Proceedings of the 57th Hawaii International Conference on System Sciences | 2024

The Positive Impact of Metric Learning on Open Set Nearest Neighbor Classification

Alexander Grote
FZI Forschungszentrum Informatik
grote@fzi.de

Wolfgang Badwitz
FZI Forschungszentrum Informatik
badewitz@fzi.de

Michael Thomas Knierim
Karlsruhe Institute of Technology (KIT)
michael.knierim@kit.edu

Christof Weinhardt
Karlsruhe Institute of Technology (KIT)
weinhardt@kit.edu

Abstract

Traditional machine classification problems assume that complete knowledge of all classes is available during training. However, this assumption does often not hold for fast-changing environments and safety-critical applications like self-driving cars or tumour detection. In our work, we assume an arguably more realistic scenario called open set recognition, where incomplete knowledge of all classes during training is assumed, and also unknown classes can occur during testing. More importantly, we simulate an open set scenario on four established datasets and show how Open Set Nearest Neighbor classification results can be improved with metric learning. Our results indicate that the prior application of the Large Margin Nearest Neighbor algorithm can consistently enhance the classification results and increase the ability to reject unknown instances, which is vital in scenarios of many unknown classes. These findings highlight the importance of metric learning and serve as a benchmark for further studies on the intersection between metric learning and open set recognition.

Keywords: Metric Learning, Open Set Recognition, Open Set Nearest Neighbor

1. Introduction

Computer-aided classification is a major application field of machine learning, with invaluable relevance for solving problems from brain tumor detection to cyber-security or self-driving cars. While it is often assumed that these classification problems are about differentiating instances from a set of finite and known objects, in many cases, the underlying multi-class classification problems rather resemble recognition problems. In recognition problems, we instead assume that we can identify classes from a larger space of unknown objects. While this has a variety of applications (e.g. digitizing hand-written documents), the problem is particularly noteworthy in detecting hazardous objects. For instance, in the case of brain tumour classification, it is unlikely that we have sufficient knowledge of all the different types of tumours at training time since there are more than 150 of them (American Association of Neurological Surgeons, 2022). Similarly, in many safety-critical environments, such as self-driving cars (Ramanagopal et al., 2018), it is essential to identify data points that fall outside the norm to ensure a secure operation. Another area where detecting previously unknown patterns is important is cyber-security. The evolution of malware poses a constant challenge as it can circumvent existing detection solutions. Therefore, it is imperative to focus on the development of autonomous countermeasures and the identification of new types of malware to effectively combat this threat (Cruz et al., 2017; Henrydoss et al., 2017).

Scheirer et al. (2013) have grouped such multi-class problems under the term ”open set recognition” (OSR). These kinds of problems are arguably a more realistic scenario than traditional multi-class classification, where a classifier assumes that only the already seen classes exist during inference time. In contrast to traditional machine learning, OSR classifiers also indicate which data points are out of distribution and, therefore, differ from the training data. In practice, these out-of-distribution data points often require manual supervision to be classified correctly.

Dealing with data points of unknown classes is a complex problem because their representation is unknown in the feature space. In particular, these

URI: https://hdl.handle.net/10125/106507
978-0-9981331-7-1

unknown data points might be relatively close to already seen ones, which makes their classification difficult. One technique that pushes data points of similar classes together while increasing the distance to data points of other classes is called metric learning (Bellet et al., 2015). Learning such a new metric in a traditional machine learning setting is known to enhance the performance of nearest neighbor classification (Domeniconi et al., 2005; Goldberger et al., 2004; Hastie and Tibshirani, 1996; Simard et al., 1992; Weinberger et al., 2005).

However, it has not yet been studied what the effect of the learned feature transformation is on unknown classes and whether the results are dependent on the type of the employed metric learning algorithm. Furthermore, it is so far also unknown how nearest neighbor classification in combination with metric learning would perform with different degrees of unknown classes. Given the key relevance of detecting unknown classes in hazard-critical applications, we argue that it is highly important to study the potential benefits of combining metric learning with open set recognition approaches. To this end, we studied the following key research questions in this work:

- **RQ1**: Does metric learning increase the performance of nearest neighbor classification in OSR?
- **RQ2**: Is the marginal contribution of metric learning positively correlated with an increasing degree of openness?
- **RQ3**: How are different metric learning algorithms contributing to nearest neighbor classification in OSR?

To answer our research questions, we evaluated the performance of the Open Set Nearest Neighbor (OSNN) classifier developed by Mendes Júnior et al. (2017) across four data sets and four degrees of openness, without and with metric learning to calculate the marginal contribution of metric learning. Specifically, we thoroughly examined two metric learning algorithms, namely Large Margin Nearest Neighbor (LMNN) and Neighborhood Components Analysis (NCA), which were developed by Weinberger et al. (2005) and Goldberger et al. (2004), respectively.

