

Novel Class Detection and Integration with Noise Removal and Classification on Acoustic Data

Aditya Madhavanur Prabhakaran
Engineering Data Sciences,
Robert Bosch Engineering and
Business Solutions Pvt Ltd
Bengaluru, India
MadhavanurPrabhakaran.Aditya
@in.bosch.com

Sheela Siddappa
Engineering Data Sciences,
Robert Bosch Engineering and
Business Solutions Pvt Ltd
Bengaluru, India
Sheela.Siddappa@in.bosch.com

Abstract—Novel class detection is employed in a wide range of applications like disease detection, fraud detection etc. In the absence of novel class detection, classification models force fit data of new classes into existing classes thereby rendering the overall solution incorrect. However, achieving novel class detection on complex, noisy, acoustic data with imbalanced classes is not trivial. In addition to discovering novel classes, minimizing the misclassification of known classes is also important. In this paper, a holistic approach to detect novel class(es) amidst noisy data yet maintaining high accuracy of classification is discussed. The technique adopts three-step approach, (i) noise removal from the data (ii) novelty detection based on grouping of known classes and (iii) integration with a multi-class classification model. The proposed method is applied on real-world, acoustic data captured from agricultural fields and realizes an accuracy of 91.6%, thus, making the overall solution apt and effective.

Keywords—novelty detection, multi-class classification, acoustic data, integration framework, acoustic noise removal

1. INTRODUCTION

Today, there exists a wide range of applications which require discovery of abnormal behavior and distinguish it from the normal state of the system. Examples include disease or disorder detection in medical data, disaster prediction from geological data, fraud detection in financial data, malicious activity detection in security systems and intrusion detection in video surveillance systems.

In some applications, domain knowledge is adapted to characterize the different abnormal states of the system and model how the corresponding signals look like - for instance, the behavior of pressure and flow rate signals is known when a pump is surging (failure mode). In some other cases, sufficient data and signals corresponding to the abnormal states are available to learn from them and characterize them. For example, events' history log can be utilized to learn how the system behaved around an event. The method is then trained on both the normal and abnormal states' signals to classify new signals. However, there exist many applications where it is either not possible to characterize all possible types of abnormalities or there exists very less abnormal data but sufficient normal data are available. In such cases, it is still required to discover abnormalities whilst learning normal data alone. Novelty detection methods achieve precisely this.

Moreover, supervised learning systems, which are trained on limited and fixed number of classes, may encounter new classes in the live data when deployed. In such scenarios, the system force fits the data into one of the trained classes and hence delivers an incorrect classification. This can be avoided if a novelty detection method is integrated to the learning system and classifies the data as normal (belonging to the trained classes of the learning system) or anomalous prior to classification.

Novelty detection is the method of recognizing a particular data or behavior, hitherto unknown, to be "abnormal". Novelty detection is often referred to synonymously with outlier detection. However, there exists a distinction between the two. In outlier detection, typically both normal (negative class instances) and anomalous data (positive class instances) are trained on by the method, but in novelty detection, only normal data are trained. Hence, novelty detection attempts to define the boundary of the normal class but does not learn the definitions of the anomalous class. So any point which lies outside the boundary of the normal class is classified as anomalous. The advantage with this approach is that it is more likely to detect different types of unseen anomalies since the method's definition of anomalies is not restricted to the characteristics of any particular types of anomalies. Novel class identification also differs from concept drift which is more extensively studied when mining large amount of streaming data [4][8][9][10]. Concept drift is the change in the pattern or behavior of the existing, known classes of data with time. As a result, the signals and features of the known classes change in such a way that the decision boundaries are crossed resulting in erroneous classification. But they do not address completely new classes of data which are, as yet, unknown to the classification model.

Many studies in literature have explored novelty detection methods and their applications. One of the earlier novelty detection methods, OLINDDA [1], employed k-means clustering to define the decision boundary of normal data as the union of all the clusters. However, as the paper itself notes, for most datasets apart from the simpler Iris and Balance Scale datasets, novelty detection was not as frequent and effective. The same authors published further improvements to the method [2][3]. In [3], the data points that lie outside the decision boundary of the normal class are held in short-term memory and a clustering approach identifies clusters in the unknown data by using cohesiveness and representativeness measures to

distinguish between outliers or noise and novel classes. However, the results betray the difficulty of the method in achieving this. In [4], the MineClass algorithm applies k-Nearest Neighbors (k-NN) and decision tree classifiers to cluster normal data and identify anomalies. However, the false alarm rates for certain data sets is still high. The method discussed in [5] also employs a decision tree based novelty detection but does not distinguish between outliers or noise and novel class. Also, it relies on attribute similarity-based clustering of points within each leaf node which can be ineffective in high-dimensional data with many classes as can be seen in the results. In [6], a semi-supervised ensemble of k-NN models identifies outliers and again a cohesiveness and separation measure (q-NSC), as proposed in [7], is applied to detect novel class among the outlier instances. However, the framework assumes that only one novel class is present in the test data at a time and does not handle the presence of multiple novel classes simultaneously. Even in the recent paper [8], the authors have defined cohesiveness and separation index MN-NSC based on Mahalanobis distance for the discovery of novel class among outlier instances identified by k-means clusters of normal data. These methods which attempt multi-class novelty detection based on cohesiveness measures are susceptible in complex, high variance data sets with many classes where the balance of cohesiveness and separation from multiple classes becomes difficult to achieve.

