Contributions

1. Formalization of Open Set Deep Networks using Meta-Recognition and OpenMax
2. Multi-Class Meta-Recognition using Activation Vectors to estimate the probability of deep network failure.
3. Proof that the proposed approach manages open space risk for deep networks.
4. Experimental analysis of the effectiveness of open set deep networks at rejecting unknown classes, fooling images and obvious errors from adversarial images, while maintaining its accuracy on testing images.

Theorem 1 (Open Set Deep Networks). A deep network extended using Meta-Recognition on activation vectors as in Alg. 2, with the SoftMax later adapted to OpenMax, as in Eq. 2, provides an open set recognition function.

Proof The Meta-Recognition probability (CDF of a Weibull) is a monotonically increasing function of $\|\cdot\|$, and hence $1 -\|\cdot\|$ is monotonically decreasing. Thus, they form the basis for a compact abating probability model. (Please refer the paper for more details about the proof.)

**Mean Activation Vectors**

**OpenMax**

**SoftMax**

**OpenMax Layer for Deep Networks**

**OpenMax Vs SoftMax**

**Failure Prediction with OpenMax**

**Mean Recognition Calibration for OSDN**

Algorithm 1 EVT Meta-Recognition Calibration for Open Set Deep Networks, with per class Weibull $\beta$ to $\gamma$ largest distance to mean activation vector. Return SoftMax model $p_2$, which includes parameters $\gamma$, for shifting the data as well as the Weibull scale and shape parameters $\beta_i$.

Require: Fullfit function from lshMR
Require: Activities levels in the penultimate network layer $V_C = (v_{1}, \ldots, v_C)$
Require: For each class $j$ let $S_j = \langle j, \alpha_j \rangle$ for each correctly classified training example $x_k$.

1. for $j = 1, \ldots, D$ do
2. compute mean AV, $\alpha_j = \text{mean}(S_j)$
3. EVT $p_j = (\tau_j, \alpha_j, \beta_j) = \text{Fullfit}(\langle S_j - \alpha_j \rangle)$
4. end for
5. Return $p_j$ and lshMR models $p_2$

Algorithm 2 OpenMax probability estimates with rejection of unknowns.

Require: Activation vector for $x_{\text{test}} = \langle x_{\text{test}}, \ldots, x_{\text{test}} \rangle$
Require: mean $\bar{\mu}$ and lshMR models $p_j = (\bar{\mu}_j, \alpha_j, \beta_j)$
Require: $n$, the number of top classes to rank
1. Let $s_{\text{test}} = \text{sigmoid}(\bar{\mu}_j)$. Let $u_j = 1$
2. for $i = 1, \ldots, D$ do
3. Compute $\mu_i = \text{softmax}(\bar{\mu}_i)$
4. end for
5. for $i$ in reverse order of $u_j$ do
6. Compute $\mu_i = \text{softmax}(\bar{\mu}_i)$
7. Define $u_j = \sum_i \mu_i(1 - \mu_i)$.

$$P_{j}(y) = \frac{\sum_{i \in S_{j}} e^{\mu_i}}{\sum_{i \in S_{j}} e^{\mu_i}}$$

**OpenMax**

**SoftMax**

**SoftMax Openset Detector**

**OpenMax Openset Detector**

**Softmax Fooling Detector**

**OpenMax Fooling Detector**

**Experiments**

**Experiment Details**
1. Dataset: Training ILSVRC’12. 1.3M training images, 1K classes.
2. Dataset: Testing 60K images total. 50K Images (1K classes) from ILSVRC’12 Validation set, 15K Fooling Images (1K classes), 15K open set images from 360 classes from ILSVRC’10 (these 360 classes are NOT present in ILSVRC’12).

**Model:** BVLC AlexNet (57.1% top-1 accuracy on ILSVRC’12 val set).


5. Performance: OpenMax performance gain is nearly 4.3% improvement accuracy over SoftMax with optimal threshold, and 12.3% over the base deep network. Putting that in context, over the test set OpenMax correctly classified 3450 more images than SoftMax with optimal threshold and 9487 more than the base deep network. Optimal F-Measure for each algorithm was SoftMax 0.58, OpenMax 0.595 and 1-vs-Set SVM 0.407.

Acknowledgements: The work supported in part by NSF Research Grant IIS-1320956 and UCCS Graduate School Fellowship.

Contact: abhijit.bendale@vast.ucsd.edu
Code/Data: http://vast.ucsd.edu/OSDN