Our results indicate that the prior application of metric learning can enhance the classification results of the OSNN classifier. While the increasing openness worsens the overall classification results, the prior application of the LMNN consistently improves the classification accuracy across all datasets for each degree of openness. Moreover, the transformed feature space by the LMNN increases the relative contribution to the classification results over an increasing number of unknown classes and enhances the ability to reject unknown data points. In the case of the NCA algorithm, we only see increases in classification accuracy on some datasets and cannot report a statistically significant relative marginal contribution across all datasets.

Our results thereby highlight the importance of learning a task-specific metric for open spaces and contribute to IS research and practice in two major ways. First, our results highlight the general benefit of combining OSNN and metric learning while simultaneously identifying limitations in some algorithmic combinations. This provides important evidence for future research to extend the understanding and development of more domain-specific feature spaces to enhance the OSNN classification performance. It also directly translates into our recommendation for practitioners to use metric learning for OSNN classification. Second, we show that learning a task-specific metric is increasingly vital for higher degrees of openness in OSNN classification, which serves as a baseline for further research on the intersection between metric learning and OSR. From a more theoretical perspective, we attribute this effect to the class-separating capability of metric learning. It can significantly enhance the ability of the OSNN to reject unfamiliar data points by creating more confined spaces.

The remainder of our paper introduces the fundamentals of OSR and metric learning in Section 2. We then present related work in Section 3. In Section 4, we elaborate on our experiments and discuss the results in Section 6. Lastly, we summarize our findings and outline directions for future research in Section 7.

2. Foundations

2.1. Open Set Recognition

Open set recognition (Scheirer et al., 2013) refers to a scenario, where we do not have full knowledge of all classes at training time and the classifier needs to reject new classes as unknown. To better describe this scenario in our work, we use the concept of known known classes (KKC) and unknown unknown classes (UUC) from Geng et al. (2021). While KKCs represent positive class samples known during training time, UUCs are entirely unknown at training time. The cases of unknown known and known unknown classes are irrelevant for our work.

The degree of openness O measures the relation between KKCs and UUCs (Geng et al., 2021). It is defined based on the number of classes seen during training (C_{TR}) and the number of classes in the test set (C_{TE}). Since C_{TE} is greater or equal than C_{TR}, the openness score ranges between zero and one. The higher the openness score, the more classes are considered unknown.

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

OSR classifiers deal with this openness by rejecting instances that are out of distribution and subsequently classifying the not rejected data points into one of the known classes. The overall risk associated with the OSR classification task is also known as the open set risk and consists of an empirical risk (R_ϵ) and an open space risk R_O component (Scheirer et al., 2013). The overall goal of OSR is to find a measurable function $f \in H$, which minimizes the open set risk. Equation 2 formalizes the open set risk, where λ represents a regularization constant and V the training data.

$$\underset{f \in H}{\operatorname{argmin}} \{R_O(f) + \lambda R_\epsilon(f(V)\} \quad (2)$$

2.2. Metric Learning

Metric learning creates task-specific metrics from supervised data (Bellet et al., 2015). The term metric hereby refers to any pairwise function that measures a distance or similarity between two objects. Most metric learning algorithms use pairwise or relative constraints to learn such a metric (Bellet et al., 2015). Pairwise constraints enforce learning the similarity on must-link / cannot-link constraints, such as $S = \{(x_i, x_j) : x_i \text{ and } x_j \text{ should be similar}\}$ and $D = \{(x_i, x_j) : x_i \text{ and } x_j \text{ should be dissimilar}\}$. On the other hand, relative constraints require that $T = \{(x_i, x_j, x_k) : x_i \text{ should be more similar to } x_j \text{ than to } x_k\}$ is fulfilled as best as possible.

Learning a metric M can be typically expressed as a minimization problem of two components, namely a loss function $l(M, S, D, R)$ penalizing the violation of the constraints and a regularization $R(M)$ with $\lambda >= 0$ as a regularization parameter (Bellet et al., 2013). Equation 3 summarizes this relationship.

$$\min_M l(M, S, D, T) + \lambda R(M) \quad (3)$$

To use metric learning for classification, a metric is learned on the training data, which is subsequentially applied to the training, validation and test data (Bellet et al., 2013). The classifier then learns on the transformed training data, validates its hyperparameters on the validation data and creates predictions on the test data.

3. Related Work

Among the first OSR algorithms are adapted versions of Support Vector Machines (Jain et al., 2014; Scheirer et al., 2013; Scheirer et al., 2014). However, also other machine learning approaches, such as distance (Bendale and Boult, 2016; Mendes Júnior et al., 2017), margin distribution (Rudd et al., 2018) and sparse representation based (Zhang and Patel, 2017) approaches exist. Similar to the nearest class mean classification, Bendale and Boult (2015) propose an extension to the SoftMax layer of neural networks that compares each input to the mean class activation vector and classifies instances that are too far away from these vectors as unknown. Shu et al. (2017) use a one-versus-rest layer that transforms a multiclass problem into multiple binary problems. The reconstruction loss of autoencoders has also been used to detect unknown classes (Lübbering et al., 2022; Oza and Patel, 2019).

