

MetaMax: Improved Open-Set Deep Neural Networks via Weibull Calibration

Zongyao Lyu, Nolan B. Gutierrez, and William J. Beksi
The University of Texas at Arlington
Arlington, TX, USA

zongyao.lyu@mavs.uta.edu, nolan.gutierrez@mavs.uta.edu, william.beksi@uta.edu

Abstract

Open-set recognition refers to the problem in which classes that were not seen during training appear at inference time. This requires the ability to identify instances of novel classes while maintaining discriminative capability for closed-set classification. OpenMax was the first deep neural network-based approach to address open-set recognition by calibrating the predictive scores of a standard closed-set classification network. In this paper we present MetaMax, a more effective post-processing technique that improves upon contemporary methods by directly modeling class activation vectors. MetaMax removes the need for computing class mean activation vectors (MAVs) and distances between a query image and a class MAV as required in OpenMax. Experimental results show that MetaMax outperforms OpenMax and is comparable in performance to other state-of-the-art approaches.

1. Introduction

Image classification with deep neural networks has made significant progress [13, 25, 8, 9]. However, the majority of the work is based on a closed-set assumption where training datasets are expected to include all the classes that may be encountered in the environments in which the vision system will be deployed. Yet, this assumption cannot be guaranteed in real-world environments where samples from unknown classes not seen during training may appear during testing and cause system failure [20, 27, 30].

To address the limitation of closed-set classification, open-set recognition (OSR) has been introduced [22]. OSR describes the scenario where incomplete knowledge of the world is present during training, and new classes can appear during testing. Not only does this require the model to maintain the capability of accurately classifying known classes, but it must also be able to effectively identify unknown classes. OSR has recently gained significant attention in the research community [7, 10, 23, 2, 3, 5, 1].

With the introduction of the formal definition of OSR,

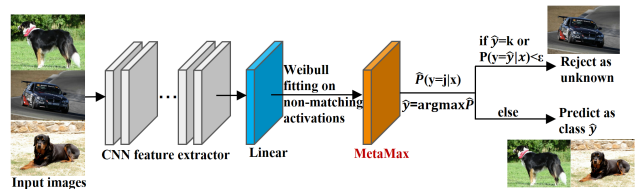


Figure 1: MetaMax utilizes extreme value theory-based meta-recognition for post-recognition analysis to discover the presence of unknown classes while preserving closed-set classification accuracy.

Scheirer et al. [22] presented a “1-vs-Set Machine” which extends the 1-class and binary SVM in a way that better supports OSR. Bendale and Boulton [3] proposed the first deep learning-based solution for OSR by introducing OpenMax as an alternative to SoftMax. OpenMax can be applied directly on a standard closed-set classification network to calibrate its output score vectors, thus enabling a closed-set classifier to perform OSR. It does not require accessing additional training data, updating the training procedure, or modification of the network architecture.

Inspired by OpenMax, we introduce MetaMax, a method that utilizes extreme value theory-based meta-recognition [24, 21] as a post-recognition analysis technique to assist the identification of unknown classes, Fig. 1. More specifically, we directly model non-match scores from the network outputs as opposed to OpenMax’s method of explicitly calculating and storing mean activation vectors for each class. By modeling the non-match distributions, we can assemble additional data points for building a Weibull model and therefore achieve more accurate parameters for each constructed model.

To the best of our knowledge, MetaMax is the *only alternative* to OpenMax. Different from most work in this area that either modifies the network architecture or requires auxiliary training data that can be hard to acquire in practice, we present a calibration tool that can be *readily plugged into various classification networks thus enabling them to perform OSR without additional overhead*.

2. Preliminaries

2.1. OpenMax

OpenMax introduced the concept of *activation vectors*, which are the output values from the penultimate layer of a neural network. Concretely, let $\mathbf{v}(x) = v_1(x), \dots, v_K(x)$ be the level of activation for each of the K classes seen during training and let the unknown classes begin at index 0. OpenMax obtains the activation vectors by saving the output of C_θ 's layer immediately before the SoftMax layer. It then separates the activation vectors into K clusters representing K known classes. For each of the clusters, the distance between every activation vector and the cluster mean (i.e., *class mean activation vector*) is calculated.

OpenMax proceeds to compute K Weibull models by fitting each model to the largest distances of the correctly classified samples. During inference, the K Weibull models are used to revise the top few highest activations to obtain a new vector $\hat{\mathbf{v}}(x)$. OpenMax then defines an additional activation for class $K + 1$ thereby producing a score vector for all $K + 1$ classes. The class corresponding to the highest probability is the prediction of the OpenMax algorithm. If the predicted class label for an input is indexed at 0, then it is considered to be an unknown class.

