ML2 - Final Project

Photoshop Detector

Sean Pili
Chirag Jhamb
Poornima Joshi

Problem Statement

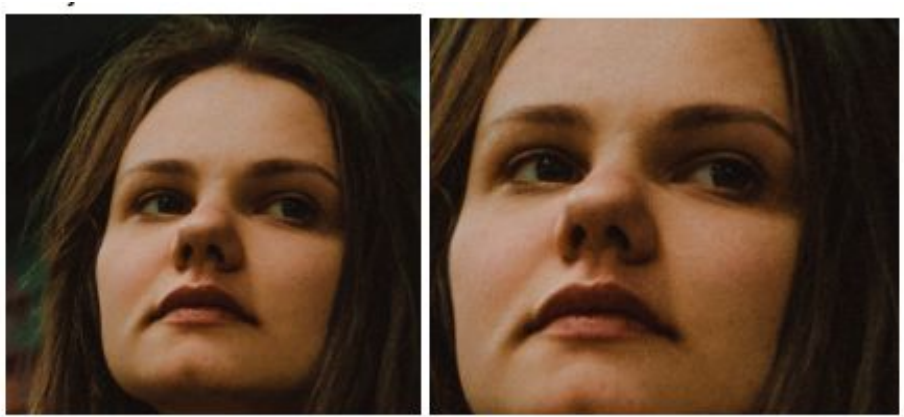
- Classify into fake and real images (Binary classification problem).
- Effort to improve digital image forensics; a field dedicated to detecting the authenticity of images.
- Verify the Adobe's model in terms of generalizability.

Dataset

- Data was extracted from kaggle (https://www.kaggle.com/ciplab/real-and-fake-face-detection).
- Curated by Computational Intelligence and Photography Lab - Department of Computer Science, Yonsei University
- Photoshop levels exist in the dataset.
- 1000 - Real & 1000 - Fake Images.

Data Preprocessing

- Sorted data so that each test has 20% of easy, medium and hard fake images. Random split into train and test set could have led to training on hard and prediction on easy, or vice-versa.
- Used a face detector: MTCNN to recognise faces and crop them.
- Customized bbox returned from MTCNN to get full face (initially MTCNN only returned face till the mouth).

