

# Recent Advances in Open Set Recognition: A Survey

Chuanxing Geng, Sheng-Jun Huang and Songcan Chen

**Abstract**—In real-world recognition/classification tasks, limited by various objective factors, it is usually difficult to collect training samples to exhaust all classes when training a recognizer or classifier. A more realistic scenario is open set recognition (OSR), where incomplete knowledge of the world exists at training time, and unknown classes can be submitted to an algorithm during testing, requiring the classifiers not only to accurately classify the seen classes, but also to effectively deal with the unseen ones. This paper provides a comprehensive survey of existing open set recognition techniques covering various aspects ranging from related definitions, representations of models, datasets, experiment setup and evaluation metrics. Furthermore, we briefly analyze the relationships between OSR and its related tasks including zero-shot, one-shot (few-shot) recognition/learning techniques, classification with reject option, and so forth. Additionally, we also overview the open world recognition which can be seen as a natural extension of OSR. Importantly, we highlight the limitations of existing approaches and point out some promising subsequent research directions in this field.

**Index Terms**—Open set recognition/classification, open world recognition, zero-shot learning, one-shot learning.

## I. INTRODUCTION

UNDER a common closed set (or static environment) assumption: the training and testing data are drawn from the same label space and the same distribution, the traditional recognition/classification algorithms have already achieved significant success in a variety of machine learning (ML) tasks. However, a more realistic scenario is usually open and non-stationary such as driverless, fault/medical diagnosis, etc., where unseen situations can emerge unexpectedly, which drastically weakens the robustness of these existing methods. To meet this challenge, several related research directions actually have been explored including lifelong learning [1], [2], transfer learning [3]–[5], domain adaptation [6], zero-shot [7]–[9], one-shot (few-shot) [10]–[16] recognition/learning and open set recognition/classification [17]–[19], and so forth.

Based on Donald Rumsfeld’s famous “There are known knowns” statement [20], we further expand the basic recognition categories of classes asserted by [18], where we restate that recognition should consider four basic categories of classes as follows:

The authors are with College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics of (NUAA), Nanjing, 211106, China. Corresponding author is Songcan Chen. E-mail: {gengchuanxing; huangsj; s.chen}@nuaa.edu.cn.

This is a preliminary and will be kept updated, any suggestions and comments are welcome.

- 1) *known known classes*, i.e., the classes with distinctly labeled positive training samples (also serving as negative samples for other known known classes), and even have the corresponding semantic/attribute information;
- 2) *known unknown classes*, i.e., labeled negative samples, not necessarily grouped into meaningful classes;
- 3) *unknown known classes*<sup>\*1</sup>, i.e., classes with no available samples in training, but available semantic/attribute information of them during training;
- 4) *unknown unknown classes*, i.e., classes without any information regarding them: not only unseen in training but also having not side-information such as the semantic/attribute information, etc., during training.

Traditional classification only considers known known classes, while including known unknown classes will result in models with an explicit “other class,” or a detector trained with unclassified negatives [18]. Unlike the traditional classification, zero-shot learning (ZSL) can identify unseen classes which have no available observations in training. However, the available semantic/attribute information shared among all classes including seen and unseen ones are needed [7]. The ZSL mainly focuses on the recognition of the unknown known classes defined above. Actually, such a setting is rather restrictive and impractical, since we usually know nothing about the testing samples which may come from known known classes or not [8]. Therefore, some researchers have begun to pay attention to the more generalized ZSL (G-ZSL) [21], where the testing samples come from both known known and unknown known classes. As a closely-related problem to ZSL, one-shot (few-shot) learning can be seen as natural extensions of zero-shot learning when limited number of samples of unseen classes during training are available [10]–[16]. Compared to zero-shot and one-shot (few-shot) learning, open set recognition (OSR) [17]–[19] probably faces more serious challenge due to the fact that only known known classes are available without any other side-information like attributes or one or few samples of unknown unknown classes.

Actually, Open set recognition [17] describes such a scenario where new classes (unknown unknown classes) unseen in training appear in testing, and requires the classifiers not only to accurately classify the known known classes, but also to effectively deal with the unknown unknown ones. Therefore, the classifiers need to have a corresponding reject option when a testing sample comes from some unknown unknown class. Fig. 1 gives a comparative demonstration of traditional

<sup>\*1</sup> represents the expanded basic recognition class by ourselves.

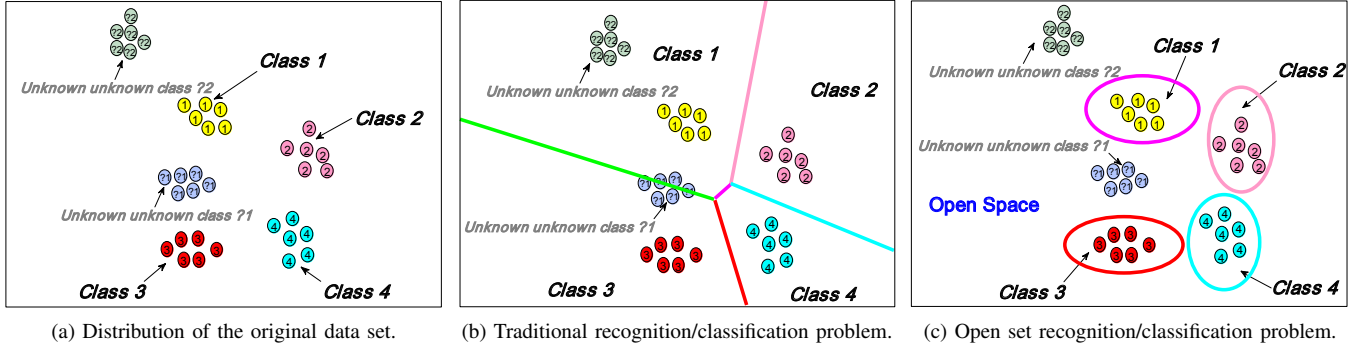


Fig. 1. Comparison of between traditional classification and open set recognition. Fig. 1(a) denotes the distribution of original dataset including known known classes 1,2,3,4 and unknown unknown classes ?1,?2, where known known classes only appear during training while unknown unknown classes may appear or not during testing. Fig. 1(b) shows the decision boundary of each class obtained by traditional classification methods, and it will obviously misclassify when unknown unknown classes ?1,?2 appear during testing. Fig. 1(c) describes the open set recognition, where the decision boundaries limit the scope of known known classes 1,2,3,4, reserving space for unknown unknown classes ?1,?2. Via these decision boundaries, samples from an unknown unknown class are labeled as “unknown” or rejected rather than misclassified as known known classes.

classification and OSR problems. It should be noted that there have been already a variety of works in the literature regarding classification with reject option [22]–[32]. Although related in some sense, this task should not be confused with open set recognition since it still works under the closed set assumption, while the corresponding classifier rejects to recognize an input sample due to its low confidence, avoiding classifying a sample of one class as a member of another one.

Additionally, the one-class classifier [33]–[40] usually used for anomaly detection seems suitable for OSR problem, in which the empirical distribution of training data is modeled such that it can be separated from the surrounding open space (the space far from known/training data) in all directions of the feature space [41]. Popular approaches for one-class classification include one-class SVM [33] and support vector data description (SVDD) [35], [42], where one-class SVM separates the training samples from the origin of the feature space with a maximum margin, while SVDD encloses the training data with a hypersphere of minimum volume. However, treating multiple known known classes as a single one in the one-class setup obviously ignores the discriminative information between these known known classes, leading to poor classification performance [19], [43]. Even if each known known class is modeled by an individual one-class classifier as proposed in [24], the classification performance is rather low [43]. Therefore, it is necessary to rebuild effective classifiers specifically for OSR problem, especially for multiclass OSR problem.

As a summary, Table I lists the differences between open set recognition and its related tasks mentioned above. In fact, OSR has been studied under a number of frameworks, assumptions, and names [44]–[49]. In a study on evaluation methods for face recognition, Phillips et al. [44] proposed a typical framework for open set identity recognition, while Li and Wechsler [45] again viewed open set face recognition from an evaluation perspective and proposed Open Set TCM-kNN (Transduction Confidence Machine-k Nearest Neighbors) method, which was suitable for multiclass authentication operational scenarios. In 2013, it is Scheirer et al. [17] that first formalized the open set

recognition problem and proposed a preliminary solution—1-vs-Set machine, which incorporates an open space risk term in modeling to account for the space beyond the reasonable support of known known classes. Afterwards, open set recognition attracted widespread attention, and simultaneously, to our best knowledge, till now there has been no systematical summary on this topic. Therefore, we here mainly provide a comprehensive review regarding the open set recognition.

The rest of this paper is organized as follows. In the next three sections, we first give the basic notation and related definitions (Section II), then we categorize the existing OSR technologies from the modeling perspective, and for each category, we review different approaches, given in Table II in detail (Section III). Lastly, we overview the open world recognition (OWR) which can be seen as a natural extension of OSR in Section IV. Furthermore, Section V reports the commonly used datasets, experiment setup and evaluation metrics, while Section VI highlights the limitations of existing approaches and points out some promising research directions in this field. Finally, Section VII gives a conclusion.

## II. BASIC NOTATION AND RELATED DEFINITION

Assume feature vector  $x \in \mathbb{R}^d$  describes the data point,  $y$  labeled by  $\mathbb{N}^+$  represents countably class, and there exists a probability measure  $P(x, y)$  over  $(x, y) \in \mathbb{R}^d \times \mathbb{N}^+$ . The fundamental multiclass recognition problem would be to find a recognition function  $f$  that minimizes the ideal risk  $R_{\mathcal{I}}$ :

$$\arg \min_f \left\{ R_{\mathcal{I}}(f) := \int_{\mathbb{R}^d \times \mathbb{N}} L(x, y, f(x)) P(x, y) \right\}, \quad (1)$$

where  $L(x, y, f(x)) \geq 0$  denotes the loss function. However, the ideal risk function is often not available since the joint distribution  $P(x, y)$  is unknown. Therefore, traditional recognition/classification approaches minimize the empirical risk instead of the ideal risk  $R_{\mathcal{I}}$  by using other knowledge, such as assuming that the label space is at least locally smooth and regularizing the empirical risk minimization [50]–[52].

