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 to not only accurately classify the seen classes, but also effectively deal with 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, evaluation criteria, and algorithm comparisons. 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 review 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-short learning, one-shot learning

1 Introduction

Under a common closed set (or static environment) assumption: the training and testing data are drawn from the same label and feature spaces, 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 topics have been explored including lifelong learning [1], [2], transfer learning [3], [4], [5], domain adaptation [6], [7], zero-shot [8], [9], [10], one-shot (few-shot) [11], [12], [13], [14], [15], [16], [17], [18], [19], [20] recognition/learning, open set recognition/classification [21], [22], [23], and so forth.

Based on Donald Rumsfeld's famous "There are known knowns" statement [24], we further expand the basic recognition categories of classes asserted by [22], where we restate that recognition should consider four basic categories of classes as follows:

1) *known known classes* (KKCs), i.e., the classes with distinctly labeled positive training samples (also serving as negative samples for other KKCs), and even have the corresponding side-information like semantic/attribute information, etc;
2) *known unknown classes* (KUCs), i.e., labeled negative samples, not necessarily grouped into meaningful classes, such as the background classes [25], the universum classes [26],[1] etc;
3) unknown known classes*[2] (UKCs), i.e., classes with no available samples in training, but available side-information (e.g., semantic/attribute information) of them during training;
4) *unknown unknown classes* (UUCs), i.e., classes without any information regarding them during training: not only unseen but also having not side-information (e.g., semantic/attribute information, etc.) during training.

Fig. 1 gives an example of visualizing KKCs, KUCs, and UUCs from the real data distribution using t-SNE [27]. Note that since the main difference between UKCs and UUCs lies in whether their side-information is available or not, we here only visualize UUCs. Traditional classification only considers KKCs, while including KUCs will result in models with an explicit "other class," or a detector trained with unclassified negatives [22]. Unlike traditional classification, zero-shot learning (ZSL) focuses more on the recognition of UKCs. As the saying goes: prediction is impossible without any assumptions about how the past relates to the future. ZSL leverages semantic information shared between KKCs and UKCs to implement such a recognition [8], [9]. In fact, assuming the test samples only from UKCs is rather restrictive and impractical, since we usually know nothing about them either from KKCs or UKCs. On the other hand, the object frequencies in natural world follow long-tailed distributions [28], [29], meaning that KKCs are more common than UKCs. Therefore, some researchers have begun to pay

• *The authors are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics of (NUAA), MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing 211106, China. E-mail: {gengchuanxing, huangsj, s.chen}@nuaa.edu.cn.*

Manuscript received 23 Nov. 2018; revised 14 Oct. 2019; accepted 8 Mar. 2020. Date of publication 18 Mar. 2020; date of current version 2 Sept. 2021.
(Corresponding author: Songcan Chen.)
Recommended for acceptance by L. Sigal.
Digital Object Identifier no. 10.1109/TPAMI.2020.2981604

1. The universum classes [26] usually denotes the samples that do not belong to either class of interest for the specific learning problem.
2. * represents the expanded basic recognition class by ourselves.

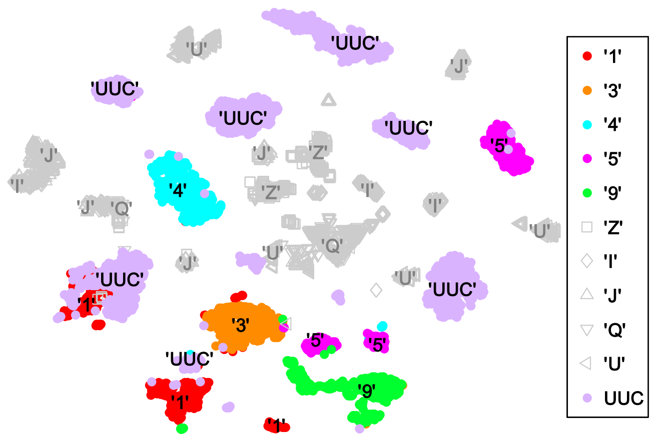


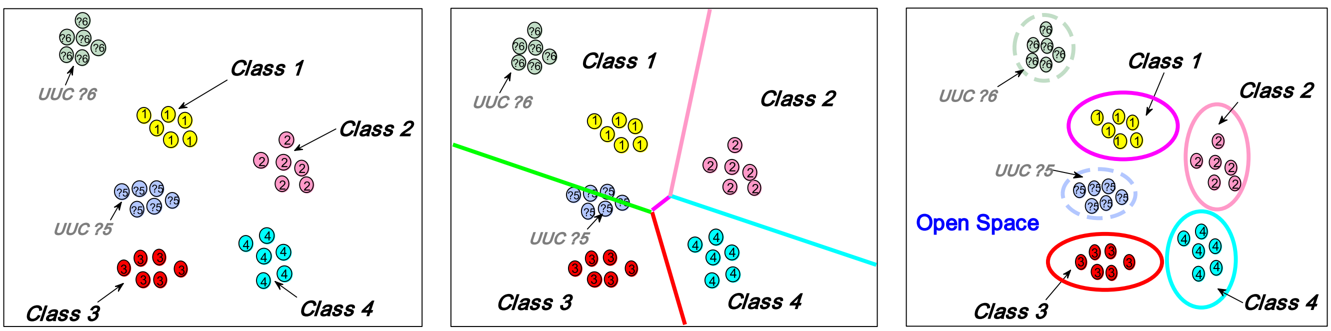
Fig. 1. An example of visualizing KKCs, KUCs, and UUCs from the real data distribution using t-SNE. Here, ‘1’,‘3’,‘4’,‘5’,‘9’ are randomly selected from PENDIGITS as KKCs, while the remaining classes in it as UUCs. ‘Z’,‘I’,‘J’,‘Q’,‘U’ are randomly selected from LETTER as KUCs. This visualization also indicates that the distribution of one class may consist of multiple subclass/subclusters, e.g., class ‘1’, ‘5’, ‘U’, etc.

attention to the more generalized ZSL (G-ZSL) [30], [31], [32], [33], where the testing samples come from both KKCs and UKCs. As a closely-related problem to ZSL, one/few-shot learning (FSL) can be seen as natural extensions of zero-shot learning when a limited number of UKCs’ samples during training are available [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. Similar to G-ZSL, a more realistic setting for FSL considering both KKCs and UKCs in testing, i.e., generalized FSL (G-FSL), is also becoming more popular [34]. Compared to (G-)ZSL and (G-)FSL, open set recognition (OSR)[21], [22], [23] probably faces a more serious challenge due to the fact that only KKCs are available without any other side-information like attributes or a limited number of samples from UUCs.

Open set recognition [21] describes such a scenario where new classes (UUCs) unseen in training appear in testing, and requires the classifiers to not only accurately classify KKCs but also effectively deal with UUCs. Therefore, the classifiers need to have a corresponding reject option when a testing sample comes from some UUC. Fig. 2 gives a comparative demonstration of traditional 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 [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45]. 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.

In addition, the one-class classifier [46], [47], [48], [49], [50], [51], [52], [53] 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. Popular approaches for one-class classification include one-class SVM [46] and support vector data description (SVDD) [48], [54], 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. Note that treating multiple KKCs as a single one in the one-class setup obviously ignores the discriminative information among these KKCs, leading to poor performance [23], [55]. Even if each KKC is modeled by an individual one-class classifier as proposed in [37], the novelty detection performance is rather low [55]. Therefore, it is necessary to rebuild effective classifiers specifically for OSR problem, especially for multiclass OSR problem.

As a summary, Table 1 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 [56], [57], [58], [59], [60], [61]. In a study on evaluation methods for face recognition, Phillips *et al.* [56] proposed a typical framework for open set identity recognition, while Li and Wechsler [57] again viewed open set face recognition from an evaluation perspective and proposed Open Set Transduction Confidence Machine-k Nearest Neighbors (TCM-kNN) method. It is Scheirer *et al.* [21] that first formalized the open set recognition problem and proposed a preliminary solution—1-versus-Set machine, which incorporates an open space risk term in modeling to account for the space beyond the reasonable support of KKCs. Afterwards, open set recognition attracted widespread attention. Note that OSR has been mentioned in the recent survey on ZSL [10],



(a) Distribution of the original data set. (b) Traditional recognition/classification problem. (c) Open set recognition/classification problem.

Fig. 2. Comparison of between traditional classification and open set recognition. Fig. 2a denotes the distribution of original dataset including KKCs 1, 2, 3, 4, and UUCs ?5,?6, where KKCs appear during training and testing while UUCs may appear or not during testing. Fig. 2b shows the decision boundary of each class obtained by traditional classification methods, and it will obviously misclassify when UUCs ?5,?6 appear during testing. Fig. 2c describes open set recognition, where the decision boundaries limit the scope of KKCs 1, 2, 3, 4, reserving space for UUCs ?5,?6. Via these decision boundaries, the samples from some UUCs are labeled as “unknown” or rejected rather than misclassified as KKCs.

TABLE 1
Differences Between Open Set Recognition and its Related Tasks

TASK \ SETTING	TRAINING	TESTING	GOAL
Traditional Classification	Known known classes	Known known classes	Classifying known known classes
Classification with Reject Option	Known known classes	Known known classes	Classifying known known classes & rejecting samples of low confidence
One-class Classification (Anomaly Detection)	Known known classes & few or none outliers from KUCs	Known known classes & few or none outliers	Detecting outliers
One/Few-shot Learning	Known known classes & a limited number of UKCs' samples	Unknown known classes	Identifying unknown known classes
Generalized Few-shot Learning	Known known classes & a limited number of UKCs' samples	Known known classes & unknown known classes	Identifying known known classes & unknown known classes
Zero-shot Learning	Known known classes & side-information[1]	Unknown known classes	Identifying unknown known classes
Generalized Zero-shot Learning	Known known classes & side-information[1]	Known known classes & unknown known classes	Identifying known known classes & unknown known classes
Open Set Recognition	Known known classes	Known known classes & unknown unknown classes	Identifying known known classes & rejecting unknown unknown classes
Generalized Open Set Recognition	Known known classes & side-information[2]	Known known classes & Unknown unknown classes	Identifying known known classes & cognizing unknown unknown classes

Note that the unknown known classes in one-shot learning usually do not have any side-information, such as semantic information, etc. The $side-information^1$ *in ZSL and G-ZSL denotes the semantic information from both KKCs and UKCs, while the* $side-information^2$ *here denotes the available semantic information only from KKCs. As part of this information usually spans between KKCs and UUCs, we hope to use it to further 'cognize' UUCs instead of simply rejecting them.*

however, it has not been extensively discussed. Unlike [10], we here provide a comprehensive review regarding OSR.

The rest of this paper is organized as follows. In the next three sections, we first give the basic notation and related definitions (Section 2). Then we categorize the existing OSR technologies from the modeling perspective, and for each category, we review different approaches, given in Table 2 in detail (Section 3). Lastly, we review the open world recognition (OWR) which can be seen as a natural extension of OSR in Section 4. Furthermore, Section 5 reports the commonly used datasets, evaluation criteria, and algorithm comparisons, while Section 6 highlights the limitations of existing approaches and points out some promising research directions in this field. Finally, Section 7 gives a conclusion.

2 Basic Notation and Related Definition

This part briefly reviews the formalized OSR problem described in [21]. As discussed in [21], the space far from known data (including KKCs and KUCs) is usually considered as *open space* $\mathcal{O}$. So labeling any sample in this space as an arbitrary KKC inevitably incurs risk, which is called *open space risk* $R_{\mathcal{O}}$. As UUCs are agnostic in training, it is often difficult to quantitatively analyze open space risk. Alternatively, [21] gives a qualitative description for $R_{\mathcal{O}}$, where it is formalized as the relative measure of open space $\mathcal{O}$ compared to the overall measure space S_o

$$R_{\mathcal{O}}(f) = \frac{\int_{\mathcal{O}} f(x)dx}{\int_{S_o} f(x)dx}, \tag{1}$$

where f denotes the measurable recognition function. $f(x) = 1$ indicates that some class in KKCs is recognized, otherwise $f(x) = 0$. Under such a formalization, the more we label samples in open space as KKCs, the greater $R_{\mathcal{O}}$ is.