Learning a metric to enhance nearest neighbour classification results has also been studied extensively outside OSR. While many machine learning based metric learning algorithms, such as the LMNN and the Neighborhood Components Analysis (NCA) proposed by Goldberger et al. (2004) learn a linear transformation with the Mahalanobis distance, also non-linear machine learning based transformations (Kedem et al., 2012; Shi and Liu, 2018) are known to increase the nearest neighbour classification. Deep metric learning algorithms mostly learn a non-linear metric and have increased the ability to recognize similar objects (Kaya and Bilge, 2019). State-of-the-art deep metric learning algorithms are SphereFace (Liu et al., 2017) or one of its derivatives (Deng et al., 2020; Deng et al., 2019; Wang et al., 2018), employing angular loss functions.

The main difference to contrastive approaches, such as the triplet (Schroff et al., 2015) or quadruplet loss (Chen et al., 2017), is that, instead of learning the Euclidean distance directly on pairs or triplets, the angular distances between the data point and the learned class representation is optimized.

The main difference to our work is that we evaluate our classifier after learning a metric under the open set assumption. Also, to the best of the authors' knowledge, the inherent link between the learned metric and its OSNN classification performance has not been studied for varying degrees of openness.

4. Methodology

4.1. Procedure

To analyse the effect of metric learning on the OSNN across varying degrees of openness, we simulate an open set scenario with four datasets and two metric learning algorithms. Figure 1 summarizes the overall experiment setup. After retrieving the data, the KKCs and UUCs are randomly selected. Subsequently, the data is split into training, validation and test datasets. Next, the metric learning algorithm is applied to the training dataset. This step is skipped if we calculate the classifier's performance without metric learning. To ensure comparability, the datasets are identical in both scenarios. The transformed data is then passed to the OSNN, and its rejection threshold is evaluated with the validation data. Finally, the macro f1 score is used to evaluate the classifier on the test set. The overall process is repeated ten times for each openness score to counteract the random class selection process.

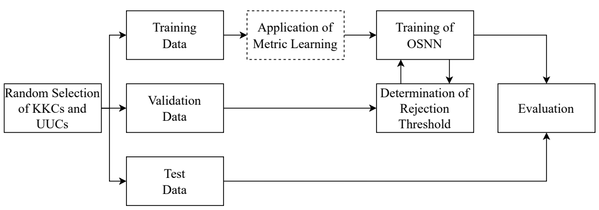

Figure 1. Data pipeline of experiments.

4.2. Datasets

We use four, frequently used datasets to evaluate OSR problems (Geng et al., 2021).

- **LETTER** (Frey and Slate, 1991): the dataset resembles images of distorted capital letters of the English alphabet. It consists of 200.000 samples from 26 classes and has 16 primitive numerical attributes, such as statistical moments and edge counts. The instances are evenly distributed among all classes and contain, on average, 769 samples.

- **PENDIGITS** (Bilenko et al., 2004): this dataset consists of handwritten digits and contains 10.992 samples from 10 classes. The average number of instances per class is 1099. Just like the LETTER data set, it also has 16 numerical features.

- **COIL20** (Nene et al., 1996): the coil dataset contains 72 images for each of the 20 classes. Each class thereby represents a distinct object, such as a rubber duck or a piggy bank. To ensure comparability, we follow the approach of Geng et al. (2021) and also downsample each image to 16×16, flatten the image, and further reduce its dimension with the Principal Component Analysis (PCA) to 55.

- **YALEB** (Georghiades et al., 2001): the dataset contains 38 classes, each of them representing the face of an individual. From each individual, on average, approximately 65 images, which represent different lighting and poses, exist. Similar to the COIL20 dataset, we use the same preprocessing as Geng et al. (2021) to ensure comparability. First, we crop and normalize the images to 32×32. Afterwards, we apply a PCA with 69 components to reduce the dimensions.

To evaluate our OSR classifier, we split our datasets into a fitting, validation and test set. We use the same splitting method as Geng et al. (2021). First, we randomly select from all available classes our known known classes Ω. The exact number of drawn KKCs depends on the degree of openness. From all data points that belong to Ω, we select a random sample of 60 % as our training set. The remaining 40 % of the KKCs and the other classes that are not contained in Ω are chosen as a test set. As a result, the test set contains not only classes the classifier has seen in training but also new classes. Both types of classes serve as a simulation of the open set scenario. A similar split is also needed for the validation set because when determining the rejection threshold of the OSNN, we need unseen classes for a proper evaluation. Depending on Ω, we choose $\frac{2}{3}\Omega + 0.5$ classes as our "KKCs" from the training set. The remaining classes are our "UUCs". We sample again 60 % of these "KKCs" and select them as our fitting set. The validation set then consists of the remaining 40 % of the "KKCs" and the "UUCs".

4.3. Metric Learning Algorithms

To demonstrate the effect of metric learning, we apply two different metric learning techniques, namely NCA and LMNN. Both of them aim at improving the k nearest neighbors classification by learning a Mahalanobis distance measure, which is a linear transformation of the feature space.

