



ESPE

**UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA**

**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES**

**CARRERA DE INGENIERÍA EN ELECTRÓNICA Y
TELECOMUNICACIONES**

**TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL TÍTULO
DE INGENIERA EN ELECTRÓNICA Y TELECOMUNICACIONES**

**TEMA: ADAPTACIÓN E IMPLEMENTACIÓN DE ALGORITMOS VOICE
ACTIVITY DETECTION PARA LA DETECCIÓN DE EVENTOS
MICROSÍSMICOS DEL VOLCÁN COTOPAXI.**

AUTORA: ROSERO JÁCOME, KAREN GISSELL

DIRECTOR: ING. LARCO BRAVO, JULIO CÉSAR

SANGOLQUI

2019



DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES
CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

CERTIFICACIÓN

Certifico que el trabajo de titulación, *"ADAPTACIÓN E IMPLEMENTACIÓN DE ALGORITMOS VOICE ACTIVITY DETECTION PARA LA DETECCIÓN DE EVENTOS MICROSÍSMICOS DEL VOLCÁN COTOPAXI"* fue realizado por la señorita *Rosero Jácome Karen Gissell*, el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 18 de diciembre de 2019

.....
Ing. Larco Bravo, Julio César
DIRECTOR
C.C: 1710638808



DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES
CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

AUTORÍA DE RESPONSABILIDAD

Yo, *Rosero Jácome, Karen Gissell*, declaro que el contenido, ideas y criterios del trabajo de titulación: *“Adaptación e implementación de algoritmos Voice Activity Detection para la detección de eventos microsísmicos del volcán Cotopaxi”* es de mi autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente el contenido de la investigación mencionada es veraz.

Sangolquí, 18 de diciembre de 2019

.....
Rosero Jácome, Karen Gissell

C.C: 0401296819



DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES
CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

AUTORIZACIÓN

Yo, Rosero Jácome, Karen Gissell autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: “Adaptación e implementación de algoritmos Voice Activity Detection para la detección de eventos microsísmicos del volcán Cotopaxi” en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 18 de diciembre de 2019

.....
Rosero Jácome, Karen Gissell

C.C: 0401296819

DEDICATORIA

Mi trabajo de titulación está dedicado principalmente a mis padres, quienes han sabido guiarme en cada momento difícil dándome sabios consejos y enseñándome siempre a dar lo mejor de mí. Me han brindado las posibilidades para realizar mis estudios y no han dudado de mi capacidad en ningún momento. A mis hermanos Ana Paula y Esteban, que son mi inspiración para ser mejor cada día ya que siempre busco darles el mejor ejemplo. Toda la confianza que han depositado en mí ha sido fuente de inspiración para no rendirme cuando las circunstancias no han sido las mejores. Con la ayuda de Dios, quiero seguir alcanzando mis metas a su lado.

Karen Gissell Rosero Jácome.

AGRADECIMIENTO

Primeramente, agradezco a Dios por haberme dado la sabiduría y fortaleza para culminar mi carrera universitaria. Agradezco a mi familia que siempre me ha brindado su apoyo incondicional y me han educado con principios y valores, inspirándome a alcanzar mis metas y a soñar en grande. Un sincero agradecimiento a mis profesores universitarios que desde los primeros niveles me han sabido formar como una estudiante perseverante y se han esforzado por transferir su conocimiento de la mejor manera. A mi tutor, Ing. Julio Larco, quien me ha guiado en la presente investigación, gracias por la paciencia y la confianza depositada en mí. Mi agradecimiento también hacia el PhD. Felipe Grijalva, docente de la EPN, quien supo encaminarme para que los resultados de este trabajo sean presentados de la mejor manera, gracias por retarme a ser mejor cada día. Finalmente, agradezco a mis compañeros y amigos con quienes he compartido esta época universitaria, nada hubiera sido igual sin ustedes.

ÍNDICE DE CONTENIDOS

CERTIFICADO DEL DIRECTOR	i
AUTORÍA DE RESPONSABILIDAD	ii
AUTORIZACIÓN	iii
DEDICATORIA	iv
AGRADECIMIENTO	v
ÍNDICE DE CONTENIDOS	vi
ÍNDICE DE TABLAS	ix
ÍNDICE DE FIGURAS	x
RESUMEN	xi
ABSTRACT	xii
CAPÍTULO I	
DESCRIPCIÓN	
1.1. Introducción.....	1
1.2. Justificación e Importancia.....	4
1.3. Alcance del Proyecto.....	6
1.4. Objetivos.....	7
1.4.1. General.....	7
1.4.2. Específicos.....	7
1.5. Trabajos Relacionados.....	8
1.6. Organización del Trabajo.....	10
CAPÍTULO II	
MARCO TEÓRICO	
2.1. Vulcanología.....	12
2.1.1. Generalidades.....	12
2.1.2. Vulcanología en el Ecuador.....	14
2.1.3. Volcán Cotopaxi.....	15

