Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 4118 – 4133
© TÜBİTAK
doi:10.3906/elk-1901-166

An automated snick detection and classification scheme as a cricket decision review system

Aftab KHAN[1*], Syed Qadir HUSSAIN[1], Muhammad WALEED[1]
Ashfaq KHAN[2], Umair KHAN[3]

[1]Department of Computer Systems Engineering, University of Engineering and Technology, Peshawar, Pakistan
[2]Department of Mechanical Engineering, University of Engineering and Technology, Peshawar, Pakistan
[3]Department of Computer Science, COMSATS Attock, Attock, Pakistan

Received: 23.01.2019 • **Accepted/Published Online:** 24.05.2019 • **Final Version:** 26.11.2019

Abstract: Umpire decisions can greatly affect the outcome of a cricket game. When there is doubt about the umpire's call for a decision, a decision review system (DRS) may be brought into play by a batsman or bowler to validate the decision. Recently, the latest technologies, including Hotspot, Hawk-eye, and Snickometer, have been employed when there is doubt among the on-field umpire, batsman, or bowlers. This research is a step forward in gaging the true class of a snick generated from the contact of the cricket ball with either (i) the bat, (ii) gloves, (iii) pad, or (iv) a combination of bat and pad. Preprocessing included noise removal from the snick audios using the Audacity program. Machine learning-based classification is achieved by training neural networks with audio features of the snick. A support vector machine was employed to enhance the classification system. Twenty-one features comprising time and frequency domain characteristics were compiled. After one-way analysis of variance was employed and multicomparison analysis was performed, a set of seventeen features was selected and utilized to increase the accuracy of the proposed DRS. The system was trained on indigenous snick data gathered by ourselves and then tested for real cricket snick scenarios. The system achieves a classification rate of 98.3 percent for the self-collected data while presenting an accuracy of 85.7 percentage for the real cricket snick scenarios. The research also developed a dataset of snick audios comprising 132 signals for the four categories to aid researchers working in the same field.

Key words: Cricket, snick detection, classification, neural networks, support vector machine, decision review system, audio processing, Snickometer, Hotspot, Hawk-eye, Ultra-edge

1. Introduction

Cricket is one of the most played games at the international level with 12 full members and 93 associate member countries. It is a game played between two teams each of 11 players. There are different formats of cricket but the International Cricket Council (ICC) main formats are Test Cricket, One Day Cricket, and the most popular, T20 cricket. In cricket, there are two on-field umpires, a match referee, and an off-field umpire, also called the third umpire. The two on-field umpires stand at the square leg and behind the stumps at the bowler's end. The decision of a leg before wicket (LBW), no ball, wide ball, bat or pad contact, and bat edge is given by the umpire at the bowling end while run outs and stumps are judged by the square leg umpire. If the on-field umpires have any doubts, a referral is given to the third umpire, who can provide a robust judgment with the aid of various technologies including slow-motion video, audio systems, and ball trajectory prediction systems.

*Correspondence: aftab.khan@uetpeshawar.edu.pk

An erroneous decision by the on-field umpire will tilt the output of the game in the favor of an opposing team. Within the previous couple of years, the ICC introduced a decision review system (DRS) that permitted a player to challenge the on-field umpire's decision and forward it for review to the third umpire. The third umpire, with the assistance of technologies like Snickometer, Hotspot, Hawk-eye, and Dart-fish, can make an improved call for a decision. Snickometer and Hotspot technology are currently being used for decisions related to the contact of the ball with either the bat or the pad. These two technologies are criticized for their accuracy within the world of cricket.

The aim of this research work is to design and investigate an efficient automated snick detection method as a cricket DRS. The focus of this paper is to employ an efficient DRS for snick detection by taking audio recordings from the contact between ball and bat, ball and pad, ball and gloves, and bat and pad and later utilizing signal processing and machine learning techniques to produce a correct decision.

Section 2 presents the related work in the field of snick detection and classification. Section 3 explains the proposed scheme with a detailed explanation of the different stages and transformations. The experimental setup and datasets of self-collected Snickometer samples and real cricket snicks are given in Section 4. Section 5 presents the feature selection, extraction, and classification details. Results for self-collected snicks and real cricket snicks are given in Section 6, followed by concluding remarks in Section 7. The following sections provide a critical analysis of the relevant literature and techniques.

2. Previous research

Snickometer technology analyzes sound captured from stump microphones to predict a snick. The snick is defined as the brief and sudden contact of the ball with the edge of the bat.

It is typically associated with slow-motion video to locate whether or not the ball touched the bat while it passed on, as shown in Figure 1. The sound varies for various material contacts, like whether with the bat, pad, helmet, glove, or the pitch itself. Sometimes it is difficult to assess the source of the sound. A player might request a review of the on-field umpire's call. In such eventualities, the third umpire might use Snickometer data to minimize the uncertainty of the sound's source.

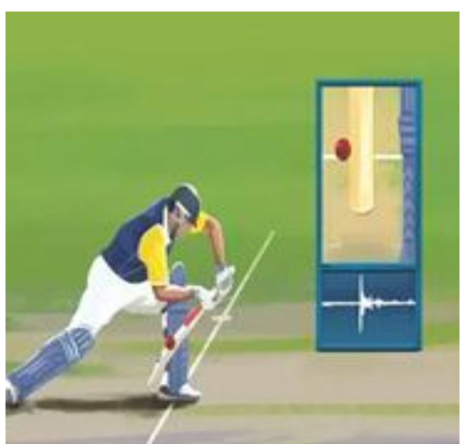


Figure 1: The Snickometer technology employed in cricket DRS.

Among the prevailing techniques, Rock et al. in [1] probed the utilization of wavelets for edge detection in cricket signals. Wavelet-based features were extracted and an artificial neural network (ANN) system was trained on them. The ANN classifier was trained to tell the varied classes apart. The accuracy of the system was 97.5 percent on raw testing data. Rock et al. in [2] supplemented their classification system's efficacy by centering their attention on extraction of wavelet domain descriptors for the short period of the snicks. Additionally, they made use of time domain-based higher order statistical features, such as skewness and kurtosis, and were able to achieve a classification rate of 100 percent on unprocessed testing data.

