Classifying Malicious Websites With Image Recognition Models

Akbar Qureshi

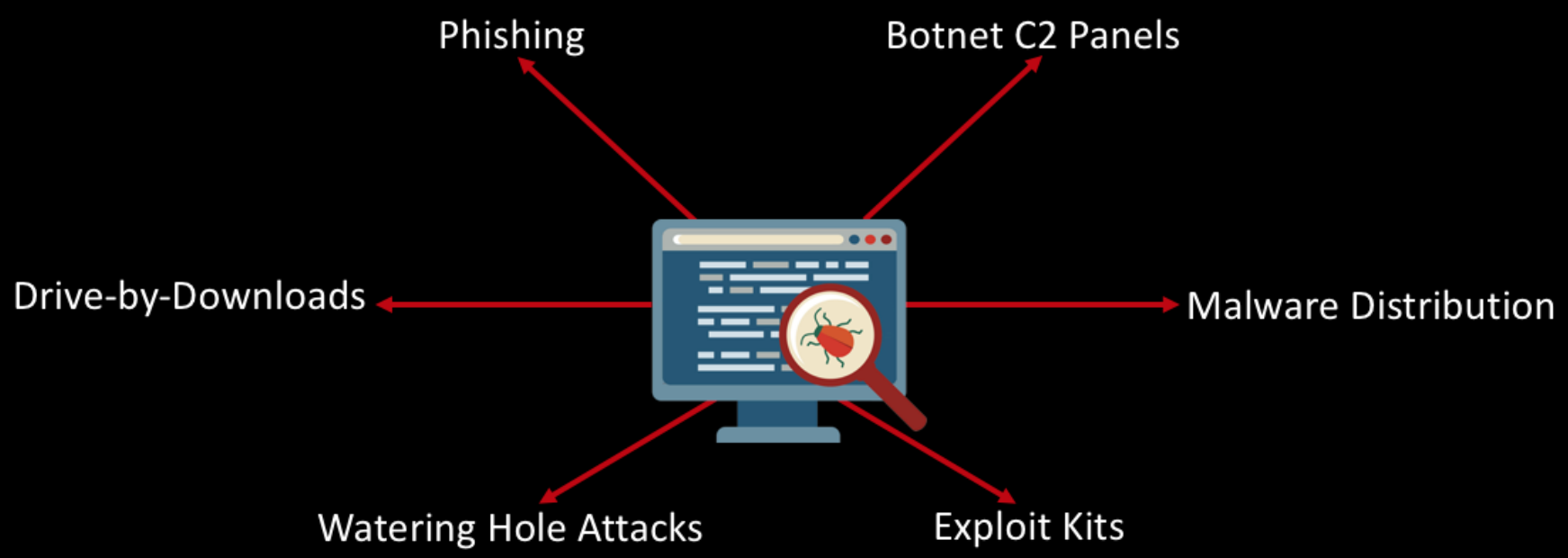
About Me

- Over 15 years of experience in cyber security
- Provided security advisory and technical services to both private and government sectors
- Interests – Cyber Counter Intelligence, Offensive Security, Artificial Intelligence, ICS \ SCADA Security, Cyber Network Defense

Agenda

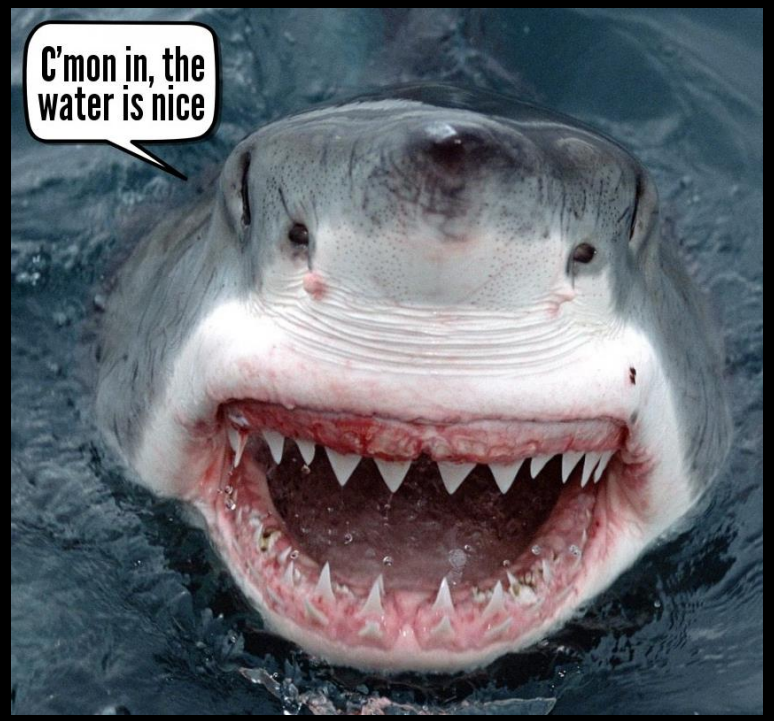
- Keep it simple
- Try and not bore you
- Not impress you with scientific formulas
- Keep the presentation interactive

Malicious Websites – They Come In All Shapes & Flavors



Legit Today Malicious Tomorrow

- Legitimate websites are attacked and compromised on a daily basis
- Unsuspecting users visit these trusted “legit” sites and get infected with malware



Who is Getting Compromised?

