Collective Decision of One-vs-Rest Networks for Open-Set Recognition

Jaeyeon Jang and Chang Ouk Kim

Abstract—Unknown examples that are unseen during training often appear in real-world pattern recognition tasks, and an intelligent self-learning system should be able to distinguish between known examples and unknown examples. Accordingly, open-set recognition (OSR), which addresses the problem of classifying knowns and identifying unknowns, has recently been highlighted. However, conventional deep neural networks (DNNs) using a softmax layer are vulnerable to overgeneralization, producing high confidence scores for unknowns. In this article, we propose a simple OSR method that is based on the intuition that the OSR performance can be maximized by setting strict and sophisticated decision boundaries that reject unknowns while maintaining satisfactory classification performance for knowns. For this purpose, a novel network structure, in which multiple one-vs-rest networks (OVRNs) follow a convolutional neural network (CNN) feature extractor, is proposed. Here, an OVRN is a simple feedforward neural network that is designed to assign confidence scores that are lower than those in the softmax layer to unknown samples so that unknown samples can be more effectively separated from known classes. Furthermore, the collective decision score is modeled by combining the multiple decisions reached by the OVRNs to alleviate overgeneralization. Extensive experiments were conducted on various datasets, and the experimental results show that the proposed method performs significantly better than the state-of-the-art methods by effectively reducing overgeneralization. The code is available at https://github.com/JaeyeonJang/Openset-collective-decision.

Index Terms—Collective decision, one-vs-rest networks (OVRNs), open-set recognition (OSR), overgeneralization, sigmoid.

I. INTRODUCTION

RECENT advancements in deep learning have greatly improved the performance of recognition systems [1]–[4], which can now surpass human-level performance in terms of classification error rates [5]. However, the vast majority of recognition systems are designed under closed-world assumptions, in which all categories are known a priori. Although these assumptions hold in many applications, the need to detect unknown objects at a given training time while classifying knowns, which is referred to as open-set recognition (OSR), has recently been highlighted [6]–[9]. Here, unknown data are unseen samples that can be categorized into meaningful classes, which are distinguished from known classes. In addition, any side information, including semantic/attribute information, cannot be provided in advance for training unknowns [10]. The ability to distinguish between knowns and unknowns has been considered a key element of intelligent self-learning systems [11].

Fig. 1 illustrates a typical example of an open-set scenario. (a) Only “koala,” “dog,” and “black stork” examples are provided during training, but “bear,” “cat,” “lion,” and “penguin,” unseen during training, appear during testing. (b) Generalized decision boundaries for closed-set classification cannot reject unknown samples in open space, which are far from any known training samples. (c) Thus, we must establish open-set classification boundaries that can reject the samples in open space.

Fig. 1. Illustration of an open-set scenario. (a) Only “koala,” “dog,” and “black stork” examples are provided during training, but “bear,” “cat,” “lion,” and “penguin,” unseen during training, appear during testing. (b) Generalized decision boundaries for closed-set classification cannot reject unknown samples in open space, which are far from any known training samples. (c) Thus, we must establish open-set classification boundaries that can reject the samples in open space.

Manuscript received 24 March 2021; revised 13 August 2021, 16 November 2021, 24 April 2022, and 23 June 2022; accepted 7 July 2022. Date of publication 14 July 2022; date of current version 6 February 2024. This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT), Korean Government, under Grant NRF-2019R1A2B5B01070358. (Corresponding author: Chang Ouk Kim.)

Jaeyeon Jang is with the Department of Data Science, The Catholic University of Korea, Bucheon 14662, Republic of Korea (e-mail: jaeyeon.jang@catholic.ac.kr).

Chang Ouk Kim is with the Department of Industrial Engineering, Yonsei University, Seoul 03722, Republic of Korea (e-mail: kimco@yonsei.ac.kr). Digital Object Identifier 10.1109/TNNLS.2022.3189996

2162-237X © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Many deep learning-based OSR methods have recently been proposed with the goal of mitigating overgeneralization [13]–[21]. Most OSR methods aim to provide regularized probability scores that can reject unknowns while classifying known samples. For this purpose, postscore analysis [22] has been applied to the recognition scores or latent features produced by learned networks. Recently, certain researchers have applied autoencoders [13], [16]–[19] and generative adversarial networks (GANs) [20], [21], in addition to deep neural networks (DNNs) using softmax layers, to provide more discriminative information for unknown identification. However, these reconstructive and generative networks can represent only a limited portion of unknowns because the diversity of unknowns is infinite. Thus, the key to OSR is to tighten the decision boundaries for known classes while maintaining satisfactory classification performance on known classes.

The commonly employed softmax activation cannot easily tighten the decision boundaries for known classes because softmax is designed to measure the relative likelihood of a known class compared with other known classes, yielding overgeneralized decision boundaries only for known classes [11]. Thus, in an open-set scenario, unknown samples are likely to be classified as belonging to one of the known classes. Moreover, a network with a sigmoid output layer is trained to tighten the decision boundary of each known class to discriminate a dissimilar example from match examples.

Based on this observation, in this article, we propose a new network structure in which a set of one-vs-rest networks (OVRNs) follows a convolutional neural network (CNN) feature extractor. Here, an OVRN is a simple feedforward neural network that uses a sigmoid output activation. Each OVRN learns class-specific discriminative features that individually distinguish between matches and nonmatches for each class, thereby assigning lower confidence scores to unknown samples than softmax CNN classifiers can. Here, the confidence score is the likelihood of belonging to a known class, usually calculated as the maximum posterior class membership probability \( \max_i P(y_i|x) \), where \( y_i \) is the \( i \)th class among \( M \) known classes [23]. In addition, given samples, the OVRNs of nontarget classes can provide useful information. For example, the lower the output of a nontarget class OVRN is, the higher the likelihood that the sample belongs to the target class. Thus, we propose a collective decision method that combines the multiple decisions reached by OVRNs to establish more sophisticated decision boundaries that reduce the open space.

