Sesión de Investigación A5: Ciberseguridad industrial y aplicaciones
http://doi.org/10.18239/jornadas_2021.34.35

Classifying Screenshots of Industrial Control System Using Transfer Learning and Fine-Tuning

Roberto A. Vasco-Carofilis *†, Pablo Blanco-Medina *†, Franciso Jáñez-Martino *†,
Guru Swaroop Bennabhaktula *‡, Eduardo Fidalgo *†, Alejandro Prieto Castro †, and Víctor Fidalgo †

*Department of Electrical, Systems and Automation, Universidad de León, León, ES
†Researcher at INCIBE (Spanish National Cybersecurity Institute), León, ES
‡Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, NL
Email:{andres.vasco, pablo.blanco, francisco.janez, eduardo.fidalgo}@unileon.es,
g.s.bennabhaktula@rug.nl, {alejandro.prieto, victor.fidalgo}@incibe.es

***Abstract*—Industrial control systems are heavily dependant on security and monitoring protocols. For this purpose, monitoring tools take screenshots of control panels for later analysis. Classifying these screenshots into specific groups can be a time-consuming process, but it is crucial for the security tasks performed by manual operators. To solve this problem, we propose a pipeline based on deep learning to classify snapshots of industrial control panels into three categories: Internet Technologies (IT), Operation Technologies (OT), and others. We compare the results obtained with transfer learning and fine-tuning on nine convolutional neuronal networks pre-trained with the ImageNet dataset, testing them on a custom CRitical INFrastructure dataset (CRINF-300). Inception-ResNet-V2 obtains the best learning result with an F1-score of 98.32% on CRINF-300, while MobileNet-V1 obtained the best performance-speed trade-off.**

***Index Terms*—Deep Learning, Image Classification; Transfer Learning; Industrial Control System; Fine-tuning**

Type of contribution: *Research already published – Detecting Vulnerabilities in Critical Infrastructures by Classifying Exposed Industrial Control Systems Using Deep Learning. Applied Sciences [1]*

I. Introduction

Critical infrastructures, namely healthcare, transportation or manufacturing, require constant monitoring in environments such as Industrial Control Systems (ICSs). An error on these infrastructures might cause serious consequences, such as equipment failure or information leak [1].

Supervisory Control and Data Acquisition (SCADA) systems are used to control physical equipment and ICS infrastructures. SCADA systems are commonly referred to as Operational Technology (OT) systems, which control and monitor specific devices. Other industrial systems used to control software, including management, and delivery of data, are known as Information Technology (IT) systems [2].

Law Enforcement Agencies (LEAs) use specialized tools known as Internet Metasearch Engines (IMEs) to monitor these exposed assets. For those services that include a graphical interface, IMEs take screenshots to log relevant information graphically. The classification of these screenshots might help to classify the devices, as well as discover vulnerabilities.

In our paper [1], we aimed to classify ICS screenshots automatically into IT and OT categories using deep learning. Our proposal might help monitor critical infrastructures in both performance and computational cost.

II. State of the Art

Image classification is the task of assigning a label to an image. In the last years, Convolutional Neural Networks (CNNs) have been established amongst the best algorithms for this task [3]. Several works have studied the use of transfer learning applied to CNNs in different fields [4]. However, these networks need to be trained on a large amount of data, and data gathering and annotation can be a complex, time-consuming process.

One of the most common problems in this task is the low number of available images for a specific task. To address this, transfer learning is a technique that allows taking a model trained for a specific application and apply it to a closely related task [5]. In cases where images for training a model are scarce, or the classification tasks are challenging, manually-crafted feature extraction can outperform the CNNs [6], [7].

III. Methodology and Experimental Settings

Our proposal for image classification is based on transfer learning and fine-tuning, due to the limited amount of data. Figure 1 depicts the proposed methodology.

For the transfer learning approach, we freeze the final layer from each pre-trained network and obtain features to train a logistic regression model and a Support Vector Machine (SVM) with a linear kernel. We measured their performance using 5-Fold cross-validation.

For the fine-tuning approach, we drop the last layer of each model and replace it with an average 2D pooling of 7x7, a Flatten layer, a Dense layer with 256 neurons and ReLU activation, a 0.5 dropout and an output layer with softmax activation. We used 70% of our images for training, 20% for testing, and 10% for validation. Our training parameters consisted of a batch size of 26 and a learning rate of $1e-4$ for 20 epochs. Finally, we applied data augmentation to our images, increasing their original number up to 5 times.

For training and evaluation, we used 337 ICS snapshots provided by the Spanish National Cybersecurity Institute (INCIBE). We manually labeled them into two categories: 74 IT and 263 OT images, resulting in a dataset we named **Cr**itical **Inf**rastructure (CRINF-300).

