

Classification of Barking Context of Domestic Dog using High-Level Descriptors

Benjamín Gutiérrez-Serafín¹, Humberto Pérez-Espinosa^{1,2},
Juan Martínez-Miranda^{1,2}, Ismael Espinosa-Curiel¹

¹ CICESE-UT3, Nayarit, Mexico

² CONACyT Research Fellow, Mexico

{hperez, jmiranda, ecuriel}@cicese.mx,
begutierrezse@outlook.com

Abstract. Barking has been a controversial topic that has been studied from different points of view. While some authors argue that dog barking is a non-communicative vocalization, others believe that barking plays a significant role in the human-dog interaction. Among the studies that take the last perspective, one of the most recent methods is to implement machine learning algorithms to classify single barks in different behavioral context by evaluating low-level descriptors. However, these research works do not incorporate the analysis of temporal structure or other dog vocalizations. In the present study, we proposed a broader approach by taking into account these relevant features that are currently not considered in the analysis of single barks for the classification of the context. By implementing an automatic process that segments long recordings of dog vocalizations and extracts both low-level and high-level descriptors, promising results were obtained for the barks' context classification from long recordings, where the highest value of F-measure was 0.71.

Keywords. Machine learning, barking classification, acoustic analysis, high-level descriptors.

1 Introduction

Within the repertoire of dog vocalizations, barking is typified by being the most characteristic sound of this species. Because of its constant presence in every context, barking has been considered as a non-functional way of communication [2].

Nevertheless, the study conducted by [5] contrasts this idea and suggests that the process of dog domestication affected to improve the barking of dogs to become the best communicative component. Similarly, the findings of [18] support the functionality of barking since they managed to prove with spectrographic analysis that bark structure varies depending on the context.

Research regarding this matter [11, 12], proved that people can identify the emotional content in which a dog barks, regardless of the age and level of experience of the participants with dogs. These results indicate that dog barking works as a communication system within the dog-human interaction. An effective identification of what the

dogs are trying to communicate can improve the one-of-a-kind relationship between humans and dogs. In the same way, this could also benefit the welfare of the dogs, as well as take advantage of their vocalizations, for instance, to use them as security systems in houses. For these reasons, further research on this topic is required.

In recent years, modern approaches have been employed, such as applying artificial intelligence techniques, to enhance the dog barking classification. For example, in [8] the authors developed an algorithm to detect and select the best feature sets for the identification of context and individual, which were later tested in a classification task. Larranaga et al. [7] went even further by adding the prediction of sex and age of domestic dogs in its research. These problems, coupled with the individual recognition, have shown better results in contrast with the ones obtained in the context classification. As determined by [9] the context classification of barks relies highly on the barking individual. It was reported that when a dog-independent classification model is used, a decrease in accuracy occurs, which explains the reason for low performance in the problem of context recognition.

Additionally, these studies have centered just on the analysis of single barks and have overlooked other dog vocalizations and temporal patterns. According to the findings of [1] when a dog vocalizes, it emits not only barks but also other types of sounds such as growls and whines among others. This could be either by consecutive emissions of two or more sort of sounds or superimposing them, as well as a mix of these two manners. Looking at research work of other dog vocalizations, T. Faragó et al. [4] confirmed that growls also convey affective and contextual content to humans. In addition, they remarked that the natural temporal structure of this vocalization influences the understanding of the emotional state of the dogs, similarly to the inter-bark time intervals in the studies carried out by Pongrácz et al. [12, 13].

Due to these last-mentioned facts, it is reasonable to consider not only the acoustic features of single barks as the single factor of assessment in the bark's context classification problem through machine learning, but also the temporal structure and other dog vocalizations present in the barking sequences. In this paper, we report our research on the analysis of long barking recordings, in which high-level descriptors providing temporal structure and number of dog vocalizations are evaluated by using an automated method. The research questions that we are aiming to solve in this analysis are:

Does the inclusion of high-level descriptors add relevant information to the context classification?

What effect does an individual-independent validation have in the different stages of our proposed method?

