Solution to Visual Domain Adaptation Challenge.

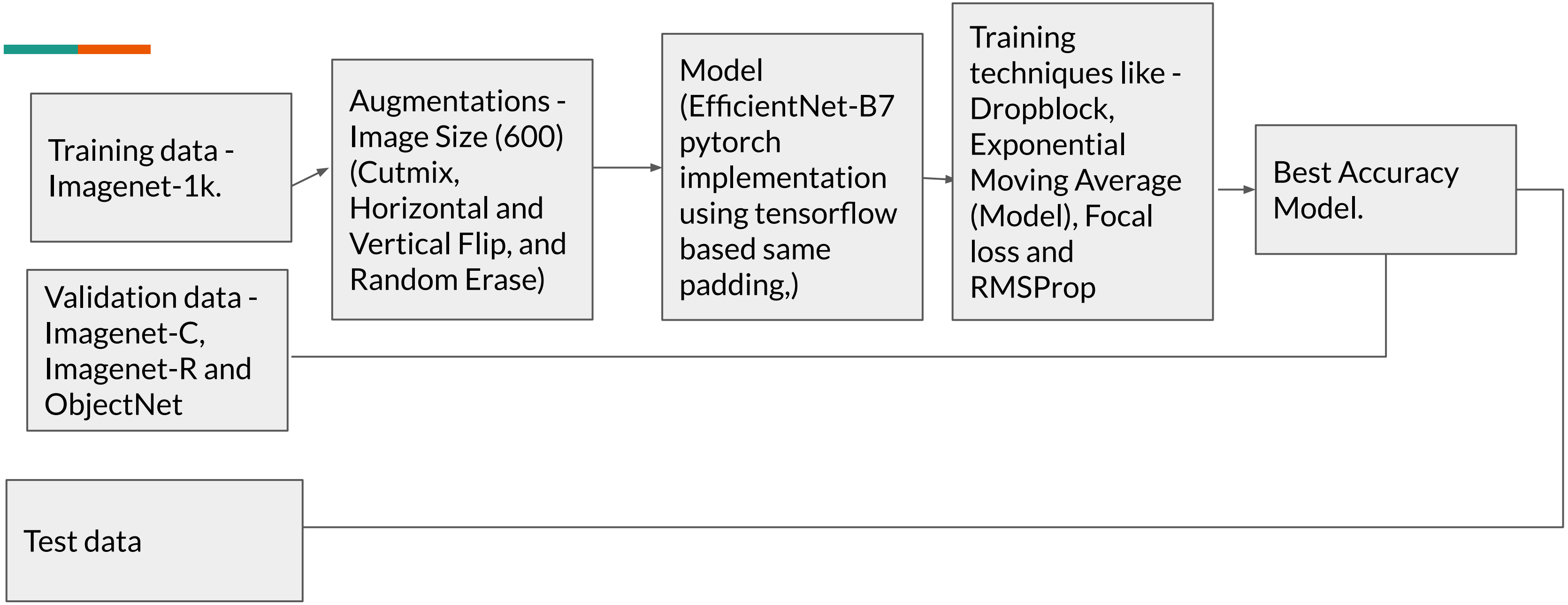
Chandramouli Rajagopalan

Overview

Most of the machine learning models today often fail to work on environments where it encounters out-of-distribution data.

We try to develop a solution that adapts to different target distribution and improve the overall accuracy score of the model.

VisDa Classification



Data Augmentations

CutMix / Mixup

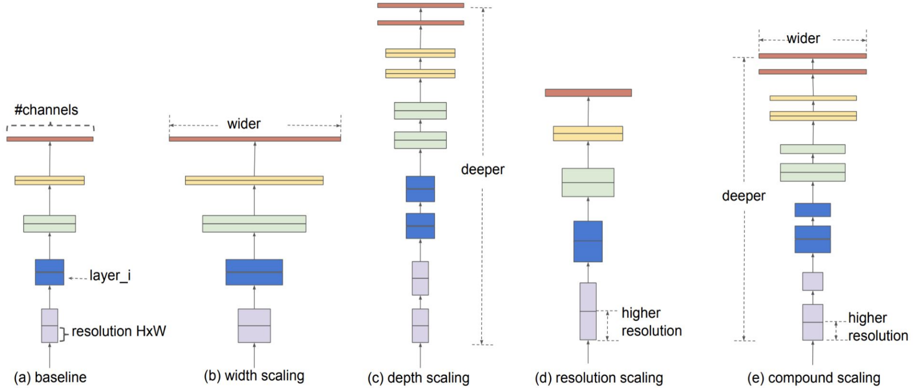
- The main idea behind cutmix / mixup augmentation is to replace a patch from another image instead of just erasing the pixels from an image. The ground truth labels are also proportionately mixed according to the area of the pixels.