- Both “big” and “small” name company websites are compromised

  Example - Online payment systems of British Airways, Feedify, and Newegg compromised by the Magecart threat group

- According to Trustwave 100% of web applications remain vulnerable to attack – *2018 Global Security Report*

Time To Fine Tune the Risk Equation

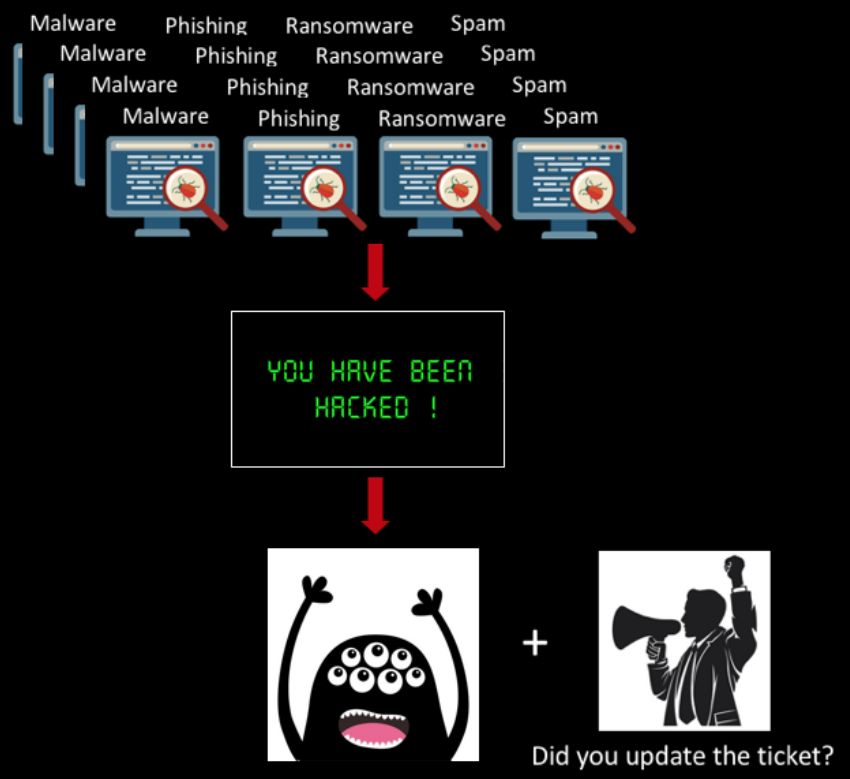
New Way

Risk = Assume we are already compromised x Now what

Life of a Cyber Warrior – With All The Maliciousness

- Continuous barrage of alerts
- Overworked
- Alert fatigue

Insufficient Incident Analysis and Reporting



How Do We Optimize Analysis?

In most cases we are:

- Hiring skilled staff
- Buying expensive security tools
- Increasing the cyber security budget and again buying more tools

What we should focus on:

- Effective use of cutting-edge technology, especially the ones that can out perform humans
- In our case we are using image recognition to classify & label malicious websites

How Does Image Recognition Help Us?

- Malicious website classification can be automated with the power of deep learning
- Incident response analysis can be optimized
- Malicious website threat label data can be enhanced

  Example - badsite.com is categorized as botnet – that's good but which botnet?

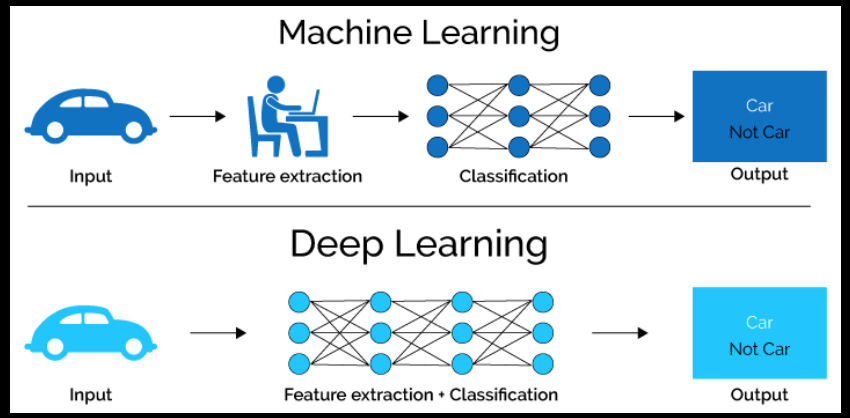
What is Deep learning?