2 Experimental Data

For this study, it was required to utilize the Mudi barking database for the classification of the contexts of barks and the UT3 dog vocalizations database for the identification of the type of dog sound.

2.1 Mudi Barking Database

The Mudi barking database was collected by Pongrácz et al. [13] from 12 Mudi dogs and consists of 244 recordings divided into 7 different behavioral contexts, recorded both inside the owners' dwelling and outdoor. These recordings last from 1 second to 6 minutes with a mean time of 44.78 seconds and a standard deviation of 48.19 seconds. They mainly include barking sequences with inter-bark intervals, as well as some other dog vocalizations and environmental sounds. Additionally, this repository also contains a total of 6614 single barks that were manually segmented from these long recordings and range from 0.1 to 0.8 seconds.

According to [13], the barks were recorded while performing a set of actions to encourage the dogs to bark. The compilation of every recording was made with a tape recorder and a microphone, where the experimenter was stood in front of the dog while holding the microphone within 1 to 4 meters from the dog. As reported in [13], the bark recordings were collected in the following 7 type of situations:

1. Alone: The dog was tied with its leash to a tree by its owner, then he/she walked away from the sight of the dog.
2. Ball: The owner held a ball or the dog's favorite toy and showed it to the dog 1.5 meters in front of it.
3. Fight: The trainer enraged the dog to bark aggressively and to bite his special glove while the owner kept the dog on a leash.
4. Food: The owner held the dog's food bowl 1.5 meters in front of the dog.
5. Play: The owner played a typical game with the dog.
6. Stranger: The experimenter presented in the owner's garden or front door in the absence of the owner.
7. Walk: The owner pretended like he or she was preparing to go for a walk with the dog.

2.2 UT3 Dog Vocalizations Database

The UT3 dog vocalization database was also employed to classify the type of sounds that the dogs emit. Audio and video of eighty dogs, mainly Chihuahuas, Schnauzers and French Poodles, were recorded to obtain the collection of audio constituted of a total of 30907 segments, of which 911 were whines, 666 growls, 5645 barks and the rest were background sounds (TV, human voices, bird sounds, etc.). The data collection protocol was designed to obtain aggressive and not aggressive vocalizations as following:

- Aggression to unfamiliar people inside the house.
 1. Normal alert due to the presence of a stranger: The experimenter (stranger) repeatedly knocked on the door to make the dog start barking.
 2. Aggression to the dog: The experimenter entered the home and provoked aggressive barks by making menacing movements in front of the dog.
 3. Aggression to the owner: An attack on the owner was simulated by making noises such as clapping and beating on the floor.
- Reaction towards enjoyment, happiness, and satisfaction.

4. Words of affection: The owner spoke in a high-pitched and affectionate tone of voice to the dog, with the intention of generating vocalizations.
 5. Scratching: Both the owner and experimenter (if the dog allowed it) caressed or scratched the back, abdomen, chest, and sides of the dog until it produced vocalizations.
 6. Play: The owner stimulated the dog using the objects or toys with which it usually plays, hoping to cause vocalizations.
 7. Arrival at home: The owner was requested to leave the house with the experimenter and, after a while, returned home and talked to the dog affectionately from the outside without opening the door.
 8. Simulation of going for a walk: The owner performed the routine that precedes taking the dog for a walk.
- Outdoor behavior.
9. Sadness/Anxiety by separation: The owner tied the dog with the leash to a tree and moved away from his sight.
 10. Aggression to the dog: While the dog was still tied to the tree, the experimenter or some other stranger threateningly approached the dog, such as in the second stimulus.
 11. Aggression to the owner: The experimenter pretended to attack the owner, such as in the third stimulus while walking his/her dog through the park.

After the audio compilation, individual sounds were automatically segmented using a Python script created by us that takes as input the energy and spectral centroid thresholds to detect audio activity. In the end, the automated method generated segments of a distinct span that range from very short individual barks (0.4 seconds) to longer segments (4 seconds) that consist of a group of a different number of individual barks emitted very quickly with the absence of pauses.

