

# Aplicación móvil para el análisis de vocalizaciones del perro doméstico (canis familiaris)

Humberto Pérez-Espinosa<sup>1,2</sup>, Yuvila M. Sanzón Rosas<sup>1</sup>,  
Emanuel Aguilar Benitez<sup>1</sup>

<sup>1</sup> Centro de Investigación Científica y de Educación Superior de Ensenada,  
Unidad de Transferencia Tecnológica,  
México

<sup>2</sup> Consejo Nacional de Ciencia y Tecnología,  
México

hperez@cicese.mx

**Resumen.** El ladrido es una vocalización muy distintiva del perro. Es muy común y un medio de interacción entre los perros y con los humanos. Sin embargo, el alcance de su análisis automatizado por técnicas computacionales, así como las posibles aplicaciones a las que pueden dar lugar, han sido poco exploradas. En este artículo, describimos el proceso para desarrollar una aplicación capaz de estimar el nivel de alarma en el ladrido de los perros. Para el entrenamiento del sistema usamos una base de datos con más de 10,000 ladridos. Dichos ladridos fueron obtenidos aplicando estímulos positivos y negativos a los perros. Las muestras de ladridos se caracterizaron acústicamente utilizando una herramienta de procesamiento de señales que extrae grandes conjuntos de características. Generamos subconjuntos óptimos de descriptores acústicos para alimentar algoritmos de aprendizaje automático y obtener modelos de clasificación entrenados. Evaluamos dichos modelos y comparamos el rendimiento de clasificación de diferentes algoritmos. La clasificación obtenida supera los resultados previamente reportados en trabajos similares. A partir de estos modelos entrenados se diseñó e implementó una arquitectura que incluye un API y una aplicación móvil que puede ser usada con propósitos de seguridad doméstica y cuidado de las mascotas.

**Palabras clave:** Aprendizaje máquina, clasificación de ladridos, aplicación móvil.

## Mobile Application for the Analysis of Domestic Dog (Canis Familiaris) Vocalizations

**Abstract.** Barking is a very distinctive vocalization of the dog. It is very common and a means of interaction between dogs and towards humans. However, the scope of its automated analysis by computational techniques, as well as the possible applications to which it can give rise, have been little explored.

In this article, we describe the process of developing an application capable of estimating the alarm level in dogs' barking. For system training, we use a database with more than 10,000 barks. These barks were obtained by applying positive and negative stimuli to the dogs. The bark samples were characterized acoustically using a signal processing tool that extracts large sets of characteristics. We generate optimal subsets of acoustic descriptors to feed machine learning algorithms and obtain trained classification models. We evaluate these models and compare the classification performance of different algorithms. The classification obtained exceeds the results previously reported in similar works. Based on these trained models, an architecture was designed and implemented that includes an API and a mobile application that can be used for home security and pet care purposes.

**Keywords:** Machine learning, dog bark classification, mobile application characterization.

## 1. Introducción

El perro ha sido compañero del hombre desde hace aproximadamente 15,000 años [5], inicialmente el propósito que tenían los perros era ayudar en la cacería, posteriormente una vez que los humanos se asentaron y comenzaron a practicar tanto la agricultura como la ganadería a los perros se les entrenó para cumplir el rol de guardianes de ganado.

En la actualidad, los perros siguen estando presentes en la sociedad principalmente como acompañantes, pero también se desempeñan como rescatistas, guardianes o asistentes entre otras importantes labores. En años recientes, se ha estudiado la manera interpretar sus ladridos para entenderlos, puesto que ellos son capaces de comunicarse con los humanos ya sea por medio de vocalizaciones o por medio de movimientos corporales.

Así mismo, los perros domésticos son capaces de entender a los humanos a través del lenguaje corporal [8]. Incluso, los perros son capaces de reaccionar a distintas situaciones y comunicarlas por medio de sus ladridos [14]. Investigaciones previas han tratado de identificar los tipos de ladridos ante contextos agradables y desagradables para el perro, otras se han enfocado en monitorear y reconocer el comportamiento del perro [7].

Con estos antecedentes el estudio aquí presentado se enfocó en usar los ladridos del perro para generar una alarma, es decir, identificar si los ladridos que emite el perro es por una situación agradable como estar jugando o si los ladridos advierten de una amenaza como estar ante la presencia de un extraño. Para esto fue necesario experimentar con distintas características que dieran mejores resultados tanto para identificar vocalizaciones del perro como para discriminar si un ladrido es amigable o de alarma.