The five features taken from the wavelet domain comprised the standard deviation, maximum correlation coefficient, pseudo-frequency, skewness, and kurtosis [1]. Ting et al. in [3] utilized a time-frequency feature set to segregate among three snick signals originating from a bat, pad, or glove. Their scheme endows a basic foundation in classifying snick signals. Nonetheless, it is marred in performance by low quality and quantity of the available features. Furthermore, the system is limited due to nonutilization of machine learning abilities and the decisions are made by empirically categorizing the ranges for features values. Nevertheless, it is envisioned that their system can be enhanced by using more features and a fuzzy decision system [4].

Though this research centers on the detection and classification of snicks in cricket audio, the literature is severely limited in this field. Only a handful of researchers have focused explicitly on cricket snick classification. Most of the research related to the research problem is in the domain of audio classification in various scenarios, e.g., classifying speech [5], environmental data [6], music [7], birds [8], vehicles [9], lungs and ECG signals [10], fruits [11] and environmental scenarios [4], to name a few. An analysis of the current literature demonstrates the limited research in the field of cricket snick detection due to the various challenges it presents, including the acquisition of snicks and real-life testing. The current research is a step forward in designing and developing a novel snick detection scheme based on time and frequency domain features for robust classification. The following section presents the details of the proposed scheme.

3. Proposed machine learning-based snick classification system

This research work presents a novel DRS for snick detection and classification using machine learning techniques such as ANNs and support vector machine (SVM) [12]. Figure 2 shows the flowchart of the proposed scheme. The snick detection and classification is performed in the followings steps.

- All the snick audio signals are loaded into the system.
- Noise profile is estimated and noise from the audios is removed.
- The filtered signals are fed to the feature extraction module, which generates the signal attributes in time, frequency, and cepstral domain.
- The attributed signals for all the snick audios are input to the classification system.
- The machine learning algorithm trains on some one portion of the data and tests the other part.
- For real-life cricket snicks, the machine learning system trained in the step above is utilized to classify the signals. No training is performed.

The experimental setup, collection of datasets, and results are as presented as follows.

4. Experimental setup and datasets

A dataset was required to gauge the accuracy of the proposed system. Unfortunately, a literature survey revealed that no such collection of snick audio signals was available. In this regard, a dataset was self-collected for training and testing purposes. The section below provides the details of the datasets used in evaluating the quality of the proposed system.

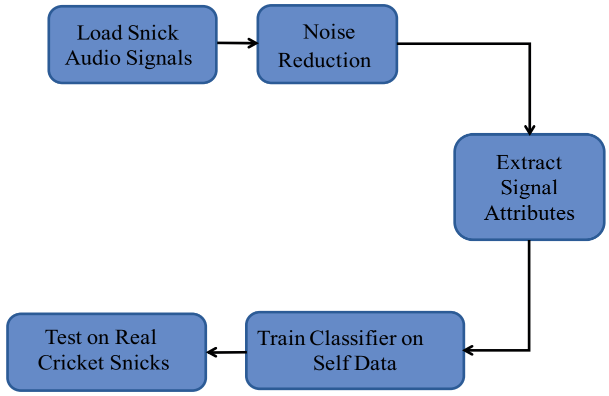


Figure 2: Overview of the proposed snick classification system.

4.1. Datasets

The Snickometer data were obtained from two sources: (i) self-collected data and (ii) data from real cricket matches.

- **Self-collected data**: This included the recording of snicks through a microphone for four classes: ball on bat, ball on pad, bat-pad mix, and ball on glove. The data were recorded in a closed room with an almost negligible level of noise. Noise removal techniques were employed to reduce the inherent noise in the recorded data.
- **Real cricket snicks**: The proposed system's robustness was evaluated by exercising it for snick audios selected randomly from different cricket matches. The aim was to test whether the proposed system's decision ability coincided with the ground truth in real scenarios and whether or not the on-field umpire's decision was correct or incorrect.

Prior to testing the snick signals, noise removal was applied as a preprocessing stage to reduce the level of noise inherent in the signal. Details regarding this are provided in the section below.

4.2. Noise removal with Audacity

Additive white Gaussian noise (AWGN) is assumed. Noise information is collected by estimating the noise properties from a specified smooth region. One such method is the mean absolute deviation technique (MADT) [13]. Assuming the noisy 1-D signal is represented by f with n samples and its mean by μ_f, then the mean deviation (MD) is presented as

$$MD = \frac{1}{n}\sum_{i=1}^{n}|f(i,j) - \mu_f|, \tag{1}$$

where

$$\mu_f = \frac{1}{n}\sum_{i=1}^{n} f(i,j). \tag{2}$$

For an AWGN model, the variance is estimated as $\sigma = 1.253 \times MD$. Once the variance-based profile of the noise is estimated, Bayesian inference is employed to reduce the noise from the complete signal. Let g represent the denoised image at this stage. The signals captured by the recording medium had inherent noise, which was removed by MADT using the Audacity program [14]. The Audacity program aids in the removal of noise from a signal by first analyzing the noise, generating a profile for the noise, and then removing it from the noisy signal. Figure 3 shows the noise removal procedure using the Audacity program. For the self-collected data, first a signal was uploaded to the Audacity program. A portion of the signal was selected and the noise reduction option was opted to analyze the noise profile as shown in Figure 3a. Once the noise profile is generated, the complete signal is selected and the noise reduction option is selected from the effect menu. This removes the noise from the noisy signal as in Figure 3b using the profile generated earlier and a noise-reduced version of the signal is obtained as shown in Figure 3c. Once the signals passed through the noise reduction stage, features were extracted and afterwards utilized for classification. The section below details the methods and techniques employed in feature selection.