Many existing approaches have introduced heavy auxiliary models, including GANs and autoencoders, in addition to the softmax CNN to achieve the capability to reject unknown samples [13], [16], [20], [21]. Alternatively, we can enhance the rejection capability by only replacing the softmax layer with a set of simple OVRNs and by introducing a simple collective decision method while minimizing the loss of generalization capability and the increase in computational complexity for training and testing. Comprehensive experiments were conducted on various datasets. The experimental results show that the proposed method outperforms state-of-the-art methods in diverse open-set scenarios and closed-set classification scenarios despite the simplicity of the implementation.

The remainder of this article is organized as follows. Section II presents related work on existing OSR methods. In Section III, we provide details of the proposed method. Section IV verifies the proposed method from various perspectives. In Section V, we conclude this study and present future research plans.

II. RELATED WORKS

OSR has been systematically investigated since the early 2010s. Initially, shallow machine learning model-based approaches were at the forefront of advancement [24]–[32]. During this period, many researchers redesigned traditional machine learning models, including support vector machines, to minimize open-set risk [24], which comprises an open space risk and an empirical risk. Here, the open space risk represents the risk of misclassifying an unknown example in open space as belonging to a known class, while the empirical risk represents the loss of classification performance for known data.

With the advancement of deep learning techniques, many deep learning-based OSR methods have been proposed. Most deep learning-based approaches aim to alleviate the overgeneralization effect of DNNs, which generally use a softmax output layer [14]–[16], [20], [21]. The first deep model introduced for OSR was OpenMax [14], which models the class-specific distribution of activation vectors in the penultimate layer of a CNN and transforms the distributional information into decision scores. Shu et al. [15] proposed a deep open classification (DOC) network, which is a CNN with OVR layers. They showed that OVR layers can further reduce the open space risk by tightening the decision boundaries via Gaussian fitting. Yoshihashi et al. [16] proposed a classification-reconstruction learning algorithm for OSR (CROS) that simultaneously implements both classification and reconstruction. The authors additionally utilized hierarchical latent representation in the OpenMax score calculation, showing that robust unknown detection is possible.

Generative adversarial learning has also been applied to generate samples to account for open space [20], [21]. Ge et al. [20] proposed a generative OpenMax (G-OpenMax) model that learns synthetic unknown samples generated by a conditional GAN (CGAN) to enhance the rejection capability. Here, CGAN generated samples based on a mixture of class labels. Neal et al. [21] proposed an OSR method using a counterfactual image (OSRCI), where an encoder–decoder GAN was utilized to produce a counterfactual sample that is near a known sample in the latent space but does not belong to any of the known classes. Similarly, Guo et al. [33] proposed a conservative novelty synthesizing network that explores and obtains marginal instances that are close to known families while falling into mimical unknown ones for malware OSR. Recently, Sabokrou et al. [34] proposed an end-to-end learning method that jointly trains a reconstructor (autoencoder) and a discriminator to generate compact representations and reconstruct realistic samples from the representations. Unlike the previous works based on generative learning, they utilized...
both networks in testing to detect out-of-distribution samples. On the other hand, Zhou et al. [35] proposed a novel idea that augments not only a synthetic unknown sample, but also classifiers by generating dummy classifiers. The authors suggested that learning multiple dummy classifiers can result in smoother decision boundaries by generating several scattered clusters. In addition, the authors applied a simple manifold mixup instead of a generative model for data augmentation. Although these methods require many hyperparameters to be tuned, they can produce synthetic samples that are limited to only a small subspace of unknowns. In other words, the aforementioned methods are not practical in a real-world application.

Some researchers have suggested two-stage methods that detect unknowns and then classify samples identified as unknown using a single softmax CNN to enhance the ability to detect unknowns. The autoencoder family that includes class-conditioned autoencoders and variational autoencoders has generally been applied for unknown detection in this category [17]–[19]. Recently, Zhang et al. [36] introduced an unsupervised flow-based model that has often been employed for out-of-distribution detection. The authors combined the flow-based model with a general closed-set classifier by adopting a multitask learning scheme during training. However, the classifier and flow-based model were separately utilized for closed-set classification and unknown detection during testing. The works in this category assumed that an open-set system can maximize the performance by well identifying unknown samples because high-performance classification on the samples detected as known can easily be achieved by adopting a state-of-the-art, closed-set classifier. However, the classification performance of certain classes with high intra-class diversity can be significantly reduced because this class is likely to have a wide positive area for unknowns when the unknown samples are misclassified by the unknown detection model.

Some researchers have proposed an OSR method that utilizes the concept of the OVR network. Yang et al. [23] proposed a convolutional prototype network (CPN) that learns several prototypes for each class and determines the class of a test sample based on the distances from the prototypes. In the CPN, an embedding layer follows a CNN feature extractor. The authors suggested an OVR discriminative loss function by combining distances from the prototypes in the embedding layer and introduced the Gaussian mixture density assumption to detect unknowns. Furthermore, Chen et al. [37] recently indicated that the prototypes that learned only the known samples may converge in the space of the unknown classes in the training, making the known and unknown classes indistinguishable. To address this problem, the authors proposed the concept of reciprocal points that constitute an instantiated representation of the extra-class space to reduce the open space risk by limiting the potential unknown space. These prototype-based models are similar to our model since they also replaced the conventional softmax output layer with a prototype-based OVR layer. However, these works simply used confidence scores for unknown detection and introduced many new parameters for the optimization of embedding space and for establishing a distance-based recognition rule. On the other hand, in this article, we show that making collective decisions through combining decision scores from multiple OVRs is sufficient for OSR, even though the proposed method has a very simple training procedure with a minor modification to the general network structure.

In the area of domain adaptation, Saito and Saenko [38] applied OVRs for universal domain adaptation. However, the authors did not consider combining multiple decisions of OVRs but proposed new techniques for the adaptation of OVRs to a new target domain, such as hard-negative sampling and open-set entropy minimization. Rather, the authors utilized the single decision of the OVR with the highest confidence for unknown detection. Similarly, Jang and Kim [39] proposed a probability model that is based on the concept of OVRs by applying a statistical extreme value theory (EVT)-based decision score calibration strategy. However, they did not find that different OVRs trained with different positive classes can assist each other in making decisions in an open-set environment. Fang et al. [40] also found the importance of OSR in the field of domain adaptation and provided learning bound for open-set domain adaptation by theoretically investigating the risk of the target classifier on unknown classes. Geng and Chen [41] proposed a method referred to as a collective decision for open-set recognition based on the hierarchical Dirichlet process, which aims to discover hidden unknown classes during testing. The name of the proposed method seems similar to our method. However, in the proposed method, “collective” means that a collective/batch operation was applied for testing. Instead, the authors proposed a single-model-based algorithm for recognition, referred to as co-clustering, that identifies multiple groups from given data by sharing mixture components among the groups.