## **2. Trabajos relacionados**

La detección y clasificación de sonidos ambientales puede ser de gran utilidad en varios escenarios, como sistemas de vigilancia, asistencia, y monitorización. En el trabajo de [11] desarrollaron un sistema de clasificación de eventos acústicos para sistemas de vigilancia. En lugar de extraer características en marcos de tiempo corto, hacen una búsqueda sobre toda la duración del evento acústico. Se usó una base de datos de 8 clases que incluía gritos humanos, disparos y ladridos de perro.

Dentro de lo que se conoce como sistemas de asistencia existen varios trabajos en la literatura. La aplicación para Android Audition [1] fue desarrollada para asistir a personas con discapacidades auditivas reconociendo los sonidos a su alrededor. Los algoritmos utilizados de aprendizaje automático para su implementación fueron AdaBoostM1 que funciona como clasificador y REPTree como un alumno débil.

La validación del sistema se realiza en cuatro condiciones ambientales diferentes (ruido alto y bajo, fuente de sonido cercana y lejana) en este trabajo no se obtuvieron resultados satisfactorios, ya que la precisión fue 26,25 % para la detección en las cuatro condiciones. Otra aplicación con la misma orientación de asistencia a usuarios con discapacidad auditiva es HomeSound [4] que consta de un micrófono y una pantalla, y utiliza varios dispositivos de captura de audio instalados en los hogares.

El prototipo muestra los sonidos acontecidos con historiales de formas de onda que representan el volumen y el tono. Una de las características del sistema es la clasificación automática de sonido. En la evaluación de desempeño se comprobó un aumento en la concientización sobre los sonidos en el hogar, pero las frecuentes clasificaciones automáticas erróneas no fueron bien recibidas por los usuarios.

Ambas aplicaciones para asistencia a usuarios con discapacidad auditiva incluyen la detección de ladridos como una característica esencial que les permite a los usuarios darse cuenta de eventos relevantes, por ejemplo si alguien está a la puerta o si el perro necesita algo. Otro trabajo que cae dentro del área de ambientes asistidos es el de [13], donde se investigó si el ladrido de un perro se puede distinguir del silencio u otros sonidos, como el ruido de un animal o del tráfico, o controlar un mecanismo de cierre de ventana en un escenario de hogar inteligente.

Estos trabajos se enfocan en detectar ladridos en el sonido ambiental. Sin embargo, existen otros trabajos que realizan un análisis más detallado de los ladridos tratando de clasificar diferentes tipos de ladrido. Recientemente se ha analizado el contexto y la emoción percibida en los ladridos explorando conjuntos de características acústicas buscando predecir la intensidad emocional de las secuencias de ladridos [3].

Los autores demuestran que el análisis acústico basado en aprendizaje automático puede alcanzar un rendimiento superior al nivel humano para clasificar el contexto en el cual fue producido un ladrido. Kim et al [6] propuso una forma de clasificar eventos acústicos de los perros eficientizando los recursos sin degradación significativa de la precisión.

Para entrenar sus modelos usaron datos de intensidad de los sonidos mediante el uso de un sensor de ruido en lugar de datos de un sensor de sonido. En [9] por medio de los ladridos crean modelos para identificar la raza, el sexo y el contexto al que fue expuesto el perro.

**Tabla 1.** Número de muestras en los conjuntos de entrenamiento y prueba.

Clases	Entrenamiento	Prueba
Ladridos	5,645	6,957
Chillidos	911	389
Gruñidos	666	688
Sonidos	23,685	8,755
<b>Total</b>	<b>30,907</b>	<b>16,789</b>

Dicho estudio compara el desempeño de los algoritmos de aprendizaje supervisado: Naive Bayes, classification trees, k-nearest neighbors y logistic regression en conjunto con características extraídas usando la herramienta Praat teniendo una precisión del 55.5 % en clasificación de contextos. Por otra parte en [10] obtiene una precisión del 43 % usando el sistema de procesamiento de señales de audio EDS para obtener en conjunto de características acústicas.

Estos trabajos muestran el interés que existe en usar la identificación y clasificación de ladridos del perro doméstico para diversas aplicaciones. En este trabajo proponemos la creación de una aplicación orientada a la seguridad doméstica que tiene como principal característica el estimar el nivel de alarma que un perro transmite a través de sus ladridos.

### 3. Datos usados

Para entrenar los modelos de clasificación automática en los cuales se basa el método propuesto se usaron dos bases de datos de vocalizaciones del perro doméstico que generó CICESE-UT en colaboración con la empresa Efecto Mescalina. Se cuenta con grabaciones de 102 perros de diferentes razas. Para obtener los datos se aplicó un conjunto de estímulos con el objetivo de generar reacciones en las que se aprecien diferentes comportamientos como agresividad, miedo, alegría, entre otros.

