Augment and Adapt: A Simple Approach to Image Tampering Detection

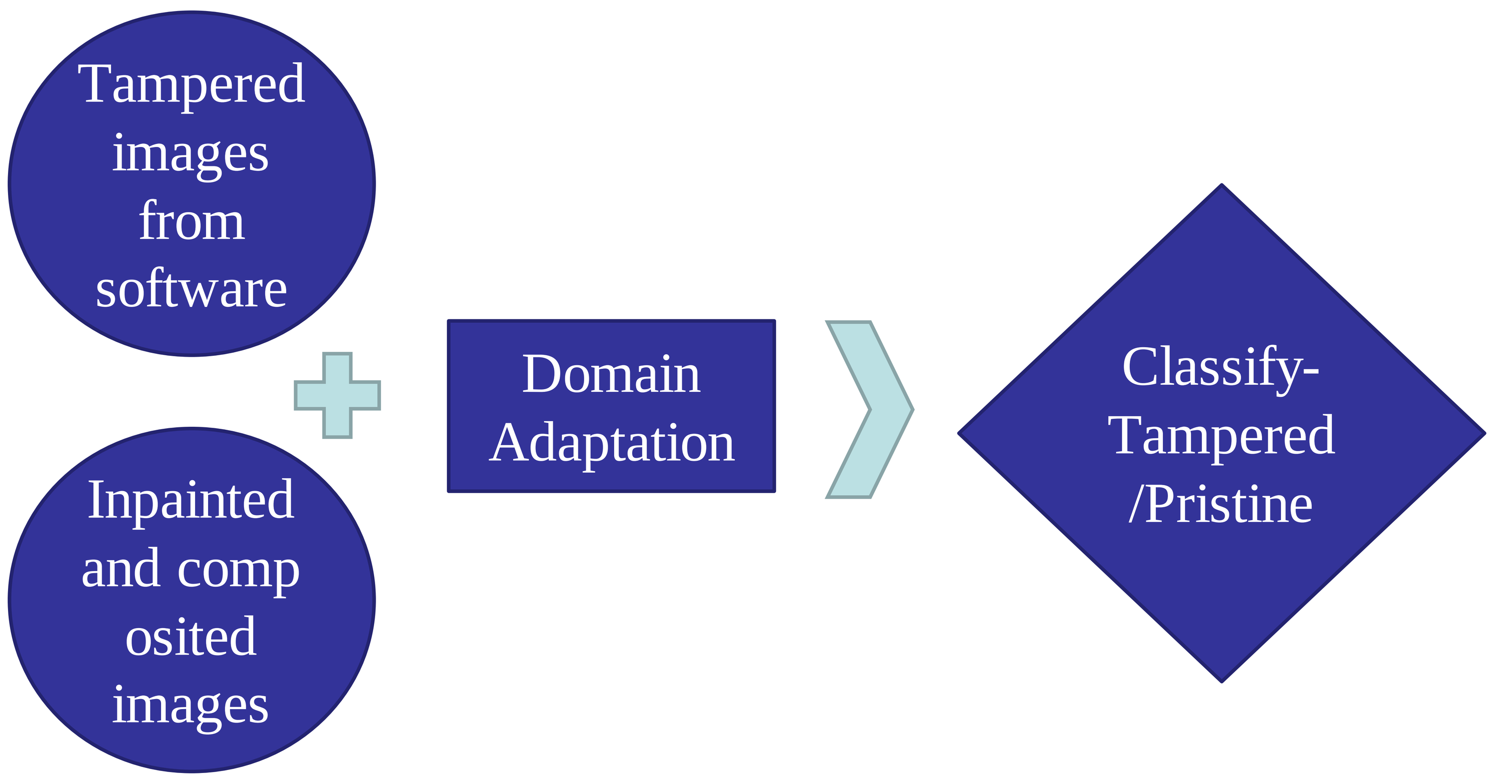
Yashas Annadani, C.V. Jawahar
Center for Visual Information Technology, Kohli Center for Intelligent Systems, IIIT Hyderabad

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD

ICPR 2018 INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION

GOAL

Classify if an image is tampered by means of *copy-paste*, *object removal* or *splicing* mechanisms.