NCA maximizes a stochastic variant of the leave-one-out k nearest neighbor score via gradient descent (Weinberger et al., 2005). Equation 4 shows the softmax function used for calculating the probability p_{ij} of point i being a neighbor to j, where M is the learned transformation.

$$p_{ij} = \begin{cases} \frac{\exp(-\|Mx_i - Mx_j\|^2)}{\sum_{k \neq i} \exp - \|Mx_i - Mx_k\|^2} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (4)$$

With p_{ij}, we can compute the probability of correct classification p_i by summing up all the probabilities p_{ij} that share the same class as i. Maximizing the expected number of points correctly classified results in a differentiable objective function, which only depends on the learned metric M. Equation 5 demonstrates this relationship.

$$f(M) = \sum_i \sum_{j \in C_i} p_{ij} = \sum_i p_i \quad (5)$$

The optimization function of the LMNN is a convex, semidefinite optimization problem (Goldberger et al., 2004) that consists of two subgoals. The first goal is to minimize the average distance between instances and their target neighbors N (Equation 6), while the second goal is to penalize imposters x_l that are less than one unit further away than the target neighbors (Equation 7). The hinge loss function prevents penalizing an imposter when it is outside of the margin.

$$\sum_{i,j \in N_i} d(x_i, x_j) \quad (6)$$

$$\sum_{i,j \in N_i, l, y_l \neq y_i} [d(x_i, x_j) - 1 + d(x_i, x_l)]_+ \quad (7)$$

Equation 8 shows the overall optimization problem, where $\lambda > 0$ is a hyperparameter, ξ_{ijl} are slack variables and $M \succeq 0$ ensures that the matrix is semi-definite.

$$\begin{aligned} \min_M \quad & \sum_{i,j \in N_i} d(x_i, x_j) + \lambda \sum_{i,j,l} \xi_{ijl} \\ \text{subject to} \quad & d(x_i, x_j) - 1 + d(x_i, x_l) \leq \xi_{ijl} \\ & \xi_{ijl} \geq 0 \\ & M \succeq 0 \end{aligned} \quad (8)$$

4.4. Open Set Nearest Neighbor

OSNN is the open set variant of the nearest neighbours classifier (Mendes Júnior et al., 2017). Unlike its traditional counterpart, it takes the similarity of the closest two classes into account. The nearest neighbor distance ratio R is given in Equation 9. While d is a distance function and serves as a similarity measurement, s, t, and u denote data instances. s is the data point we want to classify and from which we measure the distance to its nearest neighbours t and u. u, however, must be of a different class than t.

$$R = d(s,t)/d(s,u) \quad (9)$$

Since $d(s,u)$ is greater or equal to $d(s,t)$, the ratio R is smaller or equal to one. Intuitively, if a data point is equidistant from its two closest classes, R equals 1 and hence, it is uncertain to which class it belongs. Conversely, if a data point is much closer to t than to u, R is close to 0. This intuition directly translates into a classification rule: if R is smaller than a specified threshold $T \in [0,1]$, we classify s with the same label as t. Otherwise, we reject it as unknown.

4.5. Evaluation Criteria

In information retrieval and machine learning, the f1 measure is widely used to evaluate classification problems. It is defined as the harmonic mean of precision and recall. To use this measure in OSR problems, the unknown classes are not treated as an additional one (Mendes Júnior et al., 2017). Otherwise, we would count correctly classified data points as true positives (TP) despite not knowing what a representative sample of such a TP would look like. However, we still account for unknown classes because the false positive and false negative also consider the misclassification of the unknown and known classes (Mendes Júnior et al., 2017).

To evaluate how metric learning affects the performance on OSR with increasing openness, we calculate the relative contribution of metric learning. We define the relative contribution RC in Equation 10, where we divide the difference between the classification result with metric learning C_M and the results without metric learning C_{NM} through C_{NM}. The relative contribution is negative if metric learning harms the performance and positive if it enhances the classification result.

$$RC = \frac{C_M - C_{NM}}{C_{NM}} \quad (10)$$

5. Evaluation

Table 1 contains the main results of our study, which consist of an overall evaluation of the classification performance and an analysis on the rejection of unknown classes. As mentioned in Section 4.5, the adjusted f1 score for OSR does not consider the unknown classes as a separate class (Mendes Júnior

Table 1. OSNN classification results with and without the prior application of metric learning.