#### 3.1. Base de datos de perros Mescalina 2015

Para obtener los datos se pidió la autorización de los propietarios para grabar a sus mascotas y someterlos a una serie de estímulos de acuerdo a las condiciones explicadas a continuación:

1. Juego: El propietario estimula al perro utilizando los objetos o juguetes con los que normalmente juega, esperando lograr vocalizaciones.
2. Alerta normal ante presencia de un extraño: se comienza a tocar repetidamente el timbre del domicilio y dar golpes lo suficientemente fuertes a la puerta esperando provocar que el perro ladre. Esto se repite las veces que sea necesario.
3. Agresión: Una persona entra al domicilio, llamando la atención del perro y acercándose amenazadoramente a él, golpeando los pies contra el piso y dando palmadas para lograr que el perro se sienta agredido.

**Tabla 2.** Conjunto de audios para la detección de alarma.

Clases	Número de muestras
Tranquilo	12
Activo	9
Alerta	58
Alarma	33
<b>Total</b>	<b>112</b>

4. Agresión al dueño dentro del domicilio: se solicita al propietario que se acerque y se simula un ataque, haciendo también ruidos como palmadas y golpes al piso. Para lograr una mejor reacción de parte del perro, de ser necesario, se le pide al dueño que grite pidiendo auxilio o el nombre del perro.

En esta base de datos se tienen 6,077 ladridos segmentados a nivel individual con un promedio de duración 1 segundo. La base de datos Mescalina 2015 tiene treinta y siete individuos con sus nombres correspondientes. Los ladridos corresponden a las grabaciones de 11 razas de perros comunes en los hogares. En esta base de datos las edades de perros se etiquetaron en meses. Desde los 6 meses hasta 72.

### 3.2. Base de datos de perros Mescalina 2017

La base de Mescalina 2017 también la generó CICESE-UT<sup>3</sup> con el fin agregar más muestras al corpus original. Se añadieron nuevos contextos de grabación, los cuales se describen a continuación:

1. Llegada a casa: se solicita al propietario que salga del domicilio, se aleja con el experimentador, posteriormente regresan, y desde afuera de la casa le habla cariñosamente a su perro, pero sin abrir la puerta.
2. Simulación de paseo: el propietario realiza la rutina normal que precede al paseo. De ser necesaria una mejor reacción, se estimula al perro con llamadas cálidas o simulando abrir la puerta hasta que se emocione y produzca vocalizaciones.
3. Tristeza/ansiedad por separación: el propietario ata al perro con la correa a un árbol y se aleja de su vista. Para lograr un mejor resultado se le pide al dueño que se despida de él.
4. Asustado en el parque: mientras el perro continúa atado al árbol, el experimentador o algún otro extraño se acerca amenazadoramente al perro.

En esta base de datos se tienen 6,948 ladridos segmentados a nivel individual con un promedio de duración 1 segundo. La base de datos Mescalina 2017 tiene sesenta y cinco individuos. Los ladridos corresponden a las grabaciones de cuatro razas de perros comunes en los hogares mexicanos Chihuahua, French Poodle, Schnauzer y mestizos. Para este estudio se utilizaron los audios segmentados a nivel de vocalización individual. El formato del audio es WAV PCM monocal con una frecuencia de muestreo a 16 kHz.

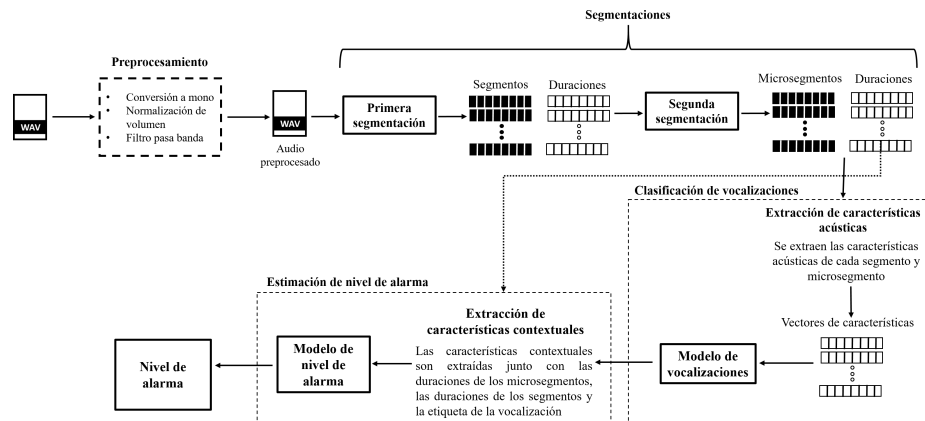


Fig. 1. Esquema de detección de un nivel de alarma.