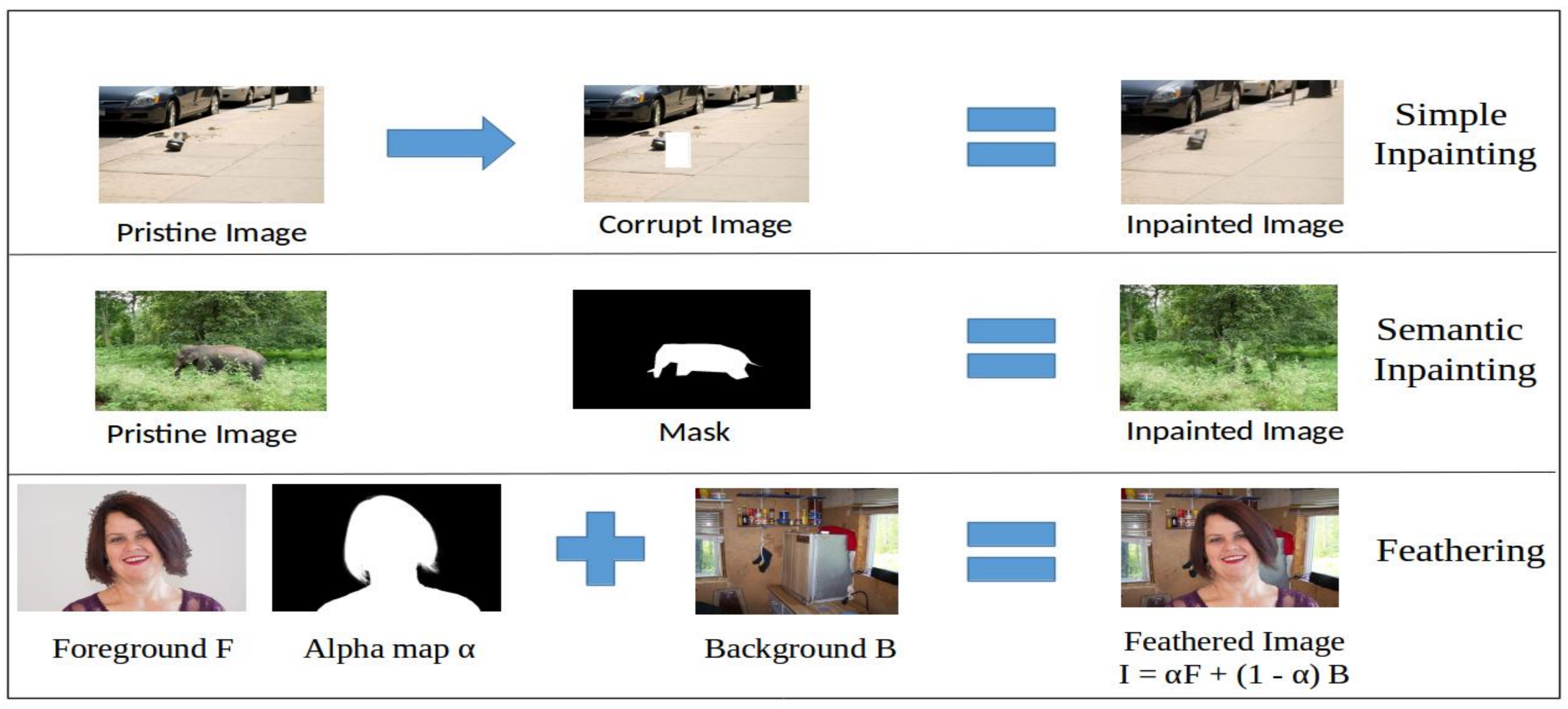
MOTIVATION

- Number of tampered images available to train a convolutional neural network is small.
- Inpainting and compositing, which are essentially forms of tampering similar to *object removal* and *splicing*, could be used to augment the data.
- Explore the possibility of performing domain adaptation between the augmented data as well as the curated data.

