Towards an In-depth Evaluation of the Performance, Suitability and Plausibility of Few-Shot Meta Transfer Learning on An Unknown Out-of-Distribution Cyber-attack Detection

Tosin Ige *, Christopher Kiekintveld , Aritran Piplai , Amy Wagler , Olukunle Kolade , Bolanle Matti

Posted Date: 10 September 2024

doi: 10.20944/preprints202409.0787.v1

Keywords: Few-Shot Learning; Meta Learning; Transfer Learning; Machine Learning; Deep Learning; Zero-Day; Malware; out-of-distribution attack; cyberattacks

Article

Towards an In-Depth Evaluation of the Performance, Suitability and Plausibility of Few-Shot Meta Transfer Learning on an Unknown Out-of-Distribution Cyber-Attack Detection

Tosin Ige [1,*], Christopher Kiekintveld [1], Aritran Piplai [1], Amy Wagler [2], Olukunle Kolade [3] and Bolanle Hafiz Matti [4]

[1] The University of Texas at El Paso, Texas, USA
[2] Dept. of Public Health Science, The University of Texas at El Paso, Texas, USA
[3] Office of Naval Research, United State Navy, Pentagon, USA
[4] Office of Network Security, Palo Alto Networks Inc, Texas, USA
* Correspondence: toige@miners.utep.edu

Abstract: The emergence of few-shot learning as a potential approach to address the problem of data scarcity by learning underlying pattern from a few training sample had so far given a mix-result especially on the suitability of model-agnostic meta learning, transfer learning, and optimization strategy to rapidly learn valid information from few sample. In this research, we did an in-depth evaluation of meta- learning to determine their plausibility and suitability for previously unknown cyberattack detection by first retrieving the original research artifacts of current state of the art meta learning to repeat the experiment with original dataset before replicating the experiment with two different malware dataset which had not been previously done with meta-transfer learning. On each of the experiments, meta-transfer learning gave good results on digital character recognition dataset but abysmal result on Malimg and Malevis malware images datasets thereby indicating its unreliability for detecting cyberattacks and the need for an improvement to the state-of-the-art meta transfer learning towards a better attack detection. Transfer learning performance is independent on imbalance and hence does not influence its performance since both malware dataset used for this experiment result in high validation loss and balancing the dataset doesn't result in reduced validation loss,the successful learning transfer seen on digital character recognition dataset is not unconnected to the fact that several languages have similar characters and digits thereby enhancing the successful learning transfer unlike malware datasets, and more importantly the finding that current meta-learning transfer approach doesn't generalize well on malware dataset and hence not suitable for detecting previously unseen out-of-distribution attack.

Keywords: few-shot learning; meta learning transfer learning; machine learning; deep learning; zero-day; malware; out-of-distribution attack; cyberattacks

1. Introduction

One of the most significant factor against state-of-the-art machine and deep learning models is the lack of proper generalization against distribution shifts [1] due to assumption of independence and identical distribution of test and validation samples which are not guarantee as distribution shifts occur over the time. We cannot ascertain that distribution of validation set used to evaluate model performance will be similar when deployed in real-time especially in the detection and classification of cyberattacks which comes in multiple dimension. This problem becomes more obvious as malware currently stands as the fastest-growing threat with 41% of enterprises witnessing a malware attack in just concluded year 2023 followed by phishing and ransomware attack. In year 2023 alone, the number of enterprises experiencing ransomware attacks increased by over 27% with only 8% of businesses attacked resorting to paying the ransom demands resulting in significant financial loss in addition to losses incurred due to downtime. There are 95 new families of malware in year 2022 alone averaging 1 new family every 4 days aside variants while year 2023 witness 43 new malware families averaging 1

 doi:10.20944/preprints202409.0787.v1

new malware family per week aside variants making emerging malware families a major threat to cybersecurity causing damages worth billions of Dollars annually.

The ease with which attacker creates new variants of malware coupled with the rate at which new variants are being release poses a real challenge both for their detection, identification and classification, reason being that machine learning and deep learning model are only effective in detecting previously seen variants during training [2,3]. To identify malware, traditional signature-based analysis requires effort of an expert to generate hand-designed signatures which is highly impossible to achieve in light of the ease and frequency at which new variants quickly emerge by the simple use of polymorphic or metamorphic techniques. Machine learning had grown to become the mainstream trend for efficient malware identification and signature generation due to its ability to learn from relevant malware features [4] from dataset and to efficiently use knowledge obtained from training for accurate prediction. In particular, deep learning based models from Recurrent Neural Network, Long-Short Term Memory, and Convolusional Neural Network had proven to be effective in the prediction of malware [5–8]