Note that traditional recognition problem is usually performed under the closed set assumption. When the assumption

TABLE I  
DIFFERENCES BETWEEN OPEN SET RECOGNITION AND ITS RELATED TASKS

Task \ Setting	Training	Testing	Goal
Traditional Classification	Known known classes	Known known classes	Discrimination between classes
Classification with Reject Option	Known known classes	Known known classes	Discrimination between classes, rejecting samples of low confidence
One-class Classification (Anomaly Detection)	Known known classes & few or none outliers	Known known classes & few or none outliers	Detecting outliers
Zero-shot Learning	Known known classes & side-information <sup>1</sup>	Unknown known classes	Identifying unknown known classes
One-shot Learning (Few-shot Learning)	Known known classes & side-information <sup>2</sup>	Unknown known classes	Identifying unknown known classes
Open Set Recognition	Known known classes	Known known classes & Unknown unknown classes	Identifying known known classes, rejecting unknown unknown classes

The side-information<sup>1</sup> in ZSL denotes the semantic/attribute information shared among all classes including known known and unknown known classes, while the side-information<sup>2</sup> in one-shot learning represents one or few available samples of unknown known classes or adding another semantic/attribute information similar to ZSL.

switches to open environment/set scenario with the open space, other things should be added since intuitively there is some risk in labeling sample in the open space as any known known classes. This gives such an insight for OSR that we do know something else: we do know where known known classes exist, and we know that in open space we do not have a good basis for assigning labels for the unknown unknown classes [17].

As the joint distribution  $P(x, y)$  in (1) is usually unknown, Scheirer et al. [17] considered the open space risk as a weak assumption: all data points in open space, all labels (including known known and unknown unknown classes) are equally likely, and risk should be computed accordingly, based on the Principle of Indifference [53]. Therefore, after assuming the probability is proportional to relative Lebesgue measure [54], they considered the measure of the open space to the full space, and gave the definition of *Open Space Risk*  $R_O(f)$ :

$$R_O(f) = \frac{\int_O f(x)dx}{\int_{S_o} f(x)dx}, \quad (2)$$

where  $S_o$  represents the full space including open space and known space (the space occupied by known known classes), and  $f$  denotes the measurable recognition function, where  $f(x) = 1$  indicates that class  $y$  of interest (known known classes) is recognized, otherwise  $f(x) = 0$ . The more we label samples in open space as the known known classes, the greater the open space risk is. Note that, Eq. (2) is just one theoretical possibility of the open space risk, which does not have a loss function, class conditional densities, or class priors. Other definitions [17] can also capture the notion of open space risk, and some may be more precise, which, however, would need more assumptions and complexity for the reason that the unknown unknown classes have unknown priors and unknown joint distributions.

Additionally, the authors in [17] also formally introduced the concept of *openness* for a particular problem or data universe.

**Definition 1.** (The openness defined in [17]) Let  $C_{TA}$ ,  $C_{TR}$ , and  $C_{TE}$  respectively represent the set of classes to be recognized, the set of classes used in training and the set of classes

used during testing. Then the openness of the corresponding recognition task  $O$  is:

$$O = 1 - \sqrt{\frac{2 \times |C_{TR}|}{|C_{TA}| + |C_{TE}|}}, \quad (3)$$

where  $|\cdot|$  denotes the number of classes in the corresponding set.

Note that the authors in [17] do not explicitly give the relationships among  $C_{TA}$ ,  $C_{TR}$ , and  $C_{TE}$ . Actually, in most existing works [18], [55]–[57], the relationship,  $C_{TA} = C_{TR} \subseteq C_{TE}$ , holds by default. Besides, the authors in [58] specifically give the following relationship:  $C_{TA} \subseteq C_{TR} \subseteq C_{TE}$ , which contains the former case. However, such a relationship is problematic for Definition 1. Consider the following simple case:  $C_{TA} \subseteq C_{TR} \subseteq C_{TE}$ , and  $|C_{TA}| = 3, |C_{TR}| = 10, |C_{TE}| = 15$ . Then we will have  $O < 0$ , which is obviously unreasonable. In fact, the  $C_{TA}$  should be a subset of  $C_{TR}$ , otherwise it would make no sense because one usually does not use the classifiers trained on  $C_{TR}$  to identify other classes which are not in  $C_{TR}$ . Intuitively, the *openness* of a particular problem should only depend on the available known known class knowledge from  $C_{TR}$  and the unknown unknown class knowledge from  $C_{TE}$  rather than  $C_{TA}$ ,  $C_{TR}$ , and  $C_{TE}$  their three. Therefore, in this paper, we redefine the concept of *openness*:

**Definition 2.** (The openness redefined in this paper) Let  $C_{TA}$ ,  $C_{TR}$ ,  $C_{TE}$  respectively denote as defined in Definition 1, and let  $C_{TA} \subseteq C_{TR} \subseteq C_{TE}$ . Then the openness of the corresponding recognition task  $O^*$  is:

$$O^* = 1 - \sqrt{\frac{2 \times |C_{TR}|}{|C_{TR}| + |C_{TE}|}}. \quad (4)$$

Note that compared to Definition 1, Definition 2 is just a relatively more reasonable form to estimate the *openness*. Actually, other definitions can also capture this notion, and some may be more precise, thus worth further exploring. Furthermore, as described in [17], larger *openness* corresponds to more open problems, while the problem is completely closed when the *openness* equals 0. Additionally, the number

of testing classes can grow rapidly with *openness* approaching 100% for almost any unconstrained real world problem. With the concepts of open space risk and *openness* in mind, the definition of OSR problem can be given as follows:

**Definition 3.** (The Open Set Recognition Problem [17]) *Let  $V$  be the training data, and let  $R_O$ ,  $R_\epsilon$  respectively denote the open space risk and the empirical risk. Then the goal of open set recognition is to find a measurable recognition function  $f \in \mathcal{H}$ , where  $f(x) > 0$  implies correct recognition, and  $f$  is defined by minimizing the following **Open Set Risk**:*

$$\arg \min_{f \in \mathcal{H}} \{R_O(f) + \lambda_r R_\epsilon(f(V))\} \quad (5)$$

where  $\lambda_r$  is a regularization constant.

The open set risk denoted in (5) balances the empirical risk and the open space risk over the space of allowable recognition functions. Although this initial definition mentioned above is more theoretical, it provides an important guidance for subsequent OSR modeling, leading to a series of OSR algorithms which will be detailed in the following section.

### III. A CATEGORIZATION OF OSR TECHNIQUES

Although Scheirer et al. [17] formalized the OSR problem, an important question is how to incorporate Eq. (2) to modeling. There is an ongoing debate between the use of generative and discriminative models in statistical learning [59], [60], with arguments for the value of each. However, as reported in [18], open set recognition introduces such a new issue, in which neither discriminative nor generative models can directly address the unknown unknown classes existing in open space unless some constraints are imposed. Thus, with some constraints, researchers have made the exploration in modeling of OSR respectively from the discriminative and generative perspective. Next, we mainly review the existing OSR models from these two perspectives.

According to the modeling forms, these models can be further categorized into five categories (Table II): Traditional ML-based, Deep Network-based, Adversarial Learning-based, EVT-based, and Dirichlet Process-based OSR models. For each category, we then review different approaches by focusing on their corresponding representative works. Additionally, several available software packages implementing those models are listed in Table V (Appendix A). Next, we first give a review from the discriminative model perspective, where almost all existing OSR algorithms are modeled from this perspective.

#### A. Traditional ML Methods-based OSR Models

As mentioned above, traditional machine learning methods (e.g., SVM, sparse representation, Nearest Neighbor, etc.) usually assume that the training and testing data are drawn from the same distribution. However, such an assumption usually does not hold any more in the OSR problem. To adapt these methods to the OSR scenario, many efforts have been made [17]–[19], [55], [61]–[70].

TABLE II  
DIFFERENT KINDS OF MODELS FOR OPEN SET RECOGNITION

Different types of OSR methods		papers
Discriminative model	Traditional ML-based	[17]–[19], [55], [61]–[70]
	Deep Network-based	[58], [71]–[77]
	Adversarial Learning-based	[56], [78]–[81]
	EVT-based	[82], [83]
Generative model	Dirichlet Process-based	[57]

1) *SVM-based*: The Support Vector Machine (SVM) [84] has been successfully used in traditional classification/recognition task. However, when unknown unknown classes appear during testing, its classification performance will decrease significantly since it divides over-occupied space for known known classes under closed set assumption. As shown in Fig. 1(b), once the unknown unknown classes' samples fall into the space divided for some known known classes, these samples will never be correctly classified. To overcome this problem, many SVM-based OSR variants have been proposed.

Using the Definition 3, Scheirer et al. [17] proposed the 1-vs-Set machine, which incorporates an open space risk term in modeling to account for the space beyond the reasonable support of known known classes. Concretely, they added another hyperplane parallelling the separating hyperplane obtained by the SVM, in score space, leading to a slab in feature space. Furthermore, the open space risk for a linear kernel slab model is defined as follows:

$$R_O = \frac{\delta_\Omega - \delta_A}{\delta^+} + \frac{\delta^+}{\delta_\Omega - \delta_A} + p_A \omega_A + p_\Omega \omega_\Omega, \quad (6)$$

where  $\delta_A$  and  $\delta_\Omega$  denote the marginal distances of the corresponding hyperplanes, and  $\delta^+$  is the separation needed to account for all positive data. Additionally, user-specified parameters  $p_A$  and  $p_\Omega$  are given to weight the importance between the margin spaces  $\omega_A$  and  $\omega_\Omega$ . In this case, a testing sample that appears between the two hyperplanes would be labeled as the appropriate class, while it is considered as non-target class or rejected, depending on which side of the slab it resides. Similar to the 1-vs-Set machine, Cevikalp [61], [62] added another constraint on the samples of positive/target class based on the traditional SVM, and proposed the Best Fitting Hyperplane Classifier (BFHC) model which also formed a slab in feature space. Additionally, the BFHC can be extended to nonlinear case by using kernel trick, and we refer reader to [62] for more details.

Although the slab models mentioned above decrease the region of the corresponding known known class for each binary SVM, the space occupied by each known known class remains unbounded, thus the open space risk still exists. To overcome this challenge, researchers further seek new ways to control the open space risk [18], [19], [63]–[65].