Dataset	Openness	Overall F1 Score			F1 Score On Unknown Classes		
		Baseline	LMNN	NCA	Baseline	LMNN	NCA
COIL	0.00	0.98±0.01	**0.99±0.01**	**0.99±0.01**	-	-	-
	0.16	0.92±0.02	**0.95±0.02**	0.93±0.03	0.92±0.06	**0.96±0.02**	0.93±0.06
	0.24	0.86±0.05	**0.93±0.04**	0.92±0.04	0.92±0.05	**0.97±0.02**	**0.97±0.02**
	0.32	0.80±0.10	**0.85±0.08**	0.81±0.07	0.89±0.07	**0.95±0.04**	0.92±0.04
LETTER	0.00	0.93±0.00	0.95±0.00	**0.96±0.00**	-	-	-
	0.16	0.80±0.01	0.82±0.01	**0.83±0.01**	0.85±0.02	**0.86±0.02**	**0.86±0.02**
	0.23	0.76±0.03	0.78±0.03	**0.78±0.02**	**0.89±0.02**	**0.89±0.02**	**0.89±0.02**
	0.31	0.65±0.04	**0.69±0.04**	**0.69±0.04**	0.88±0.03	**0.90±0.03**	**0.90±0.02**
PENDIGITS	0.00	**0.99±0.00**	**0.99±0.00**	**0.99±0.00**	-	-	-
	0.13	0.90±0.02	**0.92±0.02**	0.90±0.03	0.90±0.06	**0.92±0.02**	0.90±0.06
	0.24	0.85±0.04	**0.86±0.02**	0.85±0.04	0.93±0.02	**0.94±0.01**	0.93±0.03
	0.32	0.75±0.14	0.77±0.09	**0.82±0.08**	0.89±0.10	0.92±0.05	**0.95±0.03**
YALEB	0.00	0.53±0.02	**0.88±0.01**	0.84±0.01	-	-	-
	0.16	0.49±0.02	**0.82±0.02**	0.76±0.04	0.71±0.02	**0.88±0.03**	0.86±0.02
	0.23	0.45±0.03	**0.75±0.03**	0.74±0.06	0.76±0.05	**0.90±0.02**	0.90±0.03
	0.33	0.39±0.03	**0.68±0.02**	0.66±0.11	0.79±0.09	**0.93±0.01**	0.91±0.05

et al., 2017) and hence, the second analysis enhances our understanding of the classifier's ability to reject unknown classes. For each dataset, we exclude 0 %, 20 %, 40 % and 60 % of our classes as unknown classes, resulting in increasing degrees of openness. Depending on the number of classes in each dataset, the degree of openness can vary across the datasets for a given percentage of unknown classes. The bold numbers represent the best result for each scenario. In each scenario, we use the OSNN classification results without the prior application of metric learning as our baseline. We compare them to the OSNN performance with the previous application of either the LMNN or NCA algorithm. Regarding the overall classification performance, the LMNN algorithm consistently improves the classification results on all four datasets. The utilization of the NCA algorithm has also improved overall performance. However, upon application to the PENDIGITS, the outcomes are comparable to those obtained without prior implementation of metric learning on at least two different degrees of openness. All results show an increased variance of the f1 score for problems with a higher openness. The analysis on the binary classification of unknown classes shows that the LMNN algorithm effectively increases the ability to correctly reject unknown instances compared to the OSNN classification without metric learning. NCA, on the other hand, does not consistently enhance the rejection accuracy. On the PENDIGITS datasets, the f1 score is in two out three cases on par with no prior application of metric learning, while the f1 score shows mostly increased rejection accuracy on the remaining datasets. Moreover, the LMNN algorithm yields a much lower variance of the f1 scores compared to no prior application of metric learning.

Figure 2 shows the classification results over varying degrees of openness for each dataset and metric learning algorithm separately. The general trend is depicted with a linear regression line and its corresponding 95 % confidence interval. The degree of openness is the only independent variable for this regression model. In general, the relative contribution of metric learning increases for higher degrees of openness.

Table 2. Pearson correlation coefficients.

Dataset	Metric Learning Algorithm	
	LMNN	NCA
COIL	0.36 (p=.012)	0.17 (p=.146)
LETTER	0.30 (p=.029)	0.12 (p=.232)
PENDIGITS	0.24 (p=.044)	0.27 (p=.047)
YALEB	0.34 (p=.015)	0.27 (p=.047)

To further deepen our insights, we measure the strength of the correlation with the Pearson correlation coefficient and conduct a one-sided t-test to check for statistical significance. Our alternative hypothesis is that the correlation is greater than zero. Table 2 summarizes the results. For a significance level of 0.05,

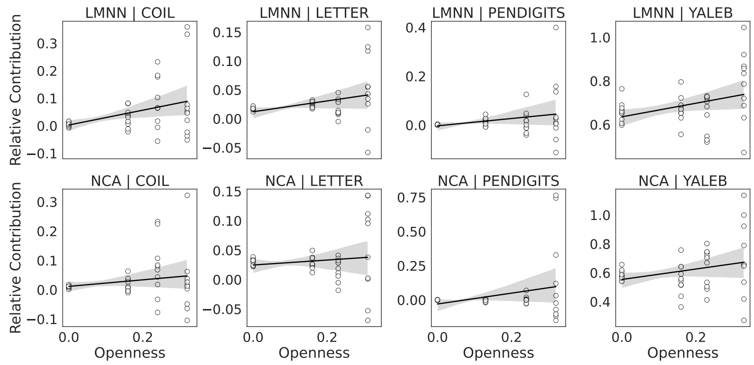

Figure 2. Correlation analysis for increasing openness.

the positive correlation of the relative contribution of the LMNN algorithm is statistically significant across all datasets. The Pearson correlation of the contribution of the NCA algorithm is positive across all datasets but only statistically significant on the PENDIGITS and YALEB datasets.