2.2. Sismicidad volcánica	19
2.2.1. Monitoreo de la sismicidad	22
2.2.2. Tipos de Eventos Sismo Volcánicos	22
2.3. Conceptos Básicos	26
2.3.1. Detección de actividad de voz	27
2.3.2. Sustracción espectral de potencia	28
2.3.3. Frecuencia de Mel	31
2.3.4. Prueba de razón de máxima verosimilitud	33
2.3.5. Modelos de entrenamiento de mezcla Gaussiana	34
2.3.6. Método de agrupamiento de k medias	36

CAPÍTULO III

METODOLOGÍA

3.1. Bases de Datos	39
3.2. Preprocesamiento	40
3.3. Algoritmo A-VAD	42
3.3.1. Extracción de MFCCs	42
3.3.2. Sustracción espectral	45
3.3.3. Modelo de entrenamiento	48
3.1. Algoritmo de decisión	50

CAPÍTULO IV

RESULTADOS

4.1. Bases de Datos	54
4.2. Medidas de Rendimiento	57
4.3. Etapa de experimentación	59
4.4. Evaluación del desempeño del algoritmo	63
4.4.1. Resultados Base de Datos 1	64
4.4.2. Resultados Base de Datos 2	67

CAPITULO V

DISCUSIÓN

5.1. Comparación con trabajos similares	74
---	----

5.2. Conclusiones.....75

5.3. Trabajos Futuros76

REFERENCIAS78

ÍNDICE DE TABLAS

Tabla 1 <i>Relaciones entre tipos de volcanes, estilos y características eruptivas.</i>	13
Tabla 2 <i>Tipos de eventos volcánicos peligrosos</i>	13
Tabla 3 <i>Parámetros de control en el algoritmo A-VAD.</i>	51
Tabla 4 <i>Número de eventos por tipo en la Base de datos 1</i>	55
Tabla 5 <i>Número de eventos por tipo en la Base de datos 1 extendida</i>	55
Tabla 6 <i>Número de eventos por tipo en la Base de datos Discreta</i>	57
Tabla 7 <i>Resultados experimentación con el tamaño de ventana</i>	60
Tabla 8 <i>Experimentación con valores de energía de los ESV.</i>	62
Tabla 9 <i>Resultados experimentación con valor mínimo de energía</i>	62
Tabla 10 <i>Resultados experimentación con porcentaje de tramas</i>	63
Tabla 11 <i>Resultados Medidas de rendimiento Base de Datos 1</i>	66
Tabla 12 <i>Resultados Medidas de rendimiento Base de Datos 2</i>	69
Tabla 13 <i>Mejora de Medidas de rendimiento por experimentación</i>	72
Tabla 14 <i>Contraste de resultados entre las Bases de Datos</i>	73
Tabla 15 <i>Comparación de resultados con un trabajo previo</i>	74

ÍNDICE DE FIGURAS

<i>Figura 1.</i> Volcán Cotopaxi visto desde el Norte.	16
<i>Figura 2.</i> Partes de un evento sísmico	20
<i>Figura 3.</i> Propagación de las ondas S y P	21
<i>Figura 4.</i> Propagación de las ondas L y R.....	21
<i>Figura 5.</i> Ejemplo de evento tipo VT.....	23
<i>Figura 6.</i> Ejemplo de evento tipo LP.	24
<i>Figura 7.</i> Ejemplo de evento tipo híbrido.....	25
<i>Figura 8.</i> Tremor de la estación CVTR del volcán Turrialba.....	26
<i>Figura 9.</i> Proceso para obtener las características de los MFCCs.....	31
<i>Figura 10.</i> Banco de filtros de la frecuencia de Mel	32
<i>Figura 11.</i> Diagrama de bloques de la metodología utilizada en la investigación.	39
<i>Figura 12.</i> Ejemplo de una señal microsísmica antes y después de ser preprocesada.	42
<i>Figura 13.</i> Diagrama de bloques del proceso de extracción de MFCCs.	43
<i>Figura 14.</i> Banco de Filtros de Mel utilizado para señales sísmicas	44
<i>Figura 15.</i> Ejemplo de señal sísmica que pasó por el proceso de Sustracción Espectral.	47
<i>Figura 16.</i> Coincidencia picos de frecuencia con vector LLR calculado.	49
<i>Figura 17.</i> Diagrama de bloques de los resultados a presentarse	53
<i>Figura 18.</i> Ejemplo de almacenamiento de ESV de la Base de Datos 1	54
<i>Figura 19.</i> Señales sísmicas en formato .mat	54
<i>Figura 20.</i> Ejemplo de EVS recortado de la señal sísmica original.	56
<i>Figura 21.</i> Ejemplo de cálculo de precisión en la etapa de experimentación.....	61
<i>Figura 22.</i> Validación de un ESV en la Base de Datos 1	64
<i>Figura 23.</i> Curva ROC resultante Base de Datos 1	67
<i>Figura 24.</i> Validación de ESV Base de Datos 2.....	68
<i>Figura 25.</i> Curva ROC resultante Base de Datos 2	70