Scheirer et al. [18] incorporated non-linear kernels into a solution that further limited open space risk by positively labeling only sets with finite measure, and they formulated a

compact abating probability (CAP) model, where probability of class membership abates as points move from known data to open space. Specifically, a novel technique called Weibull-calibrated SVM (W-SVM) was proposed, which combined the statistical extreme value theory (EVT) for score calibration with two separated SVMs. The first SVM is a one-class SVM CAP model used as a conditioner: if the one-class SVM predicts the posterior estimate  $P_O(y|x)$  of an input sample  $x$  is less than a threshold  $\delta_\tau$ , the sample will be rejected outright. Otherwise, it will be passed to the second SVM, which is a binary SVM CAP model via a fitted Weibull cumulative distribution function, yielding the posterior estimate  $P_\eta(y|x)$  for the corresponding positive known known class as well as a reverse Weibull fitting, obtaining the posterior estimate  $P_\psi(y|x)$  for the corresponding negative known known classes. Defined an indicator variable:  $\iota_y = 1$  if  $P_O(y|x) > \delta_\tau$  and  $\iota_y = 0$  otherwise, then the W-SVM model for OSR is defined as follows

$$\begin{aligned} y^* = & \arg \max_{y \in \mathcal{Y}} P_\eta(y|x) \times P_\tau(y|x) \times \iota_y \\ & \text{subject to } P_\eta(y^*|x) \times P_\tau(y^*|x) \geq \delta_R, \end{aligned} \quad (7)$$

where  $\mathcal{Y}$  denotes all the known known classes,  $\delta_R$  is the threshold of the second SVM CAP model. Additionally, the thresholds  $\delta_\tau$  and  $\delta_R$  are set empirically, e.g.,  $\delta_\tau$  is fixed to 0.001 as specified by the authors, while  $\delta_R$  is recommended to set according to the openness of the specific problem

$$\delta_R = 0.5 \times \text{openness}. \quad (8)$$

Besides, W-SVM was further used for open set intrusion recognition on the KDDCUP'99 dataset [85], while more works on intrusion detection in open set scenario can be found in [86]. Based on the intuition that one can reject the large set of unknown unknown classes even under an assumption of incomplete class knowledge if the positive data for any known known classes is accurately modeled without overfitting, Jain et al. [19] invoked the EVT to model the positive training samples at the decision boundary and proposed the  $P_I$ -SVM algorithm. Furthermore, the  $P_I$ -SVM also adopts the threshold-based classification scheme, in which the selection of corresponding threshold takes the same strategy in W-SVM.

Note that while both W-SVM and  $P_I$ -SVM effectively limit the open space risk by the threshold-based classification schemes, the selection of thresholds also gives some caveats. Specifically, first, they are assumed that all the known known classes have equal thresholds, which may be not reasonable since the distributions of classes in feature space are usually unknown. Second, the reject thresholds are recommended to set according to the problem openness [18]. However, the openness of the corresponding problem, is usually unknown as well.

To address these caveats, Scherrek et al. [63] introduced the probabilistic open set SVM (POS-SVM) classifier which could empirically determine unique reject threshold for each known known class under the Definition 3. Note that slightly different from the formula (5), POS-SVM chooses probabilistic representations respectively for open space risk  $R_O$  and empirical risk  $R_\epsilon$  (details c.f. [63]). Moreover, the authors also adopted

a new OSR evaluation metric called Youden's index which combines the true negative rate and recall, and will be detailed in Section V. Recently, to adapt to sliding window visual object detection and open set recognition tasks, Cevikalp and Triggs [64], [65] used a family of quasi-linear "polyhedral conic" functions of [87] to define the acceptance regions for positive known known classes. This choice provides a convenient family of compact and convex region shapes for discriminating relatively well localized positive known known classes from broader negative ones including negative known known classes and unknown unknown classes.

2) *Sparse Representation-based*: In recent years, the sparse representation-based techniques have been widely used in computer vision and image processing fields [88]–[90]. In particular, sparse representation-based classifier (SRC) [91] has gained a lot of attractions, which identifies the correct class by seeking the sparsest representation of the testing sample in terms of the training. The SRC method and its variants are essentially still under a closed set assumption, so in order to adapt the SRC to an open environment, Zhang and Patel [55] presented the sparse representation-based open set recognition model, briefly called SROSr.

The SROSr models the tails of the matched and sum of non-matched reconstruction error distributions using EVT due to the fact that most of the discriminative information for OSR is hidden in the tail part of those two error distributions. Furthermore, this model consists of two main stages: the first stage is to reduce the OSR problem into hypothesis testing problems by modeling the tails of error distributions using EVT, while the second stage calculates the reconstruction errors for a testing sample, then fusing the confidence scores based on the two tail distributions to determine its identity.

As reported in [55], although the SROSr outperformed many competitive OSR algorithms, it also contains some limitations. For example, in the face recognition task, the SROSr would fail in such cases that the dataset contained extreme variations in pose, illumination or resolution, where the self expressiveness property required by the SRC do no longer hold. Besides, for good recognition performance, the training set is required to be extensive enough to span the conditions that might occur in testing set. Note that while only the SROSr is currently proposed based on sparse representation, it can still be an interesting topic for future work to develop the sparse representation-based OSR algorithms.

3) *Distance-based*: Similar to other traditional ML methods mentioned above, the distance-based classifiers are usually no longer valid under the open set scenario. To meet this challenge, Bendale and Boulton [66] established a Nearest Non-Outlier (NNO) algorithm for open set recognition by extending upon the Nearest Class Mean (NCM) classifier [92], [93]. The NNO carries out classification based on the distance between the testing sample and the means of known known classes. Furthermore, it detects outliers for bounding the open space risk, and rejects an input sample when all classifiers reject it. What needs to emphasize is that this algorithm can dynamically add new classes based on manually labeled data. Additionally, the authors also introduced the concept of open world recognition which will be detailed in Section IV.

Besides, based on the traditional Nearest Neighbor classifier, Júnior et al. [67] introduced an open set version of the Nearest Neighbor classifier (OSNN) to deal with OSR problem. Different from those works which directly use a threshold on the similarity score for the most similar class, the OSNN applies a threshold on the ratio of similarity scores to the two most similar classes instead, which is called Nearest Neighbor Distance Ratio (NNDR) technique. Specifically, it first finds the nearest neighbor  $t$  and  $u$  of the testing sample  $s$ , where  $t$  and  $u$  come from different classes, then calculates the ratio

$$\text{Ratio} = d(s, t)/d(s, u), \quad (9)$$

where  $d(x, x')$  denotes the Euclidean distance between sample  $x$  and  $x'$  in feature space. If the *Ratio* is less than or equal to the pre-set threshold,  $s$  will be classified as the same label of  $t$ . Otherwise, it is considered as the unknown unknown class.

Note that OSNN is inherently multiclass, meaning that its efficiency will not be affected as the number of available classes for training increases, while it can create a bounded open space, thus gracefully protecting the classes of interest and rejecting unknown unknown classes. Moreover, the NNDR technique can be applied effortlessly to other classifiers based on the similarity score, e.g., the Optimum-Path Forest (OPF) classifier [94]. In addition, other metrics could also be used, and even the feature space considered could be a transformed one, as suggested by the authors. Complementally, one limitation of OSNN is that just selecting two reference samples coming from different classes for comparison makes the OSNN vulnerable to outliers [57].

4) *Other Traditional ML Methods-based:* Using center-based similarity (CBS) space learning, Fei and Liu [68] proposed a novel solution for text classification under OSR scenario, while Vareto et al. [69] explored the open set face recognition and proposed HPLS and HFCN algorithms by combining hashing functions, partial least squares (PLS) and fully connected networks (FCN). Neira et al. [70] adopted the integrated idea, where different classifiers and features are combined to solve the OSR problem. We refer the reader to [68]–[70] for more details. As most traditional machine learning methods for classification are under closed set assumption, it is appealing to adapt them to the open and non-stationary environment.

## B. Deep Neural Network-based OSR Models

Deep Neural Networks (DNNs) have gained significant benefits for various tasks such as visual recognition, Natural language processing, text classification, etc. However, they usually follow a typical SoftMax cross-entropy classification loss, which inevitably incurs the normalization problem, making the DNNs inherently have the closed set nature. As a consequence, DNNs often make wrong predictions, and even do so too confidently, when processing the samples of unknown unknown classes. Furthermore, the works in [95], [96] have indicated that the DNNs easily suffer from vulnerability to 'fooling' and 'rubbish' images which are visually far from the desired class but produce high confidence scores. To address these problems, researchers have made many efforts [58], [71]–[77], [97].

Replacing the SoftMax layer in the DNNs with an OpenMax layer, Bendale and Boulton [71] proposed the OpenMax model as a first solution towards open set Deep Networks. Specifically, a deep neural network is first trained with the normal SoftMax layer by minimizing the cross entropy loss. Adopting the concept of Nearest Class Mean [92], [93], each class is then represented as a mean activation vector (MAV) with the mean of the activation vectors (only for the correctly classified training samples) in the penultimate layer of that network. Next, the training samples' distances from their corresponding class MAVs are calculated and used to fit the separate Weibull distribution for each class. Further, the activation vector's values are redistributed according to the Weibull distribution, and then used to compute a pseudo-activation for unknown unknown classes. Finally, the class probabilities of known known and (pseudo) unknown unknown classes are computed by using SoftMax again on these new redistributed activation vectors.

Notably, the OpenMax effectively addressed the challenge of the recognition for fooling/rubbish and unrelated open set images. However, as discussed in [71], the OpenMax fails to recognize the adversarial images which are visually indistinguishable from training samples but are designed to make deep networks produce high confidence but incorrect answers [96], [98]. Actually, the authors in [72] have indicated that the OpenMax is susceptible to the adversarial generation techniques directly working on deep representations. Therefore, the adversarial samples are still a serious challenge for open set recognition. Furthermore, using the distance from MAV, the cross entropy loss function in OpenMax does not directly incentivize projecting class samples around the MAV. In addition to that, the distance function used in testing is not used in training, possibly resulting in inaccurate measurement in that space [73]. To address this limitation, Hassen and Chan [73] learned a neural network based representation for open set recognition, which is similar in spirit to the Fisher Discriminant, where samples from the same class are closed to each other while the ones from different classes are further apart, leading to larger space among known known classes for unknown unknown classes' samples to occupy.

Besides, Prakhyia et al. [74] followed the technical line of OpenMax to explore the open set text classification, while Shu et al. [75] replaced the SoftMax layer with a 1-vs-rest final layer of sigmoids and presented Deep Open classifier (DOC) model, in which the DOC tightens the decision boundaries of sigmoid functions with Gaussian fitting to further reduce the open space risk. Additionally, based on an elaborate distance-like computation provided by a weightless neural network, Cardoso et al. [76] proposed the tWiSARD algorithm for open set recognition, which is further developed in [58]. Recently, Shu et al. [77] paid more attention to discovering the hidden unknown unknown classes among the reject samples. Correspondingly, they proposed a joint open classification model with a sub-model for classifying whether a pair of examples belongs to the same class or not, while the sub-model can serve as a distance function for clustering to discover the hidden classes among the reject samples.

### C. Adversarial Learning-based OSR Models

Currently, the adversarial learning (AL) [99] as a novel technology has gained the striking successes, which employs a generative model and a discriminative model, where the generative model learns to generate samples that can fool the discriminative model as non-generated samples. Due to the properties of AL, some researchers also attempt to account for open space with the unknown unknown classes generated by the AL technique [78]–[81].

