融合DRAE与SVM的网页防篡改检测

周长建
东北农业大学

提纲：

1、研究背景

2、模型设计

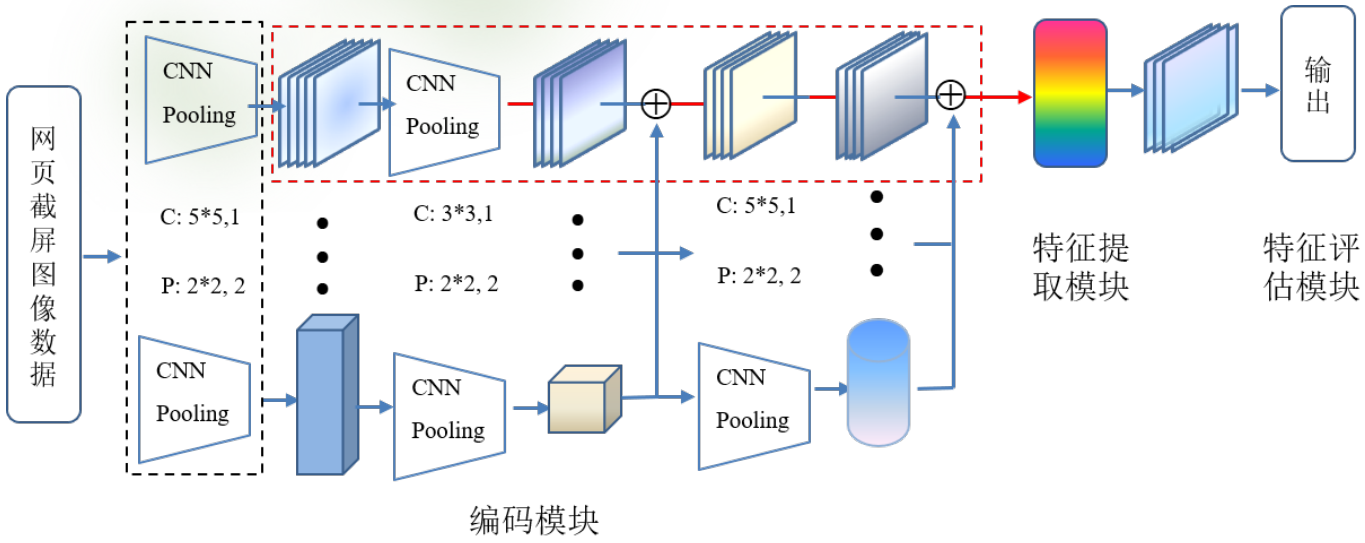
3、实验分析

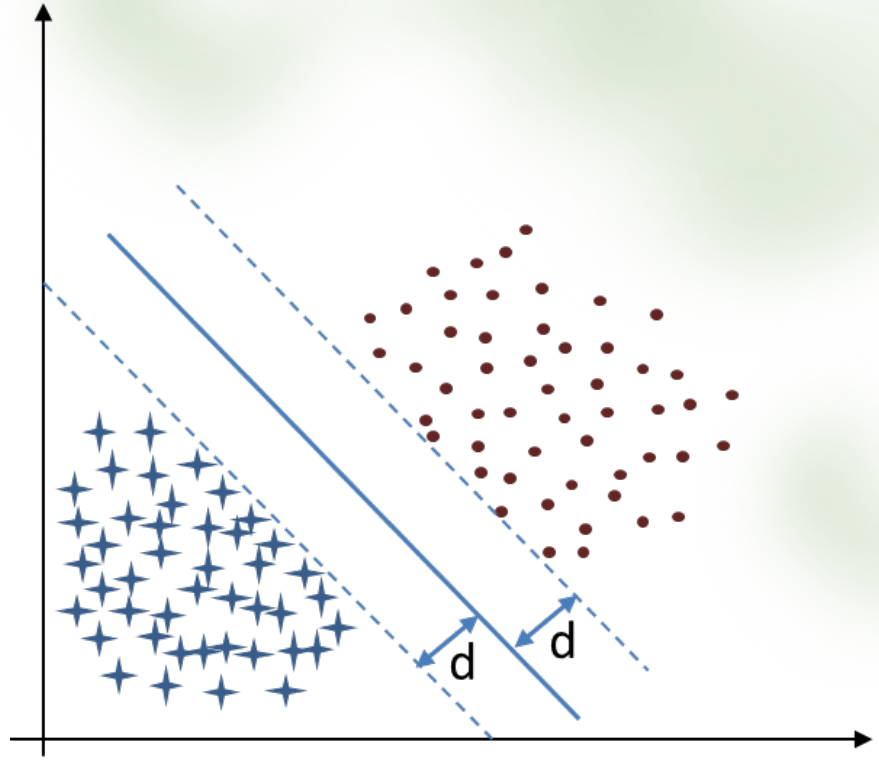
4、结果与讨论

研究背景

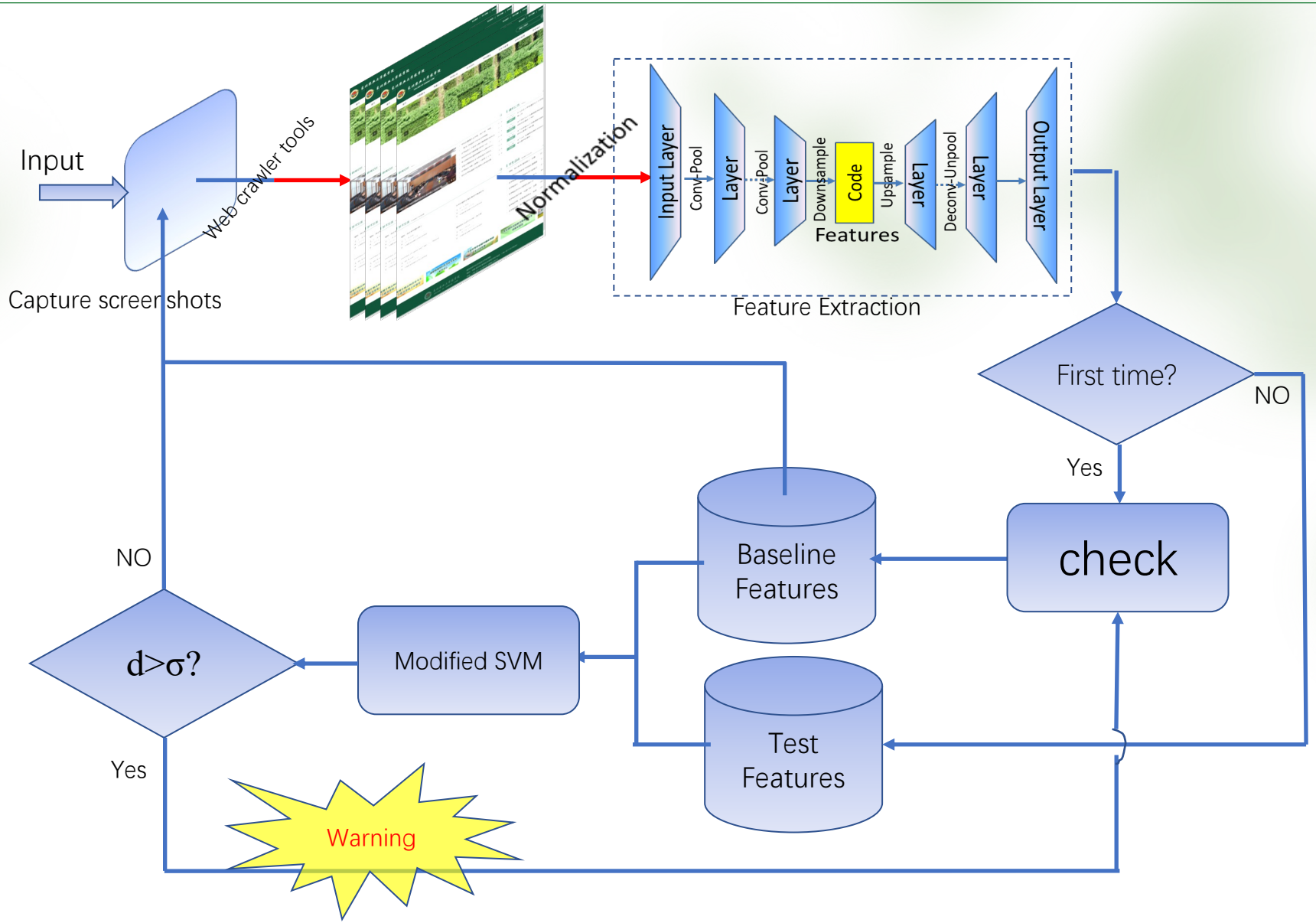
针对传统的网络安全研究如入侵检测、流量分析和主动防御等方法需要较强的网络安全相关知识、大量的网络训练数据，以及较高的研究门槛的问题，该研究提出一种基于深度残差自动编码器(deep residual auto-encoder, DRAE）与支持向量机（SVM）相结合的网页防篡改检测模型，该模型用DRAE提取网页图像特征，并输入SVM分类器判别网页是否被篡改。经过在东北农业大学范围内实验验证，结果表明，使用该模型进行网页检测的准确率高达95%，高于现有检测方法。

