Modern Steganalysis: The ALASKA Challenge

Yassine Yousfi, yyousfi1@binghamton.edu

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

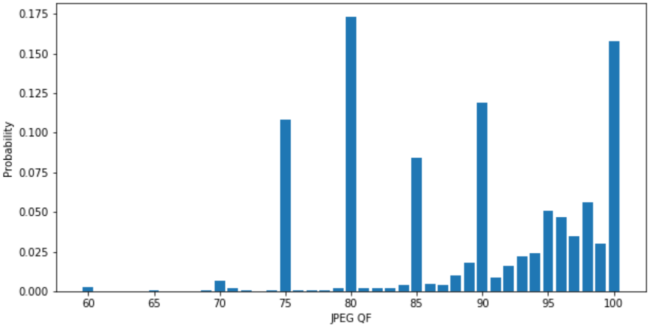
ALASKA Breakers

- Jessica Fridrich, *Binghamton*
- Jan Butora, *Binghamton*
- Quentin Giboulot, *Troyes, France*
- Yassine Yousfi, *Binghamton*

The ALASKA Challenge

- Color JPEGs, payload embedding in Y, U and V (Y, Cr and Cb)
- Multiple stego schemes: J-UNIWARD, nsF5, UED, EBS
- Variable image sizes (between, 512x512 and 1024x1024)
- Variable payload (scaled w.r.t. SRL)
- Multiple JPEG QFs 60–100
- Randomized cover image processing operations (resizing, sharpening, denoising, ...)
- Ordering images instead of hard decisions
- One submission / 4 hours

JPEG Quality factors



Distribution of 2,691,980 JPEG images downloaded from FlickR

Performance score

- ALASKArank = 5,000 images
- Ordering images allowed drawing the **ROC curve** in the back-end
- **MD5: Missed Detection rate at 5% False Alarm**
- P_{E}, and FP50 (False Alarm rate at 50% Missed Detection) are returned but not used to rank competitors
- Scores shown were on a 80% random subset of ALASKArank (to avoid competitors from using ALASKArank as a feedback loop)

Early pains

- Andreas Westfeld strikes with perfect detectors (MD5 $= 0$) !
  - Using non-handled exceptions in the website (by submitting out of range values, strings, etc.)
  - Noticing that stego images have a different timestamp than cover images
- ... We were still downloading the datasets

Diverse stego detection

- 1 Binary detector: $f : \mathcal{X} \rightarrow \mathrm{P}(\{\text{Cover}\}, \{\text{UED, EBS, J-UNI, nsF5}\})$
- 4 Binary detectors:
  $f_i : \mathcal{X} \rightarrow \mathrm{P}(\{\text{Cover}\}, \{\text{i}\})$, $i \in \{\text{UED, EBS, J-UNI, nsF5}\}$
  - How to make a final decision?
  - f_i has unpredictable behavior when given stego images from $j \neq i$
- 1 Multi-class detector:
  $f : \mathcal{X} \rightarrow \mathrm{P}(\{\text{Cover}\}, \{\text{UED}\}, \{\text{EBS}\}, \{\text{J-UNI}\}, \{\text{nsF5}\})$
  - Used as binary detector
  - P(stego) $=$ P(UED) $+$ P(EBS) $+$ P(nsF5) $+$ P(J-UNI)
  - Best strategy

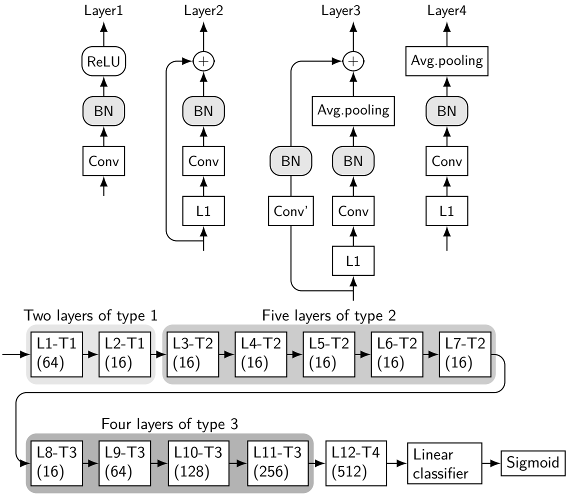
SRNet, [Boroumand et al. 2018]

- 20 Convolution (3x3) layers
- Around 4M learnable parameters
- Universal (performs well in multiple stego schemes)
- Trains in 3 to 4 days
- No pooling in early layers
  - Why?