Surveying the past investigations on sound data, a couple of methods describe traditional machine learning approaches as in [9] and [10]. In [9], density of the underlying normal distribution of normal data is estimated using a mixture of Gaussian components and abnormal instances are identified based on a threshold of difference from the normal distribution. However, the study focusses on atypical acoustic events that are vastly different from the normal class in their acoustic modality. In [10], a two-layer representation of audio stream identifies anomalous events of car crashes and tire skidding in a noisy road traffic dataset. However, the method is trained on the defined anomalous events through cross validation and hence is limited in its learning ability of unknown types of anomalies. Anomaly detection in acoustic data is dominated by deep neural network based architectures, such as in [11] and [12], which require more data than traditional machine learning approaches. These papers also do not distinguish outliers or noise from novel classes. In [12], anomalous sounds are simulated by rejection sampling but the method is not robust to unknown types of anomalies in real world applications.

In the recent survey of one-class classification methods [13], the paper notes that Support Vector Data Description (SVDD) method has the highest area under the curve (AUC) with good learning and generalization ability and a smooth decision boundary as observed from the results on two different datasets - Gauss and Banana datasets. The authors also conclude that density based methods like Gaussian Mixture Model will be subject to "Dimension Disaster" in high-dimensional estimates while reconstruction based methods like k-nearest neighbor (k-NN) suffer from over-learning problems.

Few algorithms have been tested on data sets with many classes possessing large variance and class imbalances in a high dimensional feature space. Additionally, none of the methods have attempted to achieve both noise removal and novelty

detection and integrate them with classification on real-world sound data. This paper introduces a framework for the integration of (i) acoustic noise removal, (ii) discovery of novel classes and (iii) multi-class classification to provide a complete Machine Learning based solution on acoustic data. The paper also proposes a unique approach to novelty detection by utilizing the results of the multi-class classification model to group classes and produce novelty detection models. The proposed method is demonstrated through its successful application on complex, real-world sound data in the field of agriculture which has been deployed in-service for over 500 users.

The paper is structured as follows - details of the proposed method are discussed in section 2. The practical application of the method and the results obtained are discussed in section 3. The concluding remarks and future scope are noted in section 4.

2. PROPOSED METHOD

The proposed method is, henceforth, referred to as INNC which stands for Integrated Noise Removal, Novel Class Detection and Classification. INNC utilizes a gradient boosted decision tree ensemble model for removing acoustic noise from the ingested sound data. The remaining data is then passed through One Class Support Vector Machine (OCSVM) models which identify and separate out instances of novel classes and pass through the remaining data to a gradient boosted decision tree ensemble model for multi-class classification. This is depicted in the flowchart in Fig. 1. Since the data pertains to complex acoustic data of many classes, a unique method, henceforth referred to as GCNCD (Grouping of Known Classes for Novel Class Detection), of using the misclassifications in the known classes of the multi-class classification model is adapted to group the classes to train two OCSVM models. The discovery of novel class is then developed by combining the output from both these models.

In the absence of any data pertaining to unseen classes, the data flows according to the path laid out by the thick, black arrows as shown in Fig. 1. The raw sound data are ingested and relevant sound features are extracted. First, acoustic noise data are removed by utilizing a multi-class, gradient boosted decision tree classification model (M1). Once the data are classified into broad categories of sound, the data belonging to noise categories are discarded and only the data relevant to the application are processed further. It is important to note that since this noise model effectively seeks to classify data as relevant or noise, this method is applicable when the sound signatures of the novel class is similar to the signatures of the known classes but significantly different from the signatures of the noise data. This assumption holds good for many applications.