To better understand where the learned metric places unseen data points, we have conducted an in-depth analysis of the nearest neighbor distance ratio R, which is used by the OSNN to reject data points as unknown. As mentioned in 4.4, a value close to 1 indicates that the two nearest neighbors of different classes are almost equidistant. Hence, we expect the ratios of data points from known classes to be significantly smaller than those of unknown classes. Our analysis in Table 3 confirms this expectation. For enhanced readability, we have highlighted the best ratios, which correspond to the minimum values among the known classes and the maximum values among the unknown classes. The analysis also shows that metric learning significantly lowers the ratios among the known classes. The ratios also decrease among the unknown classes. However, their relative change compared to no prior application of metric learning is less substantial than the relative change among the known classes. This finding implies that the learned metric effectively helps to separate the unknown from the known classes in a nearest-neighbor setting.

6. Discussion

The general boost in classification accuracy due to the application of LMNN and NCA is in line with the findings in their original publications, which are based on a closed set scenario (Goldberger et al., 2004; Weinberger et al., 2005). We extended the analysis for the open set scenario, which shows mixed results. While the LMNN algorithm consistently enhances the classification results, NCA does not improve the classification results on the PENDIGITS dataset. The main reason for this could be the inner working of the NCA algorithm. As outlined in 4.3, NCA uses a stochastic variant of the leave-one-out k nearest neighbor classifier. However, the optimization process with gradient descent might stop at a local optimum and, thus, never reach its global optimum. In the case of the PENDIGITS datasets, a locally optimal solution might have been encountered, and no further optimization process regarding the unknown classes is carried out. The rejection analysis of unknown classes also shows that, exactly on this dataset, no enhancement of the f1 score can be seen. On the other hand, the LMNN directly optimizes the distances of the data points to each other, which is a continuous optimization process, and separates the data points of different classes by a certain margin. It seems plausible that such a margin reduces the risk of the unknown classes occurring relatively closer to the known classes, which increases the likelihood of correct classifications. The analysis of the rejection of unknown classes confirms this claim by consistently showing enhanced results for the prior application of the LMNN algorithm. The learned margin and, as a result, the improved ability to reject unknown instances might explain why the LMNN improves the performance on all four datasets. The marginal contribution of metric learning shows positive trends for both metric learning algorithms. However, only the LMNN algorithm shows significant results on all datasets, which aligns with the overall classification results. Overall, both metric learning algorithms exhibit the potential to enhance the OSNN classification results. Yet we do not find evidence that the relative contribution of metric learning, in general, has a positive correlation with openness. Our nearest neighbor distance ratio analysis reveals that the metric learning places unknown classes closer to already known classes. However, metric learning also puts known classes closer together, compensating for the decreased ratios among unknown

Table 3. Nearest neighbor distance ratio.

Dataset	Openness	Known Classes			Unknown Classes		
		Baseline	LMNN	NCA	Baseline	LMNN	NCA
COIL	0.16	0.28±0.04	**0.16±0.03**	0.21±0.03	**0.86±0.03**	0.75±0.03	0.81±0.04
	0.24	0.23±0.05	**0.11±0.03**	0.17±0.03	**0.83±0.04**	0.71±0.05	0.77±0.03
	0.32	0.27±0.08	**0.13±0.05**	0.20±0.05	**0.82±0.04**	0.66±0.04	0.75±0.05
LETTER	0.16	0.52±0.02	0.49±0.02	**0.46±0.02**	**0.87±0.01**	0.86±0.01	0.85±0.01
	0.23	0.51±0.02	0.47±0.02	**0.45±0.02**	**0.86±0.01**	0.85±0.02	0.84±0.01
	0.31	0.47±0.04	0.42±0.05	**0.42±0.04**	**0.86±0.01**	0.84±0.01	0.82±0.02
PENDIGITS	0.13	0.35±0.02	**0.30±0.02**	0.35±0.02	**0.81±0.03**	0.80±0.04	0.81±0.04
	0.24	0.31±0.03	**0.24±0.04**	0.31±0.03	**0.79±0.03**	0.74±0.03	0.78±0.03
	0.32	0.28±0.05	**0.18±0.06**	0.28±0.05	0.74±0.05	0.60±0.09	**0.75±0.05**
YALEB	0.16	0.78±0.01	**0.51±0.01**	0.59±0.03	**0.90±0.00**	0.89±0.01	0.89±0.01
	0.23	0.77±0.01	**0.48±0.01**	0.54±0.05	**0.90±0.00**	0.85±0.01	0.88±0.01
	0.33	0.76±0.01	**0.43±0.02**	0.49±0.09	**0.89±0.01**	0.81±0.03	0.83±0.01

classes.