Further, the authors in [21] also formally introduced the concept of *openness* for a particular problem or data universe.

Definition 1 (The openness defined in [21]). *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*

$$O = 1 - \sqrt{\frac{2 \times |C_{\mathrm{TR}|}}{|C_{\mathrm{TA}}| + |C_{\mathrm{TE}}|}}, \tag{2}$$

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

Larger *openness* corresponds to more open problems, while the problem is completely closed when the *openness* equals 0. Note that [21] does not explicitly give the relationships among C_{TA}, C_{TR}, and C_{TE}. In most existing works [22], [67], [90], [91], the relationship, $C_{\mathrm{TA}} = C_{\mathrm{TR}} \subseteq C_{\mathrm{TE}}$, holds by default. Besides, the authors in [82] specifically give the following relationship: $C_{\mathrm{TA}} \subseteq C_{\mathrm{TR}} \subseteq C_{\mathrm{TE}}$, which

TABLE 2
Different Categories of Models for Open Set Recognition

Different Categories of OSR methods		Papers
Discriminative model	Traditional ML-based	[21], [22], [23], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74]
	Deep Neural Network-based	[25], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85]
Generative model	Instance Generation-based	[86], [87], [88], [89], [90]
	Non-Instance Generation-based	[91]

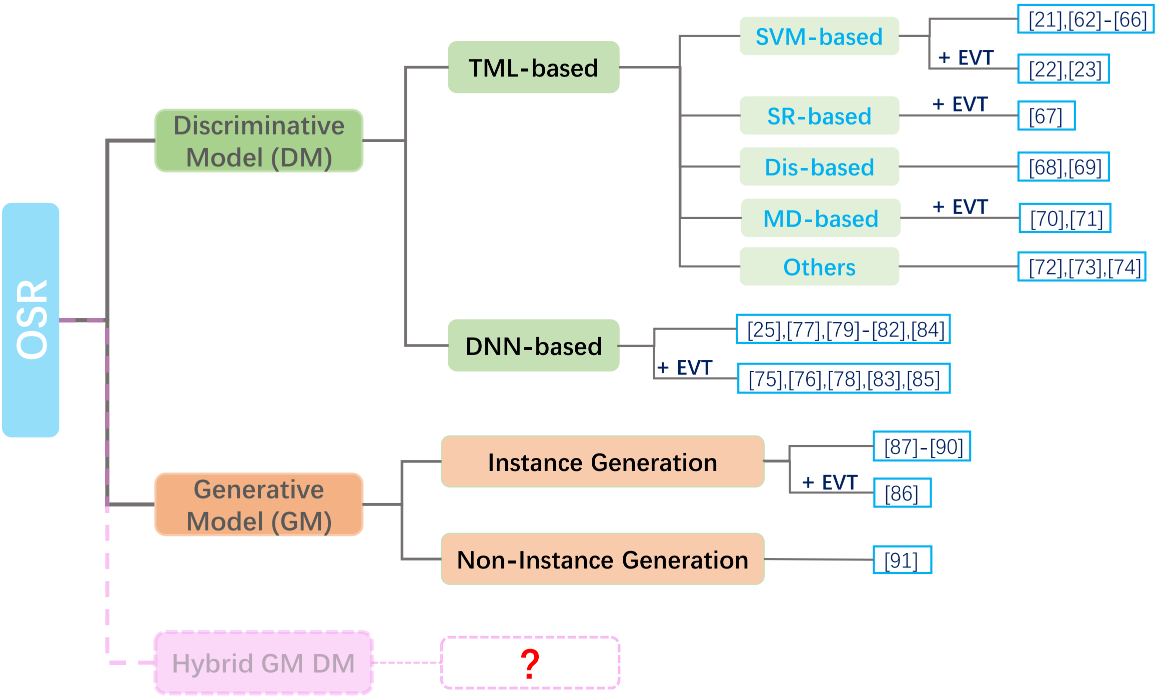


Fig. 3. A global picture on how the existing OSR methods are linked. 'SR-based', 'Dis-based', 'MD-based' respectively denote the Sparse Representation-based, Distance-based and Margin Distribution-based OSR methods, while '+EVT' represents the corresponding method additionally adopts the statistical Extreme Value Theory. Note that the pink dotted line module indicates that there is no related OSR work from the hybrid generative and discriminative model perspective at present, which seems also a promising research direction in the future.

contains the former case. However, such a relationship is problematic for Definition 1. Consider the following simple case: $C_{\mathrm{TA}} \subseteq C_{\mathrm{TR}} \subseteq C_{\mathrm{TE}}$, and $|C_{\mathrm{TA}}| = 3, |C_{\mathrm{TR}}| = 10, |C_{\mathrm{TE}}| = 15$. Then we will have $O < 0$, which is obviously unreasonable. In fact, 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 KKCs' knowledge from C_{TR} and the UUCs' knowledge from C_{TE} rather than C_{TA}, C_{TR}, and C_{TE} their three. Therefore, in this paper, we recalibrate the formula of *openness*

$$O^* = 1 - \sqrt{\frac{2 \times |C_{\mathrm{TR}}|}{|C_{\mathrm{TR}}| + |C_{\mathrm{TE}}|}}. \tag{3}$$

Compared to Eq. (2), Eq. (3) is just a relatively more reasonable form to estimate the *openness*. Other definitions can also capture this notion, and some may be more precise, thus worth further exploring. With the concepts of *open space risk* and *openness* in mind, the definition of OSR problem can be given as follows:

Definition 2 (The Open Set Recognition Problem [21]). *Let V be the training data, and let $R_{\mathcal{O}}$, R_{ε} 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

$$\arg\min_{f \in \mathcal{H}} \{R_{\mathcal{O}}(f) + \lambda_r R_{\varepsilon}(f(V))\}, \tag{4}$$

where λ_r is a regularization constant.

The open set risk denoted in formula (4) 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.

3 A Categorization of OSR Techniques

Although Scheirer *et al.* [21] formalized the OSR problem, an important question is how to incorporate Eq. (1) to modeling. There is an ongoing debate between the use of generative and discriminative models in statistical learning [92], [93], with arguments for the value of each. However, as discussed in [22], open set recognition introduces such a new issue, in which neither discriminative nor generative models can directly address UUCs 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 perspectives. Next, we mainly review the existing OSR models from these two perspectives.

According to the modeling forms, these models can be further categorized into four categories (Table 2): Traditional ML (TML)-based and Deep Neural Network (DNN)-based methods from the discriminative model perspective; Instance and Non-Instance Generation-based methods from the generative model perspective. For each category, we review different approaches by focusing on their corresponding representative works. Moreover, a global picture on how these approaches are linked is given in Fig. 3, while several available software packages' links are also listed

TABLE 3
Available Software Packages

Model	Language	Author	Link
1-vs-Set [21],W-SVM [22],P_I-SVM [23]	C/C++	Jain *et al.*	https://github.com/ljain2/libsvm-openset
BFHC [63], EPCC [66]	Matlab	Cevikalp *et al.*	http://mlcv.ogu.edu.tr/softwares.html
SROSR [67]	Matlab	Zhang *et al.*	https://github.com/hezhangsprinter/SROSR
NNO [68]	Matlab	Bendale *et al.*	https://github.com/abhijitbendale/OWR
HPLS, HFCN [73]	Python	Vareto *et al.*	https://github.com/rafaelvareto/HPLS-HFCN-openset
OpenMax [75]	Python	Bendale *et al.*	https://github.com/abhijitbendale/OSDN
DOC [79]	Python	Shu *et al.*	https://github.com/alexander-rakhlin/CNN-for-Sentence-Classification-in-Keras
RNA [25]	Python	Dhamija *et al.*	http://github.com/Vastlab/Reducing-Network-Agnostophobia
OSRCI [87]	Python	Neal *et al.*	https://github.com/lwneal/counterfactual-open-set
ASG [89]	Python	Liu *et al.*	https://github.com/eyounx/ASG
EVM [70]	Python	Rudd *et al.*	https://github.com/EMRResearch/ExtremeValueMachine

(Table 3) to facilitate subsequent research by relevant researchers. Next, we first give a review from the discriminative model perspective, where most existing OSR algorithms are modeled from this perspective.

3.1 Discriminative Model for OSR

3.1.1 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 does not hold any more in OSR. To adapt these methods to the OSR scenario, many efforts have been made [21], [22], [23], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74].

SVM-Based. The Support Vector Machine (SVM) [94] has been successfully used in traditional classification/recognition task. However, when UUCs appear during testing, its classification performance will decrease significantly since it usually divides over-occupied space for KKCs under the closed set assumption. As shown in Fig. 2b, once the UUCs' samples fall into the space divided for some KKCs, these samples will never be correctly classified. To overcome this problem, many SVM-based OSR methods have been proposed.

Using Definition 2, Scheirer *et al.* [21] proposed 1-versus-Set machine, which incorporates an open space risk term in modeling to account for the space beyond the reasonable support of KKCs. Concretely, they added another hyperplane parallelling the separating hyperplane obtained by SVM in score space, thus leading to a slab in feature space. The open space risk for a linear kernel slab model is defined as follows:

$$R_{\mathcal{O}} = \frac{\delta_\Omega - \delta_A}{\delta^+} + \frac{\delta^+}{\delta_\Omega - \delta_A} + p_A\omega_A + p_\Omega\omega_\Omega, \tag{5}$$

where δ_A and δ_Ω denote the marginal distances of the corresponding hyperplanes, and δ^+ is the separation needed to account for all positive data. Moreover, user-specified parameters p_A and p_Ω are given to weight the importance between the margin spaces ω_A and ω_Ω. In this case, a testing sample that appears between the two hyperplanes would be labeled as the appropriate class. Otherwise, it would be considered as non-target class or rejected, depending on which side of the slab it resides. Similar to 1-versus-Set machine, Cevikalp [62], [63] 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 directly formed a slab in feature space. In addition, BFHC can be extended to nonlinear case by using kernel trick, and we refer reader to [63] for more details.

Although the slab models mentioned above decrease the KKC's region for each binary SVM, the space occupied by each KKC remains unbounded. Thus the open space risk still exists. To overcome this challenge, researchers further seek new ways to control this risk [22], [23], [64], [65], [66].

Scheirer *et al.* [22] incorporated non-linear kernels into a solution that further limited open space risk by positively labeling only sets with finite measure. 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 Weibull-calibrated SVM (W-SVM) model was proposed, which combined the statistical extreme value theory (EVT) [95][3] for score calibration with two separated SVMs. The first SVM is a one-class SVM CAP model used as a conditioner: if the posterior estimate $P_O(y|x)$ of an input sample x predicted by one-class SVM is less than a threshold δ_τ, the sample will be rejected outright. Otherwise, it will be passed to the second SVM. The second one 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 KKC. Furthermore, it also obtains the posterior estimate $P_\psi(y|x)$ for the corresponding negative KKCs by a reverse Weibull fitting. Defined an indicator variable: $\iota_y = 1$ if $P_O(y|x) > \delta_\tau$ and $\iota_y = 0$ otherwise, the W-SVM recognition for all KKCs $\mathcal{Y}$ is

$$\begin{aligned} y^* = &\ \arg\max_{y\in\mathcal{Y}} P_\eta(y|x) \times P_\psi(y|x) \times \iota_y \\ &\text{subject to}\ \ P_\eta(y^*|x) \times P_\psi(y^*|x) \geq \delta_R, \end{aligned} \tag{6}$$

3. Extreme Value Theory (EVT) [95], also known as Fisher-Tippett Theorem, is a branch of statistics analyzing the distribution of data of abnormally high or low values. The application of EVT in visual tasks mainly involves post-recognition score analysis. For open set recognition, the threshold to reject/accept usually lies in the overlap region of extremes of match and non-match score distributions [67], [85]. As EVT can effectively model the tail of the match and non-match recognition scores as one of the extreme value distributions. This makes it widely used in existing OSR models. The following is the definition of EVT:

Extreme Value Theory. Let $(v_1, v_2, \ldots)$ *be a sequence of i.i.d samples. Let* $\zeta_n = \max\{v_1, \ldots, v_n\}$. *If a sequence of pairs of real numbers* (a_n, b_n) *exists such that each* $a_n > 0$ *and* $\lim_{z\to\infty} P(\frac{\zeta_n - b_n}{a_n} \leq z) = F(z)$ *then if* F *is a non-degenerate distribution function, it belongs to the Gumbel, the Fréchet or the Reversed Weibull family.*

where δ_R is the threshold of the second SVM CAP model. The thresholds δ_τ and δ_R are set empirically, e.g., δ_τ is fixed to 0.001 as specified by the authors, while δ_R is recommended to set according to the openness of the specific problem

$$\delta_R = 0.5 \times \text{openness}. \tag{7}$$

Besides, W-SVM was further used for open set intrusion recognition on the KDDCUP'99 dataset [96]. More works on intrusion detection in open set scenario can be found in [97]. Intuitively, we can reject a large set of UUCs (even under an assumption of incomplete class knowledge) if the positive data for any KKCs is accurately modeled without overfitting. Based on this intuition, Jain *et al.* [23] invoked EVT to model the positive training samples at the decision boundary and proposed the P_I-SVM algorithm. 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, their thresholds' selection also gives some caveats. First, they are assumed that all KKCs 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 [22]. However, the openness of the corresponding problem is usually unknown as well.