Data are then passed through multiple OCSVM models (M21 and M22) in the GCNCD module for discovery of novel classes. Data belonging to the known classes are then classified downstream using a multi-class, gradient boosted decision tree model (M3) and the predictions are outputted. The data belonging to novel classes are instead stored in buffer until the ground truth arrives. Then, the data belonging to both the known as well as novel classes (if present in sufficient numbers) are utilized for retraining to produce the updated version of model M3. The predictions of this updated model M3 form the basis to

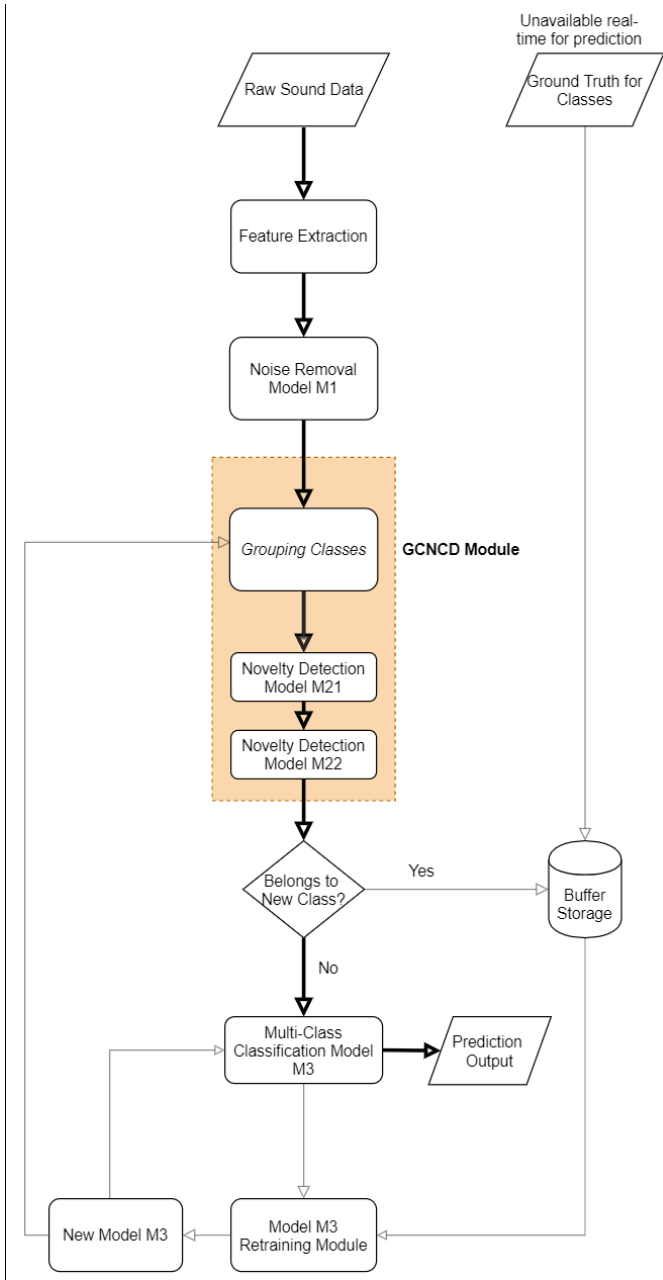


Fig. 1. Flowchart for INNC

group the classes suitably for building one OCSVM model per group. The updated models of M21 and M22 replace their existing versions in production.

2.1 Novel Class Detection

The first theoretical foundations for Support Vector Machines (SVM) were described in [14]. SVM finds the optimal hyperplane to classify linearly separable data. But SVM uses the "kernel" trick to project the data on to a higher dimensional space using the non-linear kernel function and thus achieve a non-linear decision boundary. The choice of kernels make SVM powerful and robust for problems with high dimensions of features and relatively less data. Additionally, since SVM uses

quadratic optimization, it avoids the problem of local minima unlike neural network based methods.

The One Class SVM (OCSVM) [15] adapts the SVM method to classify data belonging to a single class. The method attempts to separate the training data, belonging to the normal class, from the origin (where the abnormal class is assumed to lie) by finding the maximum margin hyperplane characterized by (w, ρ) , where w is the normal vector of the hyperplane and ρ is the bias (the distance from the hyperplane to the origin). This separation from origin is achieved by solving the below equation for every data point x_i [15],

$$\min_{w, \rho, \zeta} \quad \frac{1}{2} \|w\|^2 + \frac{1}{v_n} \sum_i \zeta_i - \rho$$

$$\text{subject to } (w \cdot \Omega(x_i)) \geq \rho - \zeta_i \text{ and } \zeta_i \geq 0$$

where $v \in (0,1)$, ζ_i are slack variables that are used to model the separation errors and Ω is a non-linear projection of the data into a high dimensional space. This is evaluated by a kernel function K , defined on two samples x_i and x_j , represented as feature vectors in some input space, through the dot product given below.

$$K(x_i, x_j) = (\Omega(x_i) \cdot \Omega(x_j))$$

A popular kernel is the Gaussian Radial Basis Function (RBF) kernel. This is given by:

$$K(x_i, x_j) = e^{-\left(\|x_i - x_j\|\right)^2 / c^2}$$

where c is referred to as kernel width parameter. It may be noted that the Gaussian RBF kernel was the one originally demonstrated in [15]. With this kernel, the method becomes equivalent to projecting the data on to a hypersphere centered at the origin in the higher dimensional feature space.