Using a conditional generative adversarial network (GAN) to synthesize mixtures of known known classes, Ge et al. [78] proposed the Generative OpenMax (G-OpenMax) algorithm, which can provide explicit probability estimation over the generated unknown unknown classes, enabling the classifier to locate the decision margin according to the knowledge of both known known and generated unknown unknown classes. Obviously, such unknown unknown classes in their setting are only limited in a subspace of the original known known class space. Moreover, although the G-OpenMax effectively detects unknown unknown classes in monochrome digit datasets, it actually has no significant performance improvement on natural images [78]. Different from G-OpenMax, Neal et al. [79] introduced a novel dataset augmentation technique, called counterfactual image generation, which adopts an encoder-decoder GAN architecture to generate the synthetic open set examples which are close to known known classes, yet do not belong to any known known classes. They further reformulated the OSR problem as classification with one additional class containing those newly generated samples. Similar in spirit to [79], Yu et al. [80] proposed the adversarial sample generation (ASG) framework for the OSR problem with the overall idea is that the OSR problem will be easily solved by standard supervised learning, if one can generate the samples of new class and put them into the training set. Furthermore, ASG framework can be applied to various learning models besides neural networks, while it can not only generate unknown unknown class data but also generate known known class data if necessary. Besides, Yang et al. [56] borrowed the generator in a typical GAN networks to produce synthetic samples that are highly similar to the target samples as the automatic negative set, while the discriminator is redesigned to output multiple classes together with an unknown unknown class. Then they explored the open-set human activity recognition based on micro-Doppler signatures. Recently, Zheng et al. [81] began to focus on the adversarial samples targeting at open set recognition systems.

Note that the main challenge for open set recognition is the incomplete class knowledge existing in training, leading to the open space risk when classifiers encounter unknown unknown classes during testing. Fortunately, the adversarial learning technique can account for open space to some extent by adversarially generating the unknown unknown class data according to the known known class knowledge, which undoubtedly provides another way to tackle the challenging multiclass OSR problem.

### D. EVT-based OSR Models

As a powerful tool to increase the classification performance, the statistical Extreme Value Theory (EVT) has recently achieved great success due to the fact that EVT can effectively model the tails of the distribution of distances between training observations using the asymptotic theory [100]. For example, the well-known RANSAC algorithm is improved by using the EVT [101]. Actually, many methods have been proposed to tackle the OSR problem by using EVT [18], [19], [55]. However, these are post hoc approaches which do not directly apply EVT during training. To change this situation, researchers have made the following exploration [82], [83].

Taking the distributional information of training observations into account at the training time, Rudd et al. [82] proposed the Extreme Value Machine (EVM) which can approximate the distribution of the margin distance of each sample point in each class using EVT. According to the density function of this distribution obtained, the probability of a new point  $x'$  associated with class  $C_l$ , i.e.,  $\hat{P}(C_l|x')$ , can then be obtained, thus resulting in the final decision function

$$y^* = \begin{cases} \arg \max_{l \in \{1, \dots, M\}} \hat{P}(C_l|x') & \text{if } \hat{P}(C_l|x') \geq \delta \\ \text{"unknown"} & \text{Otherwise,} \end{cases} \quad (10)$$

where the  $M$  denotes the number of the classes in training, and  $\delta$  represents the probability threshold which defines the boundary between the set of known known classes and unsupported open space.

Derived from EVT, the EVM has a well-grounded interpretation and can perform nonlinear kernel-free variable bandwidth incremental learning, which is further utilized to explore the open set face recognition [102] and the intrusion detection [103]. Besides, it also has some limitations as reported in [83], in which an obvious one is that the use of geometry of known known classes is risky when the geometries of known known and unknown unknown classes differ. To address these limitations, Vignotto and Engelke [83] presented the GPD and GEV classifiers relying on approximations from the EVT, which are further developed the EVT-based OSR technique. Note that, while EVT helps us effectively model the extremas, regrettably, it provides no principled means of selecting the number of samples for EVT fitting [82].

**Remark:** As mentioned above, almost all existing OSR methods adopt the threshold-based classification scheme, where recognizers in decision either reject or categorize the input samples to some known known class using empirically-set threshold. Thus the threshold plays a key role. However, the selection for it usually depends on the knowledge of known known classes, inevitably incurring risks due to lacking available information from unknown unknown classes [57]. This indicates the threshold-based OSR methods still face serious challenges.

In the following part, we will review the OSR techniques from the generative model perspective, which can be seen as another line to research the open set recognition, where only few works currently focus on this perspective.



### E. Dirichlet Process-based OSR Models

Dirichlet process (DP) [104]–[108] considered as a distribution over distributions is a stochastic process, which has been widely applied in clustering and density estimation problems as a nonparametric prior defined over the number of mixture components. Furthermore, this model does not overly depend on training samples and can achieve adaptive change as the data changes, making it naturally adapt to the open set recognition scenario. In fact, researchers have begun the related research.

With the aim to extend existing OSR methods for new class discovery while considering correlations among the testing samples, Geng and Chen [57] introduced a collective/batch decision for open set recognition, which can address both batch and individual samples. As an initial solution, they adapted the hierarchical Dirichlet process (HDP) with slight modification to the OSR scenario, and proposed the collective decision-based OSR (CD-OSR) framework. The CD-OSR first performs a co-clustering process to obtain the appropriate parameters in the training phase. In testing phase, it models each known known class data as a group of CD-OSR using a Gaussian mixture model (GMM) with an unknown number of components/subclass, while the whole testing set as one collective/batch is treated in the same way, then co-clustering all the groups under the CD-OSR framework. Thus one can obtain one or more subclasses representing the corresponding class after co-clustering. For a testing sample, it would be labeled as the appropriate known known class or unknown unknown classes, depending on whether the subclass that it is located associates with the corresponding known known class or not.

Unlike the threshold-based OSR methods, CD-OSR does not need to define the threshold and can provide explicit modeling for the unknown unknown classes appearing in testing, naturally resulting in a new class discovery function. Note that the new class discovered in CD-OSR inherently has only one subclass due to the fact that the true labels of unknown unknown classes are unknown, making it impossible to further aggregate the newly generated subclasses. Furthermore, adopting the collective/batch decision strategy also considers correlations among the testing samples obviously ignored by other existing methods. As reported by the authors, the CD-OSR is just as a conceptual proof for open set recognition towards collective decision at present, and there are still many limitations. For example, the recognition process of CD-OSR seems to have the flavor of lazy learning to some extent, where the co-clustering process will be repeated when other batch testing data arrives, resulting in higher computational overhead. Besides, the CD-OSR currently does not make full use of the information from the known known class labels. In other words, it just uses this kind of information to divide the training data into different groups.

**Remark:** Instead of addressing the OSR problem from the discriminative model perspective, CD-OSR actually reconsiders this problem from the generative model perspective due to the use of HDP, which provides another research direction for open set recognition. Furthermore, the collective decision

strategy for OSR is also worth further exploring since it not only takes the correlations among the testing samples into account but also provides a possibility for new class discovery, whereas single-sample decision strategy<sup>2</sup> adopted by other existing OSR methods can not do such a work since it can not directly tell whether the single rejected sample is an outlier or from new class.

## IV. BEYOND OPEN SET RECOGNITION

It is worth noting that the existing open set recognition was defined for a static notion of set, whilst only a reject option is adopted for the unknown unknown classes. However, that is not enough for a real-world OSR system which should neither be limited to the static set nor just rest on a reject decision but should go further, especially when the datasets are dynamic with new classes being continuously detected. Following this idea, Bendale and Boulton [66] expanded the existing open set recognition (Definition 3) to the open world recognition (OWR), where a recognition system should perform four tasks: detecting unknown unknown classes, choosing which samples to label for addition to the model, labelling those samples, and updating the classifier. Specifically, the authors give the following definition:

**Definition 4.** (Open World Recognition [66]) *Let  $\mathcal{K}_T \in \mathbb{N}^+$  be the set of labels of known known classes at time  $T$ , and let the zero label (0) be reserved for (temporarily) labeling data as unknown. Thus  $\mathbb{N}$  includes known known and unknown unknown class labels. Based on the Definition 3, a solution to open world recognition is a tuple  $[F, \varphi, \nu, \mathcal{L}, I]$  with:*

- 1) *A multi-class open set recognition function  $F(x) : \mathbb{R}^d \mapsto \mathbb{N}$  using a vector function  $\varphi(x)$  of  $i$  per-class measurable recognition functions  $f_i(x)$ , also using a novelty detector  $\nu(\varphi) : \mathbb{R}^i \mapsto [0, 1]$ . We require the per-class recognition functions  $f_i(x) \in \mathcal{H} : \mathbb{R}^d \mapsto \mathbb{R}$  for  $i \in \mathcal{K}_T$  to be open set functions that manage open space risk as Eq. (2). The novelty detector  $\nu(\varphi) : \mathbb{R}^i \mapsto [0, 1]$  determines if results from vector of recognition functions is from an unknown unknown class.*
- 2) *A labeling process  $\mathcal{L}(x) : \mathbb{R}^d \mapsto \mathbb{N}^+$  applied to novel unknown data  $U_T$  from time  $T$ , yielding labeled data  $D_T = \{(y_j, x_j)\}$  where  $y_j = \mathcal{L}(x_j), \forall x_j \in U_T$ . Assume the labeling finds  $m$  new classes, then the set of known known classes becomes  $\mathcal{K}_{T+1} = \mathcal{K}_T \cup \{i+1, \dots, i+m\}$ .*
- 3) *An incremental learning function  $I_T(\varphi; D_T) : \mathcal{H}^i \mapsto \mathcal{H}^{i+m}$  to scalably learn and add new measurable functions  $f_{i+1}(x) \dots f_{i+m}(x)$ , each of which manages open space risk, to the vector  $\varphi$  of measurable recognition functions.*

Besides, we refer the reader to [66] for more details. Ideally, all of these steps should be automated. However, the authors in [66] only presumed supervised learning with labels obtained by human labeling at present, and proposed the NNO algorithm which has been discussed in subsection III-A.

<sup>2</sup>The single-sample decision means that the classifier makes a decision, sample by sample. In fact, almost all existing OSR methods are designed specially for recognizing individual samples, even these samples are collectively coming in batch.