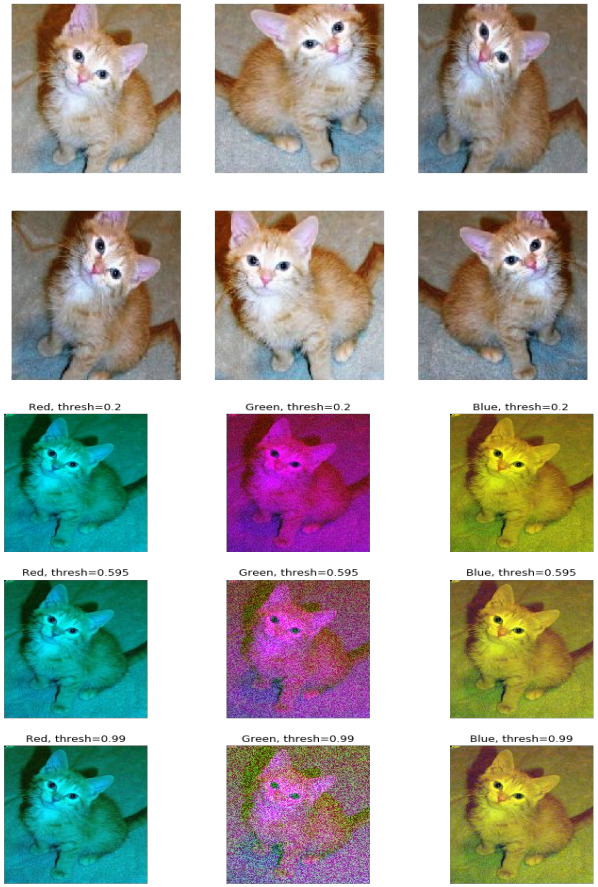


Original Image vs Cropped Image

fast.ai

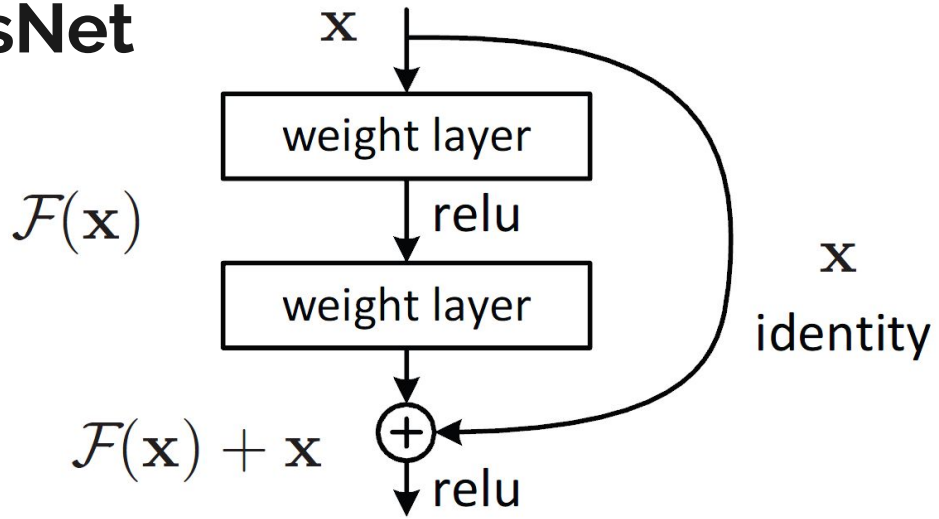
Pretrained models- FastAI

- Library that simplifies training fast and accurate neural nets using modern best practices
- Great data augmentation support. We played around with RGB, flipping and brightness
- Claims to result in significantly improved accuracy and speed over other deep learning libraries
- Always gave ~7-8% higher accuracy on transfer learning models as compared to pytorch or keras
- We used resnet, densenet and vgg16 with fast.ai



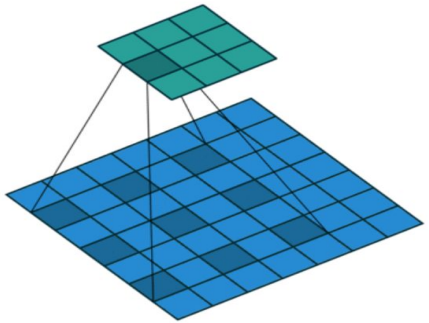
Architectures - ResNet

- Sorts Layers into Groups
- Uses **shortcut** to add input into 1st layer in group to net input of last layer in group
- Can increase accuracy with less cost than adding layers.

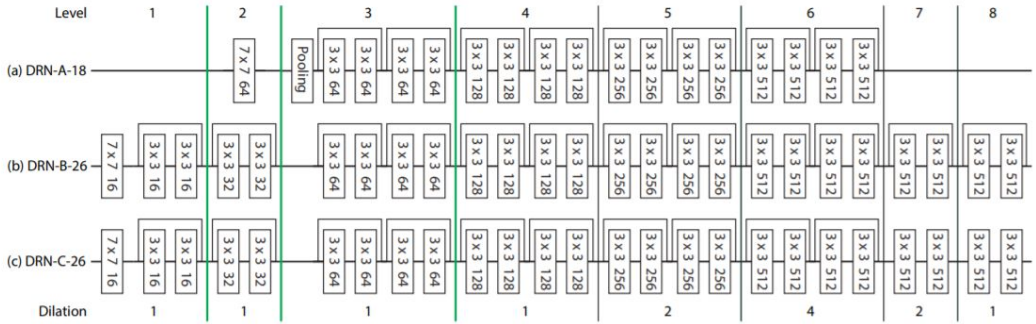


Architectures - Dilated Residual Networks

- Developed to preserve spatial resolution in Residual Network's
    - Feature-maps get progressively smaller
- Uses Dilated Convolution to increase receptive field
- Baseline Model: DRN-C-26



Dilated Convolution (l=2)



Results

- Baseline (DRN from Adobe's paper) has performed the best.
- All other came in close too.

Network	Epochs	Batch Size	Accuracy	Fake Accuracy	Real Accuracy
ResNet18	10	64	62.53%	72%	52%
FastAI-ResNet152	n/a	n/a	Out of Memory	N/A	N/A
FastAI-resnet50	14	64	63%	58%	65%
Baseline	5	17	67%	66%	68%
MobileNetV2	50	8	60%	58%	62%
FastAI-VGG16	15	64	64%	61%	65%

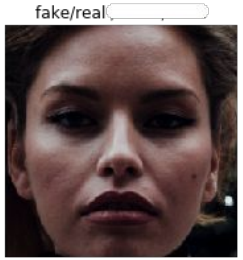
Ensembling of Models

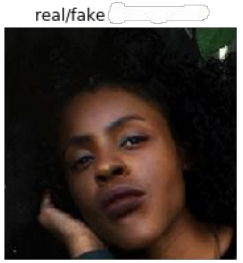
3: Approaches

1. Hard Voting: Best Result : 68% > Baseline (67%)
2. Soft Voting : (Averaging Logits) 66% < Baseline (67%)
3. Ensemble Classifier : Best Result: 67.97% ~ Basline 67.77
    a. Used sklearn's VotingClassifier to combine AdaBoost, Logistic Regression on the logits from Resnet50 and VGG16.

Key takeaways

- Data preprocessing helps more than advanced models.
- Ensembling boosts performance.
- Images on right are Predicted/Actual





Questions?

References

1. https://arxiv.org/pdf/1705.09914.pdf
2. https://peterwang512.github.io/FALdetector/
3. https://www.kaggle.com/ciplab/real-and-fake-face-detection
4. https://arxiv.org/pdf/1512.03385.pdf
5. https://docs.fast.ai/