De acuerdo al teorema de muestreo de Nyquist-Shannon, si se utilizan 16 kHz como frecuencia de muestreo, la frecuencia máxima de la señal sería de 8,000 Hz. De acuerdo al trabajo de [2] se han registrado vocalizaciones de perro que llegan a alcanzar 5,000 Hz, por lo que la frecuencia de muestreo es suficiente.

#### 4. Método para la estimación de nivel de alarma

El método propuesto para determinar el nivel de alarma a partir de ladridos, incluye dos modelos de clasificación que son aplicados uno después del otro. El primero, identifica el tipo de vocalización y el segundo modelo que estima el nivel de alarma en las vocalizaciones. Para entrenar cada uno de estos modelos se prepararon dos conjuntos de audios. El primero, al cual llamamos conjunto para clasificación de vocalizaciones, esta enfocado en diferenciar los diferentes tipos de vocalizaciones que emiten los perros utilizando las siguientes categorías: ladridos, chillidos, y gruñidos.

Además, se incluyeron en el entrenamiento todos los sonidos del ambiente capturados durante la recolección de los datos. A esta categoría se le denominó: sonido. Los segmentos de audio tienen una duración de 0.3 a 5 segundos. Este conjunto de audios se dividió en entrenamiento y prueba, ver Tabla 1.

El segundo conjunto de audios, al cual llamamos conjunto para detección de alarma, fue formado por 112 audios con una duración de 15 segundos y etiquetados con los posibles estados de nivel de alarma: tranquilo, activo, alerta y alarma, ver Tabla 2.

A diferencia del conjunto para clasificación de vocalizaciones que contiene segmentos de vocalizaciones individuales, el conjunto para detección de alarma, son una sucesión de vocalizaciones representativas de cada uno de los niveles de alarma de interés. Estos 112 audios fueron seleccionados de forma manual por las mismas personas que aplicaron los estímulos descritos en la sección anterior.

Enseguida se describe el flujo, tratamiento de los datos y construcción de modelos para encontrar el nivel de alarma de un audio (Figura 1). Este proceso consta de cuatro fases: fase de pre-procesamiento, segmentación, extracción de características acústicas, clasificación de vocalizaciones y predicción del nivel de alarma.

**Tabla 3.** Características acústicas.

<b>Características</b>	<b>No. características</b>
MFCC	12
Delta's MFCC	12
LPCC	12
Delta's LPCC	12
Zero Crossing Rate	1
Log Energy	1
<b>Total características</b>	<b>50</b>

#### 4.1. Fase de pre-procesamiento

El audio de entrada debe estar en formato WAV, tener una duración de 15 segundos a una tasa de muestreo de 16kHz. A este audio se le aplican los siguientes procedimientos:

1. Convertir el audio a un solo canal.
2. Aplicar filtro para la normalización del volumen.
3. Aplicar filtro pasa banda de 150Hz a 3500Hz.

Dado que el ancho de banda del ladrido es menor a 3,000Hz se conservan todas las frecuencias menores a este valor. En el trabajo de [12] se observa que los ladridos de distintos perros se encuentran en el intervalo de frecuencia de 1,000 Hz a 2,000 Hz en su mayoría, aunque otros tipos de vocalizaciones del perro pueden alcanzar frecuencias por arriba de ese rango.

#### 4.2. Fase de segmentación

El audio resultante de la fase de preprocesamiento, es segmentado por primera vez con el fin de encontrar grupos de vocalizaciones, es decir periodos en donde se den varias vocalizaciones continuas sin una pausa perceptible entre ellas. El resultado de esta segmentación son tuplas de vectores, donde:

- El primer elemento es un vector que contiene señal de cada segmento del audio.
- El segundo elemento es un vector que contiene los tiempos de inicio y fin de cada segmento.

Enseguida, a cada uno de los segmentos del audio se le aplica una segunda segmentación con el objetivo de obtener micro segmentos, que representan vocalizaciones individuales. Es decir, un micro segmento contiene un sólo ladrido, chillido o gruñido.

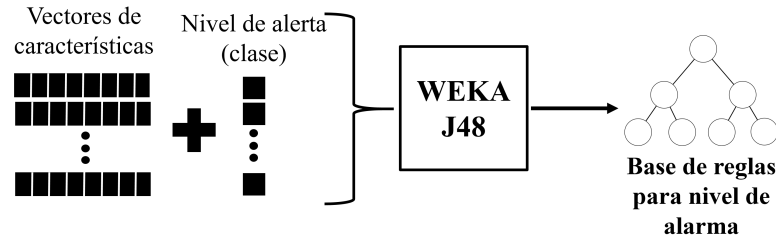


Fig. 2. Proceso de generación de reglas para un nivel de alarma.