Afterward, some researchers continued to follow up this research route. Rosa et al. [109] argued that to properly capture the intrinsic dynamic of OWR, it is necessary to append the following aspects: (a) the incremental learning of the underlying metric, (b) the incremental estimate of confidence thresholds for the unknown unknown classes, and (c) the use of local learning to precisely describe the space of classes. Towards those goals, they extended three existing metric learning methods using online metric learning. Furthermore, Doan and Kalita [110] presented the Nearest Centroid Class (NCC) model, which is similar to the online NNO [109] but differs with two main aspects. First, they adopted a specific solution to address the initial issue of incrementally adding new classes, Second, they optimized the nearest neighbor search for determining the nearest local balls. Moreover, Lonij et al. [111] tackled the OWR problem from the complementary direction of assigning semantic meaning to open-world images. To handle the open-set action recognition task, Shu et al. [112] proposed the Open Deep Network (ODN) which detects new classes by applying a multiclass triplet thresholding method, and then dynamically reconstructs the classification layer by adding predictors for new classes continually. In fact, the EVM discussed in subsection III-D also adapt to the OWR scenario due to the nature of incremental learning [82]. Recently, Xu et al. [97] proposed a meta-learning method to learn to accept new classes (unknown unknown classes) without training under the open world recognition framework.

**Remark:** As a natural extension of OSR, the OWR faces more serious challenges which require it not only to have the ability to handle the OSR task, but also to have minimal downtime, even to continuously learn, which seems to have the flavor of lifelong learning to some extent. Besides, although some progress regarding the OWR has been made, there is still a long way to go.

## V. DATASETS AND EVALUATION METRICS

### A. Datasets and Experiment Setup

1) *Datasets:* Table III summarizes the datasets commonly used in open set recognition, and these datasets are available on (<https://dx.doi.org/10.6084/m9.figshare.1097614>) and (<https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/multi-class.html>). As the OSR is essentially a multi-class classification problem, therefore, any benchmark datasets for multi-class classification task can be used for the OSR experiments. For example, the image datasets commonly used in computer vision such as ILSVRC 2012 dataset [71], YaleB dataset, UIUC attribute dataset [55], and so forth.

2) *Experiment Setup:* In open set recognition, most existing experiments are carried out on a variety of recastes multi-class benchmark datasets. Specifically, taking the Usps dataset as an example, when it is used for OSR problem, one can randomly choose  $S$  distinct labels as the known known classes, and vary openness by adding a subset of the remaining labels.

### B. Evaluation Metrics for Open Set Recognition

Here, we summarize the commonly used evaluation metrics for open set recognition, where the critical factor is that the

TABLE III  
DATASETS FOR OPEN SET RECOGNITION

Datasets	#Classes	#Features	#Examples
Usps	10	256	7291
Mnist	10	780	70000
Pendigits	10	16	10992
15-Scenes	15	1000	4485
Coil20	20	256	1440
Letter	26	16	20000
Auslan	95	22	146949
Caltech-256	256	1000	29780
ALOI	1000	128	108000
Ukbench	2550	128	10200

evaluation metrics need to take the identification of unknown unknown classes into account. Let  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  respectively denote the true positive, true negative, false positive, and false negative for known known classes, while  $TU$  and  $FU$  respectively denote the correct and false reject for unknown unknown classes. Then we can obtain the following evaluation metrics.

1) *Accuracy for OSR:* As a common choice for evaluating the performance of decision classifiers under closed set assumption, Accuracy is usually defined as follows

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}.$$

A trivial extension of Accuracy to the OSR scenario could be that a correct response should be either the correct classification (correctly classifying the positives or negatives) or correct reject if the testing sample comes from an unknown unknown class. Therefore, the Accuracy for OSR ( $\text{Accuracy}_O$ ) can be redefined as

$$\text{Accuracy}_O = \frac{(TP + TN) + TU}{(TP + TN + FP + FN) + (TU + FU)}. \quad (11)$$

However, as  $\text{Accuracy}_O$  denotes the sum of the performance of the correct classification for known known classes and the rejection for unknown unknown classes, it actually can not objectively evaluate the OSR models. Consider the following scenario: when the reject performance plays the leading role, and the testing set contains large number of samples of unknown unknown classes while only a few samples for the known known classes, the  $\text{Accuracy}_O$  then can still achieve a high value, even though the fact is that the recognizer's classification performance is really low, and vice versa. Therefore, the  $\text{Accuracy}_O$  usually used as a supplement for the evaluation of OSR methods. Besides, the authors in [67] also gave a new accuracy metric for OSR called *normalized accuracy* (NA), which considers both the accuracy on known known class samples (AKS) and the accuracy on unknown unknown class samples (AUS):

$$\text{NA} = \lambda_r \text{AKS} + (1 - \lambda_r) \text{AUS}, \quad (12)$$

where

$$\text{AKS} = \frac{TP + TN}{TP + TN + FP + FN}, \quad \text{AUS} = \frac{TU}{TU + FU},$$

and  $\lambda_r$ ,  $0 < \lambda_r < 1$ , is a regularization constant.

2) *F-measure for OSR*: The F-measure, widely applied in information retrieval and machine learning, is defined as a harmonic mean of Precision and Recall

$$\text{F-measure} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (13)$$

where

$$\text{Precision} = \frac{TP}{TP + FN}, \quad \text{Recall} = \frac{TP}{TP + FP}.$$

When the F-measure is used for the OSR scenario, it should be noted that one could not consider all the unknown unknown classes appearing in testing as one additional simple class, and obtain the F-measure value in the same way as the multiclass closed set scenario. Because once performing such an operation, the correct classification for all the unknown unknown classes samples would be considered as true positive classifications. However, such true positive classification make no sense, since we have no representative samples of unknown unknown classes to train the corresponding classifiers.

Instead, the computations of Precision and Recall in it are only for available known known classes. Additionally, the work [67] has indicated that although the computations of Precision and Recall are only for available known known classes, the  $FN$  and  $FP$  also consider the false unknown unknown classes and false known known classes by taking into account the false negative and the false positive, and we refer the reader to [67] for more details. Note that the Precision contains the macro-Precision and micro-Precision while Recall includes the macro-Recall and micro-Recall, which leads to the corresponding macro-F-measure and micro-F-measure. Nevertheless, whether it is macro-F-measure or micro-F-measure, the higher their values, the better the performance of the corresponding OSR model.

3) *Youden's index for OSR*: As the F-measure is invariant to changes in  $TN$  [113], an important factor in OSR performance, Scherrek and Rigling [63] turned to Youden's index  $J$  defined as follows

$$J = \text{Recall} + S - 1, \quad (14)$$

where  $S = TN/(TN + FP)$  represents the true negative rate [114]. Youden's index can express an algorithm's ability to avoid failure [115], and it is bounded in  $[-1, 1]$ , where higher value indicates an algorithm more resistant to failure. Furthermore, the classifier is noninformative when  $J = 0$ , whereas it tends to provide more incorrect than correct information when  $J < 0$ .

Besides, some researchers also adopted the AUC to evaluate the performance of OSR models [83]. Currently, F-measure is the most commonly used evaluation metric. In fact, as the OSR problem faces a new scenario, the new evaluation methods are also worth further exploring.

## VI. FUTURE RESEARCH DIRECTIONS

In this section, we briefly analyze and discuss the limitations of the existing OSR models, while some promising research directions in this field are also pointed out and detailed in the following aspects.

### A. About Modeling

First, as discussed above, while almost all existing OSR methods are modeled from the discriminative model perspective, only few works focus on the modeling from the generative perspective. Therefore, modeling from the later perspective has a wide space for further exploration. Second, the main challenge for OSR is that the traditional classifiers under closed set scenario dive over-occupied space for known known classes, thus once the unknown unknown classes' samples fall into the space divided for some known known classes, they will never be correctly classified. From this viewpoint, the following two modeling perspective will be promising research directions.

1) *Modeling known known classes*: Intuitively, we can effectively identify the target classes even under an assumption of incomplete class knowledge, if the target classes can be accurately modeled without overfitting. Therefore, the clustering and classification learning can be unified to achieve the best of both worlds: the clustering learning can help the target classes obtain tighter distribution areas (i.e., limited space), while the classification learning provides better discriminativeness for the target classes. Note that there actually have been some works fused the clustering and classification functions into a unified learning framework [116], [117]. Unfortunately, these works are still under a closed set assumption, thus, some serious efforts need to be done to adapt them to the OSR scenario or to specifically design this type of classifier for the OSR scenario.

2) *Modeling unknown unknown classes*: Under the open set assumption, modeling the unknown unknown classes is impossible as we only have the available knowledge of known known classes. However, properly relaxing some restrictions will make it possible, where one way is to generate the unknown unknown class data by the adversarial learning technique to account for open space to some extent like [78]–[81], in which the key is how to generate the valid unknown unknown class data. Besides, while in theory the open space should be infinitely large, it seems to be enough for us to just pay attention to the unknown unknown classes appearing in testing instead of all of them in practice. From this perspective, one can take the testing data as unlabeled data into account when training a model, which seems to have the flavor of the transductive learning. Note that the transductive learning is currently under a closed set assumption. Thus remoulding transductive learning technique to adapt it to the OSR scenario will be a promising direction in the future work. Additionally, due to the data adaptive nature of Dirichlet process, like the CD-OSR, the Dirichlet process-based OSR methods are also worth for further exploration.

### B. About Rejecting

Till now, most existing OSR algorithms mainly care about effectively rejecting unknown unknown classes, yet only a few works [66], [77] focus on the subsequent processing of reject samples, and these works usually adopt a post hoc strategy [57]. Therefore, expanding existing open set recognition with a function of new class knowledge discovery will be an

TABLE IV  
DIFFERENCES BETWEEN GENERALIZED OPEN SET RECOGNITION AND GENERALIZED ZERO-SHOT LEARNING

Task \ Setting	Training	Testing	Goal
Generalized Zero-shot Learning	Known known classes, side-information <sup>3</sup>	Known known classes & unknown known classes	Identifying known known classes & unknown known classes
Generalized Open Set Recognition*	Known known classes side-information <sup>4</sup>	Known known classes & unknown unknown classes	Identifying known known classes & 'cognizing' unknown unknown classes

The side-information<sup>3</sup> is the same to side-information<sup>1</sup> in Table I. The G-OSR here is one of the Generalized Open Set Recognition in our proposed future research work, where the side-information<sup>4</sup> denotes the available semantic/attribute information only for known known classes. As part of this information usually spans the known known and unknown unknown classes, thus we hope to use it to further 'cognize' the unknown unknown classes instead of simply rejecting them.

interesting research topic. Moreover, to our best knowledge, the interpretability of reject option seems to have not been discussed, in which a reject option may correspond to a low confidence target class, an outlier, or a new class, which is also an interesting future research direction. Besides, some related works in other research community can be found in [118]–[122].

### C. About the Decision

As discussed in subsection III-E, almost all existing OSR techniques are designed specially for recognizing individual samples, even these samples are collectively coming in batch like image-set recognition [123]. Actually, such a decision does not consider correlations among the testing samples. Therefore, the collective decision [57] seems to be a better alternative since it can not only take the correlations among the testing samples into account but also make it possible to discover new classes at the same time. We thus expect a future direction on extending the existing OSR methods by adopting the collective decision instead of the single-sample decision strategy.