The Support Vector Data Description (SVDD) proposed in [16] also adapts the SVM formulation. But instead of trying to maximize the distance from the origin, the method seeks to find a spherically shaped boundary around the data. It has been shown in [16] that for a Gaussian RBF kernel, both the SVDD formulation as well as the OCSVM formulation become similar. The paper also notes that the polynomial kernel suffers from the large influence of the norms of the object vectors but the Gaussian RBF kernel is robust and for sparse and complex data sets, it outperformed Parzen density estimation and the Nearest Neighbor method.

In this paper, OCSVM formulation [15] with Gaussian RBF kernel has been adopted. The focus of this paper is the discovery of novel classes in complex data sets with many known classes and class imbalances. For such problems, it is demonstrated that grouping the known classes of a multi-class classification model to produce one OCSVM model per group for achieving novelty detection (GCNCD) outperforms the direct, as-is application of OCSVM by considering all the known classes as one, single class.

When there are many classes of data which are closely related or having overlapping signatures and also high variance, adopting a single one-class classification method produces a very large decision boundary on account of the large spread in

the known data. This then results in many test data points being classified as known, even though they belong to novel or unknown classes. On the other hand, using a multi-class approach for novel class detection such as density based clustering results in complex, narrow decision boundaries and, thereby, more misclassifications for the known classes. Thus, an approach is needed which balances the discovery of novel classes while also minimizing the misclassification rate of known classes.

This is precisely what the grouping of classes achieves in GCNCD. The results of a classification model (M3) drive the grouping. First, the classes are sorted in increasing order of their accuracy. For the class with the least accuracy, weighted by the sample size of the class, the confusion matrix provides the paired class into which the misclassifications of this class are highest. Both these classes are then grouped together. The process is repeated for the remaining classes until all the classes are distributed into groups. The number of groups to be formed depends on the data set and the number of classes. The groups are formed such that the misclassifications among the member classes are highest. This results in each group having data from classes having similar signatures, thereby, producing a relatively smaller decision boundary derived by the OCSVM model for that group, as compared to decision boundary encompassing data from all classes. For each data point in test data, if it falls outside the decision boundary of an OCSVM model, it is deemed as an anomalous data point by that model. This determination is conducted for every OCSVM model. Only the data that are classified as anomalous by every model are finally classified as belonging to novel classes.

The process is illustrated in Fig. 2 where a sample confusion matrix with 13 classes is shown. The resulting class groups by using GCNCD for this sample data look similar to those depicted in TABLE 1. To evaluate the efficacy of the proposed method, its application on a practical data set is described in the sections below.

3. EXPERIMENTS AND RESULTS

This section describes the implementation and results of experiments with INNC as applied on a practical, real-world sound data set. The data has been captured through mobile phones from agricultural fields as part of an innovative effort to utilize sound-based analytics to predict the farming activity being performed. This improves the farmer's productivity, without the need for any additional infrastructure.

3.1 Data Description

The data set described in this paper was prepared as part of a project collaboration between Robert Bosch Engineering & Business Solutions Pvt Ltd and Mimiro for providing field assistance to farmers in Norway. The project harnesses the power of digital solutions innovatively applied in the agricultural domain to improve the efficiency, productivity and competitiveness of farming. A total of 8505 sound files were collected by using mobile devices from a set of farms in Norway. The recordings correspond to the sound from the machinery deployed to perform the farming activity. Keeping in mind battery and data consumption of the mobile devices of the end users, each sound file is just one second long. This makes

the data set and its analysis challenging. Each file has been labelled by domain experts based on the activity that was performed while recording that sound file. The details of the data set (excluding noise files) are shown in TABLE 2 where the class imbalances can be noticed. There are a total of 17 classes in the data set out of which four classes are randomly assigned to the test set. In addition to these four classes, 10% of the data belonging to the remaining 13 classes are also added to the test set. This facilitates the evaluation of the novelty detection method on both known and novel classes. The training set comprises of the remainder 90% of the 13 classes. The experiment of picking the four unseen classes in the test set is repeated for eight different, random combinations and the method's performance is evaluated and benchmarked in all the eight cases.