To address these caveats, Scherreik *et al.* [64] introduced the probabilistic open set SVM (POS-SVM) classifier which could empirically determine unique reject threshold for each KKC under Definition 2. Instead of defining $R_\mathcal{O}$ as relative measure of open and class-defined space, POS-SVM chooses probabilistic representations respectively for open space risk $R_\mathcal{O}$ and empirical risk R_ε (details c.f. [64]). Moreover, the authors also adopted a new OSR evalution metric called Youden's index which combines the true negative rate and recall, and will be detailed in Section 5.2. Recently, to address sliding window visual object detection and open set recognition tasks, Cevikalp and Triggs [65], [66] used a family of quasi-linear "polyhedral conic" functions of [98] to define the acceptance regions for positive KKCs. This choice provides a convenient family of compact and convex region shapes for discriminating relatively well localized positive KKCs from broader negative ones including negative KKCs and UUCs.

Sparse Representation-Based. In recent years, the sparse representation-based techniques have been widely used in computer vision and image processing fields [99], [100], [101]. In particular, sparse representation-based classifier (SRC) [102] has gained a lot of attentions, which identifies the correct class by seeking the sparsest representation of the testing sample in terms of the training. SRC and its variants are essentially still under the closed set assumption, so in order to adapt SRC to an open environment, Zhang and Patel [67] presented the sparse representation-based open set recognition model, briefly called SROSR.

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. This model consists of two main stages. One stage reduces the OSR problem into hypothesis testing problems by modeling the tails of error distributions using EVT, and the other first 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 [67], although 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 SROSR is currently proposed based on sparse representation, it is still an interesting topic for future work to develop the sparse representation-based OSR algorithms.

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 Boult [68] established a Nearest Non-Outlier (NNO) algorithm for open set recognition by extending upon the Nearest Class Mean (NCM) classifier [103], [104]. NNO carries out classification based on the distance between the testing sample and the means of KKCs, where it rejects an input sample when all classifiers reject it. What needs to be emphasized is that this algorithm can dynamically add new classes based on manually labeled data. In addition, the authors introduced the concept of open world recognition, which details in Section 4.

Further, based on the traditional Nearest Neighbor classifier, Júnior *et al.* [69] introduced an open set version of Nearest Neighbor classifier (OSNN) to deal with the OSR problem. Different from those works which directly use a threshold on the similarity score for the most similar class, 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

$$\text{Ratio} = d(s,t)/d(s,u), \tag{8}$$

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 UUC.

OSNN is inherently multiclass, meaning that its efficiency will not be affected as the number of available classes for training increases. Moreover, the NNDR technique can also be applied effortlessly to other classifiers based on the similarity score, e.g., the Optimum-Path Forest (OPF) classifier [105]. Other metrics could be used to replace euclidean metric as well, and even the feature space considered could be a transformed one, as suggested by the authors. Note that one limitation of OSNN is that just selecting two reference samples coming from different classes for comparison makes OSNN vulnerable to outliers [91].

Margin Distribution-Based. Considering that most existing OSR methods take little to no distribution information of the data into account and lack a strong theoretical foundation, Rudd *et al.* [70] formulated a theoretically sound classifier—the Extreme Value Machine (EVM) which stems from the concept of *margin distributions*. Various definitions and uses of margin distributions have been explored [106], [107], [108], [109], involving techniques such as maximizing the mean or median margin, taking a weighted combination margin, or optimizing the margin mean and variance. Utilizing the marginal distribution itself can provide better error bounds than those offered by a soft-margin SVM, which translates into reduced experimental error in some cases.

As an extension of margin distribution theory from a per-class formulation [106], [107], [108], [109] to a sample-wise formulation, EVM is modeled in terms of the distribution of sample half-distances relative to a reference point. Specifically, it obtains the following theorem:

Theorem 1. *Assume we are given a positive sample x_i and sufficiently many negative samples x_j drawn from well-defined class distributions, yielding pairwise margin estimates m_{ij}. Assume a continuous non-degenerate margin distribution exists. Then the distribution for the minimal values of the margin distance for x_i is given by a Weibull distribution.*

As Theorem 1 holds for any point x_i, each point can estimate its own distribution of distance to the margin, thus yielding:

Corollary 1 (Ψ Density Function). *Given the conditions for the Theorem 1, the probability that x' is included in the boundary estimated by x_i is given by*

$$\Psi(x_i, x', \kappa_i, \lambda_i) = \exp^{-\left(\frac{\|x_i - x'\|}{\lambda_i}\right)^{\kappa_i}}, \tag{9}$$

where $\|x_i - x'\|$ is the distance of x' from sample x_i, and κ_i, λ_i are Weibull shape and scale parameters respectively obtained from fitting to the smallest m_{ij}.

Prediction. Once EVM is trained, the probability of a new sample x' associated with class $\mathcal{C}_l$, i.e., $\hat{P}(\mathcal{C}_l|x')$, can be obtained by Eq. (9), thus resulting in the following decision function

$$y^* = \begin{cases} \arg\max_{l\in\{1,\ldots,M\}} \hat{P}(\mathcal{C}_l|x') & \text{if} \hat{P}(\mathcal{C}_l|x') \geq \delta \\ \text{“unknown”} & \text{Otherwise}, \end{cases} \tag{10}$$

where M denotes the number of KKCs in training, and δ represents the probability threshold which defines the boundary between the KKCs and unsupported open space.

Derived from the margin distribution and extreme value theories, 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 [110] and the intrusion detection [111]. Note that it also has some limitations as reported in [71], in which an obvious one is that the use of geometry of KKCs is risky when the geometries of KKCs and UUCs differ. To address these limitations, Vignotto and Engelke [71] further presented the GPD and GEV classifiers relying on approximations from EVT.

Other Traditional ML Methods-Based. Using center-based similarity (CBS) space learning, Fei and Liu [72] proposed a novel solution for text classification under OSR scenario, while Vareto *et al.* [73] 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.* [74] adopted the integrated idea, where different classifiers and features are combined to solve the OSR problem. We refer the reader to [72], [73], [74] for more details. As most traditional machine learning methods for classification currently are under closed set assumption, it is appealing to adapt them to the open and non-stationary environment.

3.1.2 Deep Neural Network-Based OSR Models

Thanks to the powerful learning representation ability, Deep Neural Networks (DNNs) have gained significant benefits for various tasks such as visual recognition, Natural language processing, text classification, etc. DNNs usually follow a typical SoftMax cross-entropy classification loss, which inevitably incurs the normalization problem, making them inherently have the closed set nature. As a consequence, DNNs often make wrong predictions, and even do so too confidently, when processing the UUCs' samples. The works in [112], [113] have indicated that 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 looked at different approaches [25], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85].

Replacing the SoftMax layer in DNNs with an OpenMax layer, Bendale and Boult [75] 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 [103], [104], 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 fitting score, and then used to compute a pseudo-activation for UUCs. Finally, the class probabilities of KKCs and (pseudo) UUCs are computed by using SoftMax again on these new redistributed activation vectors.

As discussed in [75], OpenMax effectively addressed the recognition challenge for fooling/rubbish and unrelated open set images, but it 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 [113], [114]. Rozsa *et al.* [76] also analyzed and compared the adversarial robustness of DNNs using SoftMax layer with OpenMax: although OpenMax provides less vulnerable systems than SoftMax to traditional attacks, it is equally susceptible to more sophisticated 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 [77]. To address this limitation, Hassen and Chan [77] learned a neural network based representation for open set recognition. In this representation, samples from the same class are closed to each other while the ones from different classes are further apart, leading to larger space among KKCs for UUCs' samples to occupy.

Besides, Prakhya *et al.* [78] continued to follow the technical line of OpenMax to explore the open set text classification, while Shu *et al.* [79] replaced the SoftMax layer with a 1-versus-rest final layer of sigmoids and presented Deep Open classifier (DOC) model. Kardan and Stanley [80] proposed the competitive overcomplete output layer (COOL) neural network to circumvent the overgeneralization of neural networks over regions far from the training data. Based on an elaborate distance-like computation provided by a weightless neural network, Cardoso *et al.* [81] proposed the tWiSARD algorithm for open set recognition, which is further developed in [82]. Recently, considering the available background classes (KUCs), Dhamija *et al.* [25] combined SoftMax with the novel Entropic Open-Set and Objectosphere losses to address the OSR problem. Yoshihashi *et al.* [83] presented the Classification-Reconstruction learning algorithm for open set recognition (CROSR), which utilizes latent representations for reconstruction and enables robust UUCs' detection without harming the KKCs' classification accuracy. Using class conditioned auto-encoders with novel training and testing methodology, Oza and Patel [85] proposed C2AE model for OSR. Compared to the works described above, Shu *et al.* [84] paid more attention to discovering the hidden UUCs in the reject samples. Correspondingly, they proposed a joint open classification model with a sub-model for classifying whether a pair of examples belong to the same class or not, where the sub-model can serve as a distance function for clustering to discover the hidden classes in the reject samples.

Remark. From the discriminative model perspective, almost all existing OSR approaches adopt the threshold-based classification scheme, where recognizers in decision either reject or categorize the input samples to some KKC using empirically-set threshold. Thus the threshold plays a key role. However, at the moment, the selection for it usually depends on the knowledge from KKCs, which inevitably incurs risks due to lacking available information from UUCs [91]. In fact, as the KUCs' data is often available at hand [25], [115], [116], we can fully leverage them to reduce such a risk and further improve the robustness of these methods for UUCs. Besides, effectively modeling the tails of the data distribution makes EVT widely used in existing OSR methods. However, regrettably, it provides no principled means of selecting the size of tail for fitting. Further, as the object frequencies in visual categories ordinarily follow long-tailed distribution [29], [117], such a distribution fitting will face challenges once the rare classes in KKCs and UUCs appear together in testing [118].

3.2 Generative Model for Open Set Recognition

In this section, we will review the OSR methods from the generative model perspective, where these methods can be further categorized into Instance Generation-based and Non-Instance Generation-based methods according to their modeling forms.

3.2.1 Instance Generation-Based OSR Models

The adversarial learning (AL) [119] 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 UUCs generated by the AL technique [86], [87], [88], [89], [90].

Using a conditional generative adversarial network (GAN) to synthesize mixtures of UUCs, Ge *et al.* [86] proposed the Generative OpenMax (G-OpenMax) algorithm, which can provide explicit probability estimation over the generated UUCs, enabling the classifier to locate the decision margin according to the knowledge of both KKCs and generated UUCs. Obviously, such UUCs in their setting are only limited in a subspace of the original KKCs' space. Moreover, as reported in [86], although G-OpenMax effectively detects UUCs in monochrome digit datasets, it has no significant performance improvement on natural images.

Different from G-OpenMax, Neal *et al.* [87] introduced a novel dataset augmentation technique, called counterfactual image generation (OSRCI). OSRCI adopts an encoder-decoder GAN architecture to generate the synthetic open set examples which are close to KKCs, yet do not belong to any KKCs. They further reformulated the OSR problem as classification with one additional class containing those newly generated samples. Similar in spirit to [87], Jo *et al.* [88] adopted the GAN technique to generate fake data as the UUCs' data to further enhance the robustness of the classifiers for UUCs. Yu *et al.* [89] proposed the adversarial sample generation (ASG) framework for OSR. ASG can be applied to various learning models besides neural networks, while it can generate not only UUCs' data but also KKCs' data if necessary. In addition, Yang *et al.* [90] 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 UUC. Then they explored the open set human activity recognition based on micro-Doppler signatures.