### D. Open Set + 'sth'

As open set scenario is a more practical assumption for the real-world classification/recognition tasks, it can naturally be combined with various fields involving classification/recognition such as semi-supervised learning, domain adaptation, active learning, multi-task learning, multi-label learning, multi-view learning, and so forth. For example, [124]–[126] recently introduced this scenario into domain adaptation, while [127] explored the open set classification in active learning field. Therefore, many interesting works are worth looking forward to.

### E. Generalized Open Set Recognition

The OSR assumes that only the knowledge of known known classes is available in training, meaning that we can also utilize a variety of side-information regarding the known known classes. Nevertheless most existing OSR methods actually just use the information of samples at feature level, leaving out other side-information such as semantic/attribute information, knowledge graph, universum etc., which is also important for improving the performance of these existing algorithms. Therefore, we give the following two promising research directions.

1) *Appending semantic/attribute information:* In fact, a lot of semantic/attribute information is shared between the known known and the unknown unknown classes. Therefore, we can fully utilize this kind of information to 'cognize' the unknown unknown classes, or at least to provide a rough semantic/attribute description for the corresponding unknown unknown classes instead of simply rejecting them. Note that this setup is different from the one in ZSL (or G-ZSL) which assumes that the semantic/attribute information of both the known known and unknown known classes are known in training. Furthermore, Table IV shows the difference between this and the generalized zero-shot recognition. Besides, some related works can be found in [111], [122], [128], [129]. In addition, there also some conceptually similar topics have been studied in other research community such as open-vocabulary object retrieval [130], [131], open-world person re-identification [132] or searching targets [133], open vocabulary scene parsing [134].

2) *Using other available side-information:* The main reason for open space risk is that the traditional classifiers trained under closed set scenario usually divide over-occupied space for known known classes, thus inevitably resulting in misclassifications once the unknown unknown class samples fall into the space divided for some known known class. From this perspective, the open space risk will be reduced as the space divided for those known known classes decreases by using other side-information like universum [135], [136] to shrink their regions as much as possible. Taking the digital identification as an example, assume the training set including the classes of interest '1', '3', '6'; the testing set including classes '1', '3', '6', '8', '9'. If we also have the data of universum classes '2', '4', '5', or English letters, and so forth, then we can use the training and universum class data in modeling to extend the existing OSR models, further reducing the open space risk. We therefore foresee a more generalized setting will be adopted by the future open set recognition.

### F. Relative Open Set Recognition

While the open set scenario is ubiquitous, there are also some real-world scenarios that are not completely open in practice. Recognition/classification in such scenarios can be called relative open set recognition. Taking the medical diagnosis as an example, the whole sample space actually can be divided into two subspace respectively for sick and not sick samples, and at such a level of detecting whether the sample is

sick or not, it is indeed a closed set problem. However, when we need to further identify the types of the diseases, this will naturally become a complete OSR problem since new disease unseen in training may appear in testing. Note that no relevant work currently explores this novel mixed scenario jointly.

### G. Knowledge Integration for Open Set Recognition

In fact, the incomplete knowledge of the world is universal, especially for the single individuals: something you know does not mean I also know. For example, the terrestrial species (sub-knowledge set) obviously are the open set for the classifiers trained on marine species. As the saying goes, “two heads are better than one”, thus how to integrate the classifiers trained on each sub-knowledge set to further reduce the open space risk will be an interesting yet challenging topic in the future work, especially for such a situation: we can only obtain the classifiers trained on corresponding sub-knowledge sets, yet these sub-knowledge sets are not available due to the privacy protection of data. This seems to have the flavor of domain adaptation having multiple source domains and one target domain (mS1T) [137]–[140] to some extent.

## VII. CONCLUSION

As discussed above, in real-world recognition/classification, it is usually impossible to model everything [141], thus the OSR scenario is ubiquitous. On the other hand, although many related algorithms have been proposed for OSR, it still faces serious challenges. In addition, till now there has been no systematic summary on this topic. Therefore, in this paper, we give a comprehensive review of existing OSR techniques, covering various aspects ranging from related definitions, representations of models, datasets, experiment setup and evaluation metrics. Note that for the sake of convenience, the categorization of existing OSR techniques in this paper is just one of the possible ways, while other ways can also effectively categorize them, and some may be more appropriate but beyond our focus here.

Furthermore, in order to avoid the reader confusing the tasks similar to OSR, we also briefly analyze the relationships between OSR and its related tasks including zero-shot, one-shot (few-shot) recognition/learning techniques, classification with reject option, and so forth. Beyond this, as a natural extension of OSR, the open world recognition is reviewed as well. More importantly, we analyze and discuss the limitations of these existing approaches, and point out some promising subsequent research directions in this field. Finally, we expect that the reader can benefit from this paper.

## ACKNOWLEDGMENT

The authors would like to thank the support from the Key Program of NSFC under Grant No. 61732006, NSFC under Grant No. 61672281, and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant No. KYCX18\_0306.

## APPENDIX A SOFTWARE PACKAGES

We present a table (Table V) of several available software packages implementing those models presented in the main text.

## REFERENCES

- [1] S. Thrun and T. M. Mitchell, “Lifelong robot learning,” *Robotics & Autonomous Systems*, vol. 15, no. 1-2, pp. 25–46, 1995.
- [2] A. Pentina and C. H. Lampert, “A pac-bayesian bound for lifelong learning,” in *International Conference on International Conference on Machine Learning*, pp. II–991, 2014.
- [3] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge & Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [4] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, “A survey of transfer learning,” *Journal of Big Data*, vol. 3, no. 1, p. 9, 2016.
- [5] L. Shao, F. Zhu, and X. Li, “Transfer learning for visual categorization: A survey,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 5, pp. 1019–1034, 2015.
- [6] V. M. Patel, R. Gopalan, R. Li, and R. Chellappa, “Visual domain adaptation: A survey of recent advances,” *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 53–69, 2015.
- [7] M. Palatucci, D. Pomerleau, G. Hinton, and T. M. Mitchell, “Zero-shot learning with semantic output codes,” in *International Conference on Neural Information Processing Systems*, pp. 1410–1418, 2009.
- [8] Y. Fu, T. Xiang, Y. G. Jiang, X. Xue, L. Sigal, and S. Gong, “Recent advances in zero-shot recognition,” *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 112–125.
- [9] C. H. Lampert, H. Nickisch, and S. Harmeling, *Learning to detect unseen object classes by between-class attribute transfer*. IEEE, 2009.
- [10] F. F. Li, R. Fergus, and P. Perona, “One-shot learning of object categories,” *IEEE transactions on pattern analysis & machine intelligence*, vol. 28, no. 4, pp. 594–611, 2006.
- [11] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, “One-shot learning by inverting a compositional causal process,” in *International Conference on Neural Information Processing Systems*, pp. 2526–2534, 2013.
- [12] F. F. Li, Fergus, and Perona, “A bayesian approach to unsupervised one-shot learning of object categories,” in *IEEE International Conference on Computer Vision, 2003. Proceedings*, pp. 1134–1141 vol.2, 2008.
- [13] Y. Amit, M. Fink, N. Srebro, and S. Ullman, “Uncovering shared structures in multiclass classification,” in *Machine Learning, Proceedings of the Twenty-Fourth International Conference*, pp. 17–24, 2007.
- [14] A. Torralba, K. P. Murphy, and W. T. Freeman, “Using the forest to see the trees: Exploiting context for visual object detection and localization,” *Communications of the Acm*, vol. 53, no. 3, pp. 107–114, 2010.
- [15] J. Snell, K. Swersky, and R. S. Zemel, “Prototypical networks for few-shot learning,” in *Advances in Neural Information Processing Systems*, pp. 4077–4087, 2017.
- [16] L. Bertinetto, J. F. Henriques, J. Valmadre, P. H. S. Torr, and A. Vedaldi, “Learning feed-forward one-shot learners,” in *Advances in Neural Information Processing Systems*, pp. 523–531, 2016.
- [17] W. J. Scheirer, A. D. R. Rocha, A. Sapkota, and T. E. Boulton, “Toward open set recognition,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 35, no. 7, pp. 1757–72, 2013.
- [18] W. J. Scheirer, L. P. Jain, and T. E. Boulton, “Probability models for open set recognition,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 36, no. 11, pp. 2317–2324, 2014.
- [19] L. P. Jain, W. J. Scheirer, and T. E. Boulton, “Multi-class open set recognition using probability of inclusion,” in *European Conference on Computer Vision*, pp. 393–409, 2014.
- [20] N. A. Ross, “Known knowns, known unknowns and unknown unknowns: A 2010 update on carotid artery disease,” *Surgeon Journal of the Royal Colleges of Surgeons of Edinburgh & Ireland*, vol. 8, no. 2, pp. 79–86, 2010.
- [21] W.-L. Chao, S. Changpinyo, B. Gong, and F. Sha, “An empirical study and analysis of generalized zero-shot learning for object recognition in the wild,” in *European Conference on Computer Vision*, pp. 52–68, Springer, 2016.
- [22] C. Chow, “On optimum recognition error and reject tradeoff,” *IEEE Transactions on information theory*, vol. 16, no. 1, pp. 41–46, 1970.