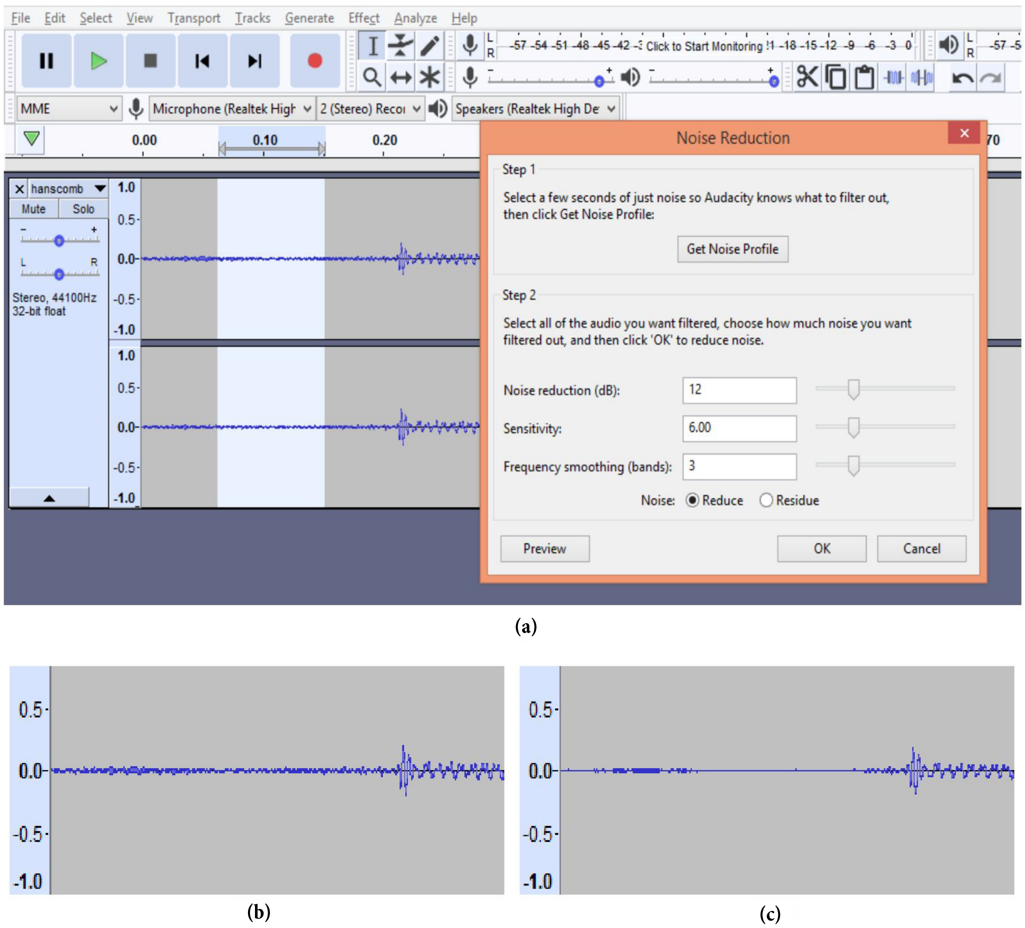


Figure 3: (a) The noise reduction menu in the Audacity program and a selected region of the noisy signal for profile analysis; (b) and (c) depict the signal before and after noise reduction, respectively.

5. Features extraction and classification

In order for a computer system to differentiate objects into various categories, it passes through a learning system that includes training and testing. Just like a child is shown pictures of apples and oranges to learn about fruits and then tested for categories, the computer brain based on ANN trains on a set of features of the objects and is then tested for validation. The section below details the features selected for identifying the objects.

5.1. Feature selection

Feature selection for the classification included a selection of time and frequency domain attributes that would support segregating the signals into various classes. Table 1 presents the time and frequency domain features used for classification in this research work. Some of the features are generic to any signal while others were selected from https://www.audiocontentanalysis.org. One way to select the best features is by picking those features that have variance higher than a particular threshold value.

Table 1: Time and frequency domain features for classification. A total of 21 features were selected, which were reduced to 17 using one-way ANOVA and multicomparison techniques. The nonnormalized mean values for each feature and category are also shown.

S. no.	Feature	Ball on bat	Ball on glove	Bat-pad mix	Ball on pad	SST (total variance)	P-value
Time domain features							
1	Energy	2.89E+02	7.84E+00	1.32E+02	2.73E+02	1.25E+06	6.03E-25
2	Autocorrelation coefficient	5.74E-01	4.17E-01	4.69E-01	6.48E-01	4.49E+00	2.24E-06
3	Maximum autocorrelation coefficient	2.80E-01	1.45E-01	2.50E-01	2.80E-01	5.15E-01	7.89E-17
4	Peak envelope	2.01E+00	8.76E+00	6.94E+00	2.68E+00	5.30E+02	5.35E-31
5	Root mean square	2.19E+01	2.89E+01	2.67E+01	2.36E+01	7.35E+02	9.72E-24
6	Standard deviation	8.27E-02	3.67E-02	4.81E-02	6.99E-02	3.21E-02	8.30E-24
7	Kurtosis	1.17E+02	3.03E+02	1.76E+02	1.52E+02	1.78E+06	3.87E-09
8	Zero crossing rate	6.49E-02	5.94E-02	5.06E-02	4.93E-02	3.43E-03	1.85E-26
Frequency domain features							
9	Power	1.20E+07	3.22E+05	7.79E+06	1.62E+07	4.35E+15	5.99E-21
10	Tonal power ratio	3.24E+02	3.20E+01	2.62E+02	2.50E+02	1.37E+06	3.69E-22
11	Spectral skewness	7.95E+00	6.52E+00	8.39E+00	8.22E+00	3.23E+02	8.83E-06
12	Spectral kurtosis (data)	8.48E+01	5.76E+01	1.02E+02	9.23E+01	1.53E+05	8.76E-06
13	Spectral kurtosis (pdf)	2.57E+05	4.16E+05	7.90E+05	3.64E+05	3.25E+12	9.20E-24
14	Spectral centroid	2.53E+02	2.31E+02	2.04E+02	2.23E+02	7.47E+05	1.04E-01
15	Spectral crest factor	2.46E-03	2.61E-03	3.16E-03	2.16E-03	1.29E-04	2.21E-03
16	Spectral decrease	2.30E-03	3.15E-03	2.10E-03	2.23E-03	1.21E-04	6.16E-05
17	Spectral flatness	1.41E-01	8.82E-02	6.55E-02	1.40E-01	1.91E-01	1.33E-14
18	Spectral roll off	1.89E+03	1.43E+03	1.25E+03	2.01E+03	1.98E+07	1.50E-13
19	Spectral slope	5.35E-08	2.29E-08	1.60E-08	2.98E-08	2.45E-14	1.36E-19
20	Spectral spread	3.76E+02	2.57E+02	2.16E+02	3.71E+02	9.07E+05	2.99E-14
21	Mel-frequency cepstral coefficients	2.15E+01	2.48E+01	2.43E+01	2.19E+01	1.87E+02	8.23E-26