APPROACH

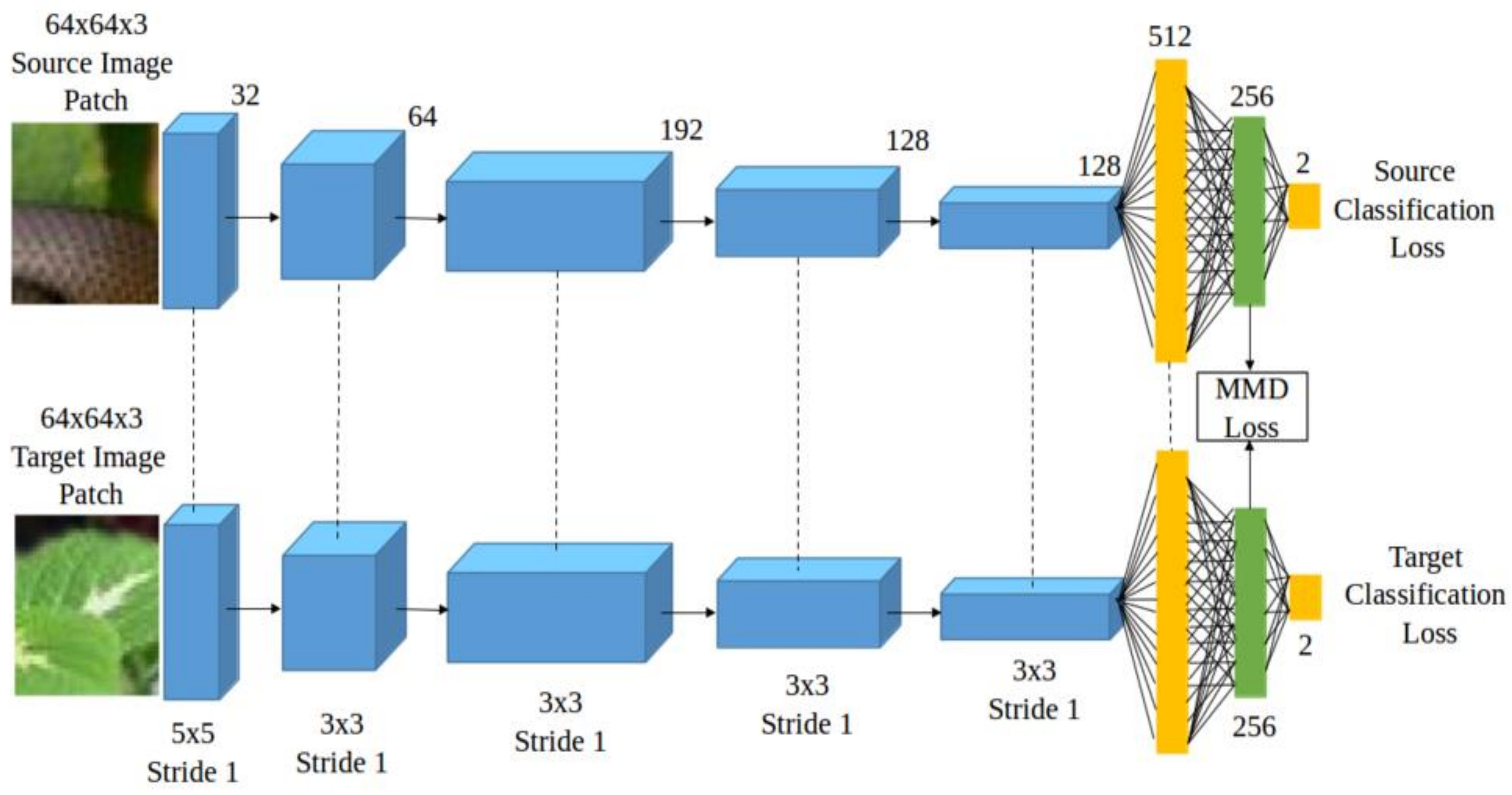
TAMPERING DATA AUGMENTATION

- Inpainting and compositing methods are employed.
- Three different augmentation schemes are used: Simple Inpainting, Semantic Inpainting and Feathering.
- Each of these schemes will help to augment atleast one of copy-paste, object removal or splicing types of tampering.



Simple Inpainting   Semantic Inpainting   Feathering

NETWORK



Classification loss: $\mathcal{L}_c = \min -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{2} 1(y^i = j)\log(p_{ij})$

Maximum Mean Discrepancy (MMD) [1] loss for domain adaptation:

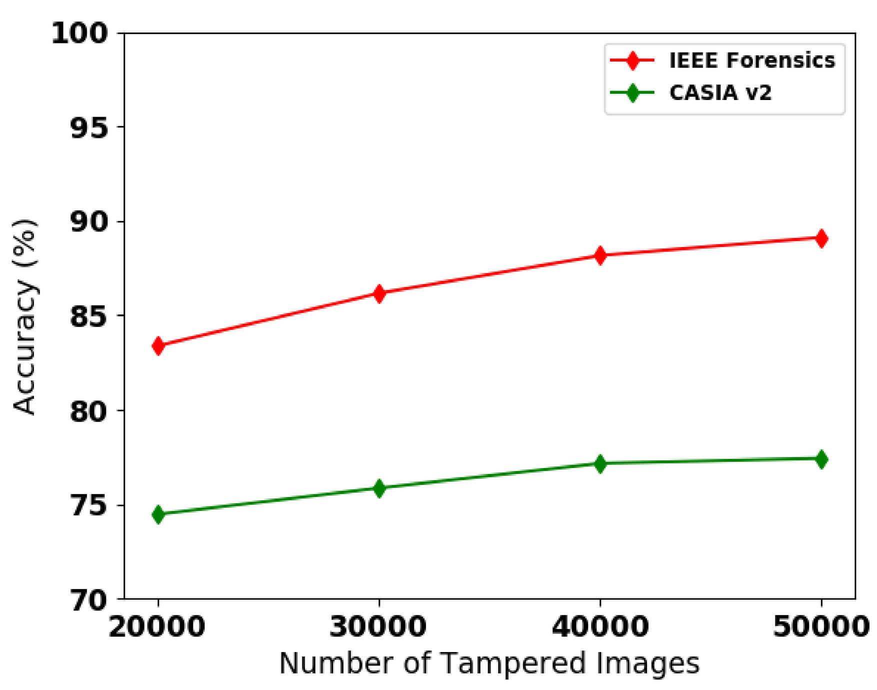
$$\mathcal{L}_M = \min \left|\left|\frac{1}{|\hat{X}_s|}\sum_{\hat{x}_s \in \hat{X}_s}\Phi(\hat{x}_s) - \frac{1}{|\hat{X}_t|}\sum_{\hat{x}_t \in \hat{X}_t}\Phi(\hat{x}_t)\right|\right|_2^2$$

Overall Loss: $\mathcal{L} = \mathcal{L}_{ct} + \lambda_1\mathcal{L}_{cs} + \lambda_2\mathcal{L}_M$

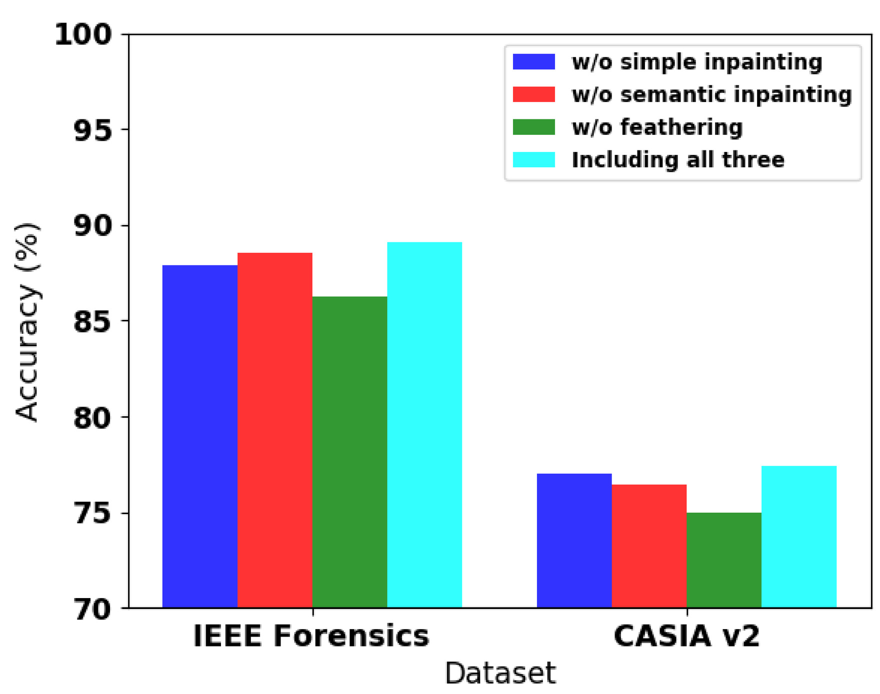
RESULTS

Method	IEEE Forensics	CASIA v2
Rota et al [2]	83.24	73.29
Bappy et al [3]	86.75	75.84
Train from scratch	82.25	71.61
Finetune on generated data	87.62	74.75
Proposed	**89.12**	**77.43**

Detection accuracy in % on two standard datasets.



Plot of number of augmented images vs accuracy



Plot indicating performance on image tampering detection due to various augmentation schemes

REFERENCES

1. E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," arXiv preprint arXiv:1412.3474, 2014.
2. P. Rota, E. Sangineto, V. Conotter, and C. Pramerdorfer, "Bad teacher or unruly student: Can deep learning say something in image forensics analysis?" in ICPR, 2016.
3. J. H. Bappy, A. K. Roy-Chowdhury, J. Bunk, L. Nataraj, and B. Manjunath, "Exploiting spatial structure for localizing manipulated image regions," in ICCV, 2017.

yashas.annadani@gmail.com, jawahar@iiit.ac.in