3 Methodology

In order to classify long barking recordings, we devised an automated method that segments the complete recordings and extracts low-level descriptors from the individual segments to classify both the type of vocalization and the context of each bark. Furthermore, high-level descriptors of each long barking audio file are also extracted to evaluate them and recognize the context in which the long recordings were induced. It should be pointed out that the number of recordings and barks used in our analysis differ from those in the original database due to the phases of automatic segmentation and dog vocalizations classification implemented in our approach. A set of the segments obtained were misclassified. Besides, some of the recordings could not be segmented properly using the optimal parameters, which precluded from being an automated process. In the end, we decided to discard this set of sub-optimal recordings and maintain those recordings useful to perform our analysis. Tables 1 and 2 show an overview of the number of recordings and single barks in the original database and those used in our experiments according to the contexts and the dogs, respectively. In the following subsections, we explain the details of the process that was carried out.

Table 1. Number of recordings and single barks for each context.

Original Database		Samples in our Analysis		Contexts
Long Recordings	Single Barks	Long Recordings	Single Barks	
22	758	19	706	Alone
53	1004	48	1000	Ball
30	1056	29	995	Fight
41	833	36	746	Food
23	752	23	977	Play
46	1425	44	1680	Stranger
29	786	28	1093	Walk
244	6614	227	7197	Total

Table 2. Number of recordings and single barks for each dog.

Original Database		Samples in our Analysis		Dogs
Long Recordings	Single Barks	Long Recordings	Single Barks	
9	275	9	317	d05
25	693	25	753	d09
5	108	5	140	d10
37	1007	31	908	d12
24	465	20	492	d14
18	336	17	336	d16
19	219	15	351	d18
29	686	29	736	d20
32	968	30	1041	d23
40	1650	40	1976	d24
2	83	2	59	d26
4	124	4	88	d27
244	6614	227	7197	Total

3.1 Overview of the Proposed Method

The general objective of this research is to take a broader approach by considering other variables that are not contemplated in the analysis of single barks. For this reason, we proposed to evaluate the long barking recordings of the Mudi database under the systematic procedure shown in Figure 1, that consists of 5 main stages, of which, three independent models had to be trained using machine learning algorithms. The functions of each of the main phases of this study are the following:

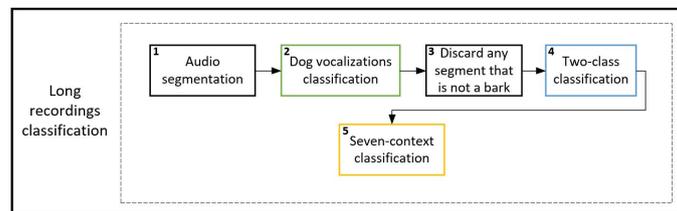


Fig. 1. Schematic figure for the proposed experimental procedure of long barking recordings classification.

1. Audio segmentation: The long barking classification model takes into account every dog vocalization within the complete recordings. Therefore, it is essential to have a segmentation stage to be able to inspect individually each of the sounds emitted by the dog. In this first step, the long barking recordings are inputted for segmentation. In addition, the span of inter-bark intervals, the duration of the vocalizations and the number of groups are registered.
2. Dog vocalizations classification: Having obtained each independent sound, the process shown in Figure 2(A) takes place to identify barks, whines, and growls with the assistance of the labeled data of the UT3 dog vocalization database. Background sounds are eliminated because they do not provide relevant information.
3. Discarding of segments that were not identified as barks: The barks are segregated from other canine vocalizations to continue with a second classification. The number of whines and growls are registered as part of other features needed in the final examination.
4. Two-class classification: To simplify the long recordings classification, a supplementary classification as illustrated in Figure 2(B), is implemented to improve the efficiency by labeling the single barks in one of the following two classes used to group the 7 different contexts according to the pleasant/unpleasant and activated/deactivated dimensions based on the circumplex model of affect [14] (see Fig. 3). The results of this assortment are added to the global features. A deeper explanation about the use of the two classes is provided in the Section 3.3.3.
5. Seven-class classification: Finally, the classification stage presented in Figure 2(C) is performed where the high-level descriptors previously identified are evaluated to complete the classification of the long recordings.