- Deep learning is a subset of machine learning
- Algorithms in deep learning mimic the structure and function of the human brain as artificial neural networks (ANN)

  Example – *Convolutional neural networks* (CNN) can classify images similar to how the brain identifies objects

- Deep learning algorithms automatically learn and extract relevant features from raw data

Automatic Feature Engineering



Source: https://images.xenonstack.com/blog/machine-learning-vs-deep-learning.png

Automatic Feature Engineering Makes Life Easy

Traditional Machine Learning

- In traditional machine learning the programmer has to manually select the features from the raw data
- In other words, the programmer has to be very specific in telling the “computer” what to look for when deciding if the image contains a car or an airplane
- Manual feature extraction is time consuming, difficult and requires expert knowledge

Deep learning

- No need to manually extract features

Why is Deep Learning Gaining Popularity?

- We now have access to large amounts of high-quality labelled data
- Complex deep learning models can be trained in less time using GPUs (graphics processing unit)
- Fairly easy to implement using open source frameworks such as Tensorflow, Caffe, Keras, PyTorch
- Deep learning breakthroughs are driving Artificial Intelligence boom

Applications of Deep Learning

- Autonomous driving (e.g. detecting traffic lights, stop signs)
- Robots (e.g. vision)
- Aerospace and Defense (e.g. identification of objects from satellites)
- Video Surveillance (e.g. facial recognition, behavior recognition)
- Image Recognition
- Natural Language Processing (NLP)
- Speech Recognition

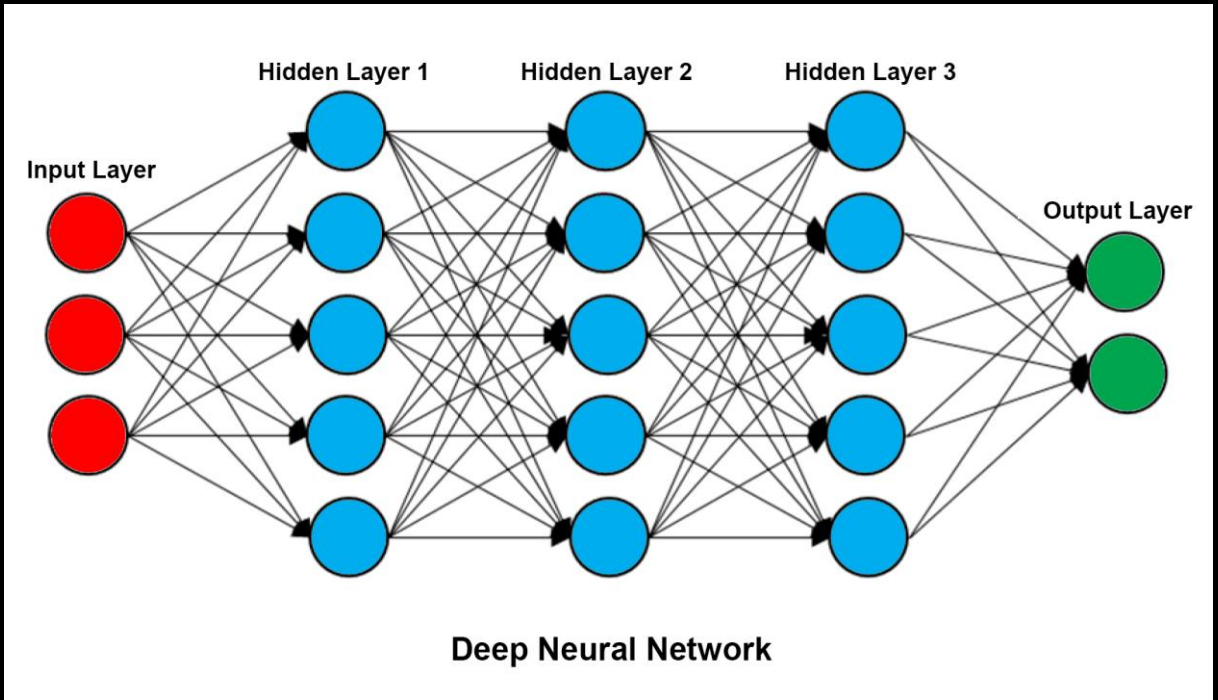
The Robot Apocalypse – Autonomous Weaponry

- The manufacturer of the famous "**AK-47**" rifle Kalashnikov has developed a range of products based on artificial neural networks
- Kalashnikov's fully automated "**combat module**" has the capacity to make decisions and identify targets on its own
- The "combat module" analyzes image data to identify targets

Source: https://kalashnikov.com/en/press-service/photo

Deep Neural Networks (DNN) – Why Deep?