3. MetaMax

MetaMax is derived from contributions in class descriptors and the modeling of non-match distributions for the purpose of calibrating OSR.

3.1. Class Representations

Object recognition can be mapped to the problem of determining *match* scores between an input image and a **class descriptor** such as the class mean activation vector (MAV) used in OpenMax. Given an activation vector for a correctly classified image by a closed-set classifier, we consider the largest score as the **match** score and the rest of the scores in the vector as the **non-match** scores. The match and non-match scores from all the training data constitute the **match distribution** and **non-match distribution**, respectively.

3.2. Modeling of Rare Events

Given a particular input, OpenMax finds the distance between each sample and its respective MAV. The farthest, i.e., the rarest, distances are used to build per class Weibull models. These models are then utilized to compute revised activation vectors that OpenMax uses to provide probability estimates for the unknown rejection. Yet, if data is limited these largest distances may misrepresent the data manifold, produce incorrect activation vectors, and thus hinder the Weibull model's accuracy.

When building the Weibull model for each class, OpenMax finds a single data point by modeling the largest dis-

tance. We eliminate this limitation by modeling the non-match distribution to provide multiple data points per input. Concretely, we model the non-match distribution via non-match scores to obtain $K - 1$ data points per vector.

The first step of MetaMax is to obtain a Weibull model fitted for each class as shown in Algorithm 1. To do this, we first train a regular classification network. Next, we collect all the activation vectors and separate them into K sets, one for each known class. Each of these sets have shape N_j by K , where N_j is the number of samples belonging to class j . We proceed to remove column j for each set, where j indicates the index of the match score (line 1 of Algorithm 1).

The next step is to concatenate all of the non-match scores to obtain \mathbf{W} resulting in $N \cdot (K - 1)$ data points (line 2). We employ the FitHigh function of LibMR [4] to find the parameters of the Weibull distribution using only the q highest activations among the non-match class activations. LibMR is an open-source library for Meta-Recognition, and FitHigh is the function in LibMR for building Weibull models. Lastly, we pass \mathbf{W} to FitHigh to obtain the model p_j for class j (line 3).

Algorithm 1 Weibull calibration for class activations.

Require: LibMR FitHigh function

Require: Activation set $\mathbf{A} \in \mathcal{R}^{N \times K}$ for inputs $\mathbf{X} \in \mathcal{R}^{H \times W \times c}$ belonging to class j

Require: q , the number of highest activations to model

Ensure: Weibull model for class j

- 1: Obtain only non-match scores by removing column j from activation set \mathbf{A} to acquire new activation vectors $\tilde{\mathbf{A}} \in \mathcal{R}^{N \times (K-1)}$
 - 2: Concatenate all non-match scores to obtain $N \cdot (K - 1)$ data points resulting in \mathbf{W}
 - 3: Weibull Fit $p_j = (\rho_j, \kappa_j, \lambda_j) = \text{FitHigh}(\mathbf{W}, q)$
-

We derived a working algorithm by analyzing how non-match activations affect calibration in MetaMax's inference step. MetaMax's inference step (Algorithm 2) uses the per-class Weibull models to make a decision on whether to accept the network's SoftMax scores or to reject the input. In lines 3 through 6, we modify the modulation vector \mathbf{m} , which has been initialized to all ones. We obtain the key second step of MetaMax by extending OpenMax to allow for the modeling of non-match activations. The main difference concerns the modulation vector \mathbf{m} 's construction, where the Weibull cumulative distribution function is applied to modify the activation vector. Since we do not model the highest class activation, we effectively remove the influence of the highest activation on the unknown activation a_0 by skipping the index associated with the highest class activation. The consequence of this modification is that the unknown class activation (defined in line 8) simply adds zero when index i is equal to the index of the highest class acti-

vation. The new activations are then passed to the standard SoftMax function to produce the class probabilities (line 9). Finally, a prediction is found and a decision to accept or reject is made (line 11).

Algorithm 2 MetaMax probability estimation using non-match activations.