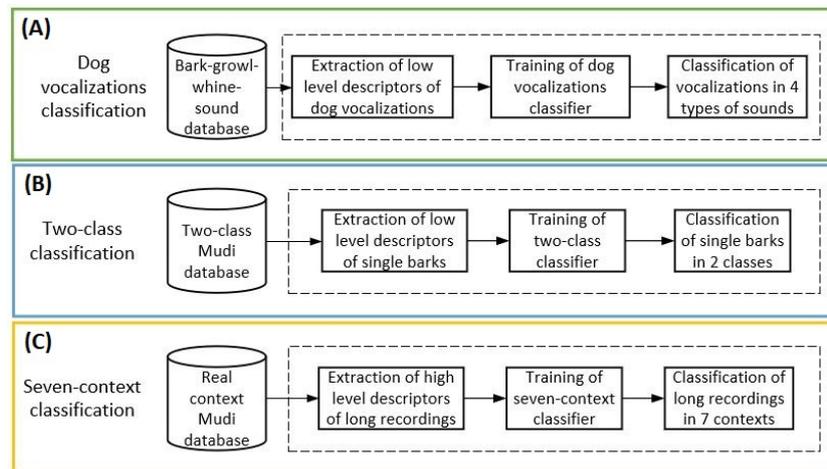


Fig. 2. Subsystems for accomplishing the classification of dog vocalizations.

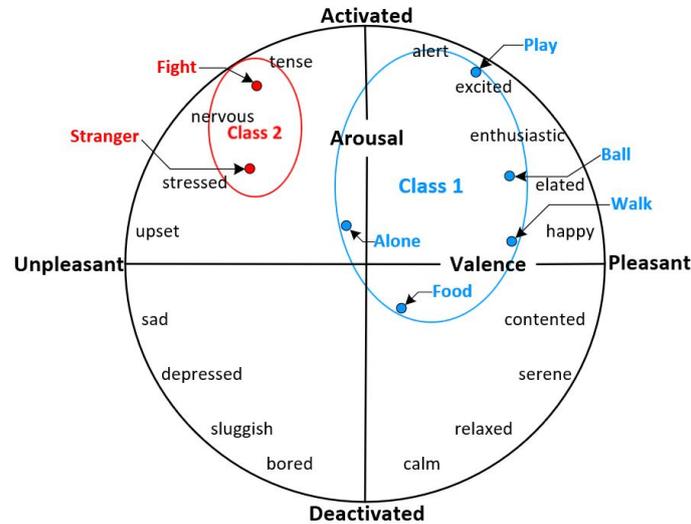


Fig. 3. Distribution of the seven barks' contexts within the circumplex model of affect.

3.2 Automatic Segmentation of Recordings

The long Mudi barking recordings were automatically segmented in the first stage by each canine vocalization emitted in their respective audio file. The duration of each recording in the Mudi database varies from a couple of seconds up to 6 minutes. An automated procedure was implemented for the segmentation process where was established an energy threshold of the signal to differentiate sound from silence and background noise. The threshold was detected by calculating the short time energy and spectral centroid. Precisely, the segmentation algorithm creates an array with the beginning and end time of the periods with sound. Then, it extracts from the long recordings the periods with sound for a further detailed examination to differentiate dog vocalizations. Figure 4 shows a fragment of a segmented long recording, in which both the periods' durations where there are no sound and the individual barks are registered to generate some of the High-Level Descriptor (HLDs). During the audio segmentation interval, the same algorithm extracts other HLDs from the long recordings in the background. More information about these features is explained in the next subsection.

3.3 Audio Characterization

Extraction of Low-Level Descriptors After segmentation stage, it was necessary to apply some signal processing techniques to acquire the acoustic properties of the segmented audio files in order to classify them, by type of vocalization (bark, growl or whine) and the context of the bark (alone, ball, fight, food, play, stranger or walk), according to the correlation of their features. To this end, the tool openSMILE [3] was used to extract Low-Level Descriptors (LLDs) for both databases. It was decided to use



Fig. 4. A recording containing five single bark sounds. Each single bark is stored in a separated audio file.

the next feature sets that have been proposed by different authors to conduct various experiments.