Remark. As most Instance Generation-based OSR methods often rely on deep neural networks, they also seem to fall into the category of DNN-based methods. But please note that the essential difference between these two categories of methods lies in whether the UUCs' samples are generated or not in learning. In addition, the AL technique does not just rely on deep neural networks, such as ASG [89].

3.2.2 Non-Instance Generation-Based OSR Models

Dirichlet process (DP) [120], [121], [122], [123], [124] 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. This model does not overly depend on training samples and can achieve adaptive change as the data changes, making it naturally adapt to the OSR scenario.

With slight modification to hierarchical Dirichlet process (HDP), Geng and Chen [91] adapted HDP to OSR and proposed the collective decision-based OSR model (CD-OSR), which can address both batch and individual samples. CD-OSR first performs a co-clustering process to obtain the appropriate parameters in the training phase. In testing phase, it models each KKC's data as a group of CD-OSR using a Gaussian mixture model (GMM) with an unknown number of components/subclasses, while the whole testing set as one collective/batch is treated in the same way. Then all of the groups are co-clustered under the HDP framework. After co-clustering, one can obtain one or more subclasses representing the corresponding class. Thus, for a testing sample, it would be labeled as the appropriate KKC or UUC, depending on whether the subclass it is assigned associates with the corresponding KKC or not.

Notably, unlike the previous OSR methods, CD-OSR does not need to define the thresholds using to determine the decision boundary between KKCs and UUCs. In contrast, it introduced some threshold using to control the number of subclasses in the corresponding class, and the selection of such a threshold has been experimentally indicated more generality (details c.f. [91]). Furthermore, CD-OSR can provide explicit modeling for the UUCs appearing in testing, naturally resulting in a new class discovery function. Please note that such a new discovery is just at subclass level. Moreover, adopting the collective/batch decision strategy makes CD-OSR consider the correlations among the testing samples obviously ignored by other existing methods. Besides, as reported in [91], 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.

Remark. The key to Instance Generation-based OSR models is generating effective UUCs' samples. Though these existing methods have achieved some results, generating more effective UUCs' samples still need further study. Furthermore, the data adaptive property makes (hierarchical) Dirichlet process naturally suitable for dealing with the OSR task. Since only [91] currently gave a preliminary exploration using HDP, thus this study line is also worth further exploring. Besides, the collective decision strategy for OSR is a promising direction as well, 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[4] adopted by other existing OSR methods cannot do such a work since it cannot directly tell whether the single rejected sample is an outlier or from new class.

4 Beyond Open Set Recognition

Please note that the existing open set recognition is indeed in an open scenario but not incremental and does not scale gracefully with the number of classes. On the other hand, though new classes (UUCs) are assumed to appear incremental in class incremental learning (C-IL) [125], [126], [127], [128], [129], [130], these studies mainly focused on how to enable the system to incorporate later coming training samples from new classes instead of handling the problem of recognizing UUCs. To jointly consider the OSR and CIL tasks, Bendale and Boult [68] expanded the existing open set recognition (Definition 2) to the open world recognition (OWR), where a recognition system should perform four tasks: detecting UUCs, 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 3 (Open World Recognition [68]). *Let $\mathcal{K}_T \in \mathbb{N}^+$ be the set of labels of KKCs at time T, and let the zero label (0) be reserved for (temporarily) labeling data as unknown. Thus $\mathbb{N}$ includes the labels of KKCs and UUCs. Based on the Definition 2, 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 $f_i(x) \in \mathcal{H} : \mathbb{R}^d \mapsto R$ for $i \in \mathcal{K}_T$ to be open set functions that manage open space risk as Eq. (1). The novelty detector $\nu(\varphi) : \mathbb{R}^i \mapsto [0, 1]$ determines if results from vector of recognition functions is from an UUC.*
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 $y_j = \mathcal{L}(x_j), \vee x_j \in U_T$. Assume the labeling finds m new classes, then the set of KKCs becomes $\mathcal{K}_{T+1} = \mathcal{K}_T \cup \{i+1, \ldots, i+m\}$.*
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 $f_{i+1}(x) \ldots f_{i+m}(x)$, each of which manages open space risk, to the vector φ of measurable recognition functions.*

For more details, we refer the reader to [68]. Ideally, all of these steps should be automated. However, [68] only presumed supervised learning with labels obtained by human labeling at present, and proposed the NNO algorithm which has been discussed in Section 3.1.1.

Afterward, some researchers continued to follow up this research route. Rosa *et al.* [131] 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 UUCs, and (c) the use of local learning to precisely describe the space of classes. Towards these goals, they extended three existing metric learning methods using online metric learning. Doan and Kalita [132] presented the Nearest Centroid Class (NCC) model, which is similar to the online NNO [131] 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. Lonij *et al.* [133] 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.* [134] proposed the Open Deep Network (ODN)

4. 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.

which first detects new classes by applying a multiclass triplet thresholding method, and then dynamically reconstructs the classification layer by adding predictors for new classes continually. Besides, EVM discussed in Section 3.1.1 also adapts to the OWR scenario due to the nature of incremental learning [70]. Recently, Xu *et al.* [135] proposed a meta-learning method to learn to accept new classes without training under the open world recognition framework.

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

5 Datasets, Evaluation Criteria, and Experiments

5.1 Datasets

In open set recognition, most existing experiments are usually carried out on a variety of recast multi-class benchmark datasets at present, where some distinct labels in the corresponding dataset are randomly chosen as KKCs while the remaining ones as UUCs. Here we list some commonly used benchmark datasets and their combinations:

LETTER [136]: has a total of 20,000 samples from 26 classes, where each class has around 769 samples with 16 features. To recast it for open set recognition, 10 distinct classes are randomly chosen as KKCs for training, while the remaining ones as UUCs.

PENDIGITS [137]: has a total of 10,992 samples from 10 classes, where each class has around 1,099 samples with 16 features. Similarly, 5 distinct classes are randomly chosen as KKCs and the remaining ones as UUCs.

COIL20 [138]: has a total of 1,440 gray images from 20 objects (72 images each object). Each image is down-sampled to 16×16, i.e., the feature dimension is 256. Following [91], we further reduce the dimension to 55 by principal component analysis (PCA) technique, remaining 95 percent of the samples' information. 10 distinct objects are randomly chosen as KKCs, while the remaining ones as UUCs.

YALEB [139]: The Extended Yale B (YALEB) dataset has a total of 2,414 frontal-face images from 38 individuals. Each individuals has around 64 images. The images are cropped and normalized to 32×32. Following [91], we also reduce their feature dimension to 69 using PCA. Similar to COIL20, 10 distinct classes are randomly chosen as KKCs, while the remaining ones as UUCs.

MNIST [140]: consists of 10 digit classes, where each class contains between 6,313 and 7,877 monochrome images with 28×28 feature dimension. Following [87], 6 distinct classes are randomly chosen as KKCs, while the remaining 4 classes as UUCs.

SVHN [141]: has ten digit classes, each containing between 9,981 and 11,379 color images with 32×32 feature dimension. Following [87], 6 distinct classes are randomly chosen as KKCs, while the remaining 4 classes as UUCs.

CIFAR10 [142]: has a total of 6,000 color images from 10 natural image classes. Each image has 32×32 feature dimension. Following [87], 6 distinct classes are randomly chosen as KKCs, while the remaining 4 classes as UUCs. To extend this dataset to larger openness, [87] further proposed the *CIFAR+10*, *CIFAR+50* datasets, which use 4 non-animal classes in CIFAR10 as KKCs, while 10 and 50 animal classes are respectively chosen from CIFAR100[5] as UUCs.

Tiny-Imagenet [143]: has a total of 200 classes with 500 images each class for training and 50 for testing, which is drawn from the Imagenet ILSVRC 2012 dataset [144] and down-sampled to 32×32. Following [87], 20 distinct classes are randomly chosen as KKCs, while the remaining 180 classes as UUCs.

5.2 Evaluation Criteria

In this subsection, we summarize some commonly used evaluation metrics for open set recognition. For evaluating classifiers in the OSR scenario, a critical factor is taking the recognition of UUCs into account. Let TP_i, TN_i, FP_i, and FN_i respectively denote the true positive, true negative, false positive, and false negative for the ith KKC, where $i \in \{1, 2, \ldots, C\}$ and C denotes the number of KKCs. Further, let TU and FU respectively denote the correct and false reject for UUCs. Then we can obtain the following evaluation metrics.

5.2.1 Accuracy for OSR

As a common choice for evaluating classifiers under closed set assumption, the accuracy $\mathcal{A}$ is usually defined as

$$\mathcal{A} = \frac{\sum_{i=1}^{C}(TP_i + TN_i)}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i)}.$$

A trivial extension of accuracy to the OSR scenario $\mathcal{A}_{\mathrm{O}}$ is that the correct response should contain the correct classification for KKCs and correct reject for UUCs

$$\mathcal{A}_{\mathrm{O}} = \frac{\sum_{i=1}^{C}(TP_i + TN_i) + TU}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i) + (TU + FU)}. \tag{11}$$

However, as $\mathcal{A}_{\mathrm{O}}$ denotes the sum of the correct classification for KKCs and the correct reject for UUCs, it can not objectively evaluate the OSR models. Consider the following case: when the reject performance plays the leading role, and the testing set contains large number of UUCs' samples while only a few samples for KKCs, $\mathcal{A}_{\mathrm{O}}$ can still achieve a high value, even though the fact is that the recognizer's classification performance for KKCs is really low, and vice versa. Besides, [69] also gave a new accuracy metric for OSR called *normalized accuracy* (NA), which weights the accuracy for KKCs (AKS) and the accuracy for UUCs (AUS)

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

where

$$\mathrm{AKS} = \frac{\sum_{i=1}^{C}(TP_i + TN_i)}{\sum_{i=1}^{C}(TP_i + TN_i + FP_i + FN_i)}, \ \mathrm{AUS} = \frac{TU}{TU + FU},$$

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

5. http://www.cs.toronto.edu/~kriz/cifar.html

5.2.2 F-Measure for OSR

The F-measure F, widely applied in information retrieval and machine learning, is defined as a harmonic mean of precision P and recall R

$$F = 2 \times \frac{P \times R}{P + R}. \tag{13}$$

Please note that when using F-measure for evaluating OSR classifiers, one should not consider all the UUCs appearing in testing as one additional simple class, and obtain F in the same way as the multiclass closed set scenario. Because once performing such an operation, the correct classifications of UUCs' samples would be considered as true positive classifications. However, such true positive classification makes no sense, since we have no representative samples of UUCs to train the corresponding classifier. By modifying the computations of Precision and Recall only for KKCs, [69] gave a relatively reasonable F-measure for OSR. The following equations detail these modifications, where Eqs. (14) and (15) are respectively used to compute the macro-F-measure and the micro-F-measure by Eq. (13)

$$P_{ma} = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i + FP_i}, R_{ma} = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i + FN_i} \tag{14}$$

$$P_{mi} = \frac{\sum_{i=1}^{C} TP_i}{\sum_{i=1}^{C}(TP_i + FP_i)}, R_{mi} = \frac{\sum_{i=1}^{C} TP_i}{\sum_{i=1}^{C}(TP_i + FN_i)}. \tag{15}$$

Note that although the precision and recall only consider the KKCs in Eqs. (14) and (15), the FN_i and FP_i also consider the false UUCs and false KKCs by taking the false negative and the false positive into account (details c.f. [69]).

5.2.3 Youden's Index for OSR

As the F-measure is invariant to changes in TN [145], an important factor in OSR performance, Scherreik and Rigling [64] turned to Youden's index J defined as follows:

$$J = R + S - 1, \tag{16}$$

where $S = TN/(TN + FP)$ represents the true negative rate [146]. Youden's index can express an algorithm's ability to avoid failure [147], 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, with the aim to overcoming the effects on the sensitivity of model parameters and thresholds, [87] adopted the area under ROC curve (AUROC) together with the closed set accuracy as the evaluation metric, which views the OSR task as a combination of novelty detection and multiclass recognition. Note that although AUROC does a good job for evaluating the models, for the OSR problem, we eventually need to make a decision (a sample belongs to which KKC or UUC), thus such thresholds seem to have to be determined.