In this research, one-way ANOVA was employed and the ANOVA P-value for the 21 features was computed against the data of four classes that we collected earlier. The P-value aids in gaging and determining the statistical significance of the means of groups. A comparison of the P-value to the significance level assesses whether equal means for a population exist or not. The significance level, usually represented by α, has a value of 0.05. A significance level of 0.05 indicates a 5 percent risk of concluding that a difference exists when there is no actual difference. If the P-value is $\leq \alpha$, then the differences between some of the means are statistically significant. Otherwise, the differences between the means are regarded as statistically insignificant.

Instead of selecting the features based on variance and P-values, a multiple comparisons technique using one-way ANOVA was conducted to select those features for which the combined multigroup P-value was less than a threshold value of 2. This way the feature set was reduced to 17 features, removing spectral crest factor, spectral skewness, spectral decrease, and spectral centroid. Table 1 presents the mean values for the features for the four classes followed by the total variance (SST) and the respective P-value. The total sum of squares, labeled SST, or total variance of the data is the accumulation of the sum of squares due to the between-groups effect (SSR) and the sum of squared errors (SSE). The SST for data y is given by Equation 3 as follows:

$$\underbrace{\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2}_{SST} = \underbrace{\sum_j n_j (\bar{y}_{.j} - \bar{y}_{..})^2}_{SSR} + \underbrace{\sum_i \sum_j (y_{ij} - \bar{y}_{.j})^2}_{SSE}, \tag{3}$$

where $\mathrm{n_j}$ is the sample size for the jth group, with $j = 1, 2, ..., k$ ensembles.

Once the features were selected, the next stage was to train a machine learning system to correctly recognize the signals and partition them into their respective classes. Such a system would then be employed for real cricket snick classification.

5.2. Classification

The aim of classification is to categorize the data or objects into respective groups though recognition, differentiation, and understanding. The classification system learns the category of each signal on the basis of its features and then groups a new signal into the category that matches it the best. Once the features are extracted from the noise-filtered snick signals, they are supplied to a classification system. In this research, the selected features were classified by ANN and SVM classifiers.

Our proposed system utilizes classification techniques from both the classical supervised models and neural networks. Among the classical supervised classification algorithms, SVM is one of the most popular algorithms; although it is mostly used for binary classification, the kernel trick transforms low-dimensional data to higher-dimensional space, creating the decision boundary between classes.

When it comes to neural networks, MLP is a deep neural network comprising three or more layers and is mostly applied to supervised learning problems. MLP classifies data points by learning the correlation between the input and output pairs, adjusting weights through a feedback mechanism. This results in a complex classification model as compared to the classical supervised learning algorithms. To draw a comparison between neural networks and classical classification models, we trained and tested our system on MLP and SVM.

ANNs are based on a collection of nodes called artificial neurons [15]. The data are processed one element at a time and learning occurs through comparison of classification of the data with the ground truth or real class.

In feedback-type ANNs, the errors are fed back into the neural network, which are used to update the neuronal weights and modify the network to match the true class. Figure 4 shows an example of the MLP-based ANN utilized in this research work.

The model comprises an input layer, output layer, and one hidden layer in between the two layers. A rectified linear function is used at the hidden layer while the softmax function is used at the output layer. The

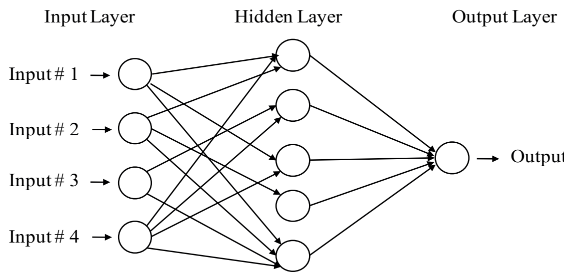


Figure 4: A hypothetical example of the multilayer perceptron (MLP) network used for classification [16].

number of nodes at the hidden layer are kept at 2/3 of the input layer size with the addition of the output layer size.

In MLP architecture, a single neuron can be mathematically represented as follows:

$$o = \tau(\beta + \sum_{i=1}^{n} w_i(\alpha)_i), \tag{4}$$

where o is the output, $(\alpha)_{\mathrm{i}}$ are n inputs, and w_i are summation weights. τ represents the activation function for the artificial neuron and β is the bias term that controls the neuron activation. By putting these neurons together, a network of neurons can be generated and regarded as a layer.

These layers then combine to form the MLP given below:

$$o_j^k = \tau_j^k(\beta_j^k + \sum_{i=1}^{n_{k-1}} w_{i,j}^k(\alpha^{k-1})_i), \tag{5}$$

where $k-1$ is the input vector and k is the output vector of the next iteration. j represents a neuron in the MLP.