TABLE V  
SOFTWARE PACKAGES

Model	Language	Author	Link
1-vs-Set [17], W-SVM [18], $P_I$ -SVM [19]	C/C++	Jain et al.	<a href="https://github.com/ljain2/libsvm-openset">https://github.com/ljain2/libsvm-openset</a>
BFHC [62], EPCC [65]	Matlab	Cevikalp et al.	<a href="http://mlcv.ogu.edu.tr/software.html">http://mlcv.ogu.edu.tr/software.html</a>
SROSR [55]	Matlab	Zhang et al.	<a href="https://github.com/hezhangsprinter/SROSR">https://github.com/hezhangsprinter/SROSR</a>
NNO [66]	Matlab	Bendale et al.	<a href="http://vast.uccs.edu/OpenWorld">http://vast.uccs.edu/OpenWorld</a>
HPLS, HFCN [69]	Python	Vareto et al.	<a href="https://github.com/rafaelvareto/HPLS-HFCN-openset">https://github.com/rafaelvareto/HPLS-HFCN-openset</a>
OpenMax [71]	Python	Bendale et al.	<a href="https://github.com/abhijitbendale/OSDN">https://github.com/abhijitbendale/OSDN</a>
DOC [75]	Python	Shu et al.	<a href="https://github.com/alexander-rakhlin/CNN-for-Sentence-Classification-in-Keras">https://github.com/alexander-rakhlin/CNN-for-Sentence-Classification-in-Keras</a>
Neal(ECCV) [79]	Python	Neal et al.	<a href="https://github.com/lwneal/counterfactual-open-set">https://github.com/lwneal/counterfactual-open-set</a>
ASG [80]	Python	Liu et al.	<a href="https://github.com/eyounx/ASG">https://github.com/eyounx/ASG</a>
EVM [82]	Python	Rudd et al.	<a href="https://github.com/EMRRResearch/ExtremeValueMachine">https://github.com/EMRRResearch/ExtremeValueMachine</a>

[23] P. L. Bartlett and M. H. Wegkamp, "Classification with a reject option using a hinge loss," <i>Journal of Machine Learning Research</i> , vol. 9, no. Aug, pp. 1823–1840, 2008.	[43] P. Bodesheim, A. Freytag, E. Rodner, M. Kemmler, and J. Denzler, "Kernel null space methods for novelty detection," in <i>Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition</i> , pp. 3374–3381, 2013.
[24] D. M. Tax and R. P. Duin, "Growing a multi-class classifier with a reject option," <i>Pattern Recognition Letters</i> , vol. 29, no. 10, pp. 1565–1570, 2008.	[44] P. J. Phillips, P. Grother, and R. Micheals, <i>Evaluation Methods in Face Recognition</i> . Handbook of Face Recognition. Springer New York, 2005.
[25] L. Fischer, B. Hammer, and H. Wersing, "Optimal local rejection for classifiers," <i>Neurocomputing</i> , vol. 214, pp. 445–457, 2016.	[45] F. Li and H. Wechsler, "Open set face recognition using transduction," <i>IEEE Transactions on Pattern Analysis &amp; Machine Intelligence</i> , vol. 27, no. 11, pp. 1686–1697, 2005.
[26] G. Fumera and F. Roli, "Support vector machines with embedded reject option," in <i>Pattern recognition with support vector machines</i> , pp. 68–82, Springer, 2002.	[46] Q. Wu, C. Jia, and W. Chen, "A novel classification-rejection sphere svms for multi-class classification problems," in <i>Natural Computation, 2007. ICNC 2007. Third International Conference on</i> , vol. 1, pp. 34–38, IEEE, 2007.
[27] Y. Grandvalet, A. Rakotomamonjy, J. Keshet, and S. Canu, "Support vector machines with a reject option," in <i>In Advances in Neural Information Processing Systems</i> , pp. 537–544, 2009.	[47] Y.-C. F. Wang and D. Casasent, "A support vector hierarchical method for multi-class classification and rejection," pp. 3281–3288, 2009.
[28] R. Herbei and M. H. Wegkamp, "Classification with reject option," <i>Canadian Journal of Statistics</i> , vol. 34, no. 4, pp. 709–721, 2006.	[48] B. Hefflin, W. Scheirer, and T. E. Boult, "Detecting and classifying scars, marks, and tattoos found in the wild," in <i>IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems</i> , pp. 31–38, 2012.
[29] M. Yuan and M. Wegkamp, "Classification methods with reject option based on convex risk minimization," <i>Journal of Machine Learning Research</i> , vol. 11, no. Jan, pp. 111–130, 2010.	[49] D. A. Pritsos and E. Stamatatos, "Open-set classification for automated genre identification," in <i>European Conference on Advances in Information Retrieval</i> , pp. 207–217, 2013.
[30] R. Zhang and D. N. Metaxas, "Ro-svm: Support vector machine with reject option for image categorization," in <i>BMVC</i> , pp. 1209–1218, Citeseer, 2006.	[50] V. Vapnik, <i>Estimation of Dependences Based on Empirical Data</i> . Springer-Verlag, 1982.
[31] M. Wegkamp, "Lasso type classifiers with a reject option," <i>Electronic Journal of Statistics</i> , vol. 1, no. 3, pp. 155–168, 2007.	[51] A. J. Smola, "Learning with kernels — support vector machines," <i>Lecture Notes in Computer Science</i> , vol. 42, no. 4, pp. 1–28, 2008.
[32] Y. Geifman and E. Y. Ran, "Selective classification for deep neural networks," pp. 4872–4887, 2017.	[52] Schölkopf, Bernhard, and A. J. Smola, <i>Learning with Kernels</i> . The MIT Press, 2002.
[33] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," <i>Neural computation</i> , vol. 13, no. 7, pp. 1443–1471, 2001.	[53] J. M. Keynes, <i>A treatise on probability</i> . Macmillan & Company, Limited, 1921.
[34] L. M. Manevitz and M. Yousef, "One-class svms for document classification," <i>Journal of machine Learning research</i> , vol. 2, no. Dec, pp. 139–154, 2001.	[54] N. Shackel, "Bertrands paradox and the principle of indifference," <i>Philosophy of Science</i> , vol. 74, no. 2, pp. 150–175, 2007.
[35] D. M. Tax and R. P. Duin, "Support vector data description," <i>Machine learning</i> , vol. 54, no. 1, pp. 45–66, 2004.	[55] H. Zhang and V. Patel, "Sparse representation-based open set recognition," <i>IEEE Transactions on Pattern Analysis &amp; Machine Intelligence</i> , vol. 39, no. 8, pp. 1690–1696, 2017.
[36] S. S. Khan and M. G. Madden, "A survey of recent trends in one class classification," in <i>Irish conference on artificial intelligence and cognitive science</i> , pp. 188–197, Springer, 2009.	[56] Y. Yang, C. Hou, Y. Lang, D. Guan, D. Huang, and J. Xu, "Open-set human activity recognition based on micro-doppler signatures," <i>Pattern Recognition</i> , vol. 85, pp. 60–69, 2019.
[37] H. Jin, Q. Liu, and H. Lu, "Face detection using one-class-based support vectors," in <i>IEEE International Conference on Automatic Face and Gesture Recognition</i> , pp. 457–462, 2004.	[57] C. Geng and S. Chen, "Collective decision for open set recognition," <i>arXiv preprint arXiv:1806.11258</i> , 2018.
[38] M. Wu and J. Ye, "A small sphere and large margin approach for novelty detection using training data with outliers," <i>IEEE Transactions on Pattern Analysis &amp; Machine Intelligence</i> , vol. 31, no. 11, pp. 2088–2092, 2009.	[58] D. O. Cardoso, J. Gama, and F. M. G. Frana, "Weightless neural networks for open set recognition," <i>Machine Learning</i> , no. 106(9-10), pp. 1547–1567, 2017.
[39] H. Cevikalp and B. Triggs, "Efficient object detection using cascades of nearest convex model classifiers," in <i>Computer Vision and Pattern Recognition</i> , pp. 3138–3145, 2012.	[59] G. Bouchard and B. Triggs, "The trade-off between generative and discriminative classifiers," <i>Proceedings in Computational Statistics Symposium of Iasc</i> , pp. 721–728, 2004.
[40] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," <i>Signal Processing</i> , vol. 99, pp. 215–249, 2014.	[60] J. A. Lasserre, C. M. Bishop, and T. P. Minka, "Principled hybrids of generative and discriminative models," in <i>Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on</i> , pp. 87–94, 2006.
[41] J. Denzler, E. Rodner, P. Bodesheim, and A. Freytag, "Beyond the closed-world assumption: The importance of novelty detection and open set recognition,"	[61] H. Cevikalp, B. Triggs, and V. Franc, "Face and landmark detection
[42] N. Görnitz, L. A. Lima, K.-R. Müller, M. Kloft, and S. Nakajima, "Support vector data descriptions and $k$ -means clustering: One class?,"	