#### 4.3. Extracción de características acústicas

Después de aplicar los procedimientos de la fase de segmentación se caracteriza el audio. Como se mencionó anteriormente, la tarea de determinar un nivel de alarma se basa en dos etapas primordiales, determinar el tipo de vocalización y estimar el nivel de alarma. Para ello, fue necesario obtener características diferentes para cada una de las dos etapas. En el caso del conjunto para clasificación de vocalizaciones la información se representó con las características acústicas de bajo nivel mencionadas en la Tabla 3.

En el caso del conjunto para detección de alarma, se determinó que las características contextuales o también llamadas de alto nivel obtenidas debían estar relacionadas con el contexto de la grabación, por lo cual se consideran características basadas en el número de vocalizaciones presentes en un periodo y en la manera en que se agrupan. A continuación se listan las nueve características de este tipo que fueron usadas:

- Número de ladridos.
- Número de chillidos.
- Número de gruñidos.
- Número de grupos de ladridos.
- Duración de los ladridos.
- Duración de silencio.
- Promedio log energy de ladridos.
- Promedio de decibelios en ladridos.
- Promedio del pitch de ladridos.

#### 4.4. Clasificación de vocalizaciones

El modelo para detección de vocalizaciones fue construido y evaluado empleando los conjuntos de datos que se describen en la Tabla 1. Ambos conjuntos de datos fueron caracterizados de acuerdo a la tabla 3 y se evaluaron los siguientes cinco algoritmos de aprendizaje automático de WEKA:

- Decision three algorithm - J48.
- Function algorithm - SVM (polynomial kernel).

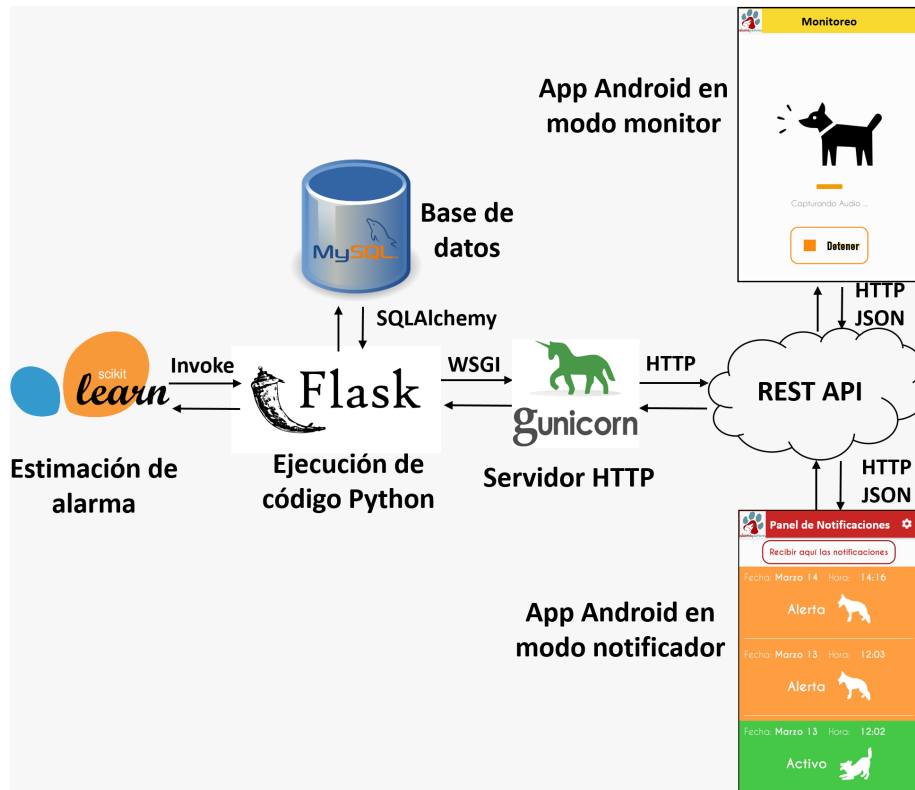


Fig. 3. Arquitectura del sistema.

- Bayesian algorithm - Naive Bayes.
- Ensemble algorithm - Random Forest.
- Ensemble algorithm - Bagging (REPTree).

Además se emplearon los siguientes cuatro algoritmos de la biblioteca de aprendizaje automático para python Scikit-Learn:

- Random Forest.
- KNN.
- Gradient Boosting.
- Extratrees.