The goal of the MLP is to reduce the error at each node by estimating the node value (NV) corresponding to the ground truth (GT) using a gradient descent scheme. Mathematically, this can be stated as:

$$\varepsilon = \frac{1}{2}\sum_j (GT - NV)^2, \tag{6}$$

and the weight update rule is:

$$w_j = w_j - \eta\frac{\partial\varepsilon}{\partial w_j}. \tag{7}$$

Initially, a statistical classification system based on the MLP ANN was utilized. The ANN system was unable to completely classify the four classes due to its inability to segregate datasets that are almost nonseparable. To overcome the inaccuracy of the ANN-based classification system, a support vector machine (SVM) type classifier was used. SVMs are more accurate when it comes to the classification of nonseparable datasets with several attributes specifically of higher dimensionality as compared to an ANN-based system [12]. Figure 5 shows the ability of the SVM to classify datasets that are not linearly separable.

SVMs for nonlinear classification were created by Boser et al. in [18]. The current method used in this research utilizes a soft-margin SVM for nonlinear classification. The nonlinear classification is done by adding

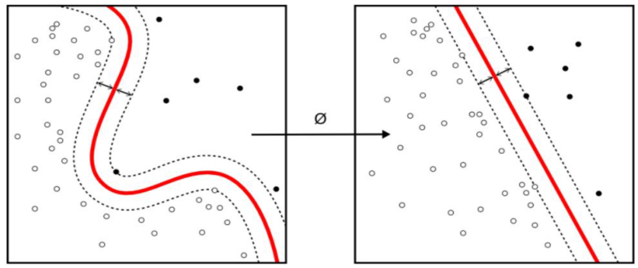


Figure 5: The ability of the SVM to classify datasets that are not linearly separable [17].

slack variables, relaxing the margin known as hinge loss function. SVM minimizes the following equation:

$$\left[\frac{1}{n}\sum_{i=1}^{n}\max(0, 1 - y_i(w.x_i - b))\right] + \lambda\|w\|^2, \tag{8}$$

where x is a data point to be classified, b is the bias, and λ is the regularization parameter controlling the hard/soft margin. n presents the total data points while the function y is a yield function with a value of -1, 1 for incorrect and correct classifications, respectively.

SVMs are limited to only classifying data into two classes and as such need to be reprogrammed for multi-class classification. There are several methods to achieve this, including the method of reducing the single multi-class SVM to a multiple binary classification [17, 19]. The multiclass SVM (MSVM) utilized here is an extension of the version coded by Neuburger, available at https://www.mathworks.com/matlabcentral/fileexchange/39352-multi-class-svm, which computes the one vs. one (OVO) classification and not the one vs. all (OVA) classification. OVA is less sensitive to imbalanced datasets as compared to OVO. This is dependent on the number of classifiers one has to train, which in turn affects the decision boundary. For example, OVA trains a single classifier for each class, assuming for a class x that x labels are positive and others are negative, rendering generic SVMs infeasible at times due to imbalanced datasets. OVO trains a distinct classifier for each label, rendering it less sensitive with the imbalanced dataset issues. The following section presents the results and analysis of the simulations of the proposed system.

6. Results and discussion

The proposed system was validated by testing it on the datasets obtained before. Results for classification of the snick signals for the self-collected and real cricket snicks are given below.

6.1. Testing on self-collected data

The proposed system was initially tested on data gathered by us for the four classes of snicks resulting from contact of the cricket ball with (i) the bat, (ii) gloves, (iii) pad, or (iv) a combination of bat and pad. To capture these sounds, a microphone was placed near the bat and snicks were generated by hitting the ball on the bat, glove, pad, and bat and pad together.

Figure 6 shows samples from the dataset. Previously no such dataset existed for evaluating a DRS as per our knowledge. This research work resulted in the development of such a type of dataset to facilitate researchers working in the same field. The dataset is available by making a request to the corresponding author.

The ANN-based classification system was trained on the self-collected data using the 17 features extracted earlier out of the 21 present. Figure 7a presents the confusion matrix for the classification of the self-dataset in

