer of knowledge from a related task already been learned [7].

In this paper, we select the supervised learning approach rather than the unsupervised learning approach when designing our deep learning model. The supervised learning technique is a function mapping an input to an output based on labeled input-output pairs. To obtain sample data of abnormal working patterns, we download 10 videos of online games from the youtube platform.

Figure 1(a) shows the difference image of two consequent frame with one second delay in a downloaded video of starcraft game downloaded. The output of this difference image is labelled as 1, which means the abnormal working pattern. To obtain sample data of normal working patterns, we download 10 videos of online classes from the online education platform of our university. Figure 1(b) shows the difference image of two consequent frame with one second delay in the recorded computer screen during displaying a power point file. The output of this difference image is labelled as 0, which means the normal working pattern. While Figure 1(a) includes irregular spontaneous changes of game characters, Figure 1(b) includes regular spontaneous changes of boxes and texts. Note that the difference image in the normal working pattern looks like the black box in most cases because there are rare changes within one second in the normal working pattern.



(a) (b)

Figure 1. (a) Difference image when playing a game video and (b) difference image when displaying a power point work

We select the deep learning model instead of the machine learning model. The machine learning model requires the feature extraction from the input images and applies extracted features to the convolution neural network, while the deep learning model directly applies the input images to the convolution neural network.

3. Proposed Scheme

The proposed scheme first captures two consequent computer screens with one second delay. With these two captured images, the scheme makes their difference image that contains the spontaneous change of computer screen. Next the scheme applies this difference image to our deep learning model that will be explained later. If our deep learning model decides that the input image is abnormal, the scheme masks all texts on the input difference image in order to protect privacy-related information, and sends it to the human resource manager to notify that the worker is distracted during on-duty hours. Otherwise, the scheme discards it. This procedure may be repeated periodically to check where the worker is distracted during on-duty hours or not. Figure 1 shows the overall flow chart of the proposed scheme.

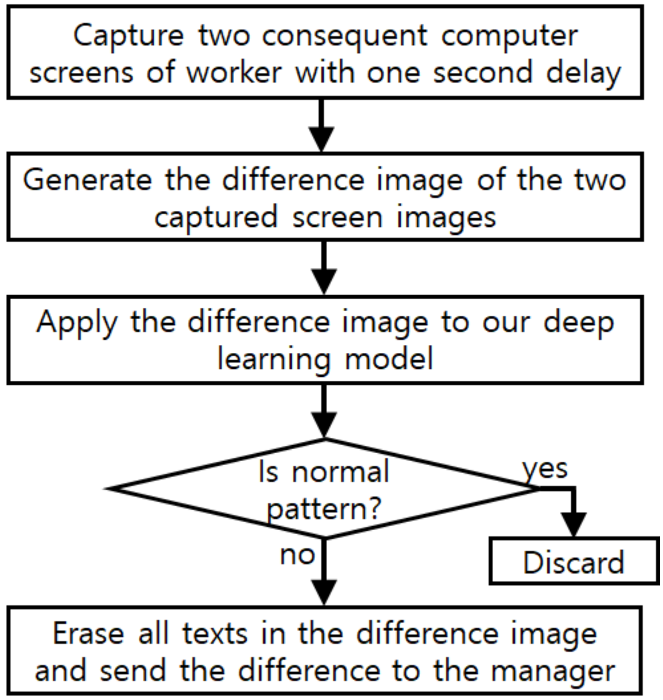


Figure 2. Overall flowchart of proposed scheme

Our deep learning model used in the proposed scheme is constructed as follows; Instead of building our own deep learning model, we use the remodeling approach of an existing machine learning model with the help of transfer learning technique because we have not enough sample images for training as explained in Section 2. The VGG16 pre-trained model [8] is selected for the transfer learning technique. Our deep learning model exploits the features generated by the VGG16 model trained on the "imagenet" dataset. Our model is implemented with python language. To import the VGG16 model, our model is built with a tensorflow high-level API keras [9], as shown in Figure 3. We choose an option "include_top = False", because this study considers one class, either the normal pattern or the abnormal pattern, while the VGG16 model predicts on lots of classes. This option implies that our model does not include fully connected layers from the VGG16 model.

```
import keras
from keras.applications import VGG16

base_model = VGG16(include_top=False, weights='imagenet')

train_class1 = base_model.predict(class1_data)
valid_class2 = base_model.predict(class2_data)
```

Figure 3. Codes for transfer learning of VGG16

In our model, we need to create a fully connected layers because we do not include fully connected layers from VGG16 model. Figure 4 shows the codes that build our deep learning model.

Output features from VGG16 model with the shape of 7*7*512 becomes the input shape for our model. There are two hidden layers with 1,024 neurons and 512 neurons respectively. One class, either the normal pattern or the abnormal pattern, becomes the output layer of our model. We insert dropout layers to make model less over-fit.

```
from keras.models import Sequential
from keras.layers import Dense, InputLayer, Dropout

model = Sequential()
model.add(InputLayer((7*7*512,)))                    # input layer
model.add(Dense(units=1024, activation='relu')) # hidden layer
model.add(Dropout(0.5))                              # adding dropout
model.add(Dense(units=512, activation='relu'))  # hidden layer
model.add(Dropout(0.5))                              # adding dropout
model.add(Dense(1, activation='sigmoid'))         # output layer

model.summary()
```