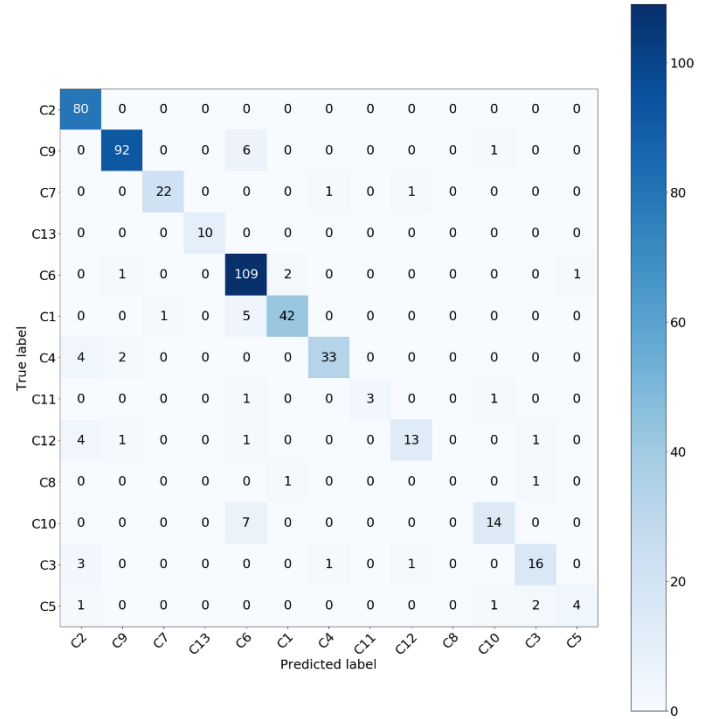


Fig. 2. Confusion matrix for sample data for illustration

TABLE 1. Grouping of classes for sample data

Group	Classes
Group 1	C1, C6, C9, C10, C11, C13
Group 2	C2, C3, C4, C5, C7, C8, C12

3.2 Feature Extraction

Once the raw sound data are ingested, it is required to extract acoustic features from the raw sound files that adequately characterize each farming activity and help distinguish between the different activities. A variety of spectral and temporal features were explored. These include Mel Frequency Cepstral Coefficients (MFCC), Chroma, Tonnetz, spectral bandwidth, spectral centroid, spectral roll-off, zero crossing rate and root mean square energy. After performing experiments with various

TABLE 2. Experiment Data Description

Class	Number of Audio Files
Fertilization	2179
Bailing	1135
Mowing	839
Raking	777
Ploughing	314
Sowing	250
Pesticide application	231
Treshing	210
Transportation bales	209
Aerate Gras	190
Harrowing	177
Self-loading wagon	155
Rolling	141
Dragging	139
Bale Wrapping	135
Combo fertilization and sowing	133
Forage wagon harvesting	50

combinations of the above features, MFCC was observed to be the most effective feature set in classifying the given data. Moreover, the addition of the other features to MFCC features produced a negligible improvement in the method's accuracy whilst increasing computation time as compared to using just the MFCC features. Hence, feature extraction is restricted to 20 MFCC features for further processing.

Mel Frequency Cepstral Coefficients (MFCC) are a very popular set of features that are commonly employed in different audio recognition applications, especially in speech recognition. MFCC was originally introduced in 1990 in [17]. MFCC is obtained by executing the following steps. First, a Short Term Fourier Transform (STFT) is applied on the audio signals sampled at a rate of 22050 Hz with buffer size (n_{fft}) set to 2048 samples and buffer overlap (hop_length) set to 512 samples. The power spectrum is then computed and filter banks are applied by using 50 triangular filters (n_{mel}) on a logarithmic scale called Mel scale. This helps to mimic the non-linear human ear perception of sound. Next, a direct cosine transform is applied on the Mel filter banks to obtain 20 MFCC features (n_{mfcc}).

3.3 Noise Removal

Before applying the GCNCD method on the data, a method to distinguish novel class data from the noise data is required. The typical approach for achieving this is to define some cohesiveness or density measure among the anomalous data - all anomalous data which are closer together in the feature space (cohesiveness) are classified as novel classes and the rest as outliers [7][8]. For sound data, acoustic noise is different from conventional outliers. The sources of noise may be varied and quite often cannot be identified by simple, rule based filtering. In this application, noise refers to sounds that are not associated with the farming activity. The data exploration revealed the main types of noise to be human speech, music, no sound (mute) and miscellaneous background noise like high frequency sounds etc. Since the acoustic characteristics of noise are quite different from those of the known and unknown classes of data, INNC adopts a two-step approach to achieve the

distinction between noise and novel classes. In the first step, a gradient boosted multi class decision tree ensemble model removes noise data. This model is trained on machine class label (useful data) and noise class labels of human speech, music, no sound and miscellaneous background noise. The data, labelled manually, contained 7264 machine files, 121 human speech files, 138 music files, 470 no sound files and 512 miscellaneous noise files. The gradient boosted decision tree ensemble, with 600 estimators, attained an accuracy of 92.7%. In the second step, only the machine sound files are sent to the GCNCD module to discover unseen classes of data as outlined in section 3.4.

3.4 Novel Class Detection with GCNCD

Once noise data are filtered out, the data are then passed through the GCNCD module. In GCNCD, two models, M21 and M22, each trained on one group of classes, check for the presence of novel class data. Only the data which are classified as novel class by both models are considered as belonging to novel classes and are kept aside while the remaining data are processed by a multi-class classification model (M3) to produce the prediction output.