III. PROPOSED METHOD

In this section, we propose a collective decision method that is based on OVRNs. Here, each OVRN uses sigmoid output activation to learn more discriminative features than the features learned using the general softmax output layer. Thus, first, we address the difference between using softmax and using sigmoid activation for the output layer in an open-set scenario. Second, a detailed description of an OVRN and the structural advantages are provided. Last, we describe the recognition rule based on collective decisions obtained by combining the decisions reached by multiple OVRNs.

A. Sigmoid Versus Softmax

Most DNN classifiers use a softmax output layer to learn categorical distributions. When $M$ known classes $y \in \mathbb{R}^M$ are given, a DNN that uses a softmax output layer learns the probability distribution over $M$ known classes for an input $x$. For the $i$th class $y_i$, the softmax layer produces the following conditional probability:

$$P(y_i | l_{y_1}, \ldots, l_{y_M}) = \frac{\exp(l_{y_i})}{\sum_{m=1}^{M} \exp(l_{y_m})}$$

(1)

where $l_{y_i}$ is the logit of class $y_i$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
A network using a softmax output layer is trained to increase \( l_{y_i} \) relative to the logits of the other classes. This training mechanism causes the activations of nontarget classes to converge to zero by increasing the logit of the target class in the normalization term (denominator), even though the logits of the nontarget classes do not decrease substantially [11], [42]. The activations that converge to zero backpropagate very small gradients through the network, and the network rarely learns latent representations that can be useful for discriminating nonmatch samples. In particular, since the softmax function is designed to measure the relative likelihood of a known class compared with the other known classes, the softmax layer gives a high confidence score to unknown samples by identifying the most similar class among all known classes. For example, in Fig. 2, softmax is highly likely to determine unknown “?” as a black circle class because the example is much closer to the black circle class than to the other classes.

In contrast, if sigmoid activation is applied to the output layer, then the sigmoid layer yields the probability for \( y_i \) conditioned only on \( l_{y_i} \), as follows:

\[
P(y_i | l_{y_i}) = \frac{1}{1 + \exp(-l_{y_i})}.
\]

Thus, each sigmoid output node is individually trained. Additionally, in contrast to softmax activation, match examples and nonmatch examples are learned equally through each sigmoid output node during training. Thus, each sigmoid node is trained to discriminate a dissimilar example from match examples.

To verify the difference between softmax and sigmoid, we produced confidence scores by inputting the MNIST [43] dataset into a CNN with a softmax layer or a sigmoid layer. MNIST was partitioned into six known classes (0–5) and four unknown classes (6–9). The experimental results are shown in Fig. 3. Since both networks provided high confidence scores greater than 0.99 for most known samples (98.9% of samples for CNN-sigmoid and 99.1% of samples for CNN-softmax), we only show the confidence scores of unknown samples. This finding means that most known samples belong to one of the known classes with a probability greater than 99%. However, the two networks show a large difference for unknown samples. Specifically, for unknown samples, the highest score bar can be reduced by approximately 20% by applying the sigmoid output layer. In addition, the CNN-sigmoid has a much longer tail of scores reaching zero, showing that the sigmoid layer alone can significantly reduce overgeneralization by assigning lower scores to unknown samples. However, it is known that neural networks with softmax output activation can handle overgeneralization by simply applying a temperature scaling technique that divides logit in (1) by the temperature during testing [44]. Thus, we additionally compare the sigmoid layer with the temperature-scaled softmax layer. Fig. 4 compares CNN-sigmoid with CNN-softmax to which the temperature scaling technique is applied. We use maximum logit values among the known classes instead of the confidence scores for CNN-sigmoid to achieve a better illustration and comparison. Because sigmoid is a monotonically increasing function, the maximum logit values provide the same separation results between known samples and unknown samples as the confidence scores. Fig. 4 shows that CNN-sigmoid shows better separation results when compared with CNN-softmax with high temperatures. This result shows that there is a limitation in temperature scaling for OSR because this technique does not affect the training mechanism of CNN-softmax.

When sigmoid activation is applied, each output node provides an individual determination. For example, in Fig. 2, the sigmoid layer can provide three individual determinations. In addition, in the presence of unknown samples, the posterior
JANG AND KIM: COLLECTIVE DECISION OF ONE-VS-REST NETWORKS FOR OPEN-SET RECOGNITION

Fig. 4. Comparison between CNN-sigmoid and CNN-softmax with temperature scaling on MNIST. $t$ denotes temperature. (a) Distribution of maximum logit values among the known classes. (b)–(d) Distributions of confidence scores with different temperatures.

probability of each class becomes supporting information. For instance, high posterior probabilities for multiple classes suggest that the example belongs to an unknown class.

B. One-vs-Rest Networks (OVRNs)

For the proposed method, the general softmax output layer is replaced with a set of OVRNs. Accordingly, the proposed network is composed of a CNN feature extractor and the following OVRNs, as shown in Fig. 5. Let $\mathcal{F}$ be the CNN feature extractor, and let $\mathcal{G}_i$ be the OVRN for class $y_i$. Next, the posterior probability of $x_j$ belonging to $y_i$ is defined as shown in (3). For each target class, one network with hidden layers and sigmoid output is constructed on the shared latent representation. Thus, compared with a single sigmoid layer, more specialized discriminative information can be learned after the shared latent representation for each target class. More sophisticated decision boundaries can be established

$$p_{ji} = P(y_i \mid x_j) = \mathcal{G}_i(\mathcal{F}(x_j)).$$

(3)

To produce $M$ individual decisions for $M$ known classes, the proposed network including both a CNN feature extractor and the following OVRNs is jointly trained to minimize the following binary cross-entropy:

$$\mathcal{L} = -\frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{M} I(t_j = y_i) \log p_{ji} + I(t_j \neq y_i) \log (1 - p_{ji}).$$

(4)


where \( N \) is the batch size, \( \mathbb{1} \) is the indicator function, and \( t_j \) is the ground-truth label of sample \( x_j \).