DNN is simply an artificial neural network with **multiple** hidden layers



Deep Neural Networks - Image Recognition

Hidden Layer 1

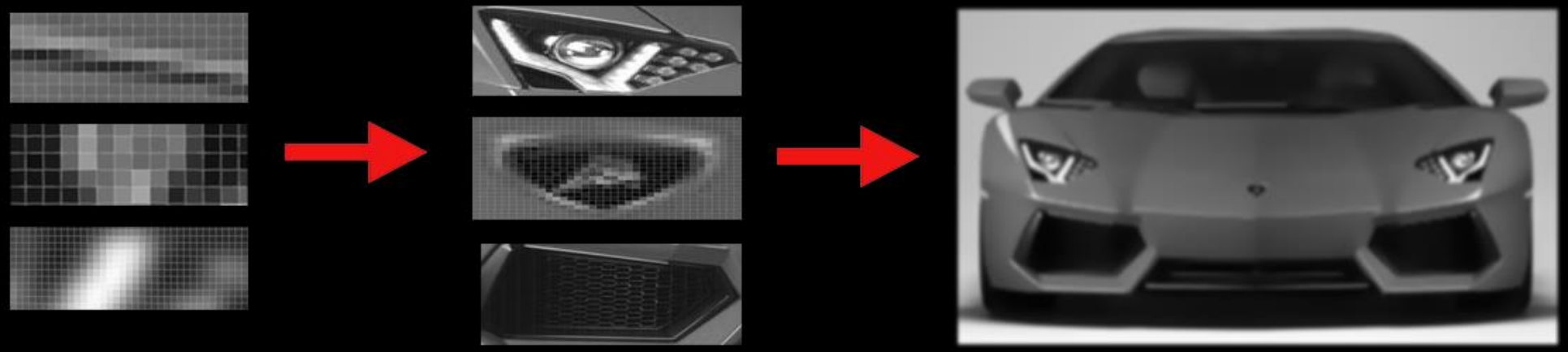
Algorithm learns to first recognize pixels and then edges and shapes

Hidden Layer 2

Learns to identify more complex features and shapes such as headlights and grill

Hidden Layer 3

Learns which shapes and objects can be used to identify a car



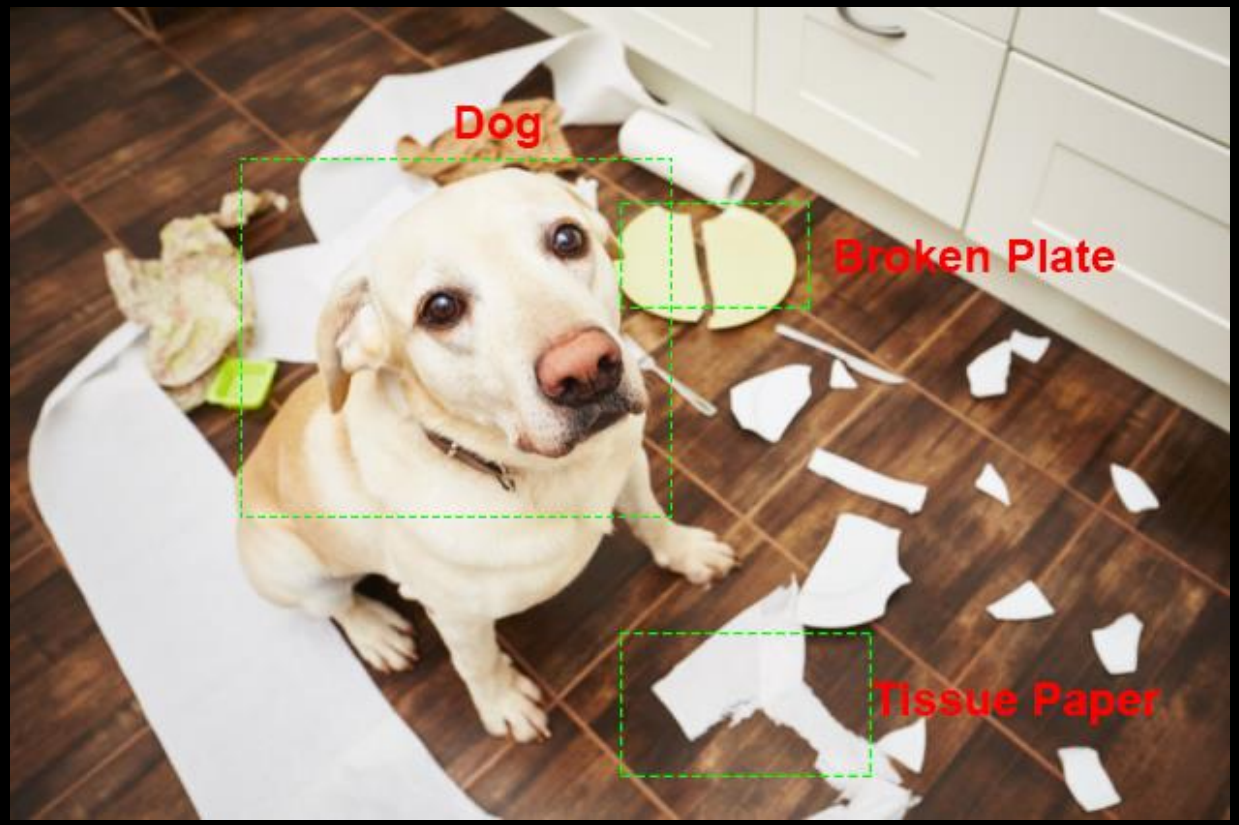
How objects are recognized in a hierarchical structure