SRNet, [Boroumand et al. 2018]

- 20 Convolution (3x3) layers
- Around 4M learnable parameters
- Universal (performs well in multiple stego schemes)
- Trains in 3 to 4 days
- No pooling in early layers
    - Why?
    - Pooling can be seen as a low-pass filter, reduces the energy of the stego signal

SRNet, [Boroumand et al. 2018]

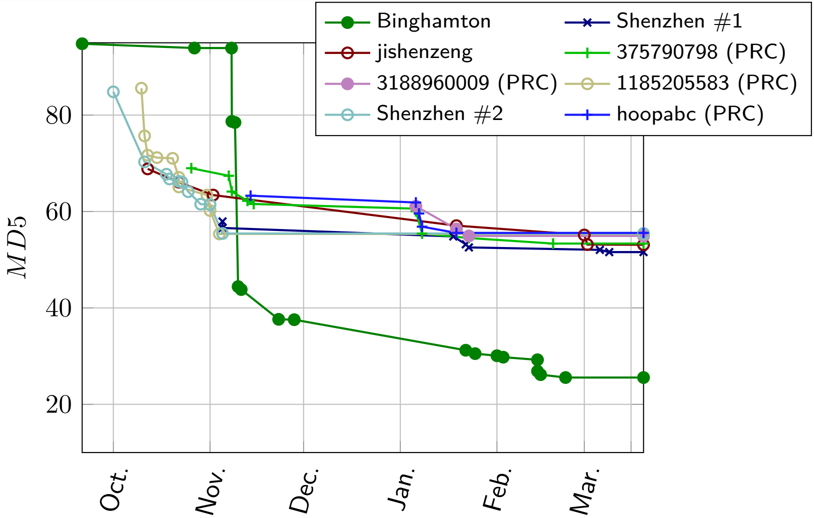


- BN = Batch Normalization was not covered = learnable scaling
- Conv' = 1x1 Conv was not covered = convolution with kernel size = 1
- (X) Shows the depth of the learned representation

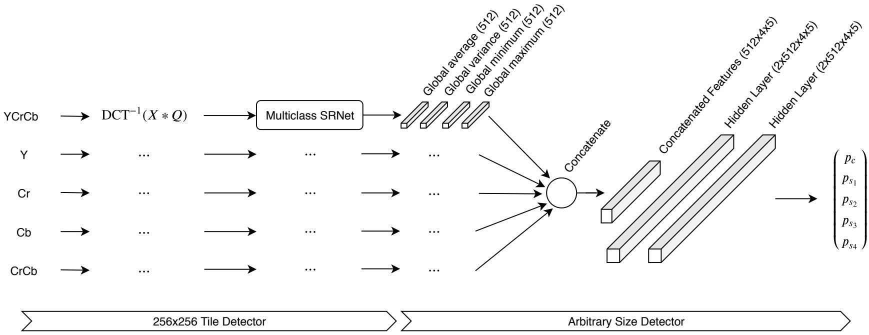
Strategy

- Training 1 detector / QF, very time-consuming
  - Double payload $\rightarrow$ Single payload $=$ **Curriculum learning**
  - Can only afford 256x256 images, GPU memory ...
  - Training in 2 stages for larger images
- Noticed only about 10% stego images per QF in ALASKArank
- The reverse JPEG compatibility attack is operational and has 99.99% accuracy for QF100, again only 10% stego images for QF100 in ALASKArank

[Cogranne et al. 2018]



Winning architecture, $QF \leq 98$



Lessons learned

- Facing multiple stego schemes = Train as Multi-Class
- Better understanding of SRNet (and CNNs)
  - Is still the state-of-the-art
  - Can contain the diversity of Alaska
  - Early merging colors is sub-optimal
  - Not "universal" failed to detect nsF5
- Still learning ...
  - Can merge certain JPEG QFs and train fewer CNNs
  - A new way of computer histograms of co-occurrences using convolutions