Despite the effectiveness of machine learning and deep learning in the identification of malware, One major problem with the resulting model is that they are only good at predicting malware that are previously seen during training provided and on the condition that the data is large enough for the model to learn from, they often performed poor against data not previously seen during training, hence their vulnerability to Zero-Day or an out-of-distribution malware or variants attacks. The ease at which new malware variants could be developed coupled with the high frequency rate in which previously unseen malware are being released to the public makes the problem more potent by worsening it more as it will be practically impossible for the rate of re-training a model to be equal to the rate at which new malware variants are being release, even if the strategy of re-training model is adopted for every new malware variants, getting enough data of every new variants will be nearly impossible to get posing another mountain of challenge considering the fact that machine learning and deep learning model learns from data hence, their heavy reliability on data.

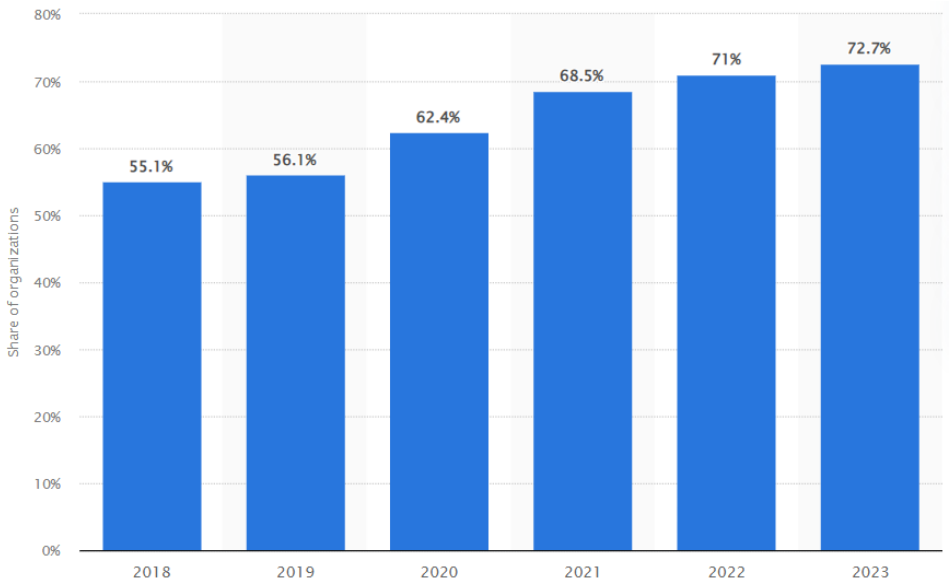

Figure 1. Annual share of organizations affected by ransomware attacks worldwide from 2018 to 2023 showing annual increase in the number of successful ransomware attack despite recent state-of-the-art research to few-shot learning

In order to resolve the problem of data scarcity for new malware variants in such a way that models will be able to learn from very few samples and improves its accuracy overtime, several state-of-the-art research had been carried out on few-short learning as a possible solution, but despite recent state-of-the-art research and development on few-shot learning approach as a cyber defense

 doi:10.20944/preprints202409.0787.v1

approaches to effectively protect computer system and critical infrastructure from malware, the ever increasing rate at which sophisticated variants of malware are created to exploit zero day vulnerabilities had continues to defy odds by making it difficult to classify previously unseen malware or its variants into correct families despite recent advancement in the state-of-the-art approach [9–11], the reason for the difficulty is being that the model had not previous seen the new variant during training thereby making the classification difficult to predict. Zero-day out-of-distribution prediction problem is exacerbated by sample scarcity due to challenges associated with the collection of a large volume of a newly detected variants or malware family to train a classifier which is extremely hard coupled with the unavoidability of over fitting as a result of using small sample of each malware family [12,13].

The increasing rate at which machine learning and deep learning models fails to predict previous unseeing out-of-distribution zero-day malware despite state-of-the-art research and development on few-shot learning put a big question mark on the effectiveness of current state-of-the-art few-shot learning (FSL) approach, and hence the importance and necessity of our investigation. In this research, we did an investigation into four (4) state-of-the-art few-shot learning to research their level of efficiency in detecting previously unseen malware variant or family. We started by pulling down the original research artifact comprising of the source code, data, etc, replicating the experiment with the original dataset, repeating the experiment with new malware dataset containing most recent malware variant that are not in existence when the original research was carried out, finally computing and comparing the MIN-MAX validation loss between training with the original dataset and new dataset containing most recent malware variants that does not previously exists. Our investigation aims the following contribution,

- To determine the plausibility and suitability of few-shot meta-transfer learning on previous unseen out-of-distribution attack
- To examine the impact of dataset imbalance on the performance of few-shot meta transfer learning
- To project future research direction on transfer learning as applicable to the area of cybersecurity as a plausible methods to address problem of unseen out-of-distribution attack