Figure 4. Codes to build our deep learning model

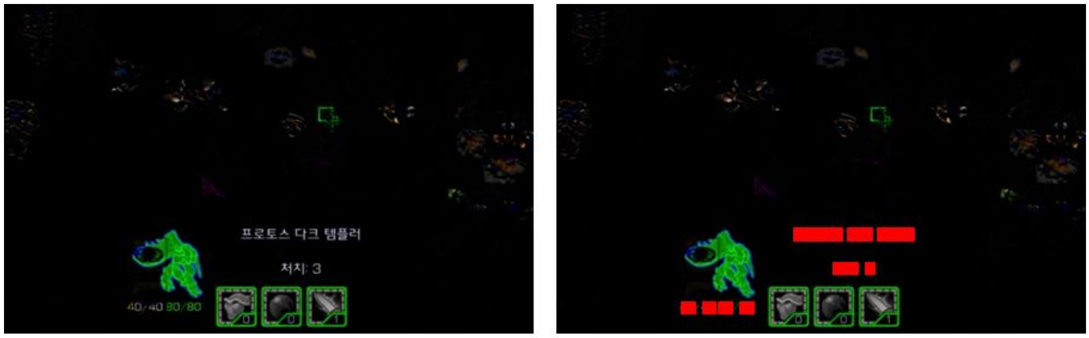
Finally we fit our model where the stochastic gradient descent is used as an optimizer and the binary cross-entropy is used as our loss function. We also save our checkpoint for the best model according to its validation dataset accuracy with a batch size of 64 and 30 epochs to train. We got an accuracy of about 98% on the training as well as the validation images. In the next section, we will evaluate the accuracy of our model in a new video.

4. Evaluation

To test our trained model, we play the starcraft online game for 30 minutes and record it into a video file to obtain the test dataset of abnormal working pattern. Also we perform an internet surfing through chrome browser for 30 minutes and record it into another video file to obtain the test dataset of normal working patterns. The difference images extracted from the game video are labelled as 1, and those extracted from the Internet surfing video are labelled as 0. The test dataset is different and independent from the train and validation dataset.

Against these test dataset, the proposed scheme shows about 96% accuracy while the accuracy on the train and validation dataset is about 98%. The reason of accuracy difference may be variances of difference images between training dataset and test dataset. We need to refine the dataset used on the train procedure to enhance the accuracy.

Figure 5(a) shows an example difference image determined as abnormal pattern by the proposed scheme. This difference image may contain privacy-related information such as bank account and login identifier. To protect the disclosure of privacy-related information, the proposed scheme hides all texts with red boxes as shown in Figure 5(b) before sending the difference image to the human resource manager. The difference image decided as the normal pattern is always dropped in the local computer and thus there is no disclosure of privacy-related information out of the local computer.



(a) (b)

Figure 5. (a) Original difference image decided as abnormal and (b) its modified difference image hiding all texts

Even though the scheme makes a wrong decision, the procedure to hide all texts on the difference images minimizes the side-effect of the wrong decision. When the scheme determines the abnormal working pattern as normal, there is penalty for worker but the human resource manager may complain it if there is no more chance to figure out the abnormal working pattern. However, the detection of abnormal working pattern is repeated periodically. So just one correct detection on the abnormal working pattern may be enough for the human resource manager even while the scheme makes many wrong detection on the abnormal working pattern. On the other hand, when the scheme determines the normal working pattern as abnormal, the human resource manager needs to filter out the difference images of normal working pattern. It is not difficult to differentiate the difference images of normal working pattern from those of abnormal working pattern.

5. Conclusions

In this paper, we proposed an automatic management scheme that detects the work distraction without privacy infringement. The proposed scheme periodically captures two consequent computer screens and makes the difference image of these two captured images. The difference image is applied to our deep learning model designed with the transfer learning technique of VGG16. Our deep learning model was designed with an idea that the difference image contains a lot of irregular spontaneous changes against playing an online game or watching a movie on the computer screen, whereas it contains a few of regular irregular spontaneous changes against performing normal jobs on the computer screen. In the evaluation, the scheme showed about 98% accuracy on the train and validation dataset and about 96% accuracy on the test dataset. To protect the disclosure of privacy-related information on the difference image decided as abnormal, the scheme hides all texts before sending the difference image to the human resource manager.

Acknowledgements

This research was supported by the Dongduk Women's University Grant, 2020.

References

[1] Alex Krizhevsky, Ilya Sutskever and Geoffery E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems (NIPS) (Dec. 2012), vol. 25, pp. 1097-1105. DOI: 10.1145/3065386

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun., "Deep Residual Learning for Image Recognition", arXiv:1512.03385, Dec. 2015.

[3] Takeshi Nakazawa and Deepak V. Kulkarni, "Anomaly Detection and Segmentation for Wafer Defect Patterns Using Deep Convolutional Encoder-Decoder Neural Network Architectures in Semiconductor Manufacturing", IEEE Transactions on Semiconductor Manufacturing, vol. 32, no. 2, pp. 250-256, May 2019. DOI: 10.1109/TSM.2019.2897690

[4] Romero Morais, Vuong Le, Truyen Tran, Budhaditya Saha, Moussa Mansour and Svetha Venkatesh, "Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11996-12004, Jun. 2019. DOI: 10.1109/CVPR.2019.01227

[5] Yi Zhu and Shawn Newsam, "Motion-Aware Feature for Improved Video Anomaly Detection", arXiv:1907.10211, Jul. 2019.

[6] Kuldeep M. Birada, Ayushi Gupta, Murari Mandal and Santosh K. Vipparthi, "Challenges in Time-Stamp Aware Anomaly Detection in Traffic Videos", arXiv:1906.04574, Jun. 2019.

[7] Sajja T. Krishna and Hemantha K. Kalluri, "Deep Learning and Transfer Learning Approaches for Image Classification", International Journal of Recent Technology and Engineering, vol. 7, no. 5S4, pp. 427-432, Feb. 2019. DOI: 10.1016/j.patrec.2017.10.037

[8] Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, Sep. 2014.

[9] Francois Chollet, Deep Learning with Python, ISBN 9781617294433, Manning Publications, Nov. 2017.