A Weak Coupling of Semi-Supervised Learning with Generative Adversarial Networks for Malware Classification

Shuwei Wang, Zhengwei Jiang, Qiuyun Wang, Xunren Wang, Rongqi Jing*

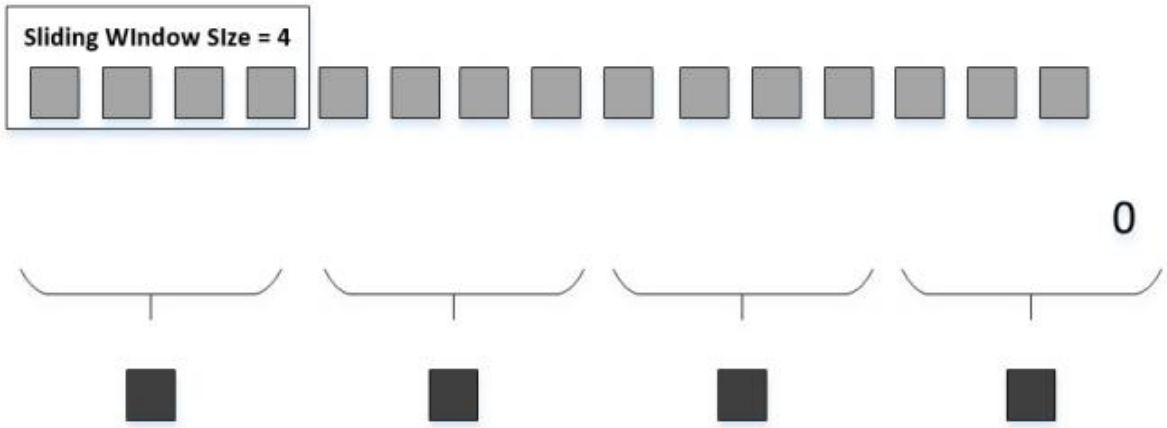
CONTENTS

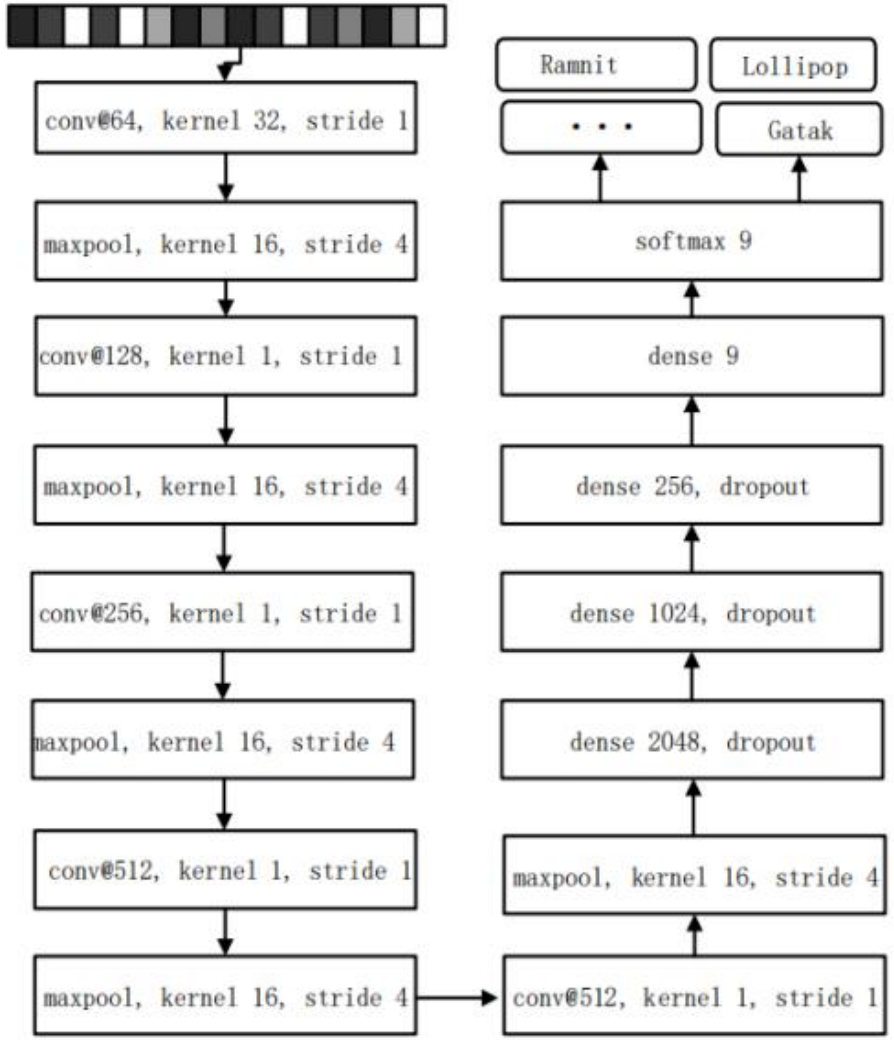
Introduction

- Malware classification helps to understand its purpose and is also an important part of attack detection.
- In this paper, we propose :
  - an improved malware image rescaling algorithm (IMIR) based on local mean algorithm.
  - a neural network structure based on VGG model, which is suitable for image classification.
  - a novel method to train the deep neural network by Semi-supervised Generative Adversarial Network (SGAN).
  - to retain the weak links of supervised part and unsupervised part of SGAN.

Proposed Method

- Improved malware image rescaling algorithm (IMIR) extends the sampling range to the edge of sliding window.
- A one-dimensional convolutional neural network (1D-CNN) is constructed by using the VGG model.