2. Related Work

One major challenge on both machine learning-based and Deep Learning-based model is data scarcity, and this directly impacts performance of these models in a proportionate manner i.e A large amount of data is needed to achieve exceptional performance from a Deep learning model [14]. DL models are extremely data-hungry models because they needed a huge amount of labelled data to automatically learnt data-representation by themselves. Unfortunately, these data are not available thereby creating a significant challenge as their scarcity have direct implication on the performance of DL models. In order to ensure that exceptional performance could be obtained from Deep learning models in the presence of data scarcity, several research work had been done, in this section, we will look at recent state-of-the-art research on Meta-Transfer Learning/ Meta-Learning (MTL/ML), Single-Shot Learning (SSL), Few-Shot Learning (FSL), and Zero-Shot Learning (ZSL).

2.1. Single-Shot Learning (SSL)

To address problem of data scarcity or scantiness, One-shot learning uses a conceptualized approach whereby machine learning and deep learning models learns from only a single sample from each represented class thereby given the model capability to recognize and generalize patterns based on previously seen single example [15–17] In deep learning, it is not unusual to use discriminative embeddings or generative models as an alternative way for one-shot learning in the presence of data scarcity, but while they may be plausible for some classification tasks, the fact that they requires large amount of data makes them unsuitable for one-shot learning. Bertinetto et al. [18] proposed the learning of parameters of deep-learning in one-shot by constructing a second deep network called learnet which has the capability to predict pupil network parameter from a single sample, hence

 doi:10.20944/preprints202409.0787.v1

were able to obtained a forward one-shot learner which minimizes the one-shot objective through an end-to-end training.

Anton et al. [19] proposed a Deep Reinforcement One-shot Learning (DeROL) framework by training a deep-Q network with the sole aim of achieving policy that would be oblivious to unseen classes in the validation set, then each state of the one-shot learning process is further map to operation actions based on the trained deep-Q network which aids the maximization of the objective function

2.2. Few-Shot Learning (FSL)

Unlike Single-shot, few-shot learning (FSL) framework aims to address problems of data scarcity or scantiness by using few samples from each class label[20–22]. It leverages a large number of similar tasks so as to adapt a base-learner to a new task for which only few samples are available. Considering the fact that this framework uses few samples and deep learning models tends to overfit in the absence of huge amount of data, meta-learning uses of shallow neural network (SNN) to address the problem of overfitting that might arise from few samples. Qianru et al. [23] proposed meta-transfer learning (MTL) as a few-shot learning method whereby the model learns to adapt a neural network to a few-shot learning task by conducting experiments with (5-class, 1-shot) and (5-class, 5-shot) recognition tasks on two different challenging few-shot learning tasks, the proposed few-shot learning is a meta-learning based whereby the deep model is trained on multiple tasks and the learning is transfered by shifting, customizing and scaling functions of DNN weights to suit individual tasks.

Based on the intuition that some internal representations are more transferable than the other, Chelsea et al. [24] proposed a few-shot learning method that can learn the parameters of any model trained with gradient descent in a way that prepares the model for a fast adaptation such that it can learn new task from few samples. For model that is agnostic, the parameters are trained in a way that small amount of gradients steps coupled with a small amount of training sets from previously unseen tasks will produce good generalization for the task.

2.3. Zero-Shot Learning (ZSL)

In Zero-shot learning (ZSL), DNN model is trained both to recognize and categorize unseen objects. During validation, the learner is made to observe and predict samples from classes which were not previous observed during training to the class they belong to [25–27]. It uses some form of auxiliary information to associate observed and non-observed classes thereby making it possible to encode observable properties that distinguish an object from another, for instance, a previously trained DNN model to recognize a horse but not seen a zebra during training can still recognize both horse and zebra through their differences i.e the previously encoded differences. Considering that previously seen train set and unseen test validation set are mutually exclusive, zero-shot learning first map a relationship between unseen and seen class and then use the relationship to determine an unseen class during validation

Vinay et al. [28] proposed a generative framework for generalized zero-shot learning with a disjointed training and test classes through a feedback-driven mechanism in which the discriminator called multivariate regressor learns to map the generated exemplars to the corresponding class attribute vectors to create an improved generator. The proposed framework is variational auto-encoder based architecture having a probabilistic encoder and emphconditional decoder thereby given the ability to generate samples from seen and unseen classes using each class attributes. The newly generated exemplar is then used to train any classification model.