To arrive at the optimal grouping of classes, the confusion matrix of model M3 is utilized to understand the misclassifications of each class. The signatures of classes, which show misclassifications into each other, possess a greater degree of overlap and similarity. Such closely related classes are grouped together as illustrated in Section 2.1. One OCSVM model is built per group and only the data that are predicted as anomalous by both the OCSVM models are finally classified as belonging to novel classes by the module. This technique (GCNCD) of grouping similar classes for novelty detection can be applied on multiple base algorithms. One-Class SVM has been chosen for illustrative purposes of this demonstration. For both the OCSVM models, Gaussian RBF kernel is adopted and the value of ν and γ are constant for all the eight test cases and are set to 0.01 and 0.001 respectively.

To benchmark the proposed method, its performance is evaluated against a single OCSVM model built on the whole training set with the same hyper-parameters (Gaussian RBF kernel with ν as 0.01 and γ as 0.001). The performance of k-means clustering, a popular algorithm for multi-class novelty detection, is also shown for comparison. The value of ' k ', which maximizes AUC as found experimentally, is set to 19. For every test case, each method is fit on the data with three different random states and the average is quoted in the results described below.

The class-wise test set accuracy is shown in TABLE 3 for GCNCD for test case 1 for illustrative purposes. The accuracy of the method for 'Known' classes is 73.1% on 651 data points and for 'Unknown' classes is 86.4% on 765 data points. The method balances the discovery of the novel classes as unknown with the recall of the known classes. However, the hyper-parameters defining the method and its decision boundary can be tuned depending on the objective - if the discovery of novel

TABLE 3. Class-wise Accuracy of GCNCD for Test Case 1

Class	Number of Anomalies	Total Points	Label	Accuracy
Raking	11	78	Known	85.9%
Combo fertilization and sowing	1	13	Known	92.3%
Rolling	2	14	Known	85.7%
Self-loading wagon	2	16	Known	87.5%
Harrowing	3	18	Known	83.3%
Ploughing	7	31	Known	77.4%
Dragging	7	14	Known	50.0%
Treshing	1	21	Known	95.2%
Forage wagon harvesting	0	5	Known	100.0%
Bailing	14	114	Known	87.7%
Mowing	22	84	Known	73.8%
Fertilization	92	218	Known	57.8%
Sowing	13	25	Known	48.0%
Bale Wrapping	106	135	Unknown	78.5%
Aerate Gras	147	190	Unknown	77.4%
Pesticide application	204	231	Unknown	88.3%
Transportation bales	204	209	Unknown	97.6%

classes or the cost of misclassification of the known classes is more important. The metric chosen for comparison of the different methods is area under the curve (AUC) since the data is imbalanced and the objective is to estimate the robustness of the method with respect to both the known as well as unknown classes. The AUC provides a single, aggregate measure to compare the performance of machine learning algorithms and is a recommended evaluation measure in literature as concluded in [18] and [19]. The AUC of GCNCD along with the single OCSVM method and k-means clustering method for the eight different test cases are depicted in Fig. 3. The AUC for the GCNCD and single OCSVM methods consistently outperform the k-means clustering method. Among the OCSVM methods, it can be seen that the GCNCD method again outperforms the single OCSVM method consistently, albeit by a smaller margin. The margin varies from 0.06 for one combination of classes in the test set to 0.02 for another. This establishes that grouping the classes to produce multiple models for novelty detection helps improve the method's performance and robustness.

The recall rate of the unknown classes, also called as detection rate, for the different methods are compared as shown in Fig. 4.