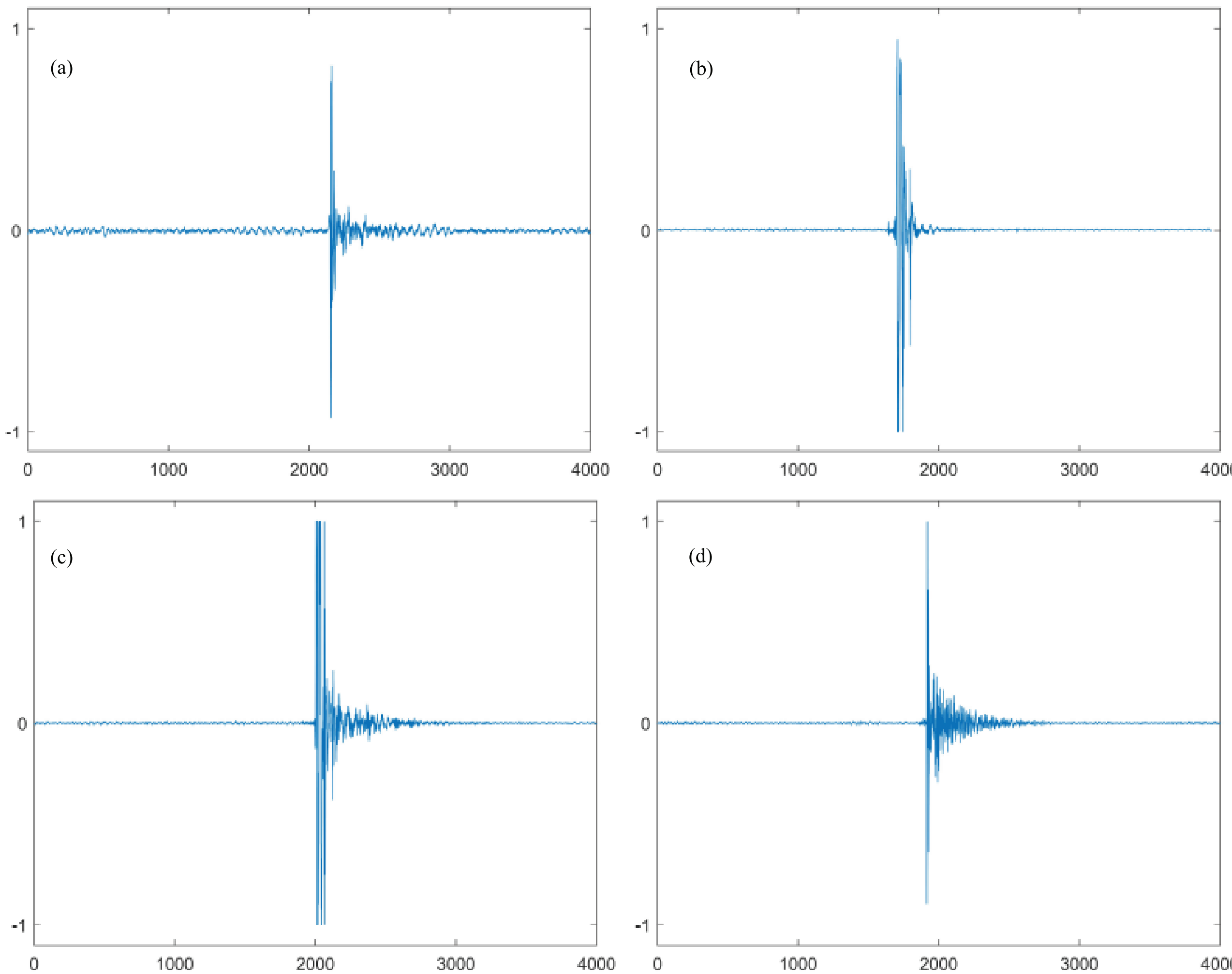
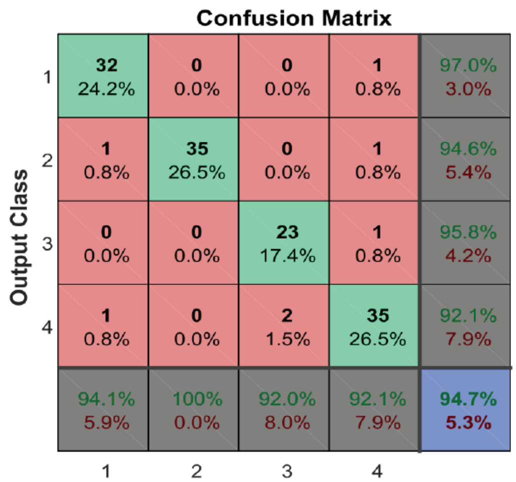


Figure 6: Samples from the self-collected snicks dataset for the four classes: (a) ball with bat, (b) ball with gloves, (c) ball with pad, (d) bat-pad mix.

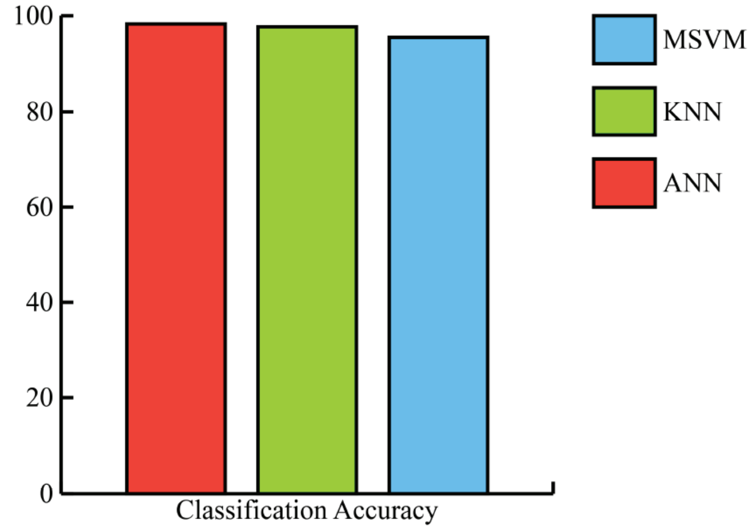
a single trial. A classification rate of 94.7 percent is achieved for the MLP-based ANN for a single trial. SVM and k nearest neighbor (KNN) were employed in the proposed system to increase the classification accuracy. KNN with 5 nearest neighbors was utilized. Figure 7b presents the classification rate as an average over 20 trials. While all the schemes are able to achieve a classification accuracy of above 95 percent, ANN tops the list with accuracy of 98.3 percent. Once the system was tested and trained on the self-collected data, it was extended for testing on snicks in real cricket data. The section below details the outcome of testing the real data.

6.2. Testing on real cricket scenarios

Once the proposed system was trained on the self-collected data, it was validated for snicks obtained from real-life cricket recordings. A total of seven different snicks were obtained relating to the different classes. All the videos were Googled and downloaded and do not come from a single source. Details for each case are presented as follows.



(a)



(b)

Figure 7: (a) Confusion matrix for single trial classification using ANN. (b) Self-collected data classification rates.

- **Ashwin LBW decision** The first case involves the ball being hit on the pad of the Indian batsman R. Ashwin. The ball appears to have hit the bat and pad simultaneously, as shown in Figure 8a. While the bowler appeals for a LBW decision, the on-field umpire calls it as not out, probably assuming that the ball hit the bat first.

(a) Ball appears to be hitting the batsman's bat and pad simultaneously. On-field umpire's call is not out.

(b) Ball appears to be hitting the batsman's thigh pad. On-field umpire's call is not out.

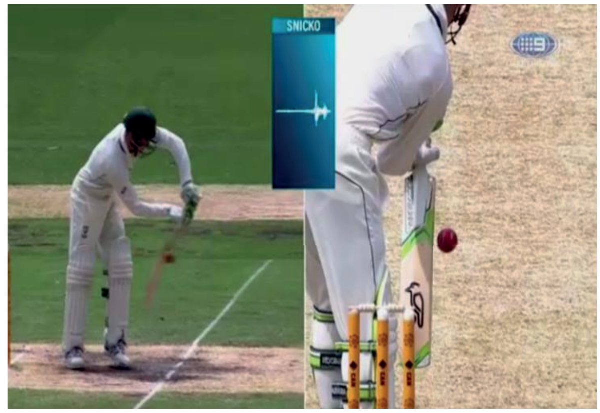
Figure 8: (a) Ashwin LBW decision. (b) George Bailey thigh pad catch case.