2.4. Meta-Transfer Learning/ Meta-Learning (MTL/ML)

and this is achieved by training the model on multiple tasks and then the learning scaling and shifting functions of DNN weights for each task as a way of transferring the learning. This enables the model not only to leverage knowledge from previous domain but to also adapt and generalize to a range of previously unseen domain.Meta-learning was introduced as a framework to better alleviate

 doi:10.20944/preprints202409.0787.v1

the problem of data scarcity and scantiness as well as a means for model to learn with few labeled samples with the ability to generalize and transfer those learning to previously unseen tasks [24,29,30] The key idea of meta learning as a framework to address problems associated with few-shot learning is to leverage a substantial number of few-shot with related tasks so as to learn how a base learner can be adjusted for a new tasks with few available label samples. Howbeit, one problem associated with deep learning is their tendency to overfit in the presence few labeled samples, meta-learning typically adopt the rule of shallow neural network (SNNs) to address this problem thereby making it ineffective.

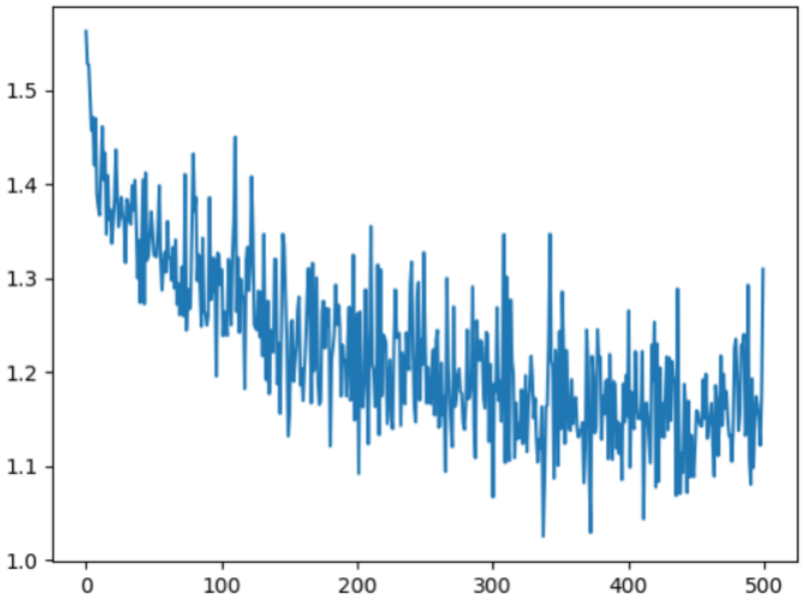

Figure 2. Unexpected high validation loss on Malevis and Malimg Dataset due to Due to failure in Learning Transfer (Model-Agnostic Meta Learning) to unseen Dataset

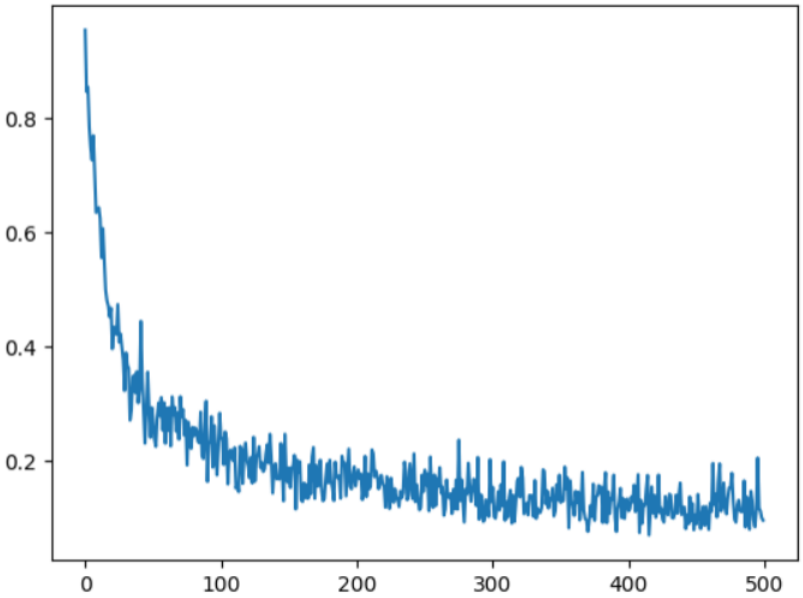

Figure 3. Low validation loss on meta-transfer on Digital Character Recognition Dataset due to successful Learning Transfer (Model-Agnostic Meta Learning)

 doi:10.20944/preprints202409.0787.v1

3. Research Methodology

3.1. Dataset

3.1.1. Malimg Dataset

Being a publicly available dataset in Kaggle, Malimg dataset [31] contains a total number of 9435 executable malwares taken from 25 malware families which were previously disarmed before being converted to 32 by 32 images based on the nearest neighbor interpolation. Each Malware family in the dataset was shaped by transforming their binaries into matrix as a result of the conversion of malware binaries to 8-bit vectors leading to a 2D matrix of malware images