Remark. Currently, F-measure and AUROC is the most commonly used evaluation metrics. As the OSR problem faces a new scenario, the new evaluation methods are worth further exploring.

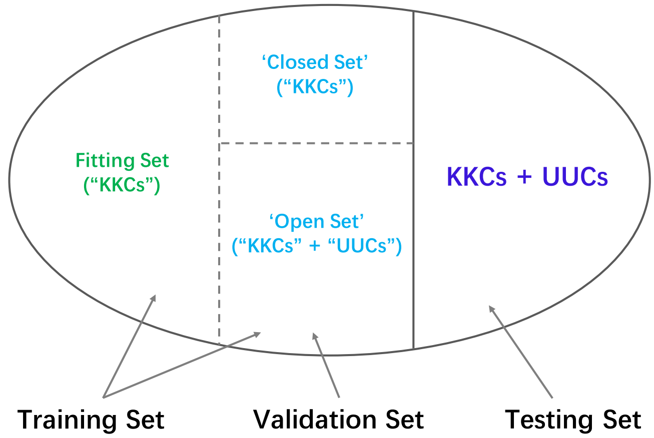


Fig. 4. Data Split. The dataset is first divided into training and testing set, then the training set is further divided into a fitting set and a validation set containing a 'closed set' simulation and an 'open set' simulation.

5.3 Experiments

This subsection quantitatively assesses a number of representative OSR methods on the popular benchmark datasets mentioned in Section 5.1. Further, these methods are compared in terms of the classification of non-depth and depth features.

5.3.1 OSR Methods Using Non-Depth Feature

The OSR methods using non-depth feature are usually evaluated on LETTER, PENDIGITS, COIL20, YALEB datasets. Most of them adopt the threshold-based strategy, where the thresholds are recommended to set according to the openness of the specific problem [22], [23], [67]. However, we usually have no prior knowledge about UUCs in the OSR scenario. Thus such a setting seems unreasonable, which is recalibrated in this paper, i.e., the decision thresholds are determined only based on the KKCs in training, and once they are determined during training, their values will no longer change in testing. To effectively determine the thresholds and parameters of the corresponding models, we introduce an evaluation protocol referring to [69], [91], as follows.

Evaluation Protocol. As shown in Fig. 4, the dataset is first divided into training set owning KKCs and testing set containing KKCs and UUCs, respectively. 2/3 of the KKCs occurring in training set are chosen as the "KKCs" simulation, while the remaining as the "UUCs" simulation. Thus the training set is divided into a fitting set $\mathcal{F}$ just containing 'KKCs' and a validation set $\mathcal{V}$ including a 'Closed-Set' simulation and an 'Open-Set' simulation. The 'Closed-Set' simulation only owns KKCs, while the 'Open-Set' simulation contains "KKCs" and "UUCs". Note that in the training phase, all of the methods are trained with $\mathcal{F}$ and evaluated on $\mathcal{V}$. Specifically, for each experiment, we

1. randomly select Ω distinct classes as KKCs for training from the corresponding dataset;
2. randomly choose 60 percent of the samples in each KKC as training set;

TABLE 4
Comparison Among the Representative OSR Methods Using Non-Depth Features

Dataset / Method		1-vs-Set	W-OSVM	W-SVM	P_I-SVM	SROSR	OSNN	EVM	CD-OSR
LETTER	$O^* = 0\%$	81.51 ± 3.94	95.64 ± 0.37	95.64 ± 0.25	96.92 ± 0.36	84.21 ± 2.49	83.12 ± 17.41	96.59 ± 0.50	**96.94 ± 1.36**
	$O^* = 15.48\%$	55.43 ± 3.18	83.83 ± 2.85	91.24 ± 1.48	90.89 ± 1.80	74.36 ± 5.10	73.20 ± 15.21	89.81 ± 0.40	**91.51 ± 1.58**
	$O^* = 25.46\%$	42.08 ± 2.63	73.37 ± 1.67	85.72 ± 0.85	84.16 ± 1.01	66.50 ± 8.22	64.97 ± 13.75	82.81 ± 2.42	**86.21 ± 1.46**
PENDIGITS	$O^* = 0\%$	97.17 ± 0.58	94.84 ± 1.46	98.82 ± 0.26	**99.21 ± 0.29**	97.43 ± 0.93	98.55 ± 0.71	98.42 ± 0.73	99.16 ± 0.25
	$O^* = 8.71\%$	78.43 ± 1.93	87.22 ± 1.71	93.05 ± 1.85	92.38 ± 2.68	96.33 ± 1.59	95.55 ± 1.30	96.97 ± 1.37	**98.75 ± 0.65**
	$O^* = 18.35\%$	61.29 ± 2.52	78.55 ± 4.91	88.39 ± 3.14	87.60 ± 4.78	93.53 ± 3.26	90.11 ± 4.15	92.88 ± 2.79	**98.43 ± 0.73**
COIL20	$O^* = 0\%$	89.59 ± 1.81	93.94 ± 1.87	86.83 ± 1.82	89.30 ± 1.45	97.12 ± 0.60	79.61 ± 7.41	97.68 ± 0.88	**97.71 ± 0.94**
	$O^* = 10.56\%$	70.21 ± 1.67	90.82 ± 2.31	85.64 ± 2.47	87.68 ± 2.02	96.68 ± 0.32	73.01 ± 6.18	95.69 ± 1.46	**97.32 ± 1.50**
	$O^* = 18.35\%$	57.72 ± 1.50	87.97 ± 5.40	84.54 ± 3.79	86.22 ± 3.34	**96.45 ± 0.66**	66.18 ± 4.49	93.62 ± 3.33	95.12 ± 2.14
YALEB	$O^* = 0\%$	87.99 ± 2.42	82.60 ± 3.54	86.01 ± 2.42	**93.47 ± 2.74**	88.09 ± 3.41	81.81 ± 8.40	68.94 ± 6.47	89.75 ± 1.15
	$O^* = 23.30\%$	49.36 ± 1.96	63.43 ± 5.33	84.56 ± 2.19	**88.96 ± 1.16**	83.99 ± 4.19	72.90 ± 9.41	54.40 ± 5.77	88.00 ± 2.19
	$O^* = 35.45\%$	34.37 ± 1.44	55.40 ± 5.26	83.44 ± 2.02	**86.63 ± 0.60**	81.38 ± 5.26	67.24 ± 7.29	46.64 ± 5.40	85.56 ± 1.07

The results report the averaged micro-F-measure (%) over 5 random class partitions. Best and the second best performing methods are highlighted in bold and underline, respectively. O^* calculated from Eq. (3) denotes the openness of the corresponding dataset.

3. select the remaining 40 percent of the samples from step 2 and the samples from other classes excluding the Ω KKCs as testing set;
4. randomly select $[(\frac{2}{3}\Omega + 0.5)]$ classes as "KKCs" for fitting from the training set, while the remaining classes as "UUCs" for validation;
5. randomly choose 60 percent of the samples from each "KKC" as fitting set $\mathcal{F}$;
6. select the remaining 40 percent of the samples from step 5 as the 'Closed Set' simulation, while the remaining 40 percent of the samples from step 5 and the ones from "UUCs" as the 'Open-Set' simulation;
7. train the models with $\mathcal{F}$ and verify them on $\mathcal{V}$, then find the suitable model parameters and thresholds;
8. evaluate the models with 5 random class partitions using micro-F-measure.

Note that the experimental protocol here is just a relatively reasonable form to evaluate the OSR methods. In fact, other protocols can also use for evaluating, and some may be more suitable, thus worth further exploring. Furthermore, since different papers often adopted different evaluation protocol before, to the best of our ability we here try to follow the parameter tuning principles in their papers. In addition, to encourage reproducible research, we refer the reader to our github[6] for the details about the datasets and their corresponding class partitions.

Under different openness O^*, Table 4 reports the comparisons among these methods, where 1-versus-Set [21], W-SVM (W-OSVM)[7] [22], P_I-SVM [23], SROSR [67], OSNN [69], EVM [70] from Traditional ML-based category and CD-OSR [91] from the Non-Instance Generation-based category.

Our First Observation is. With the increase of openness, although threshold-based methods (like W-SVM, P_I-SVM, SROSR, EVM) perform well on some datasets, there are also cases of significant performance degradation on other datasets (e.g., W-SVM performs well on LETTER, whereas underperformed on PENDIGITS significantly). This is mainly due to the fact that their decision thresholds are selected only based on the knowledge of KKCs, where once the UUCs' samples fall into the space divided for some KKCs, the OSR risk will be incurred. In contrast, due to the data adaptation characteristic of HDP, CD-OSR can effectively model the UUCs appearing in testing, making it currently achieve better performance on most datasets, especially for LETTER and PENDIGITS.

Our Second Observation is. Compared with the other methods, the performance of OSNN fluctuates greatly in terms of the standard deviation, especially for LETTER, which is likely because the NNDR strategy makes its performance heavily dependent on the distribution characteristics of the corresponding datasets. Furthermore, as the open space in 1-versus-Set is still unbounded, we can see that its performance drops sharply with the increase of openness. As a benchmark of one-class classifier, W-OSVM here works well in the closed set scenario. However, once the scenario turns to the open set, its performance also drops significantly.

A Summary. Overall, based on the data adaptation characteristic of HDP, CD-OSR currently performs relatively well compared with the other methods. However, CD-OSR is also limited by HDP itself, such as difficulty in applying it to high-dimensional data, high computational complexity, etc. As for the other methods, they are limited by the underlying models they adopt as well. For example, as SRC does not work well on LETTER, thus SROSR obtains poor performance on that dataset. Furthermore, as mentioned in the remark of Section 3.1, for the methods using EVT such as W-SVM, P_I-SVM, SROSR, EVM, they may face challenges once the rare classes in KKCs and UUCs appear together in testing. Additionally, it is also necessary to point out that this part just gives a comparison of these algorithms on all commonly used datasets, which may not fully characterize their behavior to some extent.

5.3.2 OSR Methods Using Depth Feature

The OSR methods using depth feature are often evaluated on MNIST, SVHN, CIFAR10, CIFAR+10, CIFAR+50, Tiny-Imagenet. As most of them followed the evaluation protocol[8] defined

6. https://github.com/ChuanxingGeng/Open-Set-Recognition
7. W-OSVM here denotes only using the one-clss SVM CAP model, which can be seen as benchmark of one-class classifier.
8. The datasets and class partitions in [87] can be find in https://github.com/lwneal/counterfactual-open-set.

TABLE 5
Comparison Among the Representative OSR Methods Using Depth Features

Dataset / Method		SoftMax	OpenMax	CROSR	C2AE	G-OpenMax	OSRCI
MNIST	$O^* = 13.40\%$	97.8	98.1	**99.8**	98.9	98.4	98.8
SVHN	$O^* = 13.40\%$	88.6	89.4	**95.5**	92.2	89.6	91.0
CIFAR10	$O^* = 13.40\%$	67.7	69.5	—	**89.5**	67.5	69.9
CIFAR+10	$O^* = 24.41\%$	81.6	81.7	—	**95.5**	82.7	83.8
CIFAR+50	$O^* = 61.51\%$	80.5	79.6	—	**93.7**	81.9	82.7
TinyImageNet	$O^* = 57.36\%$	57.7	57.6	67.0	**74.8**	58.0	58.6

The results report the averaged area under the ROC curve (%) over 5 random class partitions [87]. Best and the second best performing methods are highlighted in bold and underline, respectively. O^* calculated from Eq. (3) denotes the openness of the corresponding dataset. Following [85], we here only copy the AUROC values of these methods as some of the results do not provide standard deviations.

in [87] and did not provide source codes, similar to [3], [148], we here only compare with their published results. Table 5 summaries the comparisons among these methods, where SoftMax [87], OpenMax [75], CROSR [83], and C2AE [85] from Deep Neural Network-based category, while G-OpenMax [86] and OSRCI [87] from Instance Generation-based category.