Image Recognition – The Human Way

Lets see how we classify this image

Image Recognition – The Human Way

You see a dog, broken plate, tissue paper = Brain tells you “Yup the dog did it”



Deep Learning – The Brain's Visual Assessment

The point is:

- We are constantly evaluating things in the world around us
- Subconsciously making thousands of predictions everyday
- Recognizing patterns and labelling objects based on what we have learned or experienced in the past

Training machines for image classification

Convolutional Neural Networks (CNN or Convnets)

- CNNs have proven to be extremely effective in performing image recognition and classification tasks
- They perform feature identification and classification of images, text, sound, and video
- CNN takes an input image, processes it and outputs a class or a probability of classes that best describes the image (e.g. car, truck, plane)

Convolutional Neural Networks – Why So Popular?

- Automatic feature Engineering
- State-of-the-art performance in image recognition tasks e.g. accurately identifying objects such as cars in pictures
- Pre-existing networks can be retrained for new recognition tasks

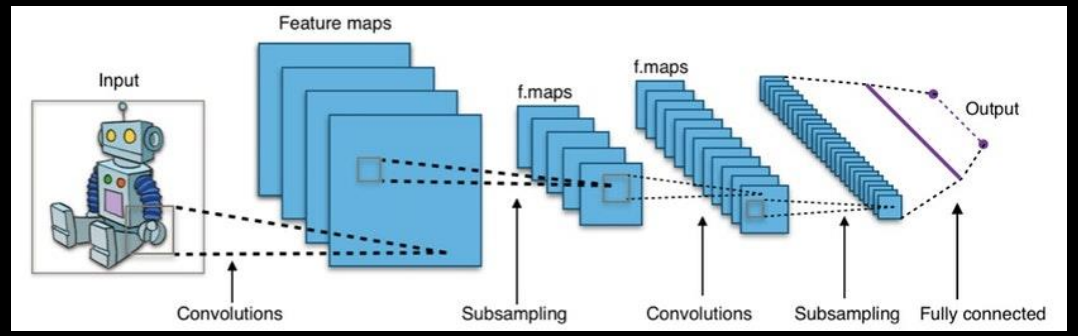
Convolutional Neural Networks – Layers

Three main types of Layers

Convolution Layer - Main task is to extract features from the input image

Sub-sampling (Pooling) Layer - Reduces the amount of parameters and computation in the network

Fully-Connected (Output) Layer - Classifies the input image



Convolutional Neural Network Architecture

You Must Train Them Right

Accuracy and performance of the classifier depends on the quality of the training data