En el proceso de detección de alarma, una vez obtenidos los micro segmentos, se extraen las características acústicas de cada uno, de manera que ahora se tienen los vectores de características acústicas. Estos son clasificados por el modelo de vocalizaciones para predecir si un segmento es un ladrido, un chillido, un gruñido o un sonido ambiental.

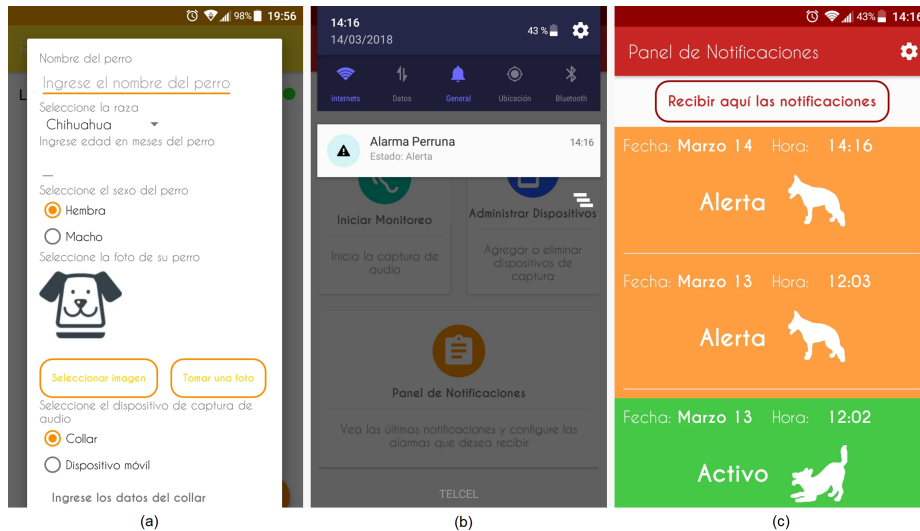


Fig. 4. Capturas de pantalla de la aplicación Alarma Perruna.

#### 4.5. Determinación del nivel de alarma

Terminada la fase de clasificación de vocalizaciones, se realiza la extracción de características contextuales. Enseguida se describe el proceso de generación de estas reglas para el cálculo de nivel de alarma. Dicho proceso se basa en obtener las características contextuales sobre el conjunto de audios para detección de alarma. A cada vector de características se le agregó el nivel de alarma correspondiente de cada audio.

Las características contextuales extraídas junto con el nivel de alarma asignado, fueron alimentados al algoritmo J48 de WEKA, para que de esta manera se obtuviese el modelo de clasificación a utilizar en la determinación del nivel de alarma correspondiente a un audio de 15 segundos (ver figura 2). Una vez que se tienen estas reglas, son aplicadas sobre las características resultantes en el paso anterior, de esta manera se obtiene el nivel de alarma correspondiente al audio de entrada.

### 5. Arquitectura del sistema

El sistema se compone de dos elementos, el primero es una Interfaz de Programación de Aplicaciones (API) bajo el paradigma de software basado en red conocido como Transferencia de Estado Representacional (REST). El segundo es una aplicación móvil (Android) que puede ejercer el rol de monitor para la captura y envío de los audios del perro, o el rol de notificador para la recepción y visualización de notificaciones por parte del API REST. Cuando la aplicación funciona en modo de monitor, se encarga de la obtención de muestras acústicas, las cuales son enviadas al API, para su procesamiento.

**Tabla 4.** Clasificación de vocalizaciones.

	<b>Clasificador</b>	<b>F-measure</b>
Weka	J48	0.762
	SVM	0.827
	Naive Bayes	0.614
	Random Forest	0.866
	Bagging	0.840
	KNN	0.833
Scikit-Learn	Random Forest	0.886
	Gradient Boosting	0.815
	ExtraTree	0.854

Cuando la aplicación móvil funciona en modo de notificador, recibe y muestra información al usuario sobre el estado actual del perro. Además, mediante la aplicación se hace el registro del usuario y el registro de otros dispositivos de captura de audio, como podrían ser collares o pecheras desarrollados por terceros.

La REST API, lleva a cabo el almacenamiento de datos de los usuarios en la base de datos, el procesamiento de audio, la clasificación del audio de entrada, la identificación de un nivel de alarma y el envío de la notificación al dispositivo notificador. En la Figura 3 se muestra la arquitectura que hace uso de un conjunto de herramientas (MySQL, Unicorn, Flask) que permiten un control flexible sobre cada componente del front y back end.