- by using cascade of classifiers,” in *IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pp. 1–7, 2013.
- [62] H. Cevikalp, “Best fitting hyperplanes for classification,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 6, p. 1076, 2017.
  - [63] M. D. Scherrek and B. D. Rigling, “Open set recognition for automatic target classification with rejection,” *IEEE Transactions on Aerospace & Electronic Systems*, vol. 52, no. 2, pp. 632–642, 2016.
  - [64] H. Cevikalp and B. Triggs, “Polyhedral conic classifiers for visual object detection and classification,” in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4114–4122, 2017.
  - [65] H. Cevikalp and H. S. Yavuz, “Fast and accurate face recognition with image sets,” in *IEEE International Conference on Computer Vision Workshop*, pp. 1564–1572, 2017.
  - [66] A. Bendale and T. Boulton, “Towards open world recognition,” in *Computer Vision and Pattern Recognition*, pp. 1893–1902, 2015.
  - [67] P. R. M. Junior, R. M. D. Souza, R. D. O. Werneck, B. V. Stein, D. V. Pazinato, W. R. D. Almeida, O. A. B. Penatti, R. D. S. Torres, and A. Rocha, “Nearest neighbors distance ratio open-set classifier,” *Machine Learning*, vol. 106, no. 3, pp. 359–386, 2017.
  - [68] G. Fei and B. Liu, “Breaking the closed world assumption in text classification,” in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 506–514, 2016.
  - [69] R. Vareto, S. Silva, F. Costa, and W. R. Schwartz, “Towards open-set face recognition using hashing functions,” in *IEEE International Joint Conference on Biometrics*, 2018.
  - [70] M. A. C. Neira, P. R. M. Junior, A. Rocha, and R. D. S. Torres, “Data-fusion techniques for open-set recognition problems,” *IEEE Access*, vol. 6, pp. 21242–21265, 2018.
  - [71] A. Bendale and T. E. Boulton, “Towards open set deep networks,” *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1563–1572, 2016.
  - [72] A. Rozsa, M. Gunther, and T. E. Boulton, “Adversarial robustness: Softmax versus openmax,” *arXiv preprint arXiv:1708.01697*, 2017.
  - [73] M. Hassen and P. K. Chan, “Learning a neural-network-based representation for open set recognition,” *arXiv preprint arXiv:1802.04365*, 2018.
  - [74] S. Prakhyia, V. Venkataram, and J. Kalita, “Open set text classification using convolutional neural networks,” in *International Conference on Natural Language Processing*, 2017.
  - [75] L. Shu, H. Xu, and B. Liu, “Doc: Deep open classification of text documents,” *arXiv preprint arXiv:1709.08716*, 2017.
  - [76] D. O. Cardoso, F. Franca, and J. Gama, “A bounded neural network for open set recognition,” in *International Joint Conference on Neural Networks*, pp. 1–7, 2015.
  - [77] L. Shu, H. Xu, and B. Liu, “Unseen class discovery in open-world classification,” *arXiv preprint arXiv:1801.05609*, 2018.
  - [78] Z. Y. Ge, S. Demyanov, Z. Chen, and R. Garnavi, “Generative openmax for multi-class open set classification,” *arXiv preprint arXiv:1707.07418*, 2017.
  - [79] L. Neal, M. Olson, X. Fern, W. Wong, and F. Li, “Open set learning with counterfactual images,” *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 613–628, 2018.
  - [80] Z. Zheng, L. Zheng, Z. Hu, and Y. Yang, “Open-category classification by adversarial sample generation,” *arXiv preprint arXiv:1705.08722*, 2017.
  - [81] Z. Zheng, L. Zheng, Z. Hu, and Y. Yang, “Open set adversarial examples,” *arXiv preprint arXiv:1809.02681*, 2018.
  - [82] E. M. Rudd, L. P. Jain, W. J. Scheirer, and T. E. Boulton, “The extreme value machine,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 40, no. 3, pp. 762–768, 2018.
  - [83] E. Vignotto and S. Engelke, “Extreme value theory for open set classification-gpd and gev classifiers,” *arXiv preprint arXiv:1808.09902*, 2018.
  - [84] C. Cortes and V. Vapnik, “Support-vector networks,” in *Machine Learning*, pp. 273–297, 1995.
  - [85] S. Cruz, C. Coleman, E. M. Rudd, and T. E. Boulton, “Open set intrusion recognition for fine-grained attack categorization,” in *IEEE International Symposium on Technologies for Homeland Security*, pp. 1–6, 2017.
  - [86] E. Rudd, A. Rozsa, M. Gunther, and T. Boulton, “A survey of stealth malware: Attacks, mitigation measures, and steps toward autonomous open world solutions,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1145–1172, 2017.
  - [87] R. N. Gasimov and G. Ozturk, “Separation via polyhedral conic functions,” *Optimization Methods & Software*, vol. 21, no. 4, pp. 527–540, 2006.
  - [88] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, “Sparse representation for computer vision and pattern recognition,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031–1044, 2010.
  - [89] R. Rubinstein, A. M. Bruckstein, and M. Elad, “Dictionaries for sparse representation modeling,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, 2010.
  - [90] J. Peng, L. Li, and Y. Y. Tang, “Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images,” *IEEE Transactions on Neural Networks and Learning Systems*, 2018.
  - [91] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.
  - [92] T. Mensink, J. Verbeek, F. Perronnin, and G. Csurka, “Distance-based image classification: Generalizing to new classes at near-zero cost,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 11, pp. 2624–2637, 2013.
  - [93] M. Ristin, M. Guillaumin, J. Gall, and L. V. Gool, “Incremental learning of ncm forests for large-scale image classification,” *Computer Vision and Pattern Recognition*, pp. 3654–3661, 2014.
  - [94] J. P. Papa and C. T. N. Suzuki, *Supervised pattern classification based on optimum-path forest*. John Wiley & Sons, Inc., 2009.
  - [95] A. Nguyen, J. Yosinski, and J. Clune, “Deep neural networks are easily fooled: High confidence predictions for unrecognizable images,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 427–436, 2015.
  - [96] I. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” in *International Conference on Learning Representations. Computational and Biological Learning Society*, 2015.
  - [97] H. Xu, B. Liu, and P. S. Yu, “Learning to accept new classes without training,” *arXiv preprint arXiv:1809.06004*, 2018.
  - [98] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, “Intriguing properties of neural networks,” in *International Conference on Learning Representations. Computational and Biological Learning Society*, 2014.
  - [99] I. Goodfellow, “Nips 2016 tutorial: Generative adversarial networks,” *arXiv preprint arXiv:1701.00160*, 2016.
  - [100] W. J. Scheirer, A. Rocha, R. J. Micheals, and T. E. Boulton, “Meta-recognition: The theory and practice of recognition score analysis,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 33, no. 8, pp. 1689–95, 2011.
  - [101] V. Fragoso, P. Sen, S. Rodriguez, and M. Turk, “Evsac: Accelerating hypotheses generation by modeling matching scores with extreme value theory,” in *IEEE International Conference on Computer Vision*, pp. 2472–2479, 2013.
  - [102] M. Gunther, S. Cruz, E. M. Rudd, and T. E. Boulton, “Toward open-set face recognition,” *Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017.
  - [103] J. Henrydoss, S. Cruz, E. M. Rudd, T. E. Boulton, et al., “Incremental open set intrusion recognition using extreme value machine,” in *Machine Learning and Applications (ICMLA), 2017 16th IEEE International Conference on*, pp. 1089–1093, IEEE, 2017.
  - [104] Y. W. Teh, “Dirichlet process,” in *Encyclopedia of machine learning*, pp. 280–287, Springer, 2011.
  - [105] S. J. Gershman and D. M. Blei, “A tutorial on bayesian nonparametric models,” *Journal of Mathematical Psychology*, vol. 56, no. 1, pp. 1–12, 2012.
  - [106] W. Fan and N. Bouguila, “Online learning of a dirichlet process mixture of beta-liouville distributions via variational inference,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 11, pp. 1850–1862, 2013.
  - [107] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, “Hierarchical dirichlet processes,” *Publications of the American Statistical Association*, vol. 101, no. 476, pp. 1566–1581, 2006.
  - [108] A. Bargi, R. Y. Da Xu, and M. Piccardi, “Adon hdp-hmm: an adaptive online model for segmentation and classification of sequential data,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 3953–3968, 2018.
  - [109] R. De Rosa, T. Mensink, and B. Caputo, “Online open world recognition,” *arXiv preprint arXiv:1604.02275*, 2016.
  - [110] T. Doan and J. Kalita, “Overcoming the challenge for text classification in the open world,” in *Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual*, pp. 1–7, IEEE, 2017.

- [111] V. P. A. Lonij, A. Rawat, and M. I. Nicolae, "Open-world visual recognition using knowledge graphs," *arXiv preprint arXiv:1708.08310*, 2017.
- [112] Y. Shi, Y. Wang, Y. Zou, Q. Yuan, Y. Tian, and Y. Shu, "Odn: Opening the deep network for open-set action recognition," in *2018 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, IEEE, 2018.
- [113] C. Anna, L. Gnter, R. H. Joachim, and M. Klaus, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
- [114] W. J. Youden, "Index for rating diagnostic tests.," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.
- [115] M. Sokolova, N. Japkowicz, and S. Szpakowicz, *Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation*. Springer Berlin Heidelberg, 2006.
- [116] W. Cai, S. Chen, and D. Zhang, "A multiobjective simultaneous learning framework for clustering and classification," *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 185–200, 2010.
- [117] Q. Qian, S. Chen, and W. Cai, "Simultaneous clustering and classification over cluster structure representation," *Pattern Recognition*, vol. 45, no. 6, pp. 2227–2236, 2012.
- [118] X. Mu, M. T. Kai, and Z. H. Zhou, "Classification under streaming emerging new classes: A solution using completely-random trees," *IEEE Transactions on Knowledge & Data Engineering*, vol. 29, no. 8, pp. 1605–1618, 2017.
- [119] S. Liu, R. Garrepalli, T. G. Dietterich, A. Fern, and D. Hendrycks, "Open category detection with pac guarantees," *arXiv preprint arXiv:1808.00529*, 2018.
- [120] A. Vyas, N. Jammalamadaka, X. Zhu, D. Das, B. Kaul, and T. L. Willke, "Out-of-distribution detection using an ensemble of self supervised leave-out classifiers," *arXiv preprint arXiv:1809.03576*, 2018.
- [121] M. Masana, I. Ruiz, J. Serrat, J. van de Weijer, and A. M. Lopez, "Metric learning for novelty and anomaly detection," *arXiv preprint arXiv:1808.05492*, 2018.
- [122] H. Dong, Y. Fu, L. Sigal, S. J. Hwang, Y.-G. Jiang, and X. Xue, "Learning to separate domains in generalized zero-shot and open set learning: a probabilistic perspective," *arXiv preprint arXiv:1810.07368*, 2018.
- [123] L. Wang and S. Chen, "Joint representation classification for collective face recognition," *Pattern Recognition*, vol. 63, pp. 182–192, 2017.
- [124] P. P. Busto and J. Gall, "Open set domain adaptation," in *IEEE International Conference on Computer Vision*, pp. 754–763, 2017.
- [125] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, "Open set domain adaptation by backpropagation," *arXiv preprint arXiv:1804.10427*, 2018.
- [126] M. Baktashmotlagh, M. Faraki, T. Drummond, and M. Salzmann, "Learning factorized representations for open-set domain adaptation," *arXiv preprint arXiv:1805.12277*, 2018.
- [127] Z.-Y. Liu and S.-J. Huang, "Active sampling for open-set classification without initial annotation," In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, 2019.
- [128] Y. Fu and L. Sigal, "Semi-supervised vocabulary-informed learning," in *Computer Vision and Pattern Recognition*, pp. 5337–5346, 2016.
- [129] Y. Fu, H. Z. Dong, Y. F. Ma, Z. Zhang, and X. Xue, "Vocabulary-informed extreme value learning," *arXiv preprint arXiv:1705.09887*, 2017.
- [130] S. Guadarrama, E. Rodner, K. Saenko, N. Zhang, R. Farrell, J. Donahue, and T. Darrell, "Open-vocabulary object retrieval.," in *Robotics: science and systems*, vol. 2, p. 6, Citeseer, 2014.
- [131] S. Guadarrama, E. Rodner, K. Saenko, and T. Darrell, "Understanding object descriptions in robotics by open-vocabulary object retrieval and detection," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 265–280, 2016.
- [132] W.-S. Zheng, S. Gong, and T. Xiang, "Towards open-world person re-identification by one-shot group-based verification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 3, pp. 591–606, 2016.
- [133] H. Sattar, S. Muller, M. Fritz, and A. Bulling, "Prediction of search targets from fixations in open-world settings," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 981–990, 2015.
- [134] H. Zhao, X. Puig, B. Zhou, S. Fidler, and A. Torralba, "Open vocabulary scene parsing," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017.
- [135] V. Cherkassky, S. Dhar, and W. Dai, "Practical conditions for effectiveness of the universum learning," *IEEE Transactions on Neural Networks*, vol. 22, no. 8, pp. 1241–55, 2011.
- [136] Z. Qi, Y. Tian, and Y. Shi, "Twin support vector machine with universum data," *Neural Networks*, vol. 36, no. 3, pp. 112–119, 2012.
- [137] S. Bendavid, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," *Machine Learning*, vol. 79, no. 1-2, pp. 151–175, 2010.
- [138] S. Sun, H. Shi, and Y. Wu, *A survey of multi-source domain adaptation*. Elsevier Science Publishers B. V., 2015.
- [139] H. Liu, M. Shao, and Y. Fu, "Structure-preserved multi-source domain adaptation," in *IEEE International Conference on Data Mining*, pp. 1059–1064, 2017.
- [140] H. Yu, M. Hu, and S. Chen, "Unsupervised domain adaptation without exactly shared categories," *arXiv preprint arXiv:1809.00852*, 2018.
- [141] T. G. Dietterich, "Steps toward robust artificial intelligence," *AI Magazine*, vol. 38, no. 3, pp. 3–24, 2017.