**C. Recognition Rule Based on Collective Decisions**

In this section, we propose a recognition rule based on the collective decisions of OVRNs. For a sample of class \( y_i \), the sample is more likely to belong to the target class when the sample has a high posterior probability for the target class and low posterior probabilities for all other classes. However, because the sigmoid function scales significantly when the input is relatively low or high, most nontarget OVRNs usually produce zero probability for a sample. Thus, the collective decision score is computed based on the logit of the sigmoid activation. Let \( l_{jy_i} \) be the logit value of a sample \( x_j \) for class \( y_i \). Next, \( \text{cds}_{jy_i} \), the collective decision score for class \( y_i \), is computed by the following simple function:

\[
\text{cds}_{jy_i} = l_{jy_i} - \frac{1}{M-1} \sum_{m \neq i} l_{jm} \tag{5}
\]

which produces relative logit values.

We can imagine three typical cases where OVRNs assign a target class to a given sample if we use the sigmoid output value as the basis for classification. cds denotes the collective decision score.

**D. Recognition Rule Using Collective Decision Scores**

We propose a recognition rule using collective decision scores for both closed-set classification and unknown detection as follows:

\[
y^* = \begin{cases} 
\arg \max_{y_i \in \{y_1, \ldots, y_M\}} \text{cds}_{jy_i}, & \text{if } \text{cds}_{jy_i} > \epsilon_{y_i} \\
"unknown", & \text{otherwise}
\end{cases} \tag{6}
\]

where \( \epsilon_{y_i} \) is the collective decision score threshold for class \( y_i \). \( \epsilon_{y_i} \) can easily be set based on the collective decision scores of the training data by class. In this article, we obtain \( \epsilon_{y_i} \) to ensure that 95% of the class \( y_i \) training data are recognized as known and classified as the target class.

**IV. EXPERIMENTS**

**A. Experimental Settings**

We conducted extensive experiments on various datasets: MNIST [43], EMNIST [45], SVHN [46], Omniglot [47], CIFAR-10 [48], CIFAR-100 [48], ImageNet [49], and LSUN [50]. In the OSR problem, the ratio of unknown classes to known classes affects the classification performance. Thus, openness was introduced to measure how open the problem setting is [24]

\[
\text{openness} = 1 - \sqrt{\frac{2C_T}{C_E + C_R}} \tag{7}
\]
where \( C_T \) is the number of classes utilized in training, \( C_E \) is the number of classes employed in evaluation (testing), and \( C_R \) is the number of classes to be recognized.

For backbone networks (CNN feature extractors), the plain CNN and redesigned VGGNet defined in [16] and ResNet-50 [1] were employed. We combined the backbone networks with single-hidden-layer OVRNs consisting of 64 hidden nodes. Features extracted by a backbone network passed through one hidden layer and were then fed to a sigmoid node for each class (refer to Fig. 5). ReLU activation was applied to the hidden nodes. Specifically, a plain CNN was utilized for MNIST, ResNet-50 was selected for large-scale ImageNet, and the redesigned VGGNet was employed for other datasets.

**B. Ablation Study**

First, we conducted a qualitative analysis to validate the effectiveness of the proposed method. For the analysis, the MNIST dataset was partitioned into six known classes (0–5) and four unknown classes (6–9). We produced confidence scores using a CNN with softmax output, a CNN with sigmoid output, and a CNN with OVRNs and normalized the confidence scores to a 0–1 scale. Fig. 7 shows histograms of the normalized confidence scores for the known and unknown classes. The histograms show that by extracting collective decision scores, unknown samples were more clearly separated from known samples compared with the results in Fig. 3 regardless of the network selected. In addition, CNN-OVRN was utilized for MNIST, ResNet-50 was selected for large-scale ImageNet, and the redesigned VGGNet was employed for other datasets.

**1) CNN-Softmax:** The first baseline model is a conventional softmax CNN. The threshold for unknown detection is set to 0.5, which is the most commonly employed value for the classwise rejection threshold, that is, a sample is rejected as unknown if the sample has an output score of less than 0.5 for all known classes.

**2) CNN-Sigmoid:** In this baseline, the softmax output layer in CNN-Softmax is replaced by a sigmoid layer. Unknown detection is implemented just as in CNN-Softmax.

**3) CNN-OVRN:** OVRNs are applied instead of the single sigmoid output layer. The training and testing procedures are the same as those in CNN-Sigmoid.

**4) Sigmoid-GF:** A Gaussian fitting (GF)-based threshold setting is added to baseline 2. This model is the DOC network, which was proposed for OSR applications in the field of natural language processing [15].

**5) OVRN-GF:** A GF-based threshold setting is added to baseline 3.

**6) Sigmoid-CD:** A collective, decision-based recognition rule is incorporated into CNN-Sigmoid.

**7) OVRN-CD (Proposed Method):** A collective decision-based recognition rule is incorporated into CNN-OVRN.

For the ablation study, ten-digit classes of MNIST and 47 letter classes of EMNIST were used as knowns and unknowns, respectively. The openness varied from 4.7% to 45.4% by randomly sampling 2–47 unknown classes in intervals of 5. For each openness, the performance of five randomized unknown class samplings was averaged, except in the case where all unknown classes were employed. To measure the performance, the macro-averaged F1 score \( \bar{F}_1 \) for the known classes and “unknown” classes was applied.

The experimental results are shown in Fig. 8. As the openness increased, the F1 scores of all baseline models generally decreased. Among the baselines, the proposed OVRN-CD achieved the best performance, except in the lowest openness case, thereby yielding the most robust results. CNN-Sigmoid performed better than CNN-OVRN when no calibrations were conducted for the confidence score. However, when the collective decision method was incorporated, OVRN-CD performed better as the openness increased. Specifically, for the highest openness case, the collective decision method improved the performance by 0.336 for CNN-OVRN. This result reveals that the proposed network structure can achieve synergy through combination with the collective decision method.

**C. Effects of Network Complexity**

In this section, we analyze the effect of network complexity when applying the proposed OVRN-CD. For this sensitivity
Fig. 8. F1 scores against openness for seven baseline methods.