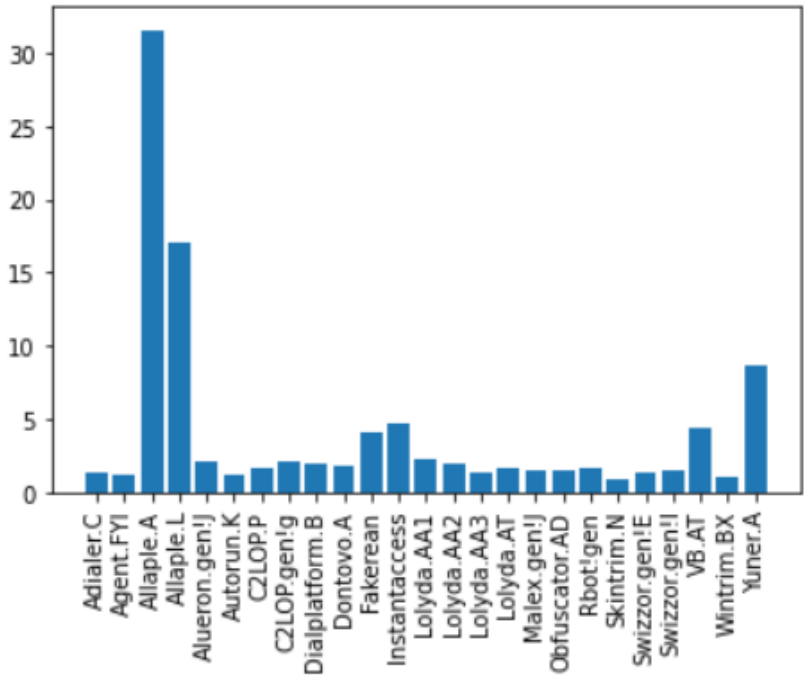

Figure 4. Malware Family Distribution in Malimg Dataset

3.1.2. Malevis Dataset

Malevis dataset consist of 26 malware family out of which one family represent "benign" or "legitimate" samples while the remaining 25 classes consist of different families of malware. The original binary images had been previously extracted from the malware files in 3 channels of RGB format before being resized into 224 by 224 and 300 by 300 dimension pixels while retaining the original number of channels in RGB format. Malevis dataset contains a total of 14,226 malware samples spanning 26 families of malware, and out of which 9100 are training samples while are 5126 validation samples in 3 channels format, the fact that the dataset makes provision for fairly larger legitimate malware samples for validation purpose makes the dataset suitable for the experiment

 doi:10.20944/preprints202409.0787.v1

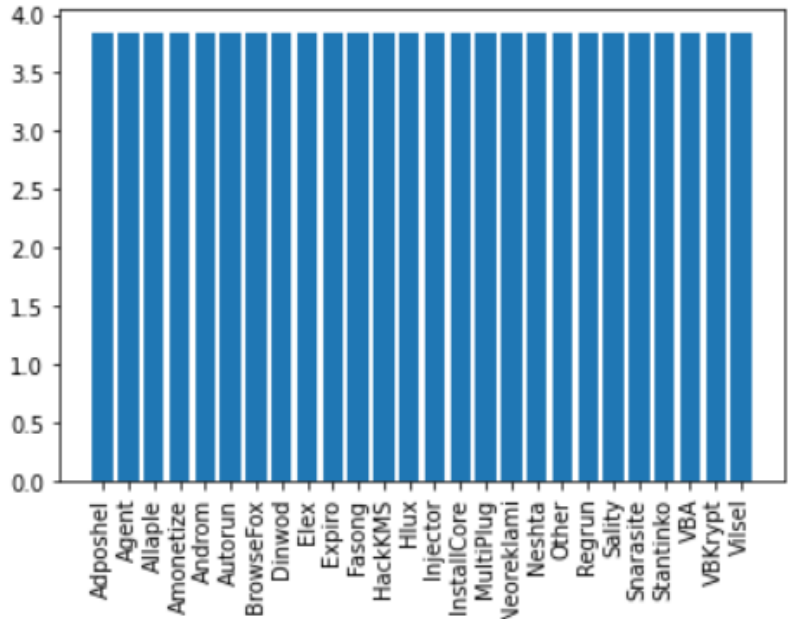

Figure 5. Malware Family Distribution in Malevis Dataset

3.2. Experimental Set-Up

As part of the experimental setup to evaluate the suitability of meta-transfer learning in transferring whatever is learned on an out-of-distribution malware to evaluate their plausibility for detecting a previously unseen attack. The following steps were carried out