Dog ✓

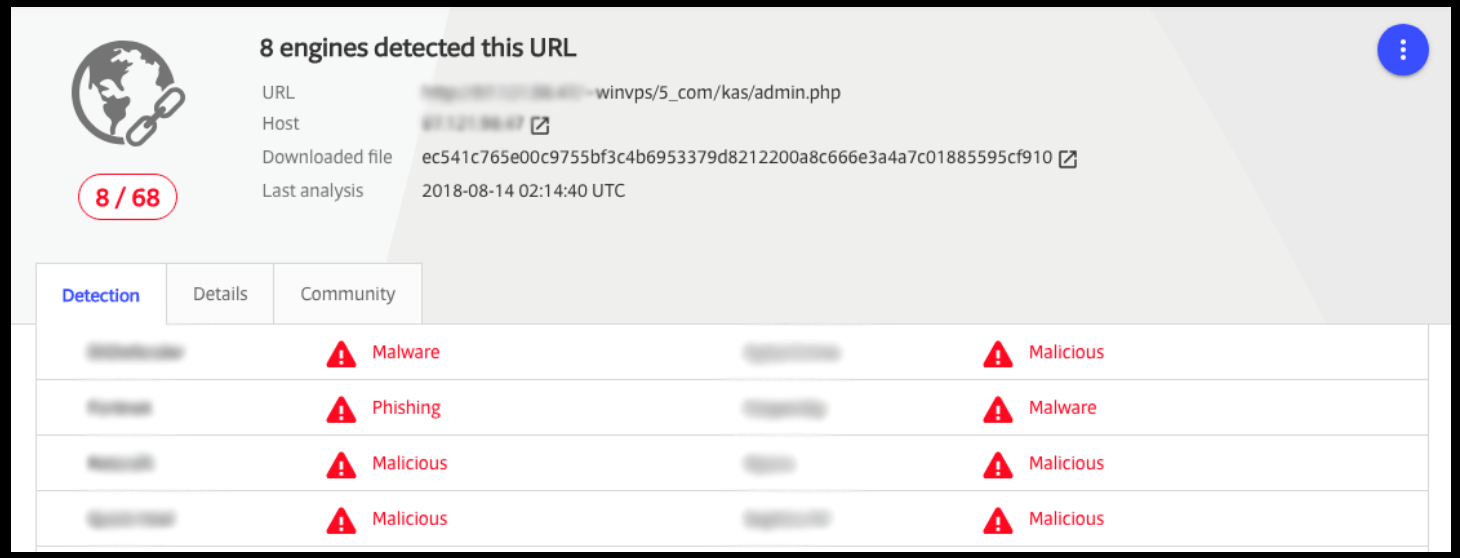
Dog ✗

Training the classifier for malicious website detection

Tasks

My tasks were simple:

- Use image recognition algorithms to classify malicious websites
- Enrich malicious website threat label data with more than just phishing, suspicious or malicious



Training A Neural Network From Scratch

Requires:

- Processing power (multiple graphics processing units)
- Large amounts of labelled data (millions of images)
- Lots of time (several days or weeks)

Solution - Transfer Learning

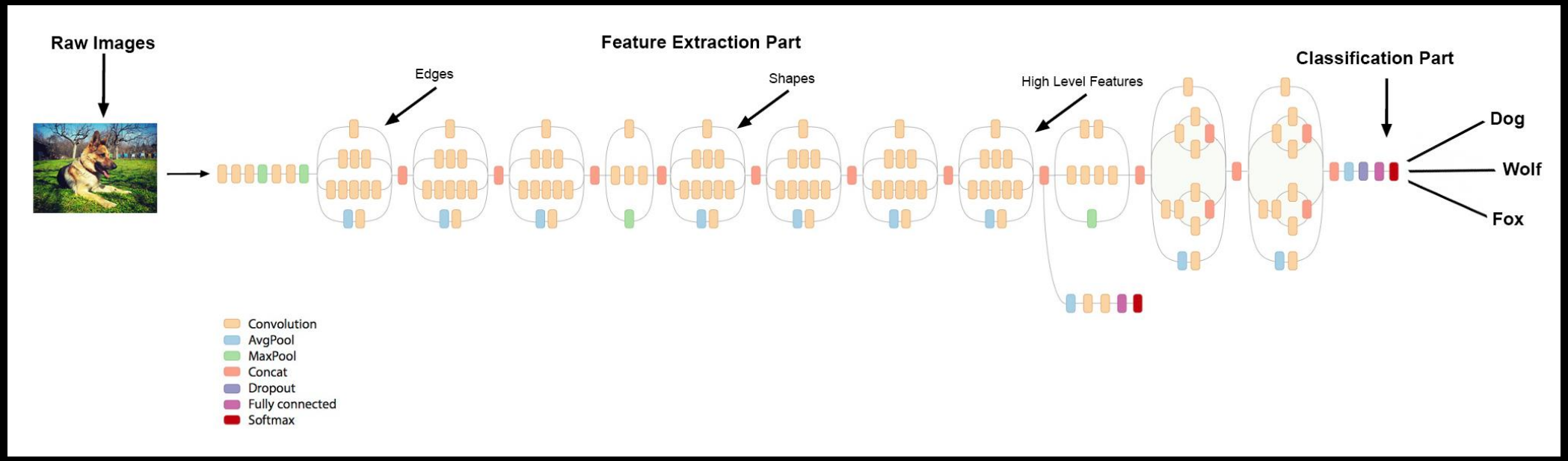
- Transfer Learning is the process of applying a trained model to a different, but related task
- Transfer learning is typically faster, since we only retrain the last layer (fully connected layer) of an existing model
- Retraining requires less resources and data
- Transfer learning has become the go-to method in many real world use cases, such as cancer detection