A limitation of this work is its narrow focus on OSNN classification. However, as shown by Shi and Liu (2018), the use of metric learning can also benefit SVMs in a closed setting. It remains an open question whether the results are transferable to other OSR algorithms, such as the OSR adapted SVM algorithms (Jain et al., 2014; Scheirer et al., 2013; Scheirer et al., 2014), the Extreme Value Machine (Rudd et al., 2018) or also OpenMax (Bendale and Boult, 2015). Further research is also needed to investigate the behaviour of more advanced metric learning algorithms, such as SphereFace (Liu et al., 2017) or ArcFace (Deng et al., 2019), for varying degrees of openness. These metric learning algorithms are typically utilized in conjunction with deep neural networks, and their application promises performance increases on more complex datasets. However, similar to the application of the NCA algorithm, their suitability for an increasing number of unknown classes could be limited.

7. Conclusion

While metric learning is known to improve nearest neighbor classification results in traditional machine learning problems, the effects of these feature space transformations are unknown for OSR problems. These open set problems assume that, at training time, there is insufficient knowledge of all classes, and hence, unknown classes could occur during inference. OSR classifiers seek to reject data points of unknown classes as unknown, instead of classifying them as one of its learned classes. The ability to reject data points is especially important in safety-critical and cyber-security-related environmental problems, where unforeseen changes can happen frequently and have hazardous effects. With our open set simulation on four established datasets for OSR, we show that two metric learning algorithms, namely the LMNN and the NCA algorithm, can improve the OSNN classification results. While the utilization of the NCA algorithm does not consistently increase classification performance, we have found statistically significant evidence that the prior application of the LMNN algorithm increases OSNN classification results consistently and with an increasing relative contribution for higher degrees of openness. We attribute this finding to learning a margin that separates the different classes from each other and have found further evidence that this increases the rejection performance on unknown classes. We thereby contribute to existing research in two major ways. First, we highlight the general benefit of combining OSNN with metric learning and identify limitations regarding the choice of the metric learning algorithm. Second, we show that learning a task-specific metric is increasingly vital for higher degrees of openness. The overall findings encourage the further application of metric learning in an open set context and establish a baseline for further research on the intersection between OSR and metric learning.

References

American Association of Neurological Surgeons. (2022). Brain Tumors - Classifications, Symptoms, Diagnosis and Treatments. Retrieved November 3, 2022, from https :

//www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Brain-Tumors

Bellet, A., Habrard, A., & Sebban, M. (2013). A survey on metric learning for feature vectors and structured data. *arXiv preprint arXiv:1306.6709*.

Bellet, A., Habrard, A., & Sebban, M. (2015). *Metric Learning*. Springer International Publishing. https://doi.org/10.1007/978-3-031-01572-4

Bendale, A., & Boult, T. (2015). Towards open world recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1893–1902.

Bendale, A., & Boult, T. E. (2016). Towards Open Set Deep Networks. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1563–1572. https://doi.org/10.1109/CVPR.2016.173

Bilenko, M., Basu, S., & Mooney, R. J. (2004). Integrating constraints and metric learning in semi-supervised clustering. *Proceedings of the twenty-first international conference on Machine learning*, 11. https://doi.org/10.1145/1015330.1015360

Chen, W., Chen, X., Zhang, J., & Huang, K. (2017). Beyond triplet loss: A deep quadruplet network for person re-identification. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 403–412.

Cruz, S., Coleman, C., Rudd, E. M., & Boult, T. E. (2017). Open set intrusion recognition for fine-grained attack categorization. *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, 1–6. https://doi.org/10.1109/THS.2017.7943467

Deng, J., Guo, J., Liu, T., Gong, M., & Zafeiriou, S. (2020). Sub-center ArcFace: Boosting Face Recognition by Large-Scale Noisy Web Faces. *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI*, 741–757. https://doi.org/10.1007/978-3-030-58621-8_43

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive Angular Margin Loss for Deep Face Recognition [ISSN: 2575-7075]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4685–4694. https://doi.org/10.1109/CVPR.2019.00482

Domeniconi, C., Gunopulos, D., & Peng, J. (2005). Large margin nearest neighbor classifiers [Conference Name: IEEE Transactions on Neural Networks]. *IEEE Transactions on Neural Networks*, *16*(4), 899–909. https://doi.org/10.1109/TNN.2005.849821

Frey, P. W., & Slate, D. J. (1991). Letter recognition using Holland-style adaptive classifiers. *Machine Learning*, *6*(2), 161–182. https://doi.org/10.1007/BF00114162

Geng, C., Huang, S.-J., & Chen, S. (2021). Recent Advances in Open Set Recognition: A Survey [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(10), 3614–3631. https://doi.org/10.1109/TPAMI.2020.2981604

Georghiades, A., Belhumeur, P., & Kriegman, D. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(6), 643–660. https://doi.org/10.1109/34.927464