- Downloading the original research artifact
- Replicating the experiment on first attempt with Digital Character Recognition Dataset
- Replication the experiment on Second attempt with Malimg malware dataset
- Replication the experiment on third attempt with Malevis malware dataset

During first replication attempt with digital character recognition dataset, some classes were intentionally omitted from the training set and added to the validation set and each represented class is represented by 5 samples each. Classes that are ommited in the validation set were placed in the training set before running our code on first replication attempt, the result of which was the successful learning transfer with an accuracy of 92% and extremely low validation loss in each epoch.

 doi:10.20944/preprints202409.0787.v1

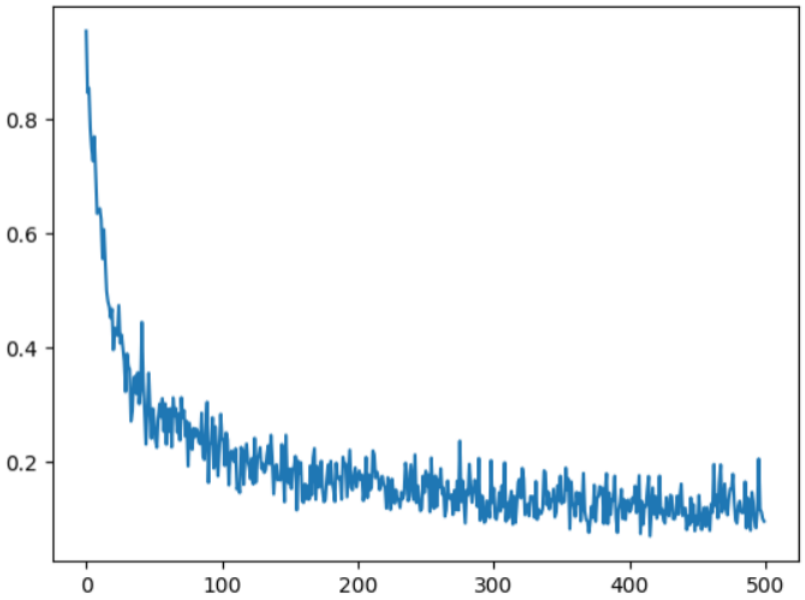

Figure 6. Low validation loss on meta-transfer on Digital Character Recognition Dataset due to successful Learning Transfer (Model-Agnostic Meta Learning)

During the second replication attempt, each of the steps and processes taken in the first replication attempt was repeated, only that we change the dataset from that of digital character recognition dataset to Malimg dataset to see if the learning could be transfer to unseen dataset and determine the plausibility and suitability for detection of an unfamiliar or previously seen cyberattacks. We unexpectedly got a very high loss of around 1 and negligible accuracy, as we kept on fine-tunning the parameters on each epoch, result got improved but the improvement was negligible as the loss was still around 1. Seeing the unexpected result obtained by changing to a cybersecurity dataset in malimg, and considering the imbalance of malimg dataset, it becones a necessity to repeat the experiment with another cybersecurity dataset in which all classes are balanced and evenly represented by same number of samples. This led to the third replication attempt for which we opted for malevis dataset. Previous steps were repeated on third attempt as usual with only differences in dataset and observation but results remains pretty the same on both Malimg and Malevis malware dataset.

 doi:10.20944/preprints202409.0787.v1

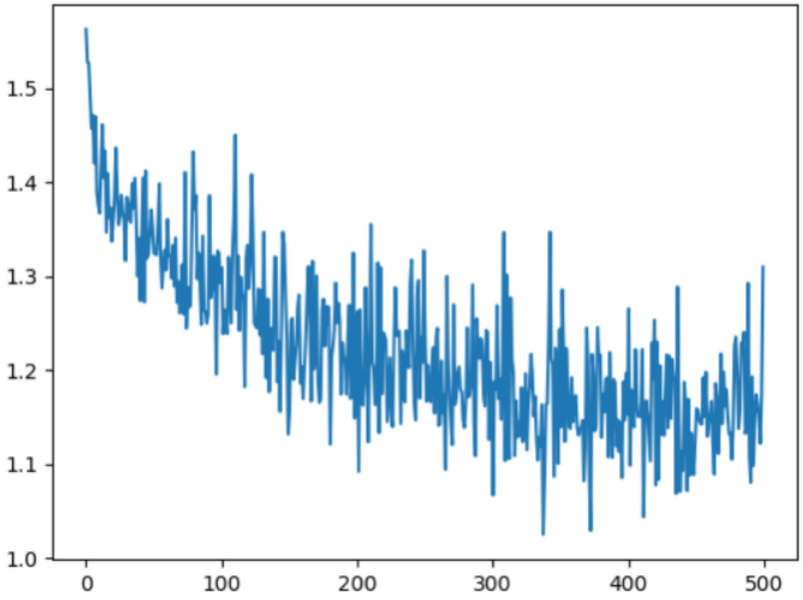

Figure 7. Unexpected high validation loss on Malevis and Malimg Dataset due to Due to failure in Learning Transfer (Model-Agnostic Meta Learning) to unseen Dataset