El servidor que realiza el procesamiento del audio interactúa con las aplicaciones mediante una API basada en Python que se encarga de ejecutar los algoritmos y métodos necesarios para segmentar, caracterizar y clasificar el audio. Una vez que el servidor recibe un archivo de audio de la aplicación monitor, se activa el procesamiento, y se envía un mensaje asincrónico a la aplicación notificadora.

## 6. Aplicación alarma perruna

Como se mencionó anteriormente cuando la aplicación funciona en modo de monitor se encarga de la obtención de muestras acústicas, las cuales son enviadas por medio del API, para su tratamiento. Las principales funcionalidades de la aplicación son:

1. Registro: La primera vez que se usa la aplicación se debe registrar el usuario con un correo electrónico y contraseña. Una vez registrado el usuario, inicia sesión y deberá acceder a la opción Registro de perros en la pantalla principal para registrar a uno o más perros. Por cada perro se debe especificar su nombre, raza, sexo, y fotografía (ver Figura 4 (a)). Además, por cada perro registrado se debe dar de alta el dispositivo que capturaré sus vocalizaciones, que puede ser un dispositivo móvil en el cual se ejecute la aplicación en modo monitor o un dispositivo vestible, como un collar o una pechera. Mediante la REST API se lleva a cabo el almacenamiento de la información de registro en la base de datos.

**Tabla 5.** Matriz de confusión clasificación de nivel de alarma.

	<b>Tranquilo</b>	<b>Activo</b>	<b>Alerta</b>	<b>Alarma</b>
<b>Tranquilo</b>	12	0	0	0
<b>Activo</b>	1	7	1	0
<b>Alerta</b>	0	1	48	9
<b>Alarma</b>	0	0	14	19

2. **Monitoreo:** para iniciar el monitoreo del sonido ambiental, en la pantalla principal de la aplicación se pulsa el botón Iniciar monitoreo. En ese momento la aplicación empieza a capturar el sonido ambiental, y lo envía al sistema de procesamiento cada 15 segundos.

Durante el monitoreo la aplicación muestra en pantalla un mensaje que dice capturando audio, una barra de progreso que se reinicia cada 15 segundos y un botón con la etiqueta Detener (esta pantalla se muestra en la parte superior de la Figura 3). El dispositivo deberá colocarse en un lugar de la casa donde se espere que el perro se mantenga cerca y las vocalizaciones sean capturadas con suficiente intensidad.

3. **Notificaciones:** En este apartado de la aplicación se reciben las notificaciones (ver Figura 4 (c)). A esta pantalla se accede pulsando la opción Panel de notificaciones en la pantalla principal. Cuando la aplicación está en modo notificador puede estar trabajando en segundo plano y notificar al usuario como se muestra en la Figura 4 (b).

En este apartado la aplicación cuenta con opciones de configuración para indicar que niveles de alarma quiere recibir el usuario. En dicha pantalla el usuario marca los estados del perro que desea que se le notifiquen. Por ejemplo, únicamente el nivel de alarma máximo.

## 7. Resultados

El sistema fue evaluado en las dos etapas de clasificación de las que consta. En la Tabla 4 se muestra el desempeño en la primera etapa, es decir en la tarea de clasificación de vocalizaciones y sonidos ambientales correspondiente a cada algoritmo de aprendizaje y de acuerdo al valor F-measure obtenido.

Como se puede observar el mejor desempeño se obtuvo con el algoritmo Random Forest de Scikit-Learn. Para la evaluación de la segunda etapa, debido a que el tamaño del corpus es reducido, sólo 112 muestras, se optó por evaluar el modelo realizando 10-fold cross-validation. Usando este esquema de validación se obtuvo un F-measure de 0.764.

A continuación se muestra la matriz de confusión resultante de aplicar este proceso sobre los datos descritos en la Tabla 5. En esta matriz de confusión se puede apreciar que existe una mayor confusión entre las clases alerta y alarma.

## 8. Conclusiones y trabajo futuro

En este trabajo se desarrolló un sistema para la estimación del nivel de alarma transmitido en los ladridos del perro doméstico. El sistema está compuesto por una aplicación móvil y un sistema de procesamiento, y clasificación de la señal de audio que se comunican por medio de un API. Usando descriptores acústicos de bajo nivel como MFCCs, LPCC, Zero Crossing Rate y Log Energy, que son usados para entrenar un modelo de clasificación de tipos de vocalizaciones con el algoritmo de aprendizaje máquina Random Forest, se logró un F-measure de 0.886. Lo cual puede considerarse como un muy buen resultado para este tipo de tareas de clasificación.