模型设计

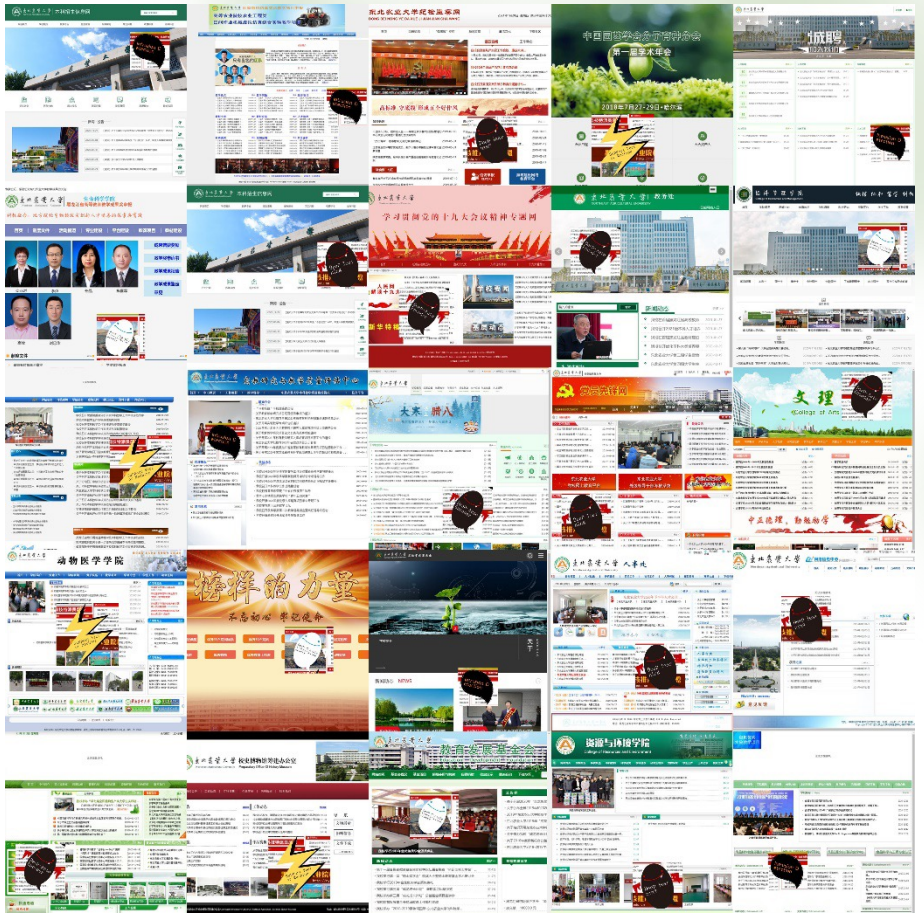

网页截屏图像数据
CNN Pooling
C: 5*5,1
P: 2*2, 2
CNN Pooling
C: 3*3,1
P: 2*2, 2
CNN Pooling
C: 5*5,1
P: 2*2, 2
CNN Pooling
CNN Pooling
CNN Pooling
编码模块
特征提取模块
输出
特征评估模块


d
d

模型设计


Input
Web crawler tools
Capture screen shots
Normalization
Input Layer
Conv-Pool
Layer
Conv-Pool
Layer
Downsample
Code
Features
Upsample
Layer
Deconv-Unpool
Layer
Output Layer
Feature Extraction
First time?
NO
Yes
check
Baseline Features
Test Features
Modified SVM
d>σ?
NO
Yes
Warning

实验设计



实验设计

表 1　计算资源

Table 1 Computing Resource

软硬件资源	参数
操作系统	Cent OS 7.5
CPU	8×Intel(R) Xeon(R) Silver 4216 CPU @ 2.10GHz
开发语言	Python 3.8
GPU	NVIDIA GeForce RTX 2080 Ti (11G)
内存	32G
硬盘	320G
深度学习框架	Tensorflow2.3.1, Cuda10.1

实验设计

表 2　训练参数

Table 2 Training parameters

参数	值
输入尺寸	512×512×3
训练批次	16
优化函数	RMSProp
激活函数	ReLU, Sigmoid
损失函数	binary_crossentropy
迭代次数	150
SVM 核函数	Linear

结果分析

表 3　实验结果

Table 3 Experimental results

模型	P/%	R/%	FS	Acc/%
KNN	68	61	0.64	65
AlexNet	73	64	0.68	68
DenseNet-121	76	66	0.70	70
PCA-SVM	92	85	0.88	89
Auto-encoder+SVM	97	87	0.92	92
.SnapCatch[10]	90	92	0.91	92
IntruDTree[11]	89	94	0.91	94
Dehghani, M [12]	91	94	0.92	93
本方法	96	94	0.94	95

结果分析

表 4　消融实验

Table 4 Ablation Experiment

模块	P/%	R/%	FS	Acc/%
Deep ResNet-50	72	67	0.69	71
SVM	100	83	0.82	85
DRAE + SVM	96	94	0.94	95

总结

网络安全已经上升为国家安全战略，与国土安全同等重要[13]，一套可在线值班系统以节约人力成本，提升工作效率，具有一定的应用价值。该研究提出的基于DRAE与SVM相结合的网页防篡改检测模型有效的利用DRAE和SVM 的优势，即利用了Auto–Encoder的自编码架构，Deep Residual Network 的特征表达能力和SVM强鲁棒性分类算法，这三种方法相结合进行小范围域名范围内网页防篡改检测，具有较高的准确性和使用价值。但由于大规模网站网络结构复杂，该模型难以适应大规模网站情况。其次，该模型对网页脚本飞扬的情况表现不够友好，需要人工干预的次数增加。为此，项目组成员在未来工作中会在这两个方向继续进行深入研究。

谢谢