Fig. 9. F1 scores according to the diverse complexities of OVRNs.

Fig. 10. F1 scores according to the diverse complexities of feature extractors.

Fig. 11. Samples from MNIST, Omniglot, Noise, and MNIST-Noise.

| Method              | Omniglot | Noise | MNIST-Noise |
|---------------------|----------|-------|-------------|
| Softmax             | 0.592    | 0.826 | 0.641       |
| OpenMax [14]        | 0.680    | 0.890 | 0.720       |
| LadderNet+Softmax [16]| 0.588  | 0.828 | 0.772       |
| LadderNet+OpenMax [16]| 0.764  | 0.826 | 0.821       |
| DHRNet+Softmax [16] | 0.595    | 0.829 | 0.801       |
| DHRNet+OpenMax [16] | 0.780    | 0.826 | 0.816       |
| CROSR [16]          | 0.793    | 0.826 | 0.827       |
| DOC [15]            | 0.863    | 0.921 | 0.892       |
| CGDL [18]           | 0.850    | 0.859 | 0.887       |
| PROSER [35]         | 0.862    | 0.874 | 0.882       |
| OVRN-CD (Ours)      | 0.918    | 0.953 | 0.926       |

Analysis, we applied the same setting of the ablation study, where CIFAR-10 and CIFAR-100 were used as knowns and unknowns, respectively. First, we analyzed the effect of OVRN complexity according to the number of hidden layers and nodes in the OVRN, as shown in Fig. 9. We set the linear mapping network that directly produces the logit value as the simplest model.

Fig. 9 shows that OVRNs outperform the linear mapping network in most cases. However, a single hidden layer OVRN consisting of two hidden nodes shows performance similar to the linear mapping network as the openness increases. This finding implies that the OVRN may not be effective if the network is too simple. The OVRN with four hidden nodes decreases more rapidly in the F1 score than other more complex networks as the openness increases. However, if an OVRN has eight or more hidden nodes, the performance gap is not significant, even if more nodes are added. Furthermore, the addition of hidden layers does not have a significant effect. In summary, an OVRN needs to provide only minimal complexity.

In most machine learning tasks, the complexity of the feature extractor is one of the key factors that influence performance. Thus, we analyzed the effect of feature extractor complexity. We incorporated backbone networks of various complexities into single-layer OVRNs. Specifically, reduced VGGNet 1, reduced VGGNet 2, ResNet-20, ResNet-32, ResNet-44, and ResNet-56 were selected for this sensitivity analysis, with the redesigned VGGNet as the baseline. The results are shown in Fig. 10. Here, reduced VGGNet 1 and VGGNet2 removed the last convolutional block and the last two convolutional blocks from the baseline, respectively. The last convolutional block and the second-to-last convolutional block are composed of four convolutional layers with one max-pooling layer and two convolutional layers with one max-pooling layer, respectively. ResNet-x denotes a residual
network with $x$ layers. Details about the ResNet family are provided in [1].

The comparison among the reduced VGGNet 1, reduced VGGNet 2, and baseline reveals that the feature extractor must be complex enough to provide a good OSR performance. In addition, the performance gap widens as the openness increases. Interestingly, the F1 scores of the reduced VGGNet 1 and VGGNet 2 are even lower than those of the linear mapping network with a more complex feature extractor in Fig. 9 for all openness values. This finding demonstrates the importance of introducing a sufficiently complex feature extractor. However, increasing the network complexity does not guarantee better performance if the complexity of a backbone network already exceeds the minimally required level.

### D. Experiments Involving Open-Set Recognition

We compared the proposed method with the state-of-the-art methods in terms of OSR performance. The training samples of the MNIST or CIFAR-10 dataset were used for training, but for testing, another dataset was added to the MNIST or CIFAR-10 test samples as unknowns. Specifically, when the MNIST data were used for training, we utilized three datasets of grayscale images Omniglot, Noise, and MNIST-Noise as the unknowns (refer to Fig. 11). Here, the noise is a set of images synthesized by independently sampling each pixel value from a uniform distribution on $[0, 1]$, and MNIST-Noise is a synthesized dataset created by superimposing the test images of the MNIST on the Noise. Each “unknown” dataset contains 10 000 samples, the same as the MNIST test dataset, making the known to unknown ratio 1:1. The F1 score comparison results are shown in Table I. Details about LadderNet and DHRNet are presented in [51] and [16]. In the table, the best performance and second-best performance are highlighted in bold and underlined, respectively. PROSER provided a slightly higher F1 score than ours when CIFAR-10 was trained. However, the proposed OVRN-CD made a vast improvement in OSR performance when MNIST was selected for training. Considering that OVRN-CD requires much simpler training and testing procedures than state-of-the-art methods, including PROSER, introducing a collective decision method based on OVRNs is very practical and effective.

### E. Experiments on Closed-Set Classification and Unknown Detection

Enhancing the rejection capability may cause a loss of the generalization capability that is critical in classifying known class samples. Thus, a good OSR model must show high performance in both closed-set classification and unknown detection, which represent generalization and rejection capabilities, respectively. To evaluate OVRN-CD from this point of view, we conducted additional experiments to compare the proposed model with state-of-the-art one-stage OSR methods in terms of the closed-set accuracy and area under the receiver operating curve (AUROC). Here, the closed-set accuracy measures the classification performance on known data, and the AUROC measures the performance in discriminating unknown samples from known classes. We followed the experimental protocol suggested in [21] with the six datasets MNIST, SVHN, CIFAR-10, CIFAR + 10, CIFAR + 50, and Tiny-ImageNet. The MNIST, SVHN, and CIFAR-10 datasets were randomly partitioned into six known classes and four unknown classes. For these three datasets, the closed-set classification