A partir de la clasificación del tipo de vocalización y de otros descriptores de alto nivel en segmentos de audio de 15 segundos se entrenó un árbol de decisión que determina el nivel de alarma detectado en las vocalizaciones para ser notificado al usuario. Esta estimación mostró un F-measure de 0.764 y un mayor nivel de confusión en las clases de alerta y alarma, lo cual podría no ser tan grave en la práctica, ya que de acuerdo a la orientación de la aplicación, lo importante es identificar cuando el perro puede haber detectado algo inusual en su entorno.

Como trabajo futuro es necesario realizar una evaluación de la aplicación en un ambiente real, ya que en la evaluación reportada se usaron sólo muestras de la base de datos previamente colectada, pero no se usaron muestras obtenidas a través de la aplicación y enviadas por medio del API.

La aplicación Alarma Perruna muestra un desempeño en la clasificación de vocalizaciones caninas suficiente para ser usada de manera eficaz por dueños de perros que dejan a sus mascotas en casa y desean saber que sucede mientras están fuera. Mediante el uso de esta aplicación los dueños de los perros pueden saber si existe algo inusual en su hogar, por ejemplo si un extraño intenta entrar, o el perro tiene algún problema y tomar alguna medida al respecto.

**Agradecimientos.** Este trabajo de investigación se ha llevado a cabo en el contexto del programa “Cátedras CONACyT ”financiado por el Consejo Nacional de Ciencia y Tecnología (CONACyT).

## Referencias

1. Ayu Melati, A. I. S., Karyono, K. K.: Audio detection (audition): Android based sound detection application for hearing-impaired using AdaBoostM1 classifier with REPTree weaklearner. In: Proceedings of the Asia-Pacific Conference on Computer Aided System Engineering, pp. 136–140 (2014) doi: 10.1109/apcase.2014.6924487
2. Frommolt, K. H., Gebler, A.: Directionality of dog vocalizations. The Journal of the Acoustical Society of America, no. 1, pp. 561–565 (2004) doi: 10.1121/1.1758971
3. Hantke, S., Cummins, N., Schuller, B.: What is my dog trying to tell me? the automatic recognition of the context and perceived emotion of dog barks. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5134–5138 (2018) doi: 10.1109/icassp.2018.8461757

4. Jain, D., Mack, K., Amrous, A., Wright, M., Goodman, S., Findlater, L., Froehlich, J. E.: Homesound: An iterative field deployment of an in-home sound awareness system for deaf or hard of hearing users. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–12 (2020) doi: 10.1145/3313831.3376758
5. Kevin, S. J.: The Domestication, Behaviour and Use of the Dog. Springer Netherlands (2006) doi: 10.1007/978-1-4020-4362-8\_1
6. Kim, Y., Sa, J., Chung, Y., Park, D., Lee, S.: Resource-efficient pet dog sound events classification using lstm-fcn based on time-series data. *Sensors*, vol. 18, no. 11, pp. 4019 (2018) doi: 10.3390/s18114019
7. Ladha, C., Hammerla, N., Hughes, E., Olivier, P., Ploetz, T.: Dog’s life: Wearable activity recognition for dogs. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, pp. 415–418 (2013) doi: 10.1145/2493432.2493519
8. Lakatos, G., Gácsi, M., Topál, J., Miklósi, Á.: Comprehension and utilisation of pointing gestures and gazing in dog–human communication in relatively complex situations. *Animal Cognition*, vol. 15, no. 2, pp. 201–213 (2012) doi: 10.1007/s10071-011-0446-x
9. 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*, vol. 18, no. 2, pp. 405–421 (2015) doi: 10.1007/s10071-014-0811-7
10. Molnár, C., Kaplan, F., Roy, P., Pachet, F., Pongrácz, P., Dóka, A., Miklósi, Á.: Classification of dog barks: A machine learning approach. *Animal Cognition*, vol. 11, no. 3, pp. 389–400 (2008) doi: 10.1007/s10071-007-0129-9
11. Nguyen, Q., Choi, J.: Matching pursuit based robust acoustic event classification for surveillance systems. *Computers & Electrical Engineering*, vol. 57, pp. 43–54 (2017) doi: 10.1016/j.compeleceng.2016.11.007
12. Sakamoto, S.: Sound attenuation devices for dogs barking (estimation of aperture ratio and experimental study of silencer). *International Journal of Mechanical Engineering and Applications*, vol. 2, no. 1, pp. 18–24 (2014), doi: 10.11648/j.ijmea.20140201.14
13. Schneider, M., Jokisch, O.: Towards a robust analysis and classification of dog barking. *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung*, pp. 117–124 (2020)
14. Yin, S.: A new perspective on barking in dogs (*cannis familiaris*). *Journal of Comparative Psychology*, vol. 16, pp. 189–193 (2002)