Goldberger, J., Hinton, G. E., Roweis, S., & Salakhutdinov, R. R. (2004). Neighbourhood Components Analysis. *Advances in Neural Information Processing Systems*, *17*. Retrieved November 1, 2022, from https://papers.nips.cc/paper/2004/hash/42fe880812925e520249e808937738d2-Abstract.html

Hastie, T., & Tibshirani, R. (1996). Discriminant adaptive nearest neighbor classification [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(6), 607–616. https://doi.org/10.1109/34.506411

Henrydoss, J., Cruz, S., Rudd, E. M., Gunther, M., & Boult, T. E. (2017). Incremental Open Set Intrusion Recognition Using Extreme Value Machine. *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1089–1093. https://doi.org/10.1109/ICMLA.2017.000-3

Jain, L. P., Scheirer, W. J., & Boult, T. E. (2014). Multi-class Open Set Recognition Using Probability of Inclusion. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 393–409). Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-319-10578-9_26

Kaya & Bilge. (2019). Deep Metric Learning: A Survey. *Symmetry*, *11*(9), 1066. https://doi.org/10.3390/sym11091066

Kedem, D., Tyree, S., Sha, F., Lanckriet, G., & Weinberger, K. Q. (2012). Non-linear metric learning. *Advances in neural information processing systems*, *25*.

Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). Sphereface: Deep hypersphere embedding for face recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 212–220.

Lübbering, M., Gebauer, M., Ramamurthy, R., Bauckhage, C., & Sifa, R. (2022). Bounding open space risk with decoupling autoencoders in open set recognition. *International Journal of Data Science and Analytics*, *14*(4), 351–373. https://doi.org/10.1007/s41060-022-00342-z

Mendes Júnior, P. R., de Souza, R. M., Werneck, R. d. O., Stein, B. V., Pazinato, D. V., de Almeida, W. R., Penatti, O. A. B., Torres, R. d. S., & Rocha, A. (2017). Nearest neighbors distance ratio open-set classifier. *Machine Learning*, *106*(3), 359–386. https://doi.org/10.1007/s10994-016-5610-8

Nene, S. A., Nayar, S. K., & Murase, H. (1996). Columbia Object Image Library (COIL-20) [Technical Report CUCS-005-96].

Oza, P., & Patel, V. M. (2019). C2ae: Class conditioned auto-encoder for open-set recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2307–2316.

Ramanagopal, M. S., Anderson, C., Vasudevan, R., & Johnson-Roberson, M. (2018). Failing to Learn: Autonomously Identifying Perception Failures for Self-Driving Cars. *IEEE Robotics and Automation Letters*, *3*(4), 3860–3867. https://doi.org/10.1109/LRA.2018.2857402

Rudd, E. M., Jain, L. P., Scheirer, W. J., & Boult, T. E. (2018). The Extreme Value Machine [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(3), 762–768. https://doi.org/10.1109/TPAMI.2017.2707495

Scheirer, W. J., de Rezende Rocha, A., Sapkota, A., & Boult, T. E. (2013). Toward Open Set Recognition [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(7), 1757–1772. https://doi.org/10.1109/TPAMI.2012.256

Scheirer, W. J., Jain, L. P., & Boult, T. E. (2014). Probability Models for Open Set Recognition [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(11), 2317–2324. https://doi.org/10.1109/TPAMI.2014.2321392

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering [ISSN: 1063-6919]. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823. https://doi.org/10.1109/CVPR.2015.7298682

Shi, B., & Liu, J. (2018). Nonlinear metric learning for kNN and SVMs through geometric transformations. *Neurocomputing*, *318*, 18–29. https://doi.org/10.1016/j.neucom.2018.07.074

Shu, L., Xu, H., & Liu, B. (2017). DOC: Deep Open Classification of Text Documents. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2911–2916. https://doi.org/10.18653/v1/D17-1314

Simard, P., LeCun, Y., & Denker, J. (1992). Efficient Pattern Recognition Using a New Transformation Distance. In S. Hanson, J. Cowan, & C. Giles (Eds.), *Advances in Neural Information Processing Systems*. Morgan-Kaufmann. https://proceedings.neurips.cc/paper/1992/file/26408ffa703a72e8ac0117e74ad46f33-Paper.pdf

Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., & Liu, W. (2018). CosFace: Large Margin Cosine Loss for Deep Face Recognition [ISSN: 2575-7075]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5265–5274. https://doi.org/10.1109/CVPR.2018.00552

Weinberger, K. Q., Blitzer, J., & Saul, L. (2005). Distance Metric Learning for Large Margin Nearest Neighbor Classification. *Advances in Neural Information Processing Systems*, *18*. Retrieved November 1, 2022, from https://proceedings.neurips.cc/paper/2005/hash/a7f592cef8b130a6967a90617db5681b-Abstract.html

Zhang, H., & Patel, V. M. (2017). Sparse Representation-Based Open Set Recognition [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(8), 1690–1696. https://doi.org/10.1109/TPAMI.2016.2613924