- IS-09 [15]: This feature set is extracted using a frame size of 25 ms and a frame step of 10 ms applying a Hamming window. Then, data contours are smoothed by a moving average filter to process the signal. Subsequently, 39 statistical functions are calculated over the values of the LLDs and their deltas and double deltas coefficient in each frame of the audio samples. In the end, a total of 5148 attributes are obtained.
- IS-10 [16]: The LLDs values are extracted using a frame size of 60 ms, and a frame of 25 ms applying a Gauss window for some features and a Hamming window for others. Then, the signal is processed by a moving average filter for smoothing data contours. In the next stage, it computes 21 statistical functions in some features and 19 in others over the values of the LLDs and their deltas and double deltas coefficient in each frame of the audio samples. As a result, an amount of 1582 attributes is obtained.
- IS-11 [17]: LDDs are computed using a frame size of 60 ms, and a frame of 25 ms applying a Gauss window for some features and a Hamming window for others. Then, the signal is processed by a moving average filter for smoothing data contours. After that, it calculates 37 statistical functions in some features and 36 in others over the values of the LLDs and their deltas and double deltas coefficient in each frame of the audio samples. Eventually, 4368 acoustic features are obtained.

Low-Level Descriptors Selection To improve the performance of the classifiers, it was determined to apply the Subset Evaluation method to every feature set with the assistance of the software WEKA [6]. This process evaluated all the original attributes from the IS-09, IS-10, and IS-11 sets to reduce the dimensionality of their feature vectors to 192, 140 and 217, respectively.

Extraction of High-Level Descriptors The classification model of long barking recordings was created based on the high-level characteristics registered during some processes that form our automated method. The feature vector of each long recording is constituted of 9 values, of which, 7 of them were extracted in the stages of segmentation and dog vocalizations classification. Moreover, these features were normalized due to the recordings do not have the same time duration. The other two remaining HLDs

were obtained in the two-class classification phase and are represented in percentage, with respect to the number of barks classified in their respective class. The arrangement of the two classes was decided after evaluating the segments of our analysis in different combinations of the seven barks' contexts according to their distribution within the circumplex model of affect (see Fig. 3). We grouped the barks induced by the stimuli of stranger and fight in one class, and the barks caused by the rest of the stimuli in a second class. This grouping, as shown in Table 3, obtained the best classification results with 80/20 validation and using the machine learning algorithms mentioned in the next subsection. The HLDs of the classification of dog barking recordings are listed below:

1. Number of barks: The collection of segments that were identified as barks during the dog vocalizations classification.
2. Number of whines: The number of segments that were recognized as whines during the second stage of the automated process.
3. Number of growls: The quantity of segments that were labeled as growls in the phase two.
4. Number of groups: The number of segments grouped that were detected in the segmentation process according to their temporal proximity.
5. Number of pauses: The number of gaps that exists between the segments that were perceived in the segmentation process.
6. Total duration of barks: The sum time in seconds of each bark that was calculated during the segmentation process.
7. Total duration of pauses: The total amount of time in seconds of each space between the segments that was calculated during the first part of the process.
8. Class 1: The percentage of barks classified in the two-class classification as a context within the class 1 (alone, ball, food, play and walk).
9. Class 2: The proportion of barks classified in the two-class classification as a context within the class 2 (fight and stranger).

Table 3. Average F-measure of every experiment carried out for the Two-class classification using the barks obtained in our analysis.