Require: β , the number of activation scores to revise

Require: LibMR models p_j

Require: Activation vector $\mathbf{a} \in \mathcal{R}^{K \times 1}$

- 1: Obtain sorted indices $\mathbf{b} = \text{argsort}(\mathbf{a})$
 - 2: **for** $i = 0, \dots, K - 1$ **do** $m_i = 1$
 - 3: **for** $i = 1, \dots, \beta$ **do**
 - 4: **if** $a_{b_i} > a_j \forall j \neq i$ **then** continue
 - 5: **else**
 - 6:
$$m_{b_i} = 1 - \frac{\beta - i}{\beta} e^{-\left(\frac{a_{b_i}}{\lambda_{b_i}}\right)^{\kappa_{b_i}}}$$
 - 7: Revise activation vector $\hat{\mathbf{a}} = \mathbf{a} \cdot \mathbf{m}$
 - 8: Define $a_K = \sum_i (a_i - a_i \cdot m_i)$
 - 9:
$$\hat{P}(y = j | \mathbf{x}) = \frac{e^{a_j}}{\sum_{i=0}^K e^{a_i}}$$
 - 10: Let $\hat{y} = \underset{j}{\text{argmax}} \hat{P}(y = j | \mathbf{x})$
 - 11: **if** \hat{y} is K **then** reject input
-

4. Experiments

In this section, we present the experimental details for testing MetaMax, compare it against OpenMax and other recent work, and report the measured performance on several public datasets. Additionally, to demonstrate the applicability of MetaMax to enable closed-set classifiers to perform OSR, we apply and report the performance of MetaMax to the following popular classification networks: DenseNet [9], ResNet [8], and VGGNet [25].

4.1. Experimental Setup

Following experimental setup in [28], we use DenseNet [9] as the backbone network for classification. We evaluate our method on the following common benchmark datasets: MNIST [15], SVHN [17], CIFAR10 [12], and TinyImageNet (TIN) [14]. To test under open-set conditions, we follow the most common data partition protocol [16]. Specifically, for MNIST, SVHN, and CIFAR10, we split each dataset at random such that 6 classes are chosen to be known and the remaining 4 classes to be unknown. For the TinyImageNet dataset, we train on $K = 20$ known classes and test on the full 200-class set. We repeat the experiment over 5 runs and report the average score. For the evaluation metrics, we use the area under receiver operating characteristic (AUROC) curve and the F1-scores.

Method	MNIST	SVHN	CIFAR10	TIN
G-OpenMax [6]	0.984	0.896	0.675	0.580
OSRCI [16]	0.988	0.910	0.699	0.586
CROSR [28]	0.991	0.899	0.883	0.589
C2AE [18]	0.989	0.922	0.895	0.748
GDFR [19]	-	0.935	0.807	0.608
CGDL [26]	0.994	0.935	0.903	0.762
OpenHybrid [29]	0.995	0.947	0.950	0.793
PROSR [31]	-	0.943	0.891	0.693
OpenGAN [11]	0.999	0.988	0.973	0.907
MetaMax (Ours)	0.997	0.977	0.938	0.846

Table 1: AUROC scores of MetaMax and the compared methods using DenseNet. TIN stands for TinyImageNet.

4.2. Results

In the first set of experiments, we demonstrate the effectiveness of MetaMax. We compare the AUROC score of MetaMax against recent methods in this area. We also show the multiclass ROC curves of our method on each dataset in the supplementary material. The ROC curves demonstrate the magnitude of difficulty with each dataset.

As shown in Table 1, MetaMax outperforms most recent methods and is on par with the state of the art. Note that most of these OSR methods either need to *modify the network architecture* or require *auxiliary training data*, which can be *nontrivial* to obtain in practice. On the contrary, MetaMax is *much simpler* and it can be *directly applied* post-hoc to a classifier with little-to-no extra overhead.

In the second set of experiments, to verify the wide applicability of allowing various classification networks to perform OSR using MetaMax, we apply and report the F1-scores for two other networks: VGGNet [25] and ResNet [8]. The results can be found in the supplementary material, which shows that MetaMax outperforms OpenMax and the baseline network consistently. This demonstrates the significance of our work in that MetaMax can potentially be applied to any classification network and enable it to operate under open-set conditions.

5. Conclusion

In this work we introduced MetaMax, an approach to calibrate deep neural network-based classifiers by modeling non-match class activations for OSR. Experiments on four standard image datasets demonstrate the effectiveness of the proposed method. As a simpler and more effective alternative to OpenMax, the modularization and general applicability of our method can have a wide impact in the community and benefit future research by applying MetaMax to calibrate standard closed-set classification networks for open-set conditions.