---

**TABLE II**

Open-Set Recognition Results on CIFAR-10 With Various “Unknown” Datasets

| Method               | ImageNet-crop | ImageNet-resize | LSUN-crop | LSUN-resize | Avg.    |
|----------------------|---------------|-----------------|-----------|-------------|---------|
| Softmax              | 0.639         | 0.653           | 0.642     | 0.647       | 0.645   |
| OpenMax [14]         | 0.660         | 0.684           | 0.657     | 0.668       | 0.667   |
| LadderNet+Softmax [16]| 0.640         | 0.646           | 0.644     | 0.647       | 0.644   |
| LadderNet+OpenMax [16]| 0.535         | 0.670           | 0.652     | 0.659       | 0.659   |
| DHRNet+Softmax [16]  | 0.645         | 0.649           | 0.650     | 0.649       | 0.648   |
| DHRNet+OpenMax [16]  | 0.655         | 0.675           | 0.656     | 0.664       | 0.663   |
| CROSSR [16]          | 0.721         | 0.735           | 0.720     | 0.749       | 0.731   |
| DOC [15]             | 0.760         | 0.753           | 0.748     | 0.764       | 0.756   |
| MLOSIR [17]          | 0.837         | 0.826           | 0.783     | 0.801       | 0.812   |
| CGDL [18]            | 0.840         | 0.832           | 0.806     | 0.812       | 0.823   |
| PROSER [35]          | **0.849**     | 0.824           | **0.867** | **0.856**   | **0.849** |
| OVRN-CD (Ours)       | 0.835         | 0.825           | 0.846     | 0.839       | 0.836   |

---

**TABLE III**

Closed-Set Classification Results on Various Datasets

| Method   | MNIST   | SVHN    | CIFAR-10 |
|----------|---------|---------|----------|
| Softmax  | 995 ± 0.02 | 947 ± 0.06 | 801 ± 0.03 |
| OpenMax [14] | 995 ± 0.02 | 947 ± 0.06 | 801 ± 0.03 |
| G-OpenMax [20] | 996 ± 0.01 | 948 ± 0.08 | 816 ± 0.05 |
| OSRCl [21] | 996 ± 0.01 | 951 ± 0.06 | 821 ± 0.02 |
| CROSSR [16] | 992 ± 0.01 | 945 ± 0.05 | 930 ± 0.02 |
| CPN [23] | 997 ± 0.01 | 967 ± 0.04 | 927 ± 0.02 |
| OVRN-CD   | **998 ± 0.01** | **975 ± 0.03** | **932 ± 0.01** |
TABLE IV
UNKNOWN DETECTION RESULTS ON VARIOUS DATASETS

| Method   | MNIST | SVHN  | CIFAR-10 | CIFAR+10 | CIFAR+50 | Tiny-ImageNet |
|----------|-------|-------|----------|----------|----------|---------------|
| Softmax  | .978 ± .006 | .886 ± .014 | .667 ± .038 | .816 ± .016 | .805 ± .005 | .577           |
| OpenMax [14] | .981 ± .005 | .894 ± .013 | .695 ± .044 | .817 ± .017 | .796 ± .056 | .576           |
| G-OpenMax [20] | .984 ± .005 | .896 ± .017 | .675 ± .044 | .827 ± .019 | .819 ± .058 | .580           |
| OSRC [21] | .988 ± .004 | .910 ± .010 | .699 ± .038 | .838 ± .017 | .827 ± .058 | .586           |
| CROSR [16] | .991 ± .004 | .899 ± .018 | -         | -        | -         | -              |
| CPN [23]  | .990 ± .002 | .926 ± .006 | .828 ± .021 | .881 ± .079 | .879 ± .069 | .639           |
| OVRN-CD  | .989 ± .003 | .941 ± .004 | .903 ± .020 | .907 ± .011 | .902 ± .003 | .730 ± .018   |

accuracy and AUROC were measured. For CIFAR + 10 and CIFAR + 50, four nonanimal classes were utilized, but ten and 50 animal classes were randomly drawn from the more diverse CIFAR-100. Twenty classes were randomly chosen from the Tiny-ImageNet dataset as known classes, and the remaining 180 classes were set as unknown. CIFAR + 10, CIFAR + 50, and Tiny-ImageNet were used to measure the AUROC. For all datasets, the random class split was repeated five times, and the averaged accuracy and AUROC are reported with the standard deviation in Tables III and IV. For CIFAR + 10, CIFAR + 50, and Tiny-ImageNet, the standard deviation was not reported in the literature.

The tables show that the proposed OVRN-CD significantly outperforms the other approaches on most datasets in both closed-set classification and unknown detection tasks. Although CROSR and CPN provide superior unknown detection performance on MNIST, they are significantly inferior when the other datasets are utilized. Therefore, we can state that the proposed model is much more effective than the other models despite the simplicity of implementing OVRN-CD.

F. Experiments Using Large-Scale Datasets

To compare our method with conventional CNNs in harsh but realistic environments where large-scale datasets are considered, we selected the ImageNet dataset, which includes 1000 classes with more than 1200000 training images and 50000 validation images. Specifically, we selected 50 and 100 classes as known classes, and the remaining classes were considered unknown. As the backbone network, ResNet-50 [1] was adopted for the CNN feature extractor. We compared OVRN-CD with CNN-softmax, CNN-sigmoid, and DOC [15] in terms of the closed-set classification accuracy and with AUROC for unknown detection. In addition, we introduce a new evaluation metric, the open-set classification rate (OSCR) [52], which fairly combines the accuracy of known classes and the unknown detection performance. Specifically, the OSCR measures the correct classification rate (CCR) given the false positive rate (FPR).

The CCR is the fraction of samples where the target known class \( y_k \) has the maximum probability and has a probability greater than \( \delta \). Let \( D_k \) and \( D_U \) be the set of class \( y_k \) samples and set of unknown samples, respectively. Then, CCR is defined as follows:

\[
\text{CCR}(\delta) = \frac{\{x \in D_k \land \arg\max_{y_k} P(y_k|x) = y_k \land P(y_k|x) \geq \delta\}}{|D_k|}.
\]  

(8)

The FPR is the fraction of unknown samples that are classified as any known class \( y_k \) with a probability greater than \( \delta \)

\[
\text{FPR}(\delta) = \frac{\{|x \in D_U \land \max_{y} P(y|x) \geq \delta\}|}{|D_U|}.
\]  

(9)

A larger value of the OSCR indicates better OSR performance. The comparison results for closed-set classification and unknown detection are shown in Table V. In the table, DOC shows the same closed-set accuracy as sigmoid CNN since it applies the same recognition rule using sigmoid CNN for closed-set classification. According to the table, OVRN-CD outperforms softmax CNN in both closed-set classification and unknown detection tasks, although softmax is the de facto activation for closed-set classification. In particu-
lar, OVRN-CD can achieve high-performance improvement for unknown detection. OVRN-CD provides slightly lower closed-set accuracy compared with sigmoid CNN and DOC. However, interestingly, OVRN-CD obtains a much higher AUROC even though all of the methods use the same sigmoid activation. In addition, OVRN-CD achieves the lowest drop, 0.007, in unknown detection performance after increasing the number of known classes from 50 to 100. In addition, Fig. 12 shows that OVRN-CD achieved a much better OSCR curve. These results demonstrate that OVRN-CD works well on open-set scenarios using a large-scale real-world dataset with a deep feature extractor, showing the generality of the proposed method.