IS-09			80/20						IS-11			Arrangements
RF	NB	SVM	IS-10		IS-10		IS-11		RF	NB	SVM	
0.33	0.35	0.35	0.38	0.40	0.35	0.35	0.36	0.39	0.36	0.39	0.39	Alone, Ball, Play Food, Walk Fight, Stranger
0.45	0.42	0.46	0.42	0.40	0.43	0.52	0.41	0.47	0.41	0.47	0.47	Alone, Food Play, Ball, Walk Fight, Stranger
0.37	0.38	0.38	0.41	0.39	0.35	0.43	0.43	0.42	0.42	0.42	0.42	Alone, Food, Walk Play, Ball Fight, Stranger
0.52	0.61	0.64	0.62	0.58	0.63	0.66	0.47	0.69	0.69	0.69	0.69	Ball, Food, Play, Walk Alone, Fight, Stranger
0.60	0.52	0.59	0.66	0.51	0.56	0.70	0.70	0.67	0.67	0.67	0.67	Alone, Ball, Food, Play, Walk Fight, Stranger

3.4 Classification and Validation Methods

Due to the positive results in previous similar classification analyzes [8–10], we used Support Vector Machine (SVM), Random Forest (RF) and Naive Bayes (NB) to conduct

every experimentation in this study and store their results within Table 4 to Table 7. In addition, we also validated the stability of these machine learning algorithms with two validation methods:

- 10-Fold Cross-Validation (10FCV): In this technique, a 90% of the samples in the dataset is trained to create a model that is then tested with the remaining 10%. This process is repeated ten times using different training and testing sets on each occasion.
- 80/20 Validation: In this validation scheme, a classification model is trained using 80% of the samples in the dataset. The 20% left is used to test the model. In this validation process, we created a training model in which were not included the dogs belonging to the test model, generating as a consequence, a dog-independent training model.

Table 4. F-measure of single barks from the Mudi database for each 7 contexts in 10FCV settings.

10FCV										Context
IS-09			IS-10			IS-11				
RF	NB	SVM	RF	NB	SVM	RF	NB	SVM		
0.80	0.44	0.75	0.82	0.46	0.74	0.81	0.46	0.76		Alone
0.69	0.34	0.61	0.69	0.33	0.60	0.67	0.26	0.61		Ball
0.81	0.72	0.82	0.82	0.67	0.82	0.81	0.67	0.84		Fight
0.72	0.51	0.63	0.70	0.51	0.61	0.67	0.47	0.60		Food
0.69	0.57	0.68	0.69	0.59	0.68	0.65	0.54	0.66		Play
0.77	0.59	0.74	0.76	0.62	0.73	0.75	0.53	0.75		Stranger
0.63	0.44	0.62	0.65	0.46	0.60	0.60	0.45	0.61		Walk
0.74	0.52	0.70	0.74	0.53	0.69	0.72	0.49	0.70		Weighted Average

4 Results

It has been reported in previous studies related to the analysis of barking classification with a machine learning approach that, when samples of the same dog are used in the training and test sets, favorable results are obtained. However, the opposite occurs when both sets have samples from different dogs. To prove this, we conducted two experiments. Using the single context audio recordings of the Mudi database, a classification model was created, in which the barks of the 12 dogs were considered to train it. Looking at the results of the first experiment in Table 4, a good classification performance can be observed as it more than half of the outcomes shown 0.70 or higher in terms of F-measure. On the other hand, when using the same data but this time evaluating a barking test set of the dogs d18, d23 and d27 with a training set comprised from barks of the remaining dogs, poor results were obtained. As it is presented in Table 5, the experiment of the dog-independent classification model provided us with F-measure values less than 0.30 in all the cases.

When the number of contexts that had to be classified was reduced, an increase in F-measure values was presented. This occurred even if the training and test sets

Table 5. F-measure of single barks from the Mudi database for each 7 contexts in 80/20 validation settings.

80/20											Context
IS-09			IS-10			IS-11					
RF	NB	SVM	RF	NB	SVM	RF	NB	SVM			
0.08	0.13	0.06	0.02	0.10	0.02	0.05	0.29	0.06			Alone
0.06	0.10	0.18	0.08	0.04	0.09	0.11	0.06	0.13			Ball
0.61	0.63	0.49	0.50	0.46	0.34	0.56	0.63	0.55			Fight
0.29	0.18	0.14	0.27	0.24	0.14	0.07	0.15	0.09			Food
0.30	0.35	0.27	0.39	0.19	0.47	0.33	0.23	0.16			Play
0.35	0.30	0.28	0.39	0.38	0.32	0.31	0.37	0.31			Stranger
0.22	0.18	0.27	0.28	0.19	0.28	0.28	0.19	0.22			Walk
0.27	0.26	0.24	0.27	0.22	0.22	0.24	0.27	0.22	Weighted Average		