We selected nine architectures commonly used in image classification, pre-trained with the Imagenet dataset; Inception-V3 [8], MobileNet-V1, MobileNet-V2 [9],

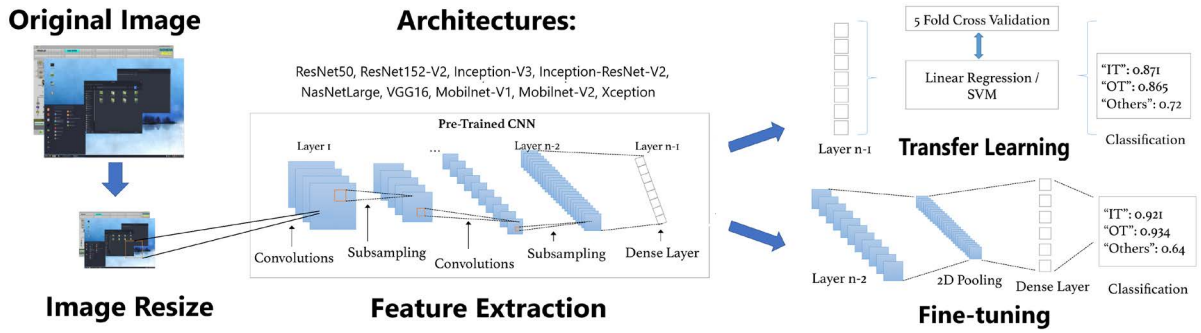

Fig. 1. Proposed pipeline for classifying ICS snapshots into three categories –IT, OT and Others– using transfer learning and fine-tuning

ResNet50, ResNet152v2 [10], VGG16 [11], NasNetLarge [12], Inception-ResNet-V2 [13], and Xception [14].

IV. Experimental Results

The F1-Score-based results of our experiments are presented in Table I. Inception-ResNet-V2 obtained the best results with an F1-score of up to 98.32% using transfer learning and logistic regression. In fine-tuning, the best architecture was VGG16, with an F1-score of 93.73%.

We found that the fastest model was VGG16, with an average of 0.03s to process each image in GPU with both SVM and logistic regression, and 0.14s with fine-tuning. CPU processing achieved the best speed of up to 0.14s with transfer learning and 0.21s with fine-tuning. On the other hand, MobileNet-V1 obtained the best performance-speed balance with a CPU speed of 0.06s and a GPU speed of 0.04s.

Although our fine-tuning additions perform well on top of ResNet50 and VGG16, they do not achieve good performance on the rest of the architectures, obtaining lower results than those of the transfer learning based models.

TABLE I
F1-Score results in the three proposed strategies

Architecture	Logistic Reg.	SVM	Fine-tuning
ResNet50	87.67 (±0.35)	87.67(±0.35)	92.16
VGG16	87.67(±0.35)	87.67(±0.35)	**93.73**
Xception	89.46(±0.67)	90.08(±1.51)	80.98
Inception-V3	97.02(±1.50)	97.20(±1.31)	82.01
Mobilenet-V1	97.58(±0.70)	97.56(±0.75)	86.51
Mobilenet-V2	97.55(±0.49)	97.92(±0.74)	48.78
NasNetLarge	96.44(±1.29)	96.78(±1.04)	81.20
Inception-ResNet-V2	**98.32**(±0.70)	**98.13**(±0.84)	76.22
ResNet152v2	97.41(±0.90)	97.59(±0.94)	71.22

V. Conclusions

This paper presented a pipeline for classifying ICS images as belonging to IT, OT or Others systems, using transfer learning and fine-tuning based approaches on CRINF-300, a custom dataset. We used nine CNN architectures pre-trained on the Imagenet dataset from two perspectives: transfer learning, with a logistic regression classifier and an SVM classifier, and fine-tuning. We chose these techniques due to our short supply of images.

The best CNN architectures to solve the proposed problem are Inception-ResNet-V2 and Mobilenet-V1. The first obtained the best performance with an F1-score of up to 98.32%, while the latter achieved the best performance speed trade-off, with an F1-score of 97.58 and a speed of 0.06s on CPU. We conclude that transfer learning with logistic regression is the best approach, based on the performance of the networks.

Acknowledgements

This work was supported by the framework agreement between the Universidad de León and INCIBE (Spanish National Cybersecurity Institute) under Addendum 01. We acknowledge NVIDIA Corporation with the donation of the TITAN Xp and Tesla K40 GPUs used for this research.

References

[1] P. Blanco-Medina, E. Fidalgo, E. Alegre, R. A. Vasco-Carofilis, F. Jañez-Martino, and V. F. Villar, "Detecting vulnerabilities in critical infrastructures by classifying exposed industrial control systems using deep learning," *Applied Sciences*, vol. 11, no. 1, 2021.

[2] W. A. Conklin, "It vs. ot security: A time to consider a change in cia to include resilienc," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*. IEEE, 2016, pp. 2642–2647.

[3] N. Sharma, V. Jain, and A. Mishra, "An analysis of convolutional neural networks for image classification," *Procedia Computer Science*, vol. 132, pp. 377–384, 2018.

[4] Z. Xiao, Y. Tan, X. Liu, and S. Yang, "Classification method of plug seedlings based on transfer learning," *Applied Sciences*, vol. 9, no. 13, 2019. [Online]. Available: https://www.mdpi.com/2076-3417/9/13/2725

[5] M. Hussain, J. J. Bird, and D. R. Faria, "A study on cnn transfer learning for image classification," in *UK Workshop on Computational Intelligence*. Springer, 2018, pp. 191–202.

[6] E. Fidalgo, E. Alegre, V. Gonzalez-Castro, and L. Fernández-Robles, "Boosting image classification through semantic attention filtering strategies," *Pattern Recognition Letters*, vol. 112, pp. 176–183, 2018.

[7] E. Fidalgo, E. Alegre, L. Fernández-Robles, and V. González-Castro, "Classifying suspicious content in tor darknet through semantic attention keypoint filtering," *Digital Investigation*, vol. 30, pp. 12–22, 2019.

[8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[9] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[12] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8697–8710.

[13] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.

[14] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.