Our First Observation is. First, the performance of all methods on MNIST are comparable, which is mainly because results on MNIST are almost saturated. Second, compared with the earlier methods like SoftMax, OpenMax, G-OpenMax and OSRCI, CROSR and C2AE currently achieve better performance on the benchmark datasets. The main reason for their successes perhaps are: for CROSR, training networks for joint classification and reconstruction of KKCs makes the representation learned for KKCs more discriminative and tight (making the KKCs obtain tighter distribution areas); for C2AE, dividing OSR into closed set classification and open set identification allows it to avoid performing these two subtasks, concurrently, under a single score modified by the SoftMax scores (finding such a single score measure usually is extremely challenging [85]).

Our Second Observation is. As a state-of-the-art Instance Generation-based OSR method, OSRCI currently does not win CROSR and C2AE (two state-of-the-art Deep Neural Network-based OSR methods) on almost all datasets mentioned above, this seems a bit counter-intuition, since OSRCI gains the additional information from UUCs. But this just right indicates (from another side) that the performance of Instance Generation-based methods still have more room for improvement, deserving further exploration, while also showing the effectiveness of the strategies in CROSR and C2AE.

Remark. As mentioned previously, due to using EVT, OpenMax, CROSR, C2AE and G-OpenMax may also face challenges when the rare classes in KKCs and UUCs appear together in testing. In addition, it is also worth mentioning that the Instance Generation-based methods are orthogonal to the other three categories of methods, meaning that it can be combined with those methods to achieve their best.

6 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.

6.1 About Modeling

First, as shown in Fig. 3, while almost all existing OSR methods are modeled from the discriminative or generative model perspective, a natural question is: can construct OSR models from the hybrid generative discriminative model perspective? Note that to our best knowledge, there is no OSR work from this perspective at the moment, which deserves further discussions. Second, the main challenge for OSR is that the traditional classifiers under closed set scenario dive over-occupied space for KKCs, thus once the UUCs' samples fall into the space divided for KKCs, they will never be correctly classified. From this viewpoint, the following two modeling perspectives will be promising research directions.

6.1.1 Modeling Known Known Classes

To moderate the over-occupied space problem above, we usually expect to obtain better discrimination for each target class while confining it to a compact space with the help of clustering methods. To achieve this, 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 them. In fact, there have been some works fusing the clustering and classification functions into a unified learning framework [149], [150]. 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 specially design this type of classifiers for OSR.

6.1.2 Modeling Unknown Unknown Classes

Under the open set assumption, modeling UUCs is impossible, as we only have the available knowledge from KKCs. However, properly relaxing some restrictions will make it possible, where one way is to generate the UUCs' data by the adversarial learning technique to account for open space to some extent like [87], [88], [89], in which the key is how to generate the valid UUCs' data. Besides, due to the data adaptive nature of Dirichlet process, the Dirichlet process-based OSR methods, such as CD-OSR [91], are worth for further exploration as well.

6.2 About Rejecting

Until now, most existing OSR algorithms mainly care about effectively rejecting UUCs, yet only a few works [68], [84] focus on the subsequent processing for the reject samples, and these works usually adopt a post-event strategy [91]. Therefore, expanding existing open set recognition together with new class knowledge discovery will be an interesting research topic. Moreover, to our best knowledge, the interpretability of reject option seems to have not been discussed as well, 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. Some related works in other research communities can be found in [116], [151], [152], [153], [154].

6.3 About the Decision

As discussed in Section 3.2.2, almost all existing OSR techniques are designed specially for recognizing individual samples, even these samples are collectively coming in batch like image-set recognition [155]. In fact, such a decision does not consider correlations among the testing samples. Therefore, the collective decision [91] seems to be a better alternative as 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 such a collective decision.

6.4 Open Set + Other Research Areas

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-view learning, multi-label image classification problem, and so forth. For example, [156], [157], [158] have introduced this scenario into domain adaptation, while [159] introduced it to the semantic instance segmentation task. Recently, [160] explored the open set classification in active learning field. It is also worthy mentioning that the datasets NUS-Wide and MS COCO have been used for studying multi-label zero-shot learning [161], which are suitable to the study of multi-label OSR problem as well. Therefore, many interesting works are worth looking forward to.

6.5 Generalized Open Set Recognition

OSR assumes that only the KKCs' knowledge is available in training, meaning that we can also utilize a variety of side-information regarding the KKCs. Nevertheless most existing OSR methods just use the feature level information of KKCs, leaving out their other side-information such as semantic/attribute information, knowledge graph, the KUCs' data (e.g., the universum data), etc., which is also important for further improving their performances. Therefore, we give the following promising research directions.

6.5.1 Appending Semantic/Attribute Information

With the exploration of ZSL, we can find that a lot of semantic/attibute information is usually shared between KKCs and unknown class data. Therefore, such information can fully be used to 'cognize' UUCs in OSR, or at least to provide a rough semantic/attribute description for the UUCs' samples instead of simply rejecting them. Note that this setup is different from the one in ZSL (or G-ZSL) which assumes that the semantic/attibute information of both the KKCs and UUCs are known in training. Furthermore, the last row of Table 1 shows this difference. Besides, some related works can be found in [133], [154], [162], [163]. There also some conceptually similar topics have been studied in other research communities such as open-vocabulary object retrieval [164], [165], open world person re-identification [166] or searching targets [167], open vocabulary scene parsing [168].

6.5.2 Using Other Available Side-Information

For the over-occupied space problem mentioned in Section 6.1, the open space risk will also be reduced as the space divided for those KKCs decreases by using other side-information like the KUCs data (e.g., universum data [169], [170]) to shrink their regions as much as possible. As shown in Fig. 1, taking the digital identification as an example, assume the training set including the classes of interest '1', '3', '4', '5', '9'; the testing set including all of the classes '0'-'9'. If we also have the available universum data—English letters 'Z', 'I', 'J', 'Q', 'U', we can fully use them 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.

6.6 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 can be divided into two subspace respectively for sick and healthy 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. There are few works currently exploring this novel mixed scenario jointly. Please note that under such a scenario, the main goal is to limit the scope of the UUCs appearing in testing, while finding the most specific class label of a novel sample on the taxonomy built with KKCs. Some related work can be found in [171].

6.7 Knowledge Integration for Open Set Recognition

In fact, the incomplete knowledge of the world is universal, especially for 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 data privacy protection. This seems to have the flavor of domain adaptation having multiple source domains and one target domain (mS1T) [172], [173], [174], [175] to some extent.

7 Conclusion

As discussed above, in real-world recognition/classification tasks, it is usually impossible to model everything [176], thus the OSR scenario is ubiquitous. On the other hand, although many related algorithms have been proposed for OSR, it still faces serious challenges. As there is no systematic summary on this topic at present, this paper gives a comprehensive review of existing OSR techniques, covering various aspects ranging from related definitions, representations of models, datasets, evaluation criteria, and algorithm comparisons. 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.

Further, in order to avoid the reader confusing the tasks similar to OSR, we also briefly analyzed 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 was reviewed as well. More importantly, we analyzed and discussed the limitations of these existing approaches, and pointed out some promising subsequent research directions in this field.

Acknowledgments

The authors would like to thank the support from NSFC under Grant No. 61672281, the Key Program of NSFC under Grant No. 61732006 and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under grant no. KYCX18_0306.

References

[1] S. Thrun and T. M. Mitchell, “Lifelong robot learning,” *Robot. Auton. Syst.*, vol. 15, no. 1/2, pp. 25–46, 1995.
[2] A. Pentina and C. H. Lampert, “A PAC-Bayesian bound for lifelong learning,” in *Proc. Int. Conf. Int. Conf. Mach. Learn.*, 2014, pp. II-991–II-999.
[3] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
[4] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, “A survey of transfer learning,” *J. Big Data*, vol. 3, no. 1, 2016, Art. no. 9.
[5] L. Shao, F. Zhu, and X. Li, “Transfer learning for visual categorization: A survey,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 5, pp. 1019–1034, May 2015.
[6] V. M. Patel, R. Gopalan, R. Li, and R. Chellappa, “Visual domain adaptation: A survey of recent advances,” *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 53–69, May 2015.
[7] M. Yamada, L. Sigal, and Y. Chang, “Domain adaptation for structured regression,” *Int. J. Comput. Vis.*, vol. 109, no. 1/2, pp. 126–145, 2014.
[8] M. Palatucci, D. Pomerleau, G. Hinton, and T. M. Mitchell, “Zero-shot learning with semantic output codes,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2009, pp. 1410–1418.
[9] C. H. Lampert, H. Nickisch, and S. Harmeling, “Learning to detect unseen object classes by between-class attribute transfer,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 951–958.
[10] Y. Fu, T. Xiang, Y.-G. Jiang, X. Xue, L. Sigal, and S. Gong, “Recent advances in zero-shot recognition: Toward data-efficient understanding of visual content,” *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 112–125, Jan. 2018.
[11] F. F. Li, R. Fergus, and P. Perona, “One-shot learning of object categories,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 594–611, Apr. 2006.
[12] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, “One-shot learning by inverting a compositional causal process,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2013, pp. 2526–2534.
[13] L. Fe-Fei, Fergus, and Perona, “A Bayesian approach to unsupervised one-shot learning of object categories,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2008, pp. 1134–1141.
[14] Y. Amit, M. Fink, N. Srebro, and S. Ullman, “Uncovering shared structures in multiclass classification,” in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp. 17–24.
[15] A. Torralba, K. P. Murphy, and W. T. Freeman, “Using the forest to see the trees: Exploiting context for visual object detection and localization,” *Commun. ACM*, vol. 53, no. 3, pp. 107–114, 2010.
[16] Y. Fu, T. M. Hospedales, T. Xiang, and S. Gong, “Learning multimodal latent attributes,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 2, pp. 303–316, Feb. 2014.
[17] O. Vinyals *et al.*, “Matching networks for one shot learning,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 3630–3638.
[18] J. Snell, K. Swersky, and R. S. Zemel, “Prototypical networks for few-shot learning,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 4077–4087.
[19] L. Bertinetto, J. F. Henriques, J. Valmadre, P. H. S. Torr, and A. Vedaldi, “Learning feed-forward one-shot learners,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 523–531.
[20] Z. Chen, Y. Fu, Y. Zhang, Y. G. Jiang, X. Xue, and L. Sigal, “Semantic feature augmentation in few-shot learning,” 2018. [Online]. Available: https://arxiv.org/abs/1804.05298
[21] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, “Toward open set recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1757–1772, Jul. 2013.
[22] W. J. Scheirer, L. P. Jain, and T. E. Boult, “Probability models for open set recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 11, pp. 2317–2324, Nov. 2014.
[23] L. P. Jain, W. J. Scheirer, and T. E. Boult, “Multi-class open set recognition using probability of inclusion,” in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 393–409.
[24] N. A. Ross, “Known knowns, known unknowns and unknown unknowns: A 2010 update on carotid artery disease,” *Surgeon J. Roy. Colleges Surgeons Edinburgh Ireland*, vol. 8, no. 2, pp. 79–86, 2010.
[25] A. R. Dhamija, M. Günther, and T. Boult, “Reducing network agnostophobia,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 9157–9168.
[26] J. Weston, R. Collobert, F. Sinz, L. Bottou, and V. Vapnik, “Inference with the Universum,” in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 1009–1016.
[27] L. V. D. Maaten and G. Hinton, “Visualizing data using t-SNE,” *J. Mach. Learn. Res.*, vol. 9, no. Nov., pp. 2579–2605, 2008.
[28] R. Salakhutdinov, A. Torralba, and J. Tenenbaum, “Learning to share visual appearance for multiclass object detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 1481–1488.
[29] X. Zhu, D. Anguelov, and D. Ramanan, “Capturing long-tail distributions of object subcategories,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 915–922.
[30] R. Socher, M. Ganjoo, C. D. Manning, and A. Ng, “Zero-shot learning through cross-modal transfer,” in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2013, pp. 935–943.
[31] 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 *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 52–68.
[32] Y. Xian, C. H. Lampert, B. Schiele, and Z. Akata, “Zero-shot learning—A comprehensive evaluation of the good, the bad and the ugly,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 9, pp. 2251–2265, Sep. 2019.
[33] R. Felix, V. B. Kumar, I. Reid, and G. Carneiro, “Multi-modal cycle-consistent generalized zero-shot learning,” in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 21–37.
[34] S. Gidaris and N. Komodakis, “Dynamic few-shot visual learning without forgetting,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 4367–4375.
[35] C. Chow, “On optimum recognition error and reject tradeoff,” *IEEE Trans. Inf. Theory*, vol. 16, no. 1, pp. 41–46, Jan. 1970.
[36] P. L. Bartlett and M. H. Wegkamp, “Classification with a reject option using a hinge loss,” *J. Mach. Learn. Res.*, vol. 9, no. Aug., pp. 1823–1840, 2008.
[37] D. M. Tax and R. P. Duin, “Growing a multi-class classifier with a reject option,” *Pattern Recognit. Lett.*, vol. 29, no. 10, pp. 1565–1570, 2008.