References

- [1] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. Concrete problems in ai safety. *arXiv preprint arXiv:1606.06565*, 2016.
- [2] Abhijit Bendale and Terrance Boulton. Towards open world recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1893–1902, 2015.
- [3] Abhijit Bendale and Terrance E Boulton. Towards open set deep networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1563–1572, 2016.
- [4] Terrance E Boulton, Walter J Scheirer, Vijay Iyer, Ethan Rudd, and Manuel Gunther. *LibMR*, 2018. <https://pypi.org/project/libmr>.
- [5] Thomas G Dietterich. Steps toward robust artificial intelligence. *AI Magazine*, 38(3):3–24, 2017.
- [6] Zongyuan Ge, Sergey Demianov, Zetao Chen, and Rahil Garnavi. Generative openmax for multi-class open set classification. In *Proceedings of the British Machine Vision Conference*. British Machine Vision Association and Society for Pattern Recognition, 2017.
- [7] Chuanxing Geng, Sheng-jun Huang, and Songcan Chen. Recent advances in open set recognition: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.
- [9] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4700–4708, 2017.
- [10] Lalit P Jain, Walter J Scheirer, and Terrance E Boulton. Multi-class open set recognition using probability of inclusion. In *Proceedings of the European Conference on Computer Vision*, pages 393–409. Springer, 2014.
- [11] Shu Kong and Deva Ramanan. Opengan: Open-set recognition via open data generation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 813–822, 2021.
- [12] Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images. Technical report, University of Toronto, Toronto, 2009.
- [13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105, 2012.
- [14] Ya Le and Xuan Yang. Tiny imagenet visual recognition challenge. *CS 231N*, 7(7):3, 2015.
- [15] Yann LeCun, Corinna Cortes, and Chris Burges. Mnist handwritten digit database, 2010.
- [16] Lawrence Neal, Matthew Olson, Xiaoli Fern, Weng-Keen Wong, and Fuxin Li. Open set learning with counterfactual images. In *Proceedings of the European Conference on Computer Vision*, pages 613–628, 2018.
- [17] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bisacco, Bo Wu, and Andrew Y Ng. Reading digits in natural images with unsupervised feature learning. In *Proceedings of the Conference on Neural Information Processing Systems Workshops*, 2011.
- [18] Poojan Oza and Vishal M Patel. C2ae: Class conditioned auto-encoder for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2307–2316, 2019.
- [19] Pramuditha Perera, Vlad I Morariu, Rajiv Jain, Varun Manjunatha, Curtis Wigington, Vicente Ordonez, and Vishal M Patel. Generative-discriminative feature representations for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11814–11823, 2020.
- [20] Nicolas Pinto, David D Cox, and James J DiCarlo. Why is real-world visual object recognition hard? *PLOS Computational Biology*, 4(1):e27, 2008.
- [21] Walter J Scheirer, Anderson de Rezende Rocha, Jonathan Parris, and Terrance E Boulton. Learning for meta-recognition. *IEEE Transactions on Information Forensics and Security*, 7(4):1214–1224, 2012.
- [22] Walter J Scheirer, Anderson de Rezende Rocha, Archana Sapkota, and Terrance E Boulton. Toward open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(7):1757–1772, 2012.
- [23] Walter J Scheirer, Lalit P Jain, and Terrance E Boulton. Probability models for open set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11):2317–2324, 2014.
- [24] Walter J Scheirer, Anderson Rocha, Ross J Micheals, and Terrance E Boulton. Meta-recognition: The theory and practice of recognition score analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1689–1695, 2011.
- [25] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [26] Xin Sun, Zhenning Yang, Chi Zhang, Keck-Voon Ling, and Guohao Peng. Conditional gaussian distribution learning for open set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13480–13489, 2020.
- [27] A Torralba and AA Efros. Unbiased look at dataset bias. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1521–1528, 2011.
- [28] Ryota Yoshihashi, Wen Shao, Rei Kawakami, Shaodi You, Makoto Iida, and Takeshi Naemura. Classification-reconstruction learning for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4016–4025, 2019.
- [29] Hongjie Zhang, Ang Li, Jie Guo, and Yanwen Guo. Hybrid models for open set recognition. In *Proceedings of the European Conference on Computer Vision*, pages 102–117. Springer, 2020.

- [30] Peng Zhang, Jiuling Wang, Ali Farhadi, Martial Hebert, and Devi Parikh. Predicting failures of vision systems. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3566–3573, 2014.
- [31] Da-Wei Zhou, Han-Jia Ye, and De-Chuan Zhan. Learning placeholders for open-set recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4401–4410, 2021.