V. CONCLUSION

Alleviating the overgeneralization inherent in the closed-set DNN classifier is the key to a high-performance OSR system. Thus, in this article, we proposed a collective decision method that combines decisions reached by different OVRNs. The proposed method improves open-set recognition performance by producing class-specific decisions for each OVRN and by combining them into the collective decision score, as verified through ablation studies. In addition, extensive comparison experiments were conducted on multiple standard datasets. The experimental results demonstrated that the proposed method outperforms state-of-the-art methods in many OSR scenarios, despite its simplicity.

The collective decision method should be extended so that we can provide rigorous multiclass probabilities, including the probability of unknowns. We are considering the rules of combining multiple classifiers suggested in [53]. In addition, the OSR system should be extended so that new unknown classes can be incorporated into the trained system. Based on the framework proposed in [54], we plan to add a new feature of incremental class learning to the proposed method. These research themes will be addressed in our future work.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 770–778.

[2] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep face recognition,” in Proc. Brit. Mach. Vis. Conf., 2015, pp. 1–12.

[3] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proc. Int. Conf. Learn. Represent., 2015, pp. 1–14.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1106–1114.

[5] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification,” in Proc. IEEE/CVF Int. Conf. Comput. Vis., Dec. 2015, pp. 1026–1034.

[6] J. Jang, M. Seo, and C. O. Kim, “Support weighted ensemble model for open set recognition of wafer map defects,” IEEE Trans. Semicond. Manuf., vol. 33, no. 4, pp. 635–643, Nov. 2020.

[7] D. Jung, “Data-driven open-set fault classification of residual data using Bayesian filtering,” IEEE Trans. Control Syst. Technol., vol. 28, no. 5, pp. 2045–2052, Sep. 2020.

[8] J. Henrydoss, S. Cruz, E. M. Rudd, M. Gunther, and T. E. Boul, “Incremental open set intrusion recognition using extreme value machine,” in Proc. Int. Conf. Mach. Learn. App., 2017, pp. 1089–1093.

[9] S. Cruz, C. Coleman, E. M. Rudd, and T. E. Boul, “Open set intrusion recognition for fine-grained attack categorization,” in Proc. IEEE Int. Symp. Homeland Secur. (HST), Apr. 2017, pp. 1–6.

[10] C. Geng, S.-J. Huang, and S. Chen, “Recent advances in open set recognition: A survey,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 10, pp. 1–19, Oct. 2020.

[11] T. E. Boul, S. Cruz, A. R. Dhamija, M. Gunther, J. Henrydoss, and W. J. Scheirer, “Learning and the unknown: Surveying steps toward open world recognition,” in Proc. AAAI Conf. Artif. Intell., vol. 33, Jul. 2019, pp. 9801–9807.

[12] S.-M. Moosavi-Dezfooli, A. Fawzi, J. Usato, and P. Frossard, “Robustness via curvature regularization and vice versa,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2019, pp. 9070–9078.

[13] G. Spigler, “Denoising autoencoders for overgeneralization in neural networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 4, pp. 998–1004, Apr. 2020.

[14] A. Bendale and T. E. Boul, “Towards open set deep networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 1563–1572.

[15] L. Shu, H. Xu, and B. Liu, “DOC: Deep open classification of text documents,” in Proc. Conf. Empirical Methods Natural Lang. Process., 2017, pp. 2911–2916.

[16] R. Yoshihashi, W. Shao, R. Kawakami, S. You, M. Iida, and T. Naemura, “Classification-reconstruction learning for open-set recognition,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 4011–4020.

[17] P. Oza and V. M. Patel, “Deep CNN-based multi-task learning for open-set recognition,” 2019, arXiv:1903.03161.

[18] X. Sun, Z. Yang, C. Zhang, K.-V. Ling, and G. Peng, “Conditional Gaussian distribution learning for open set recognition,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 13477–13486.

[19] P. Oza and V. M. Patel, “C2AE: Class conditioned auto-encoder for open-set recognition,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 2307–2316.

[20] Z. Ge, S. Demyanov, and R. Garnavi, “Generative OpenMax for multi-class open set classification,” in Proc. Brit. Mach. Vis. Conf., 2017, pp. 1–12.

[21] L. Neal, M. Olson, X. Fern, W.-K. Wong, and F. Li, “Open set learning with counterfactual images,” in Proc. Eur. Conf. Comput. Vis., 2018, pp. 613–628.

[22] W. J. Scheirer, A. Rocha, R. J. Micheals, and T. E. Boul, “Meta-recognition: The theory and practice of recognition score analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 8, pp. 1689–1695, Aug. 2011.

[23] H.-M. Yang, X.-Y. Zhang, F. Yin, Q. Yang, and C.-L. Liu, “Convolutional prototype network for open set recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 5, pp. 1311–1323, May 2020.

[24] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, “Toward open set recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 7, pp. 1757–1772, Jul. 2013.

[25] H. Cevikalp, “Best fitting hyperplanes for classification,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1076–1088, Jun. 2017.

[26] W. J. Scheirer, L. P. Jain, W. J. Scheirer, and T. E. Boul, “Probability models for open set recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 11, pp. 2317–2324, Nov. 2014.

[27] L. P. Jain, W. J. Scheirer, and T. E. Boul, “Multi-class open set recognition using probability of inclusion,” in Proc. Eur. Conf. Comput. Vis., 2014, pp. 393–409.

[28] H. Zhang and V. M. Patel, “Sparse representation-based open set recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 8, pp. 1690–1696, May 2017.