Transfer Learning - Retraining Inception v3 CNN Model

- Inception-v3 is a convolutional neural network that is trained on more than a million images from the ImageNet database
- The model achieves state-of-the-art accuracy for recognizing general objects with 1000 classes, such as ”Laptop", ”Cheeseburger”, ”Beer” etc.
- The model extracts general features from input images and classifies them based on those features

My Goal = Retrain Inception-v3 to classify images of malicious websites

Inception v3 Architecture



Transfer Learning With TensorFlow

- TensorFlow is an opensource software library for high performance numerical computation
- Originally developed by researchers and engineers from the Google Brain team
- Strong support for machine learning and deep learning
- Flexible architecture allows easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs)
- URL: https://www.tensorflow.org

My TensorFlow Setup

- The “retrain.py” script in TensorFlow retrains the image classifier
- This script loads the pre-trained Inception v3 model
- Removes the old final layer and trains a new one on the downloaded images

```
aq@u1:~/TF$ ls
classify.py  inception  retrain.py  tf_files  training_dataset  train.sh
```

classify.py does the following:

- Takes a screenshot of the website using Selenium WebDriver
- Saves the screenshot as “out.png”
- Predicts whether “out.png” is a phishing or botnet site

- Contains training data
- E.g. Images of phishing sites etc.

Training Data

- Used opensource data to download images of phishing and botnet websites

- Moved downloaded images to the appropriate sub folders in the *training_dataset* directory

```
aq@u1:~/TF$ ls training_dataset/office365_phishing_site/
0aff413c.png  343a5343.png  7a5b0e91.png  a9638eed.png
0edbe414.png  3472e0a6.png  7aa29722.png  abdd5e5c.png
1004c8d5.png  3546eb9a.png  7cf8c43f.png  b037deab.png
```

- Sub folder names define what label is applied to each image during the inception retraining process

```
aq@u1:~/TF$ ls training_dataset/
bank_of_america_phishing_site  lokibot_botnet           pony_botnet
dropbox_phishing_site          office365_phishing_site
aq@u1:~/TF$
```

Retraining Inception v3

- Once “retrain.py” script is executed with all the parameters, it starts the retraining process
- The script downloads the previously trained Inception-v3 model (if not already installed)
- Looks for images in the “training_dataset” folder
- Proceeds to generate bottleneck files

```
aq@u1:~/TF$ python retrain.py \
>   --bottleneck_dir=tf_files/bottlenecks \
>   --model_dir=inception \
>   --summaries_dir=tf_files/training_summaries/basic \
>   --output_graph=tf_files/retrained_graph.pb \
>   --output_labels=tf_files/retrained_labels.txt \
>   --image_dir=training_dataset
>> Downloading inception-2015-12-05.tgz 100.0%
Successfully downloaded inception-2015-12-05.tgz 88931400 bytes.
Looking for images in 'bank_of_america_phishing_site'
Looking for images in 'lokibot_botnet'
Looking for images in 'office365_phishing_site'
Looking for images in 'dropbox_phishing_site'
Looking for images in 'pony_botnet'
Creating bottleneck at tf_files/bottlenecks/dropbox_phishing_site/5284710a.png.txt
Creating bottleneck at tf_files/bottlenecks/dropbox_phishing_site/8.png.txt
Creating bottleneck at tf_files/bottlenecks/dropbox_phishing_site/44de6043.png.txt
Creating bottleneck at tf_files/bottlenecks/dropbox_phishing_site/28d002d8.png.txt
```

Retraining Inception v3 – cont'd

- Once the bottleneck files are generated for all the training data, the script begins training of the new top layer
- '***Bottleneck***' is an informal term used for the layer just before the final output layer that actually does the classification

```
Creating bottleneck at tf_files/bottlenecks/office365_phishing_site/b0a51409.png.txt
Creating bottleneck at tf_files/bottlenecks/office365_phishing_site/1cbdd365.png.txt
Creating bottleneck at tf_files/bottlenecks/office365_phishing_site/3e8bc157.png.txt
```

Retraining Inception v3 – cont'd

In each step three values are calculated:

Training Accuracy - Percentage of training images that were classified correctly

Cross Entropy - A Loss function which tells us how well the learning process is progressing

Validation Accuracy - True measure of how well the model is performing

```
2018-08-24 13:46:29.904036: Step 10: Train accuracy = 91.0%
2018-08-24 13:46:29.904143: Step 10: Cross entropy = 0.925858
2018-08-24 13:46:30.025306: Step 10: Validation accuracy = 82.0% (N=100)
```