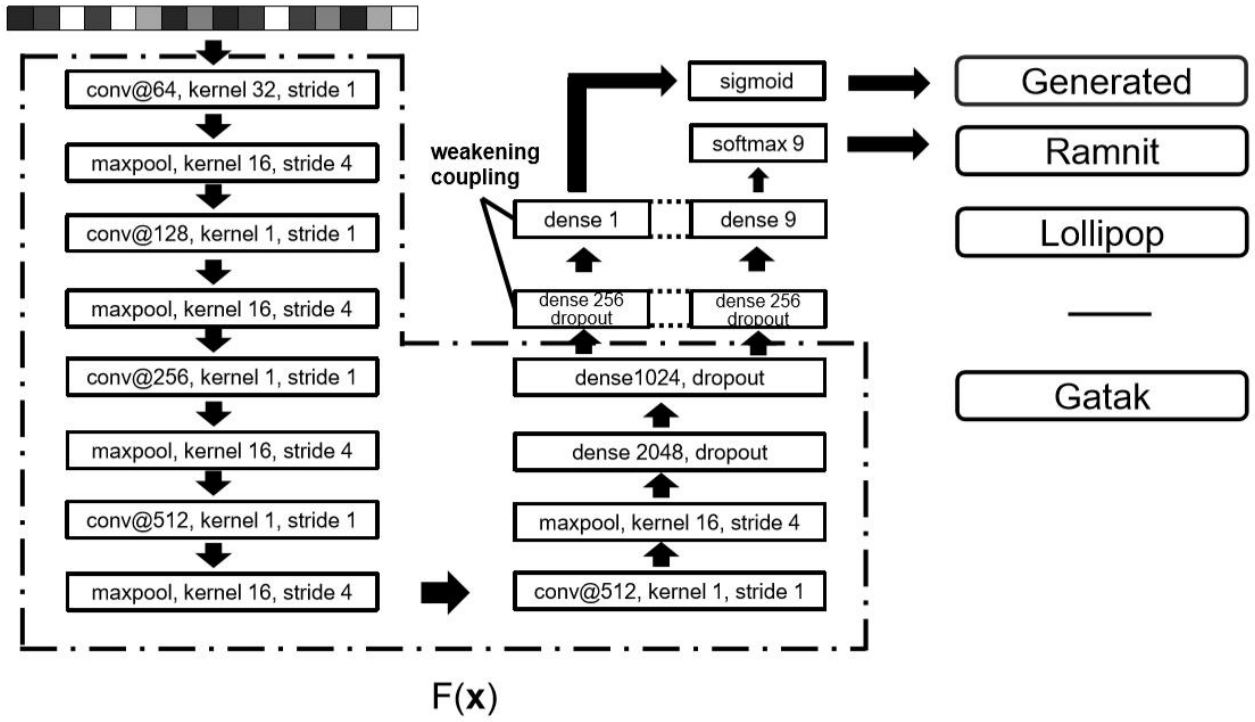




Proposed Method

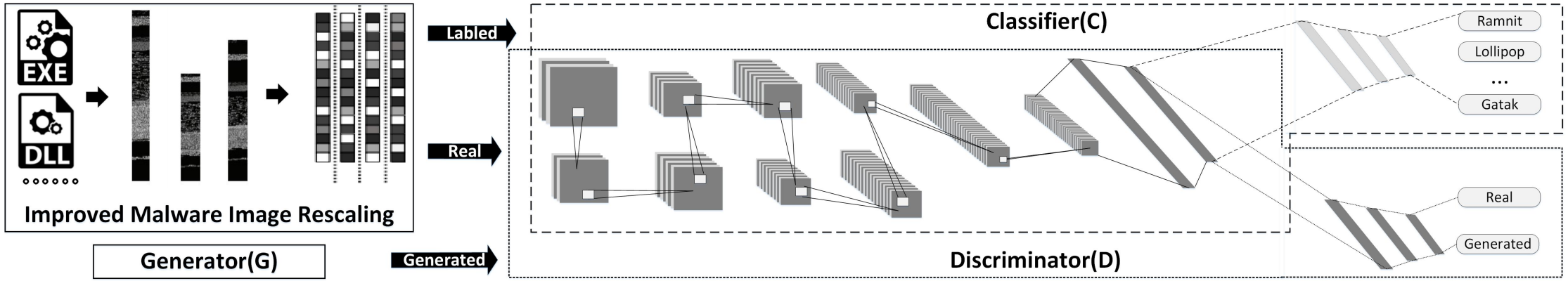
- SGAN uses the supervised learning with labeled samples to train model for judging categories, and the unsupervised learning with unlabeled samples to train model for judging true or false.

- We found the separation of classifier **C** and discriminator **D** can further strengthen the effect of classifier **C** based on semi-supervised learning.



Proposed Method

- We present our improved method to build a malware classifier with a weak coupling of semi-supervised learning based on GAN.



- And the equation represents the goal of training in our method.

$$L_{supervised} = -\mathrm{E}_{x,y\sim labeled} \ln P_C(y|F(x))$$

$$L_{unsupervised} = -\mathrm{E}_{x\sim real} \ln P_D\big(F(x)\big) - E_{z\sim N(0,1)} \ln\big(1 - P_D(F(G(z)))\big)$$

$$\max_G \min_{C,D,F} L = \begin{cases} \min\limits_{C,F} L_{supervised} \\ \min\limits_{D,F} L_{unsupervised} \\ \max\limits_{G} -\mathrm{E}_{z\sim N(0,1)} \ln\big(1 - P_D\ F(G(z))\big) \end{cases}$$

Environment

- **Server Hardware:**

✓ **Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz × 2.**

✓ **94GB memory.**

✓ **Nvidia Tesla P100 graphics card × 1.**

- **Adam Optimizer:**

➢ **batch size: 64.**

➢ **epochs: 200.**

➢ **keep prob: 0.85.**

➢ **learning rate: 2e-4.**

➢ **bn momentum: 0.9.**

➢ **bn epsilon: 1e-5.**

Experiment

- In the experiment group, we evaluated the performance of SGAN and the weak coupling method by 5-fold cross-validation on the adjusted dataset.