TABLE V
CLOSED-SET CLASSIFICATION AND UNKNOWN DETECTION PERFORMANCE ON IMAGE NET

| Method | # Known classes: 50 | | # Known classes: 100 | |
|----------------|----------------|----------------|----------------|----------------|
| | Accuracy | AUROC | Accuracy | AUROC |
| Softmax | 0.780 | 0.906 | 0.752 | 0.889 |
| Sigmoid | 0.800 | 0.928 | 0.768 | 0.919 |
| DOC | 0.800 | 0.865 | 0.768 | 0.834 |
| OVRN-CD | 0.796 | 0.960 | 0.767 | 0.953 |
[29] A. Bendale and T. Boul, “Towards open world recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 1893–1902.

[30] P. R. Mendes et al., “Nearest neighbors distance ratio open-set classifier,” Mach. Learn., vol. 106, pp. 359–386, Mar. 2017.

[31] E. M. Rudd, L. P. Jain, W. J. Scheirer, and T. E. Boul, “The extreme value machine,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 3, pp. 762–768, Mar. 2017.

[32] L. Livi, A. Sadeghian, and W. Pedrycz, “Entropic one-class classifiers,” IEEE Trans. Neural Netw. Learn. Syst., vol. 26, no. 12, pp. 3187–3200, Dec. 2015.

[33] J. Guo, S. Guo, S. Ma, Y. Sun, and Y. Xu, “Conservative novelty synthesizing network for malware recognition in an open-set scenario,” IEEE Trans. Neural Netw. Learn. Syst., early access, Aug. 2, 2021.

[34] M. Sabokrou, M. Fathy, G. Zhao, and E. Adeli, “Deep end-to-end one-class classifier,” IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 2, pp. 675–684, Feb. 2021.

[35] H. Zhang, A. Li, J. Guo, and Y. Guo, “Hybrid models for open set recognition,” in Proc. Eur. Conf. Comput. Vis., 2020, pp. 1–16.

[36] G. Chen, P. Peng, X. Wang, Y. Tian, and S. Member, “Adversarial reciprocal points learning for open set recognition,” IEEE Trans. Pattern Anal. Mach. Intell., early access, Aug. 24, 2021, doi: 10.1109/TPAMI.2021.3106743.

[37] K. Saito and K. Saenko, “OVANet: One-vs-all network for universal domain adaptation,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2021, pp. 9000–9009.

[38] J. Jang and C. Ouk Kim, “One-vs-rest network-based deep probability model for open set recognition,” 2020, arXiv:2004.08067.

[39] Z. Fang, J. Lu, F. Liu, J. Xuan, and G. Zhang, “Open set domain adaptation: Theoretical bound and algorithm,” IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 10, pp. 4309–4322, Oct. 2021.

[40] C. Geng and S. Chen, “Collective decision for open set recognition,” IEEE Trans. Knowl. Data Eng., vol. 34, no. 1, pp. 192–204, Jan. 2022.

[41] A. Oland, A. Bansal, R. B. Dannenberg, and B. Raj, “Be careful what you backpropagate: A case for linear output activations & gradient boosting,” 2017, arXiv:1707.04199.

[42] L. Deng, “The MNIST database of handwritten digit images for machine learning research,” IEEE Signal Process. Mag., vol. 29, no. 6, pp. 141–142, Nov. 2012.

[43] S. Liang, Y. Li, and R. Srikant, “Enhancing the reliability of out-of-distribution image detection in neural networks,” in Proc. Int. Conf. Mach. Learn., 2018, pp. 1–15.

[44] G. Cohen, S. Afshar, J. Tapson, and A. Van Schaik, “EMNIST: Extending MNIST to handwritten letters,” in Proc. IJCNN, 2017, pp. 2921–2926.

[45] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, “Reading digits in natural images with unsupervised feature learning,” in Proc. Adv. Neural Inf. Process. Syst. Deep Learn. Unsupervised Feature Learn., 2011, pp. 1–9.

[46] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, “Reading digits in natural images with unsupervised feature learning,” in Proc. Adv. Neural Inf. Process. Syst. Deep Learn. Unsupervised Feature Learn., 2011, pp. 1–9.

[47] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, “Human-level concept learning through probabilistic program induction,” Science, vol. 350, no. 6266, pp. 1332–1338, 2015.

[48] A. Krizhevsky, V. Nair, and G. Hinton. (2009). CIFAR-10 and CIFAR-100 Datasets. [Online]. Available: https://www.cs.toronto.edu/kriz/cifar.html

[49] O. Russakovsky et al., “ImageNet large scale visual recognition challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015.

[50] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao, “LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop,” 2015, arXiv:1506.03365.

[51] D. Zhou, “From neural PCA to deep unsupervised learning,” in Advances in Independent Component Analysis and Learning Machines. Amsterdam, The Netherlands: Elsevier, 2015, pp. 143–171.

[52] A. R. Dhamija, M. Günther, and T. E. Boul, “Reducing network agnosophobia,” in Proc. Adv. Neural Inf. Process. Syst., 2018, pp. 9157–9168.

[53] C.-L. Liu, “Classifier combination based on confidence transformation,” Pattern Recognit., vol. 38, no. 1, pp. 11–28, Jan. 2005.

[54] D.-W. Zhou, Y. Yang, and D.-C. Zhan, “Learning to classify with incremental new class,” IEEE Trans. Neural Netw. Learn. Syst., vol. 33, no. 6, pp. 2429–2443, Jun. 2022.

Jaeyeon Jang received the Ph.D. degree in industrial engineering from Yonsei University, Seoul, South Korea, in 2021. From 2021 to 2022, he was a Post-Doctoral Scholar with the Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, USA. In 2022, he joined the Department of Data Science, The Catholic University of Korea, Bucheon, South Korea, where he is currently an Assistant Professor. His current research interests include industrial data analytics, deep learning, machine learning, and reinforcement learning.

Chang Ouk Kim received the Ph.D. degree in industrial engineering from Yonsei University, Seoul, South Korea, in 2011. He currently teaches at the Department of Industrial Engineering, Yonsei University, Seoul, South Korea. He has published more than 100 papers in journals and conference proceedings. His current research interests include pattern recognition, machine learning, and data science for manufacturing and defense analysis.