Retraining Inception v3 – cont'd

- In each step 10 images are randomly chosen from the training set
- Bottlenecks are fed into the final classification layer to perform predictions
- Predictions are checked for accuracy by comparing them to the actual labels
- Depending on the accuracy of the predictions, the final layer's weights are updated through the back-propagation process

Retraining Inception v3 – cont'd

Once the training process is complete, the script outputs the final test accuracy for the model

```
Final test accuracy = 94.5% (N=55)
```

Time To Predict

Dropbox - Sign in - Mozilla Firefox

Dropbox - Sign in

satcomit-bd.com/TNA/dropb/dpbx/dpbx/

Select your email provider

Now , you can sign in to dropbox with your email

Gmail by Google

YAHOO!

Windows Live

AOL

Other Emails

Bank of America | Online Banking | Sign In | Online ID - Mozilla Firefox

Bank of America | Onlin

dmendestransportes.com.br/bankofamerica/0f01c10e/signOnV2Screen.php

Personal | Small Business | Wealth Management | Businesses & Institutions | About Us

Bank of America

Locations | Contact Us | Help | En español | Search Bank of America

Secure Sign-in

Online ID | Passcode | Sign In

Save this Online ID | Security & Help

Forgot ID | Forgot Passcode | Enroll

Banking | Credit Cards | Loans | Investments | Learning

Checking + Mobile Banking = Convenience

Get a checking account that goes wherever you go.

Get started

Information for: Select a state Go

Advertising Practices

Online Banking

Bank from home or on the go.

Enroll now »

Online Bill Pay

One site. Easy access. Total convenience.

Get started »

Fighting hunger

A food bank is developing new ways to help those in need.

See how »

Locations

Enter city, state or ZIP code Go

More search options

Privacy & security

Report phishing email Go

Get a $100 cash rewards online bonus offer

The BankAmericard Cash Rewards™ credit card gives you more cash back on the things you buy most with no hoops to jump through. Earn 1% on purchases, 2% at grocery stores and 3% on gas—and get a $100 cash rewards online bonus offer.

Learn more—BankAmericard Cash Rewards™ »

Need help with your home loan payments?

If you're a homeowner struggling with your loan payments, you may want to learn about our home loan assistance programs. Bank of America is committed to helping homeowners and is a participant in the national mortgage settlement agreement.

Learn more about home loan assistance »

Learn about the Department of Justice settlement »

Popular links

- Order checks
- Order a debit card
- Order foreign currency
- Schedule an appointment

Announcements

Bank of America was awarded the 2014 Digital Banking Experience Leader in Mobile Banking by Javelin Strategy & Research. Learn more about Mobile Banking »

CyberCrime - Mozilla Firefox

CyberCrime

https://cybercrime-tracker.net

Set sail for...

sq2

[C&C: 15001] - [ZbotScan+CCAM: 4451]

[RSS] - [Full List] - [Tracker] - [ZbotScan] - [Submit C&C] - [Tools] - [VX] - [Stats] - [Relax] - [About]

<< Start ··· < Previous ··· 1 ··· Next > ··· End >>

Search (URL or TYPE):

-::DATE	-::URL	-::IP	-::TYPE
25-09-2018	ark.treassurebank.org/Panel/admin.php	68.65.122.50	Pony
25-09-2018	193.29.187.201/ember/workflow/PvqDq929BSx_A_D_M1n_a.php		Lokibot
25-09-2018	89.46.222.135/lewy/lewy/sun/bros/solar/gem/PvqDq929BSx_A_D_M1n_a.php		Lokibot
25-09-2018	deloilte.com/nc_assets/fonts/five/PvqDq929BSx_A_D_M1n_a.php	198.187.29.21	Lokibot
25-09-2018	deloilte.com/wp-admin/includes/five/PvqDq929BSx_A_D_M1n_a.php	198.187.29.21	Lokibot
25-09-2018	hamon.ir/ogbu/five/PvqDq929BSx_A_D_M1n_a.php	185.159.153.121	Lokibot
25-09-2018	strutitinca.ro/cas/PvqDq929BSx_A_D_M1n_a.php	89.45.173.63	Lokibot
25-09-2018	support-office365.date/youngboss/Panel/five/PvqDq929BSx_A_D_M1n_a.php	93.157.63.185	Lokibot
25-09-2018	theonlygoodman.com/fit/PvqDq929BSx_A_D_M1n_a.php	49.51.84.226	Lokibot
25-09-2018	theonlygoodman.com/hnd/PvqDq929BSx_A_D_M1n_a.php	49.51.84.226	Lokibot

Questions?

Thank you