[38] L. Fischer, B. Hammer, and H. Wersing, "Optimal local rejection for classifiers," *Neurocomputing*, vol. 214, pp. 445–457, 2016.
[39] G. Fumera and F. Roli, "Support vector machines with embedded reject option," in *Pattern Recognition with Support Vector Machines*. Berlin, Germany: Springer, 2002, pp. 68–82.
[40] Y. Grandvalet, A. Rakotomamonjy, J. Keshet, and S. Canu, "Support vector machines with a reject option," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2009, pp. 537–544.
[41] R. Herbei and M. H. Wegkamp, "Classification with reject option," *Can. J. Statist.*, vol. 34, no. 4, pp. 709–721, 2006.
[42] M. Yuan and M. Wegkamp, "Classification methods with reject option based on convex risk minimization," *J. Mach. Learn. Res.*, vol. 11, no. Jan., pp. 111–130, 2010.
[43] R. Zhang and D. N. Metaxas, "RO-SVM: Support vector machine with reject option for image categorization," in *Proc. Brit. Mach. Vis. Conf.*, 2006, pp. 1209–1218.
[44] M. Wegkamp, "Lasso type classifiers with a reject option," *Electron. J. Statist.*, vol. 1, no. 3, pp. 155–168, 2007.
[45] Y. Geifman and E. Y. Ran, "Selective classification for deep neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 4872–4887.
[46] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, no. 7, pp. 1443–1471, 2001.
[47] L. M. Manevitz and M. Yousef, "One-class SVMs for document classification," *J. Mach. Learn. Res.*, vol. 2, no. Dec., pp. 139–154, 2001.
[48] D. M. Tax and R. P. Duin, "Support vector data description," *Mach. Learn.*, vol. 54, no. 1, pp. 45–66, 2004.
[49] S. S. Khan and M. G. Madden, "A survey of recent trends in one class classification," in *Proc. Irish Conf. Artif. Intell. Cogn. Sci.*, 2009, pp. 188–197.
[50] H. Jin, Q. Liu, and H. Lu, "Face detection using one-class-based support vectors," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.*, 2004, pp. 457–462.
[51] M. Wu and J. Ye, "A small sphere and large margin approach for novelty detection using training data with outliers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 2088–2092, Nov. 2009.
[52] H. Cevikalp and B. Triggs, "Efficient object detection using cascades of nearest convex model classifiers," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 3138–3145.
[53] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Process.*, vol. 99, pp. 215–249, 2014.
[54] N. Görnitz, L. A. Lima, K.-R. Müller, M. Kloft, and S. Nakajima, "Support vector data descriptions and k-means clustering: One class?," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 3994–4006, Sep. 2018.
[55] P. Bodesheim, A. Freytag, E. Rodner, M. Kemmler, and J. Denzler, "Kernel null space methods for novelty detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 3374–3381.
[56] P. J. Phillips, P. Grother, and R. Micheals, "Evaluation methods in face recognition," in, *Handbook of Face Recognition*. New York, USA: Springer, 2005.
[57] F. Li and H. Wechsler, "Open set face recognition using transduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1686–1697, Nov. 2005.
[58] Q. Wu, C. Jia, and W. Chen, "A novel classification-rejection sphere SVMs for multi-class classification problems," in *Proc. 3rd Int. Conf. Natural Comput.*, 2007, pp. 34–38.
[59] Y.-C. F. Wang and D. Casasent, "A support vector hierarchical method for multi-class classification and rejection," in *Proc. Int. Joint Conf. Neural Netw.*, 2009, pp. 3281–3288.
[60] B. Heflin, W. Scheirer, and T. E. Boult, "Detecting and classifying scars, marks, and tattoos found in the wild," in *Proc. IEEE 5th Int. Conf. Biometrics: Theory Appl. Syst.*, 2012, pp. 31–38.
[61] D. A. Pritsos and E. Stamatatos, "Open-set classification for automated genre identification," in *Proc. Eur. Conf. Advances Inf. Retrieval*, 2013, pp. 207–217.
[62] H. Cevikalp, B. Triggs, and V. Franc, "Face and landmark detection by using cascade of classifiers," in *Proc. IEEE Int. Conf. Workshops Autom. Face Gesture Recognit.*, 2013, pp. 1–7.
[63] H. Cevikalp, "Best fitting hyperplanes for classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1076–1088, Jun. 2017.
[64] M. D. Scherreik and B. D. Rigling, "Open set recognition for automatic target classification with rejection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 2, pp. 632–642, Apr. 2016.
[65] H. Cevikalp and B. Triggs, "Polyhedral conic classifiers for visual object detection and classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4114–4122.
[66] H. Cevikalp and H. S. Yavuz, "Fast and accurate face recognition with image sets," in *Proc. IEEE Int. Conf. Comput. Vis. Workshop*, 2017, pp. 1564–1572.
[67] H. Zhang and V. Patel, "Sparse representation-based open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 8, pp. 1690–1696, Aug. 2017.
[68] A. Bendale and T. Boult, "Towards open world recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1893–1902.
[69] P. R. M. Júnior *et al.*, "Nearest neighbors distance ratio open-set classifier," *Mach. Learn.*, vol. 106, no. 3, pp. 359–386, 2017.
[70] E. M. Rudd, L. P. Jain, W. J. Scheirer, and T. E. Boult, "The extreme value machine," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 3, pp. 762–768, Mar. 2018.
[71] E. Vignotto and S. Engelke, "Extreme value theory for open set classification-GPD and GEV classifiers," 2018. [Online]. Available: https://arxiv.org/abs/1808.09902
[72] G. Fei and B. Liu, "Breaking the closed world assumption in text classification," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Technol.*, 2016, pp. 506–514.
[73] R. Vareto, S. Silva, F. Costa, and W. R. Schwartz, "Towards open-set face recognition using hashing functions," in *Proc. IEEE Int. Joint Conf. Biometrics*, 2017, pp. 634–641.
[74] 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.
[75] A. Bendale and T. E. Boult, "Towards open set deep networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1563–1572.
[76] A. Rozsa, M. Günther, and T. E. Boult, "Adversarial robustness: Softmax versus openmax," 2017. [Online]. Available: https://arxiv.org/abs/1708.01697
[77] M. Hassen and P. K. Chan, "Learning a neural-network-based representation for open set recognition," 2018. [Online]. Available: https://arxiv.org/abs/1802.04365
[78] S. Prakhya, V. Venkataram, and J. Kalita, "Open set text classification using convolutional neural networks," in *Proc. Int. Conf. Natural Lang. Process.*, 2017, pp. 466–475.
[79] L. Shu, H. Xu, and B. Liu, "DOC: Deep open classification of text documents," 2017. [Online]. Available: https://arxiv.org/abs/1709.08716
[80] N. Kardan and K. O. Stanley, "Mitigating fooling with competitive overcomplete output layer neural networks," in *Proc. Int. Joint Conf. Neural Netw.*, 2017, pp. 518–525.
[81] D. O. Cardoso, F. Franca, and J. Gama, "A bounded neural network for open set recognition," in *Proc. Int. Joint Conf. Neural Netw.*, 2015, pp. 1–7.
[82] D. O. Cardoso, J. Gama, and F. M. G. França, "Weightless neural networks for open set recognition," *Mach. Learn.*, vol. 106, no. 9/10, pp. 1547–1567, 2017.
[83] R. Yoshihashi, W. Shao, R. Kawakami, S. You, M. Iida, and T. Naemura, "Classification-reconstruction learning for open-set recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 4016–4025.
[84] L. Shu, H. Xu, and B. Liu, "Unseen class discovery in open-world classification," 2018. [Online]. Available: https://arxiv.org/abs/1801.05609
[85] P. Oza and V. M. Patel, "C2AE: Class conditioned auto-encoder for open-set recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2307–2316.
[86] Z. Y. Ge, S. Demyanov, Z. Chen, and R. Garnavi, "Generative OpenMax for multi-class open set classification," 2017. [Online]. Available: https://arxiv.org/abs/1707.07418
[87] L. Neal, M. Olson, X. Fern, W. Wong, and F. Li, "Open set learning with counterfactual images," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 613–628.
[88] I. Jo, J. Kim, H. Kang, Y.-D. Kim, and S. Choi, "Open set recognition by regularising classifier with fake data generated by generative adversarial networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2018, pp. 2686–2690.
[89] Y. Yu, W.-Y. Qu, N. Li, and Z. Guo, "Open-category classification by adversarial sample generation," 2017. [Online]. Available: https://arxiv.org/abs/1705.08722
[90] Y. Yang, C. Hou, Y. Lang, D. Guan, D. Huang, and J. Xu, "Open-set human activity recognition based on micro-doppler signatures," *Pattern Recognit.*, vol. 85, pp. 60–69, 2019.

[91] C. Geng and S. Chen, "Collective decision for open set recognition," 2018. [Online]. Available: https://arxiv.org/abs/1806.11258

[92] G. Bouchard and B. Triggs, "The trade-off between generative and discriminative classifiers," in *Proc. Comput. Statist. Symp. Iasc*, 2004, pp. 721–728.

[93] J. A. Lasserre, C. M. Bishop, and T. P. Minka, "Principled hybrids of generative and discriminative models," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 87–94.

[94] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, 1995.

[95] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications*. Singapore: World Scientific, 2000.

[96] S. Cruz, C. Coleman, E. M. Rudd, and T. E. Boult, "Open set intrusion recognition for fine-grained attack categorization," in *Proc. IEEE Int. Symp. Technol. Homeland Secur.*, 2017, pp. 1–6.

[97] E. Rudd, A. Rozsa, M. Gunther, and T. Boult, "A survey of stealth malware: Attacks, mitigation measures, and steps toward autonomous open world solutions," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1145–1172, Second Quarter 2017.

[98] R. N. Gasimov and G. Ozturk, "Separation via polyhedral conic functions," *Optim. Methods Softw.*, vol. 21, no. 4, pp. 527–540, 2006.

[99] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," *Proc. IEEE*, vol. 98, no. 6, pp. 1031–1044, Jun. 2010.

[100] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.

[101] J. Peng, L. Li, and Y. Y. Tang, "Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 6, pp. 1790–1802, Jun. 2019.

[102] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[103] T. Mensink, J. Verbeek, F. Perronnin, and G. Csurka, "Distance-based image classification: Generalizing to new classes at near-zero cost," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2624–2637, Nov. 2013.

[104] M. Ristin, M. Guillaumin, J. Gall, and L. V. Gool, "Incremental learning of NCM forests for large-scale image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3654–3661.

[105] J. P. Papa and C. T. N. Suzuki, *Supervised Pattern Classification Based on Optimum-Path Forest*. Hoboken, NJ, USA: Wiley, 2009.

[106] A. Garg and D. Roth, "Margin distribution and learning," in *Proc. 20th Int. Conf. Mach. Learn.*, 2003, pp. 210–217.

[107] L. Reyzin and R. E. Schapire, "How boosting the margin can also boost classifier complexity," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 753–760.

[108] F. Aiolli, G. Da San Martino, and A. Sperduti, "A kernel method for the optimization of the margin distribution," in *Proc. Int. Conf. Artif. Neural Netw.*, 2008, pp. 305–314.

[109] K. Pelckmans, J. Suykens, and B. D. Moor, "A risk minimization principle for a class of Parzen estimators," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2008, pp. 1137–1144.

[110] M. Günther, S. Cruz, E. M. Rudd, and T. E. Boult, "Toward open-set face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2017, pp. 573–582.