- **George Bailey thigh pad catch** This case involves the ball being hit on the thigh pad of the Australian batsman George Bailey. The ball appears to have hit the thigh pad and been caught behind, as shown in Figure 8b. While the bowler appealed for a catch by the wicket keeper, the on-field umpire called it as not out assuming that the ball hit the thigh pad and did not have any contact with the gloves or the bat.

- **Chris Gayle catch** This case involves the ball getting a definite hit and changing direction from the bat of West Indian batsman Chris Gayle. The ball hits the face of the bat and is caught behind, as shown in Figure 9a. While the bowler appealed for a catch by the wicketkeeper, the on-field umpire made no call. Surprisingly, the batsman does not leave the ground and depicts that the ball did not have any contact with his bat whatsoever.

- **Peter Handscomb unusual snick** This interesting case involves the ball seemingly hitting the bat of Australian batsman Peter Handscomb. The ball appears to have no contact with the bat; however, a

(a) Ball appears to be hitting the batsman's bat. On-field umpire's call is not out.



(b) Ball appears to be distant from the bat however Snickometer shows a snick.

Figure 9: (a) Chris Gayle case. (b) Peter Handscomb unusual snick case.

snick is produced on the Snickometer as shown in Figure 9b. This was probably given as out by the third umpire.

- **Stuart Broad catch** This case involves the ball seemingly hitting the edge of the bat of English batsman Stuart Broad. The ball appears to have hit the upper edge of the bat and was caught behind by the keeper as shown in Figure 10. The batsman refused to leave the ground, the umpire did not give an out, and since the opposing Australian team did not have any reviews left, the batsman kept playing.

Figure 10: Ball appears to have hit the edge of the bat and is caught behind. However, the batsman and umpire believe it is not out.

The SVM classifier is able to classify accurately for the trained classes for the cases of Ashwin, Broad, Gayle, and George. Both classifiers were unable to classify accurately for the cases where the bat hit the ground. The case of the bat hitting the ground was not considered earlier in the research and as such no training samples were obtained. For the case of Handscomb, both classifiers give inaccurate classification. However, in this case, the Snickometer-captured audio may be erroneous as the ball has no contact with the bat, as visible to the naked eye. Overall, the proposed system with the SVM classifier is able to offer improved classification as compared to the ANN-based system. The classification results depict the proposed system's finer ability to gauge the Snickometer signal source accurately to a large extent. This can be attributed to the two facts that the proposed system was incorporated with added features from the time and frequency domain as well as extensive training on source data collected by us.

Figure 11 shows the classification results for the three classifiers, ANN, KNN, and MSVM, for the real cricket snicks as 77.1, 71.4, and 85.7 percent, respectively.

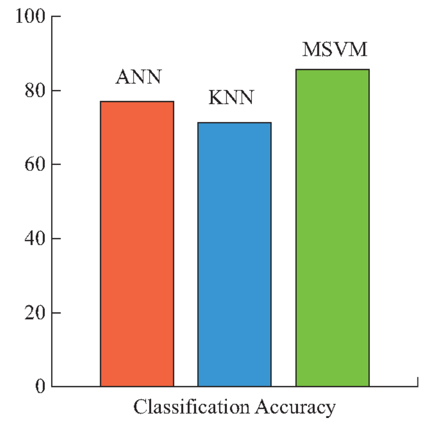


Figure 11: Real snick classification rates.

Table 2 shows the classification results for the real snicks obtained from the videos for the cases presented above. The ANN and SVM classifiers, pretrained on the 132 snick audios presented before, were tested to review the decisions made by the on-field umpires and to verify the sounds' sources. In order to gage the robustness of the proposed scheme, it was compared with benchmarks schemes. This included the works of Ting et al. and Rock et al. Ting et al. in [3] proposed a time-frequency features-based snick analysis scheme. While their method aims at classification among various snick audio classes, it utilizes only six features as compared to the proposed scheme. The classes comprise (i) ball and bat contact, (ii) ball and glove contact, (iii) ball and pad contact, and (iv) background noise. The features include average power, energy, max amplitude, max frequency, peak power, and RMS amplitude. The classes are not statistically significant and a threshold-based hard classification is employed. The proposed scheme uses machine learning-based classification for a more robust segregation. This scheme is among the first in the literature to design a basic snick detection system. It followed the research work of Rock et al.

Rock et al. in [1] proposed an automated edge detection system based on wavelets and neural networks. The system can detect two classes of sounds, i.e. bat on pad and bat on ball (edge). The proposed scheme has four classes. Five features extracted from the wavelet transform included the maximum correlation coefficient and its associated pseudo-frequency for several CWT scale ranges, along with the standard deviation, kurtosis, and skewness of the said correlation coefficients. ANNs were used to produce the final results. While the scheme achieves 97.5 percent accuracy, it is limited to only the two classes and utilizes ANNs for these linearly separable classes, making its performance questionable for real-life cases. The performance results of the proposed schemes are compared against these two benchmark schemes in Table 3. Concluding remarks on the research work are presented below.

7. Conclusions

A novel method for snick detection and classification as part of a DRS was proposed. The system is based on time and frequency domain features extraction and classification into four groups, namely ball on bat, ball on glove, ball on pad, and bat-pad mix. The system was trained initially using an ANN-based classifier and achieved a classification accuracy of 98.3 percent. To further enhance its accuracy, a SVM-based classifier was

Table 2: Classification results for real cricket snick audios. Pretrained ANN and SVM classifiers were utilized to classify the snick audio among the four sources.