- The recalls of each family are all higher than 93.75%, most of which exceed 98.00%.

Family Name	Number
Ramnit	1541
Lolipop	2478
Kelihoss.ver3	2942
Vundo	475
Simda	42
Tracur	751
Kelihos._verl	398
ObfuscalorACY	1228
Gatak	101

True\Pred	Ramnit	Lolipop	Kelihoss.ver3	Vundo	Simda	Tracur	Kelihos._verl	ObfuscalorACY	Gatak
Ramnit	**97.73%**	0.00%	0.00%	0.00%	0.00%	1.29%	0.64%	0.32%	0.00%
Lolipop	0.00%	**98.80%**	0.00%	0.00%	0.00%	0.20%	0.20%	0.20%	0.60%
Kelihoss.ver3	0.00%	0.00%	**99.66%**	0.00%	0.00%	0.00%	0.17%	0.17%	0.00%
Vundo	0.00%	0.11%	0.00%	**98.94%**	0.00%	0.00%	0.00%	0.00%	0.00%
Simda	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%
Tracur	0.00%	0.00%	0.00%	0.66%	0.00%	**98.68%**	0.66%	0.00%	0.00%
Kelihos._verl	1.25%	0.00%	0.00%	0.00%	0.00%	1.25%	**93.75%**	0.00%	3.75%
ObfuscalorACY	2.03%	0.81%	0.00%	0.41%	0.41%	0.41%	0.00%	**95.12%**	0.81%
Gatak	0.00%	0.50%	0.00%	0.00%	0.00%	0.00%	0.00%	0.99%	**98.52%**

Conclusion

- We improve the effective with deep learning from three aspects: sample feature extraction, neural network structure, and data labeling.

- After experimental verification, the three work together to reduce the time cost of model construction and use, improve update efficiency, and enhance timeliness.

- It is commonly known that the loss of GAN is hard to optimize, while WGAN guarantees the stability of training, so we hope to adapt the loss of WGAN to SGAN in future.

- In addition, considering that the “mean” operation in IMIR may be a bit rough, a better summarizing operation for every sampling window needs further study.


2020
Thanks