differ from barking individuals. We conducted experimentation regarding this matter in a dog-independent classification model using the same data sets of the previous independent analysis. The improvement of the results can be noticed in Table 6, wherein almost all of the cases are above 0.75. In this experiment, the seven contexts were regrouped into two classes with the purpose of facilitating the classification process. Due to its efficiency, this procedure was included as the two-class classification stage in our method to allocate the segmented barks of long recordings in Class 1 and Class 2 features for the subsequent classification of long barking recordings in 7 contexts.

Table 6. F-measure of single barks from the Mudi database for each 2 contexts in 80/20 validation settings.

80/20											Context
IS-09			IS-10			IS-11					
RF	NB	SVM	RF	NB	SVM	RF	NB	SVM			
											Alone
											Ball
0.84	0.75	0.81	0.85	0.76	0.83	0.83	0.83	0.84	Class 1		Food
											Play
											Walk
0.56	0.61	0.60	0.55	0.50	0.63	0.59	0.64	0.68	Class 2		Fight
											Stranger
0.76	0.71	0.75	0.76	0.69	0.77	0.76	0.78	0.79	Weighted Average		

The last part of this study was to implement the automatic method explained in Section 3. We used the long bark audio recordings of the Mudi database as the input files of the process. As in the past two experiments, the same arrangement of barking individuals was implemented to create the dog-independent classification model. In this case, the training set was used to train the two-class classifier to designate the single barks in the long recordings to their respective category. Once the process concluded, the features extracted from the complete recordings of the dogs d18, d23 and d27 were evaluated. It was complicated to create again effective training and test sets for an

individual-independent validation due to the small number of instances and the variation of samples of each dog in this step. Nevertheless, we managed to get some promising results by carrying out a 10FCV to the final data frame. It can be contemplated in Table 7 that we obtained a positive outcome when we trained all the classification models with the Random Forest and Naive Bayes algorithms. However, poor performance was presented in the experiments where Support Vector Machine algorithm was used for training and caused that some of the values could not be captured.

Table 7. F-measure of long barking audio files from the Mudi database for each 7 contexts in 10FCV settings.

10FCV										Context
IS-09			IS-10			IS-11				
RF	NB	SVM	RF	NB	SVM	RF	NB	SVM		
0.67	0.67	-	0.73	0.75	-	0.73	0.80	-		Alone
0.55	0.56	0.44	0.50	0.58	0.36	0.67	0.58	0.46		Ball
0.33	0.40	0.00	0.33	0.46	-	0.50	0.31	-		Food
0.92	1.00	0.83	0.86	0.80	0.67	0.83	0.77	0.73		Play
0.77	0.77	-	0.80	0.86	-	0.86	0.86	-		Walk
1.00	0.80	0.50	0.89	0.75	0.00	1.00	0.80	0.22		Fight
0.44	0.43	0.00	0.44	0.38	0.00	0.44	0.43	0.00		Stranger
0.65	0.64	-	0.62	0.64	-	0.71	0.64	-		Weighted Average

5 Conclusion and Future Work

In our study, we explored the idea of evaluating high-level descriptors of long recordings of domestic dogs for context classification. Our findings demonstrate that the inclusion of these descriptors provide useful information for this classification problem. Additionally, our results support the idea that the presence of other dog vocalizations and temporal structure contain valuable patterns to assist machine learning algorithms in determining the real context of barking sequences. Concerning the model independence affair, a satisfactory performance with a dog-independent model was presented during most part of the proposed method. Nonetheless, the last part of the evaluation of this process was difficult to perform, since the data sample did not have a required diversity of contexts to demonstrate the effectiveness of the method. Thus, the results indicated poor accuracy in an individual-independent validation. Even though the final results were obtained only with a dependent classification model because of the lack of data, further study with this approach and more recordings may provide better results in the dog-independent model classification. Therefore, future research will concentrate on addressing this problem with a more extensive database, as well as the possibility of including more features to add more useful high-level information to the data frame to obtain better results.