Having similar result from the two malware dataset but different from the digital character recognition dataset which gives a low validation loss as compared against the second and third replication attempt calls for further inspection considering that all the dataset used for the experiment are image datasets. Further observation leads to the following assertions;

- Transfer learning performance is independent on imbalance and hence does not influence its performance since both malware dataset (Malimg and Malevis) has high validation loss
- By image inspection, We were able to assert that successful learning transfer on digital character recognition dataset is not unconnected to the fact that several languages have similar characters and digits thereby enhancing the successful transfer unlike malware datasets
- Current meta-transfer learning approach doesn't generalize well on malware dataset and hence not suitable for detecting previously unseen out-of-distribution attack

4. Conclusion

In this research, we did an in-depth evaluation of meta- learning to determine their plausibility and suitability for previously unknown cyberattack detection by first retrieving the original research artifacts of current state of the art meta learning to repeat the experiment with original dataset before replicating the experiment with two different malware dataset which had not been previously done with meta-transfer learning. On each of the experiments, meta-transfer learning gave good results on digital character recognition dataset but abysmal result on the two malware images datasets. Hence, we asserts that (1) Transfer learning performance is independent on imbalance and hence does not influence its performance since both malware dataset used for this experiment result in high validation loss and balancing the dataset doesn't result in reduced validation loss (2) the successful learning transfer seen on digital character recognition dataset is not unconnected to the fact that several languages have similar characters and digits thereby enhancing the successful learning transfer unlike malware datasets, and (3) more importantly the finding that current meta-learning transfer approach doesn't generalize well on malware dataset and hence not suitable for detecting previously unseen out-of-distribution attack.

While current meta-transfer learning approach might be suitable for certain tasks, they are not reliable for detecting previously unseen out-of-distribution attacks and the need for an improvement

 doi:10.20944/preprints202409.0787.v1

to the existing state-of-the-art meta transfer learning approach towards an out-of-distribution attack detection as a future research direction.