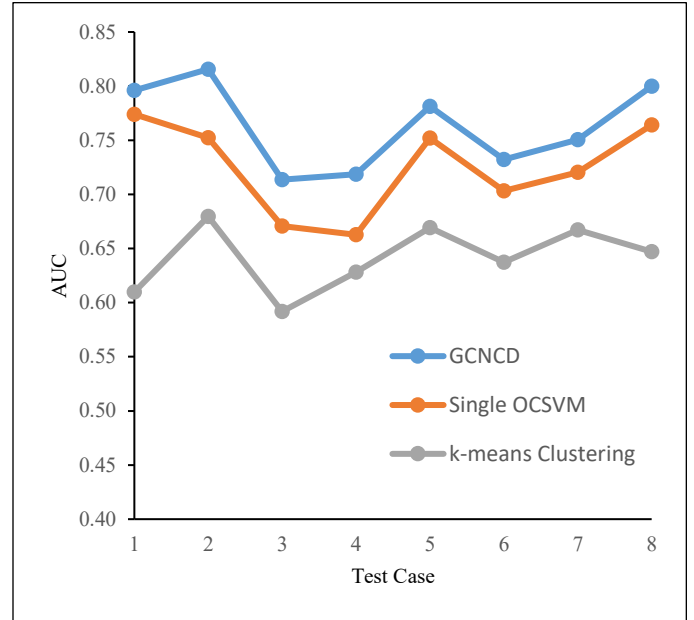


Fig. 3. AUC score for the different novelty detection methods

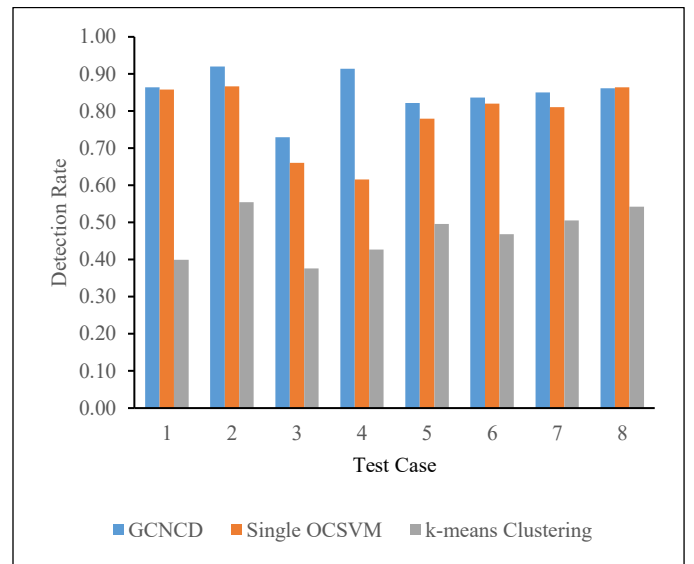


Fig. 4. Detection rate for the different novelty detection methods

The detection rate measures the effectiveness of the method in accurately discovering data points as belonging to novel classes. The detection rate of GCNCD is 0.85 averaged across eight test cases while that of the single OCSVM method and k-means clustering method are 0.78 and 0.48 respectively. This shows that the SVM-based methods are much more effective at discovering and classifying new data points as belonging to novel classes compared to the clustering method. The clustering method struggles to differentiate the closely related sound signatures of the known and the unknown classes in this data set. Also, the grouping technique of GCNCD introduced in this paper helps boost the average detection rate by 0.07 in comparison to the vanilla OCSVM method. This result, coupled

with the average outperformance of 0.03 in AUC, establishes the benefit of using the results of a multi-class classification model to optimally group classes and achieve novelty detection in such challenging data sets.

3.5 Multi-Class Classification

Finally, the data predicted as belonging to the known classes by the novel class detection module are classified as one among the known 13 classes for the final prediction output. A gradient boosted decision tree ensemble with 900 estimators achieves the multi-class classification. The model is tuned using the popular stratified cross-validation approach. To illustrate the significance of the noise removal and novelty detection achieved prior to the classification using INNC, the results of this classification model (M3) are compared with and without the noise removal method (M1) and novelty detection method (M21 and M22). The classification accuracy for these two scenarios for all the eight test cases is shown in Fig. 5. It is clearly seen that INNC, by virtue of filtering out the acoustic noise data and also the novel classes' data, achieves an average improvement of 14.2% in the prediction output accuracy of the classification model M3 over simply applying the classification model on the features extracted from raw sound data. This result testifies the importance of a framework to integrate these different steps and their impact on the accuracy of the overall solution.

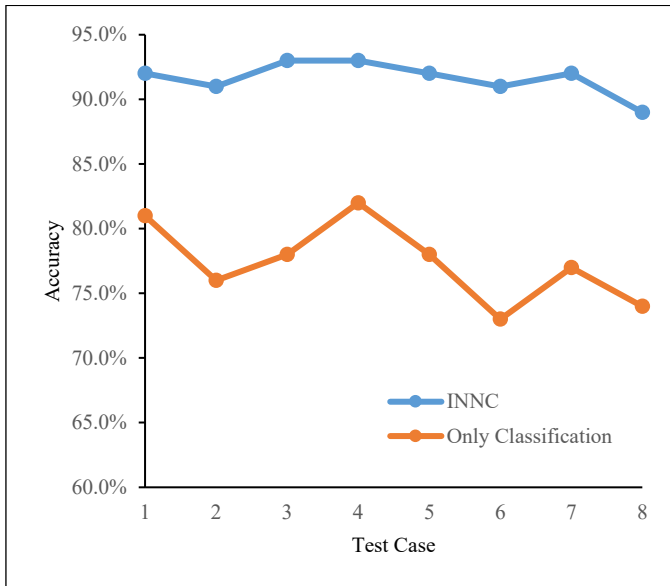


Fig. 5. Classification accuracy with and without the Noise Removal and Novelty Detection methods