Acknowledgment. This research work has been carried out in the context of the “Catedras CONACyT” programme funded by the Mexican National Research Council (CONACyT).

References

1. Cohen, J., Fox, M.: Vocalizations in wild canids and possible effects of domestication. *Behavioural Processes* 1(1), 77–92 (1976)
2. Coppinger, R., Feinstein, M.: Hark! hark! the dogs do bark...’and bark and bark. *Smithsonian* 21(10), 119–127 (1991)
3. Eyben, F., Wöllmer, M., Schuller, B.: Opensmile: the munich versatile and fast open-source audio feature extractor. In: *Proceedings of the 18th ACM international conference on Multimedia*. pp. 1459–1462. ACM (2010)
4. Faragó, T., Takács, N., Miklósi, Á., Pongrácz, P.: Dog growls express various contextual and affective content for human listeners. *Royal Society open science* 4(5), 170134 (2017)
5. Feddersen-Petersen, D.: Vocalization of european wolves (*canis lupus lupus* l.) and various dog breeds (*canis lupus f. fam.*). *Archives Animal Breeding* 43(4), 387–398 (2000)
6. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. *ACM SIGKDD explorations newsletter* 11(1), 10–18 (2009)
7. Larranaga, A., Bielza, C., Pongrácz, P., Faragó, T., Bálint, A., Larranaga, P.: Comparing supervised learning methods for classifying sex, age, context and individual mudi dogs from barking. *Animal cognition* 18(2), 405–421 (2015)
8. Molnár, C., Kaplan, F., Roy, P., Pachet, F., Pongrácz, P., Dóka, A., Miklósi, Á.: Classification of dog barks: a machine learning approach. *Animal Cognition* 11(3), 389–400 (2008)
9. Pérez-Espinosa, H., Pérez-Martínez, J.M., Durán-Reynoso, J.Á., Reyes-Meza, V.: Automatic classification of context in induced barking. *Research in Computing Science* 100(6), 63–74 (1976)
10. Pérez-Espinosa, H., Reyes-Meza, V., Aguilar-Benitez, E., Sanzón-Rosas, Y.M.: Automatic individual dog recognition based on the acoustic properties of its barks. *Journal of Intelligent & Fuzzy Systems* 34(5), 3273–3280 (2018)
11. Pongrácz, P., Molnár, C., Dóka, A., Miklósi, Á.: Do children understand man’s best friend? classification of dog barks by pre-adolescents and adults. *Applied animal behaviour science* 135(1), 95–102 (2011)
12. Pongrácz, P., Molnár, C., Miklósi, Á.: Acoustic parameters of dog barks carry emotional information for humans. *Applied Animal Behaviour Science* 100(3), 228–240 (2006)
13. Pongrácz, P., Molnár, C., Miklósi, A., Csányi, V.: Human listeners are able to classify dog (*canis familiaris*) barks recorded in different situations. *Journal of Comparative Psychology* 119(2), 136 (2005)
14. Russell, J.A.: A circumplex model of affect. *Journal of personality and social psychology* 39(6), 1161–1178 (1980)
15. Schuller, B., Steidl, S., Batliner, A.: The interspeech 2009 emotion challenge. In: *Tenth Annual Conference of the International Speech Communication Association* (2009)
16. Schuller, B., Steidl, S., Batliner, A., Burkhardt, F., Devillers, L., Müller, C., Narayanan, S.: The interspeech 2010 paralinguistic challenge. In: *Proc. INTERSPEECH 2010, Makuhari, Japan*. pp. 2794–2797 (2010)
17. Schuller, B., Steidl, S., Batliner, A., Schiel, F., Krajewski, J.: The interspeech 2011 speaker state challenge. In: *Twelfth Annual Conference of the International Speech Communication Association* (2011)
18. Yin, S.: A new perspective on barking in dogs (*canis familiaris*). *Journal of Comparative Psychology* 116(2), 189 (2002)