- Mixup trains a neural network on convex combinations of pairs of examples and their labels.

Architecture of the Neural Network

EfficientNet B7

- We use EfficientNet as our base architecture for classification as it scales uniformly across width, resolution and depth of the network. Given the resolution of the image size to be 600x600, the compound scaling helps increase the receptor field and more channels to capture fine grained details in the image.



[EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks]

Learning Rate Optimizer - Adam

Adam or Adaptive Moment Optimization algorithms accelerates our search in direction of minima, and also impedes our search in direction of oscillations.

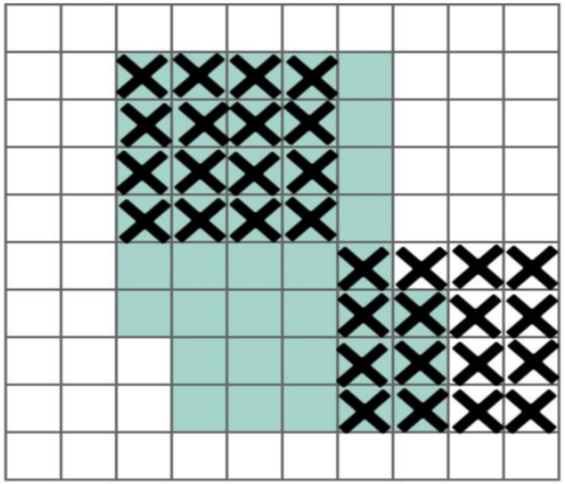
Generalization Techniques.

- DropBlock
- Exponential Moving Average
- Dropout

[Dropout: A Simple Way to Prevent Neural Networks from Overfitting]

Dropblock

Using dropblock, we try to make the neural networks look for elsewhere in the feature space to fit the data. Instead of randomly dropping out random features, we drop out continuous section so that the model learns to identify using remaining features.

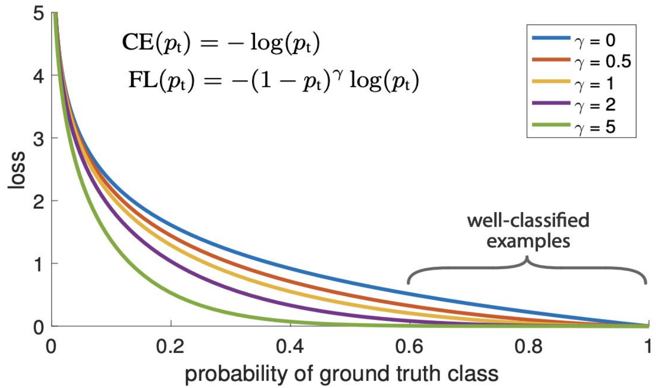


[DropBlock: A regularization method for convolutional networks]

Loss Function - Focal Loss

- Focal loss applies a modulating term to the cross entropy loss in order to focus learning on hard negative examples.

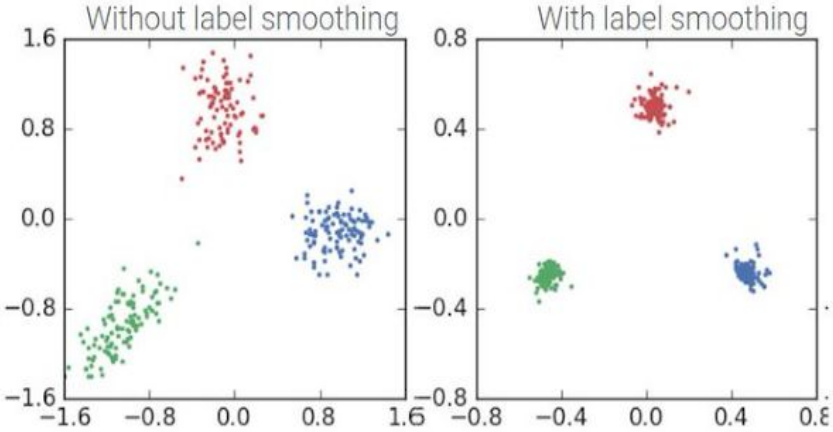
$$\mathrm{FL}(p_t) = -(1 - p_t)^{\gamma} \log(p_t)$$



[Focal Loss for Dense Object Detection]

Label Smoothing

- Label Smoothing
  - It introduces noise in the labels, which accounts for any errors in the labels. Helped the model generalize on the data better.



Thank you