4. CONCLUSIONS

In this paper, a framework was introduced to integrate and achieve noise removal and novelty detection with multi-class classification. The proposed method, INNC, is demonstrated on a challenging, real-world acoustic data set pertaining to agricultural activity with many classes, class imbalances and related sound signatures. In the novelty detection module, a new

technique, GCNCD, of grouping the known classes based on the misclassification results of the multi-class classification model was introduced to produce one novelty detection model per group. The technique is applied on the one-class SVM model and benchmarked against the standard OCSVM implementation as well as k-means clustering method. This new method, GCNCD, results in a modest improvement in AUC compared to the single OCSVM method whilst achieving an enhanced novel class detection rate of 0.07 on average over the single OCSVM method. Both the SVM-based methods achieve better AUC scores than the clustering method and the margin of outperformance is significantly wider for the detection rate. The results show the efficacy of the GCNCD method in discovering novel classes. Finally, the prediction accuracy of the multi-class classification model is compared with and without the noise removal and novelty detection methods. An average increase of 14.2% in the prediction accuracy establishes the need and significance of the proposed framework INNC on noisy acoustic data sets in the presence of unseen classes. In future work, the INNC framework will be extended to other domains and data sets and the GCNCD method will be explored on different algorithms including neural-network based approaches.

ACKNOWLEDGMENTS

The authors express their gratitude to Robert Bosch Engineering & Business Solutions Pvt Ltd, India and Mimiro, Norway for their support, collaboration and encouragement.

REFERENCES

- [1] E.J. Spinosa, A.P.L.F. de Carvalho and J. Gama, OLINDDA: A Cluster-Based Approach for Detecting Novelty and Concept Drift in Data Streams. In 22nd Annual ACM Symposium on Applied Computing (SAC 2007), pages 448–452. ACM, 2007.
- [2] E.J. Spinosa, A.P.L.F. de Carvalho and J. Gama, Learning Novel Concepts: Beyond One-Class Classification with OLINDDA, In International Workshop on Knowledge Discovery from Ubiquitous Data Streams (IWKDUDS 2007), during the 18th European Conference on Machine Learning (ECML 2007) and the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2007), pages 13–24, 2007.
- [3] E.J. Spinosa, A.P.L.F. de Carvalho and J. Gama, Novelty detection with application to data streams, In Intelligent Data Analysis 2009.
- [4] M. M. Masud, J. Gao, L. Khan, J. Han, and B. Thuraisingham, Integrating novel class detection with classification for concept-drifting data streams, in Machine Learning and Knowledge Discovery in Databases, pp. 79–94, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [5] Farid DM, Rahman CM, Novel class detection in concept-drifting data stream mining employing decision tree. In Proceedings of the 7th international conference on electrical computer engineering (ICECE), pp 630–633, 2012.
- [6] A. Haque, L. Khan and M. Baron, SAN: Semi-supervised adaptive novel class detection and classification over data stream, 30th AAAI Conf. Artif. Intell., pp. 1652–1658, 2016.
- [7] M. M. Masud, J. Gao, L. Khan, J. Han, and B. Thuraisingham, Classification and novel class detection in concept-drifting data streams under time constraints. IEEE Trans. Knowl. Data Eng. 23(6):859–874, 2011.
- [8] X. Li, Y. Zhou, Z. Jin, P. Yu, and S. Zhou, A classification and novel class detection algorithm for concept drift data stream based on the cohesiveness and separation index of mahalanobis distance. In Journal of Electrical and Computer Engineering, Hindawi, 2020.

- [9] S. Ntalampiras, I. Potamitis and N. Fakotakis, Probabilistic novelty detection for acoustic surveillance under real-world conditions, *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 713-719, 2011.
- [10] P. Foggia, N. Petkov, A. Saggese, N. Strisciuglio and M. Vento, Audio surveillance of roads: A system for detecting anomalous sounds, *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 279-288, 2016.
- [11] Marchi, E., Vesperini, F., Eyben, F., Squartini, S., Schuller, B., A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional lstm neural networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1996–2000, IEEE 2015.
- [12] Y. Koizumi, S. Saito, H. Uematsu, Y. Kawachi, and N. Harada, Unsupervised detection of anomalous sound based on deep learning and the Neyman–Pearson lemma, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 1, pp. 212–224, 2018.
- [13] S. Wenzhu, H. Wenting, X. Zufeng and C. Jianping, "Overview of one-Class Classification," *IEEE 4th International Conference on Signal and Image Processing (ICSIP)*, pp. 6-10, 2019.
- [14] C. Cortes and V. Vapnik, Support-vector network, *Machine Learning*, vol. 20, pp. 273-297, 1995.
- [15] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt. Support vector method for novelty detection. In *Proceeding of the 12th International Conference on Neural Information Processing Systems*, pages 582–588, 1999.
- [16] R.P.W. Tax, D.M.J. and Duin. Support vector data description. *Machine learning*, 54(1):45–66, 2004.
- [17] Steven B Davis and Paul Mermelstein, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences In *Readings in speech recognition*, Elsevier, pages 65–74, 1990.
- [18] J. Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms", *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 3, pp. 299-310, Mar. 2005.
- [19] A.P. Bradley, "The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms," *Pattern Recognition*, vol. 30, pp. 1145-1159, 1997.