References

1. Yu, H.; Liu, J.; Zhang, X.; Wu, J.; Cui, P. A survey on evaluation of out-of-distribution generalization. *arXiv preprint arXiv:2403.01874* **2024**.
2. Pillai, S.E.V.S.; Polimetla, K. Mitigating DDoS Attacks using SDN-based Network Security Measures. 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS). IEEE, 2024, pp. 1–7.
3. Vallabhaneni, R.; Vaddadi, S.A.; Pillai, S.; Addula, S.R.; Ananthan, B. Detection of cyberattacks using bidirectional generative adversarial network. *Indonesian Journal of Electrical Engineering and Computer Science* **2024**, *35*, 1653–1660.
4. Ye, Y.; Li, T.; Adjeroh, D.; Iyengar, S.S. A survey on malware detection using data mining techniques. *ACM Computing Surveys (CSUR)* **2017**, *50*, 1–40.
5. Vasan, D.; Alazab, M.; Wassan, S.; Safaei, B.; Zheng, Q. Image-Based malware classification using ensemble of CNN architectures (IMCEC). *Computers & Security* **2020**, *92*, 101748.
6. Wang, P.; Tang, Z.; Wang, J. A novel few-shot malware classification approach for unknown family recognition with multi-prototype modeling. *Computers & Security* **2021**, *106*, 102273.
7. Okomayin, A.; Ige, T. Ambient Technology & Intelligence. *arXiv preprint arXiv:2305.10726* **2023**.
8. Adewale, S.; Ige, T.; Matti, B.H. Encoder-decoder based long short-term memory (lstm) model for video captioning. *arXiv preprint arXiv:2401.02052* **2023**.
9. Ige, T.; Kiekintveld, C.; Piplai, A.; Wagler, A.; Kolade, O.; Matti, B.H. An In-Depth Investigation into the Performance of State-of-the-Art Zero-Shot, Single-Shot, and Few-Shot Learning Approaches on an Out-of-Distribution Zero-Day Malware Attack Detection **2024**.
10. Ige, T.; Marfo, W.; Tonkinson, J.; Adewale, S.; Matti, B.H. Adversarial sampling for fairness testing in deep neural network. *arXiv preprint arXiv:2303.02874* **2023**.
11. Ige, T.; Kiekintveld, C. Performance comparison and implementation of bayesian variants for network intrusion detection. 2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings). IEEE, 2023, pp. 1–5.
12. Ige, T.; Kiekintveld, C.; Piplai, A. An investigation into the performances of the state-of-the-art machine learning approaches for various cyber-attack detection: A survey. *arXiv preprint arXiv:2402.17045* **2024**.
13. Ige, T.; Kiekintveld, C.; Piplai, A. Deep Learning-Based Speech and Vision Synthesis to Improve Phishing Attack Detection through a Multi-layer Adaptive Framework. *arXiv preprint arXiv:2402.17249* **2024**.
14. Song, Y.; Wang, T.; Cai, P.; Mondal, S.K.; Sahoo, J.P. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities. *ACM Computing Surveys* **2023**, *55*, 1–40.
15. Wang, S.; Xu, M.; Sun, Y.; Jiang, G.; Weng, Y.; Liu, X.; Zhao, G.; Fan, H.; Li, J.; Zou, C.; others. Improved single shot detection using DenseNet for tiny target detection. *Concurrency and Computation: Practice and Experience* **2023**, *35*, e7491.
16. Zhu, W.; Zhang, H.; Eastwood, J.; Qi, X.; Jia, J.; Cao, Y. Concrete crack detection using lightweight attention feature fusion single shot multibox detector. *Knowledge-Based Systems* **2023**, *261*, 110216.
17. Lew, A.J.; Buehler, M.J. Single-shot forward and inverse hierarchical architected materials design for nonlinear mechanical properties using an Attention-Diffusion model. *Materials Today* **2023**, *64*, 10–20.
18. Bertinetto, L.; Henriques, J.F.; Valmadre, J.; Torr, P.; Vedaldi, A. Learning feed-forward one-shot learners. *Advances in neural information processing systems* **2016**, *29*.
19. Puzanov, A.; Zhang, S.; Cohen, K. Deep reinforcement one-shot learning for artificially intelligent classification in expert aided systems. *Engineering Applications of Artificial Intelligence* **2020**, *91*, 103589.
20. Jeong, J.; Zou, Y.; Kim, T.; Zhang, D.; Ravichandran, A.; Dabeer, O. Winclip: Zero-/few-shot anomaly classification and segmentation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 19606–19616.
21. Dooley, S.; Khurana, G.S.; Mohapatra, C.; Naidu, S.V.; White, C. Forecastpfn: Synthetically-trained zero-shot forecasting. *Advances in Neural Information Processing Systems* **2024**, *36*.

 doi:10.20944/preprints202409.0787.v1

22. Luo, X.; Wu, H.; Zhang, J.; Gao, L.; Xu, J.; Song, J. A closer look at few-shot classification again. International Conference on Machine Learning. PMLR, 2023, pp. 23103–23123.
23. Sun, Q.; Liu, Y.; Chua, T.S.; Schiele, B. Meta-transfer learning for few-shot learning. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 403–412.
24. Finn, C.; Abbeel, P.; Levine, S. Model-agnostic meta-learning for fast adaptation of deep networks. International conference on machine learning. PMLR, 2017, pp. 1126–1135.
25. Wang, Q.; Liu, L.; Jing, C.; Chen, H.; Liang, G.; Wang, P.; Shen, C. Learning conditional attributes for compositional zero-shot learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 11197–11206.
26. Liu, M.; Li, F.; Zhang, C.; Wei, Y.; Bai, H.; Zhao, Y. Progressive semantic-visual mutual adaption for generalized zero-shot learning. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023, pp. 15337–15346.
27. Guo, J.; Guo, S.; Zhou, Q.; Liu, Z.; Lu, X.; Huo, F. Graph knows unknowns: Reformulate zero-shot learning as sample-level graph recognition. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, Vol. 37, pp. 7775–7783.
28. Verma, V.K.; Arora, G.; Mishra, A.; Rai, P. Generalized zero-shot learning via synthesized examples. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4281–4289.
29. Sun, Q.; Liu, Y.; Chen, Z.; Chua, T.S.; Schiele, B. Meta-transfer learning through hard tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2020**, *44*, 1443–1456.
30. Park, S.J.; Han, S.; Baek, J.W.; Kim, I.; Song, J.; Lee, H.B.; Han, J.J.; Hwang, S.J. Meta variance transfer: Learning to augment from the others. International conference on machine learning. PMLR, 2020, pp. 7510–7520.
31. Nataraj, L.; Karthikeyan, S.; Jacob, G.; Manjunath, B.S. Malware images: visualization and automatic classification. Proceedings of the 8th international symposium on visualization for cyber security, 2011, pp. 1–7.