Scenario name	Umpire decision	Ground truth (action type)	Classification result		Verdict on classification
			ANN classifier	SVM classifier	
Ashwin	Not out (LBW)	Possible bat-pad mix contact	Ball on glove	Ball on pad	SVM classifier decides accurately
Broad	Not out (catch behind)	Definite catch behind	Ball on bat	Ball on bat	Both classifiers decide accurately
Gayle	Not out (catch behind)	Definite catch behind	Bat on glove	Ball on bat	SVM classifier decides accurately
George	Not out (catch behind)	Hitting thigh pad	Bat on glove	Bat on pad	SVM classifier decides accurately
Handscomb	Out snick (catch behind)	Erroneous snick capture	Bat pad mix	Bat pad mix	While both classifiers depict same result, the classification is inadmissible as ground truth was erroneous

Table 3: Classification results of the proposed scheme compared to benchmark schemes.

Characteristics		Proposed scheme	Ting et al.	Rock et al.
Accuracy w.r.t snick class	Ball on bat	97.00	76.00	78.78
	Ball on glove	94.60	71.40	NA
	Ball on pad	95.80	79.16	NA
	Bat-Pad mix	92.10	NA	66.00
Number of snicks utilized for training	Ball on bat	34	84	130
	Ball on glove	35	88	NA
	Ball on pad	25	80	NA
	Bat-pad mix	38	NA	130
Number of features		17	6	5
Features' domain		Time - Frequency - Cepstral	Time - Frequency	Time - Wavelet
Average accuracy (own dataset)		98.3 percent	86 percent	97.5 percent
Datasets tested		Self-collected Real scenarios	Self-collected	Self-collected
Classification method		ANN and SVM	Hard threshold	ANN

employed resulting in classification accuracy of 85.7 percent. This was achieved by utilizing 17 features of the snick audio signals compiled. Features were reduced using one-way ANOVA and multicomparison techniques.

Noise removal as a preprocessing stage enhanced the proposed system by aiding in improving classification accuracy. A dataset for snick detection and classification has been developed, which can serve as a benchmark for comparison. Future directions of the research work including the investigation of methods for background noise removal in real time, development of audio sensors and vibration sensors that can be placed on the surface of the bat, and the possibility of including sensors within the ball for sensing its direction, vibration, and possible sound recording.

References

[1] Rock R, Als A, Gibbs P, Hunte C. The 5^{th} umpire: crickets edge detection system. Systemics, Cybernetics and Informatics 2012; 11 (1): 4-9.

[2] Rock R, Als A, Gibbs P. Optimizing a cricket edge detection system using feature extraction from wavelets over the time domain. In: Proceedings of the International Conference on Scientific Computing; Las Vegas, NV, USA; 2012. pp. 338-344.

[3] Ting S, Chilukuri MV. Novel pattern recognition technique for an intelligent cricket decision making system. In: IEEE Instrumentation and Measurement Technology Conference; Singapore; 2009. pp. 978-982.

[4] Lughofer E, Pratama M, Skrjanc I. Incremental rule splitting in generalized evolving fuzzy systems for autonomous drift compensation. IEEE Transactions on Fuzzy Systems 2018; 26 (4): 1854-1865.

[5] Singh L. Speech signal analysis using FFT and LPC. International Journal of Advanced Research in Computer Engineering and Technology 2015; 4 (14): 1658-1660.

[6] Sigtia S, Stark AM, Krstulović S, Plumbley MD. Automatic environmental sound recognition: performance versus computational cost. IEEE/ACM Transactions on Audio, Speech, and Language Processing 2016; 24 (11): 2096-2107.

[7] Zhang F, Meng H, Li M. Emotion extraction and recognition from music. In: 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery; Changsha, China; 2016. pp. 1728-1733.

[8] Kong Q, Xu Y, Plumbley MD. Joint detection and classification convolutional neural network on weakly labelled bird audio detection. In: 25th European Signal Processing Conference; Kos Island, Greece; 2017. pp. 1749-1753.

[9] George J, Mary L, Riyas KS. Vehicle detection and classification from acoustic signal using ANN and KNN. In: International Conference on Control Communication and Computing; Thiruvananthapuram, India; 2013. pp. 436-439.

[10] Aykanat M, Kılıç Ö, Kurt B, Saryal S. Classification Of lung sounds using convolutional neural networks. EURASIP Journal on Image and Video Processing 2017; (1): 65.

[11] Rubio JJ. A method with neural networks for the classification of fruits and vegetables. Soft Computing 2017; 21 (23): 7207-7220.

[12] Zanaty EA. Support vector machines (SVMs) versus multilayer perception (MLP) in data classification. Egyptian Informatics Journal 2012; 13 (3): 177-183.

[13] Sari S, Roslan H, Shimamura T. Noise estimation by utilizing mean deviation of smooth region in noisy image. In: Fourth International Conference on Computational Intelligence, Modelling and Simulation Computational Intelligence, Modelling and Simulation; Kuantan, Malaysia; 2012. pp. 232-236.

[14] Mazzoni D, Dannenberg R. Audacity - Audio Software for Multi-track Recording and Editing. 2019. https://www.audacityteam.org/.

[15] de Jesús Rubio J. Stable Kalman filter and neural network for the chaotic systems identification. Journal of the Franklin Institute 2017; 354 (16): 7444-7462.

[16] Hassan M, Hassan A, Mohamed Z, Oliver S. Assessment of artificial neural network for bathymetry estimation using high resolution satellite imagery in shallow lakes: case study El Burullus Lake. In: Eighteenth International Water Technology Conference; Sharm El-Sheikh, Egypt; 2015. pp. 248-259.

[17] Duan KB, Keerthi SS. Which is the best multiclass SVM method? An empirical study. Lect Notes Comp Sci 2005; 3541: 278-285.

[18] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory; Pittsburgh, PA, USA; 1992. pp. 144-152.

[19] Chih-Wei H, Chih-Jen L. A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks 2002; 13 (2): 415-425.