[111] J. Henrydoss, S. Cruz, E. M. Rudd, M. Gunther, and T. E. Boult, "Incremental open set intrusion recognition using extreme value machine," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl.*, 2017, pp. 1089–1093.

[112] A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 427–436.

[113] I. Goodfellow, J. Shelns, and C. Szegedy, "Explaining and harnessing adversarial examples," 2015. [Online]. Available: https://arxiv.org/abs/1412.6572

[114] C. Szegedy *et al.*, "Intriguing properties of neural networks," 2014. [Online]. Available: https://arxiv.org/abs/1312.6199

[115] Q. Da, Y. Yu, and Z.-H. Zhou, "Learning with augmented class by exploiting unlabeled data," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014, pp. 1760–1766.

[116] M. Masana, I. Ruiz, J. Serrat, J. van de Weijer, and A. M. Lopez, "Metric learning for novelty and anomaly detection," 2018. [Online]. Available: https://arxiv.org/abs/1808.05492

[117] W. J. Reed, "The Pareto, Zipf and other power laws," *Econ. Lett.*, vol. 74, no. 1, pp. 15–19, 2001.

[118] Z. Liu, Z. Miao, X. Zhan, J. Wang, B. Gong, and S. X. Yu, "Large-scale long-tailed recognition in an open world," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2537–2546.

[119] I. Goodfellow, "NIPS 2016 tutorial: Generative adversarial networks," 2016. [Online]. Available: https://arxiv.org/abs/1701.00160

[120] Y. W. Teh, "Dirichlet process," in *Encyclopedia of Machine Learning*. Berlin, Germany: Springer, 2011, pp. 280–287.

[121] S. J. Gershman and D. M. Blei, "A tutorial on Bayesian nonparametric models," *J. Math. Psychol.*, vol. 56, no. 1, pp. 1–12, 2012.

[122] W. Fan and N. Bouguila, "Online learning of a Dirichlet process mixture of beta-liouville distributions via variational inference," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 11, pp. 1850–1862, Nov. 2013.

[123] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet processes," *Publications Amer. Statist. Assoc.*, vol. 101, no. 476, pp. 1566–1581, 2006.

[124] A. Bargi, R. Y. Da Xu, and M. Piccardi, "AdOn HDP-HMM: An adaptive online model for segmentation and classification of sequential data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 3953–3968, Sep. 2018.

[125] G. Cauwenberghs and T. Poggio, "Incremental and decremental support vector machine learning," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2001, pp. 409–415.

[126] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, "Online passive-aggressive algorithms," *J. Mach. Learn. Res.*, vol. 7, no. Mar., pp. 551–585, 2006.

[127] M. Fink, S. Shalev-Shwartz, Y. Singer, and S. Ullman, "Online multiclass learning by interclass hypothesis sharing," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 313–320.

[128] T. Yeh and T. Darrell, "Dynamic visual category learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.

[129] M. D. Muhlbaier, A. Topalis, and R. Polikar, "Learn++ NC: Combining ensemble of classifiers with dynamically weighted consult-and-vote for efficient incremental learning of new classes," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 152–168, Jan. 2009.

[130] I. Kuzborskij, F. Orabona, and B. Caputo, "From N to N+ 1: Multiclass transfer incremental learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 3358–3365.

[131] R. De Rosa, T. Mensink, and B. Caputo, "Online open world recognition," 2016. [Online]. Available: https://arxiv.org/abs/1604.02275

[132] T. Doan and J. Kalita, "Overcoming the challenge for text classification in the open world," in *Proc. IEEE 7th Annu. Comput. Commun. Workshop Conf.*, 2017, pp. 1–7.

[133] V. P. A. Lonij, A. Rawat, and M. I. Nicolae, "Open-world visual recognition using knowledge graphs," 2017. [Online]. Available: https://arxiv.org/abs/1708.08310

[134] Y. Shi, Y. Wang, Y. Zou, Q. Yuan, Y. Tian, and Y. Shu, "ODN: Opening the deep network for open-set action recognition," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2018, pp. 1–6.

[135] H. Xu, B. Liu, and P. S. Yu, "Learning to accept new classes without training," 2018. [Online]. Available: https://arxiv.org/abs/1809.06004

[136] P. W. Frey and D. J. Slate, "Letter recognition using Holland-style adaptive classifiers," *Mach. Learn.*, vol. 6, no. 2, pp. 161–182, 1991.

[137] M. Bilenko, S. Basu, and R. J. Mooney, "Integrating constraints and metric learning in semi-supervised clustering," in *Proc. 21st Int. Conf. Mach. Learn.*, 2004, Art. no. 11.

[138] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (COIL-20)," Tech. Rep., Columbia Univ., 1996.

[139] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.

[140] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[141] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, "Reading digits in natural images with unsupervised feature learning," 2011. [Online]. Available: http://ufldl.stanford.edu/housenumbers

[142] A. Krizhevsky *et al.*, "Learning multiple layers of features from tiny images," Tech. Rep., Univ. Toronto, 2009.

[143] Y. Le and X. Yang, "Tiny ImageNet Visual Recognition 1694 Challenge," 2015. [Online]. Available: http://tiny-imagenet.herokuapp.com

[144] O. Russakovsky *et al*., "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis*., vol. 115, no. 3, pp. 211–252, 2015.

[145] C. Anna, L. Günter, R. H. Joachim, and M. Klaus, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage*., vol. 45, no. 4, pp. 427–437, 2009.

[146] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.

[147] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-Score and ROC: A family of discriminant measures for performance evaluation," in *Proc. Australas. Joint Conf. Artif. Intell*., 2006, pp. 1015–1021.

[148] M. Wang and W. Deng, "Deep visual domain adaptation: A survey," *Neurocomputing*, vol. 312, pp. 135–153, 2018.

[149] W. Cai, S. Chen, and D. Zhang, "A multiobjective simultaneous learning framework for clustering and classification," *IEEE Trans. Neural Netw*., vol. 21, no. 2, pp. 185–200, Feb. 2010.

[150] Q. Qian, S. Chen, and W. Cai, "Simultaneous clustering and classification over cluster structure representation," *Pattern Recognit*., vol. 45, no. 6, pp. 2227–2236, 2012.

[151] X. Mu, M. T. Kai, and Z. H. Zhou, "Classification under streaming emerging new classes: A solution using completely-random trees," *IEEE Trans. Knowl. Data Eng*., vol. 29, no. 8, pp. 1605–1618, Aug. 2017.

[152] S. Liu, R. Garrepalli, T. G. Dietterich, A. Fern, and D. Hendrycks, "Open category detection with PAC guarantees," 2018. [Online]. Available: https://arxiv.org/abs/1808.00529

[153] 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," 2018. [Online]. Available: https://arxiv.org/abs/1809.03576

[154] 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," 2018. [Online]. Available: https://arxiv.org/abs/1810.07368

[155] L. Wang and S. Chen, "Joint representation classification for collective face recognition," *Pattern Recognit*., vol. 63, pp. 182–192, 2017.

[156] P. P. Busto and J. Gall, "Open set domain adaptation," in *Proc. IEEE Int. Conf. Comput. Vis*., 2017, pp. 754–763.

[157] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, "Open set domain adaptation by backpropagation," 2018. [Online]. Available: https://arxiv.org/abs/1804.10427

[158] M. Baktashmotlagh, M. Faraki, T. Drummond, and M. Salzmann, "Learning factorized representations for open-set domain adaptation," 2018. [Online]. Available: https://arxiv.org/abs/1805.12277

[159] T. Pham, B. Vijay Kumar, T.-T. Do, G. Carneiro, and I. Reid, "Bayesian semantic instance segmentation in open set world," 2018. [Online]. Available: https://arxiv.org/abs/1806.00911

[160] Z.-Y. Liu and S.-J. Huang, "Active sampling for open-set classification without initial annotation," in *Proc. 33rd AAAI Conf. Artif. Intell*., 2019, pp. 4416–4423.

[161] C.-W. Lee, W. Fang, C.-K. Yeh, and Y.-C. Frank Wang, "Multi-label zero-shot learning with structured knowledge graphs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., 2018, pp. 1576–1585.

[162] Y. Fu and L. Sigal, "Semi-supervised vocabulary-informed learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., 2016, pp. 5337–5346.

[163] Y. Fu, H. Z. Dong, Y. F. Ma, Z. Zhang, and X. Xue, "Vocabulary-informed extreme value learning," 2017. [Online]. Available: https://arxiv.org/abs/1705.09887

[164] S. Guadarrama *et al*., "Open-vocabulary object retrieval," *Robot.: Sci. Syst*., vol. 2, 2014, Art. no. 6.

[165] S. Guadarrama, E. Rodner, K. Saenko, and T. Darrell, "Understanding object descriptions in robotics by open-vocabulary object retrieval and detection," *Int. J. Robot. Res*., vol. 35, no. 1–3, pp. 265–280, 2016.

[166] W.-S. Zheng, S. Gong, and T. Xiang, "Towards open-world person re-identification by one-shot group-based verification," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 38, no. 3, pp. 591–606, Mar. 2016.

[167] H. Sattar, S. Muller, M. Fritz, and A. Bulling, "Prediction of search targets from fixations in open-world settings," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., 2015, pp. 981–990.

[168] H. Zhao, X. Puig, B. Zhou, S. Fidler, and A. Torralba, "Open vocabulary scene parsing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., 2017, pp. 2021–2029.

[169] V. Cherkassky, S. Dhar, and W. Dai, "Practical conditions for effectiveness of the universum learning," *IEEE Trans. Neural Netw*., vol. 22, no. 8, pp. 1241–1255, Aug. 2011.

[170] Z. Qi, Y. Tian, and Y. Shi, "Twin support vector machine with universum data," *Neural Netw*., vol. 36, no. 3, pp. 112–119, 2012.

[171] K. Lee, K. Lee, K. Min, Y. Zhang, J. Shin, and H. Lee, "Hierarchical novelty detection for visual object recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*., 2018, pp. 1034–1042.

[172] S. Bendavid, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, "A theory of learning from different domains," *Mach. Learn*., vol. 79, no. 1/2, pp. 151–175, 2010.

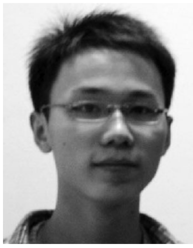
[173] S. Sun, H. Shi, and Y. Wu, *A Survey of Multi-Source Domain Adaptation*. Amsterdam, The Netherlands: Elsevier, 2015.

[174] H. Liu, M. Shao, and Y. Fu, "Structure-preserved multi-source domain adaptation," in *Proc. IEEE Int. Conf. Data Mining*, 2017, pp. 1059–1064.

[175] H. Yu, M. Hu, and S. Chen, "Unsupervised domain adaptation without exactly shared categories," 2018. [Online]. Available: https://arxiv.org/abs/1809.00852

[176] T. G. Dietterich, "Steps toward robust artificial intelligence," *AI Mag*., vol. 38, no. 3, pp. 3–24, 2017.

Chuanxing Geng received the BS degree in mathematics from Liaocheng University, Liaocheng, China, in 2013, and the MS degree in applied mathematics from Ningbo University, Ningbo, China, in 2016. Currently, he is working toward the PhD degree in the College of Computer Science & Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests include pattern recognition and machine learning.



Sheng-Jun Huang received the BSc and PhD degrees in computer science from Nanjing University, Nanjing, China, in 2008 and 2014, respectively. He joined the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics as an assistant professor in 2014. His main research interests include machine learning and patter recognition. He has won the Microsoft Fellowship Award (2011) and the Best Poster Award at KDD'12. He has been a Program Committee member of several conferences including KDD15, IJCAI15, ICDM15, etc.

Songcan Chen received the BS degree in mathematics from Hangzhou University (now merged into Zhejiang University), Hangzhou, China, in 1983, and the MS degree in computer applications from Shanghai Jiaotong University, Shanghai, China, in 1985, and then worked with NUAA in January 1986. He received the PhD degree in communication and information systems from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 1997. Since 1998, as a full-time professor, he has been with the College of Computer Science & Technology, NUAA. His research interests include pattern recognition, machine learning, and neural computing. He is also an IAPR fellow.