RESUMEN

Los sistemas de alerta temprana se vuelven elementos imprescindibles contemplando la probabilidad de ocurrencia de desastres naturales, ya que se busca salvaguardar un número máximo de vidas, de esta manera, este tema toma protagonismo a nivel mundial en el marco de la monitorización de los Eventos Sismo Volcánicos (ESV) que son un indicador de una eventual erupción volcánica. Ecuador cuenta con el Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN) que se encarga de realizar la monitorización, análisis y clasificación de la actividad volcánica en el país; una red de sismómetros de banda ancha brinda los datos obtenidos del volcán Cotopaxi, para que esta información sea procesada y analizada. El comportamiento de las señales y los eventos obtenidos por el sismómetro tiene similitud con el de las señales de voz, debido a que la presencia de segmentos de habla se manifiesta con un aumento de frecuencia y amplitud de la señal, entre otros aspectos sensibles a ser analizados; por esto, en este trabajo se genera la adaptación de un algoritmo de detección de actividad de voz como herramienta para la detección de los ESV, mismos que, al aumentar su ocurrencia arrojan una mayor probabilidad de erupción volcánica. Además, se realiza un estudio comparativo del rendimiento del algoritmo implementado, en retrospectiva con trabajos realizados anteriormente en cuanto al mismo volcán. Los resultados obtenidos demuestran el pleno funcionamiento del detector, que opera con precisión del 100% y exactitud de 99.1039%, considerando una base de datos de señales sísmicas del volcán Cotopaxi.

PALABRAS CLAVE:

- **VOICE ACTIVITY DETECTION**
- **DETECCIÓN DE ESV**
- **SISMICIDAD VOLCÁNICA**

ABSTRACT

Early warning systems become essential elements contemplating the probability of occurrence of natural disasters, since it seeks to safeguard a maximum number of lives, in this way, this issue takes center stage worldwide. Ecuador has the Geophysical Institute of the National Polytechnic School (IGEPN) that is responsible for monitoring, analysis and classification of volcanic activity in the country; a network of seismometers provides the data obtained from the Cotopaxi volcano, so that this information can be processed and analyzed. The behavior of the signals and the events obtained by the seismometer is similar to that of the voice signals, because the presence of speech segments manifests itself with an increase in the frequency and amplitude of the signal, among other sensitive aspects to be analyzed; for this reason, the adaptation of a speech activity detection algorithm is generated as a tool for the detection of microseisms, which, by increasing their occurrence, show a greater probability of volcanic eruption. In addition, a comparative study of the performance of the algorithm to be implemented is done, in retrospect with works previously carried out regarding the same volcano. The results show the full operation of the detector, which works with 100% of precision and 99.1039% accuracy, considering a database of seismic signals from the Cotopaxi volcano.

KEYWORDS:

- **VOICE ACTIVITY DETECTION**
- **DETECTION OF MICROSISMIC EVENTS**
- **VOLCANIC SISMICITY**

CAPÍTULO I

1. DESCRIPCIÓN

1.1. Introducción

El Cinturón de Fuego de los Andes es un área de 40000 km donde existe intensa actividad volcánica y sísmica, en su territorio se encuentran el 75% de los volcanes activos del mundo y se han producido el 90% de los terremotos documentados. Su área se divide en cuatro zonas volcánicas, dentro de las cuales el volcán Cotopaxi, Cerro Azul y Ojos del Salado son considerados los más activos y peligrosos. (Rosenberg, 2018).

En Ecuador está ubicado uno de los volcanes más peligrosos del mundo, el volcán Cotopaxi. Mismo que se posiciona en la Cordillera Oriental entre las ciudades de Latacunga y Quito, a una altura de 5897 m sobre el nivel del mar. Es un volcán en estado activo con un cráter de diámetro de 800 m en sentido Norte-Sur y 650 m en sentido Este-Oeste. Cercanos al él, se encuentran otros volcanes como el Sincholagua, Quilindaña y Rumiñahui. Los cinco episodios eruptivos que ha presentado el volcán desde el año 1532 han generado grandes pérdidas socio – económicas en el país, ya que su erupción da lugar a la formación de lahares (flujos de lodo y escombros) que fluyen por zonas que actualmente son densamente pobladas como el Valle Interandino entre Mulaló y Latacunga, y una parte del valle de los Chillos, afectando así a más de 300000 personas. Debido al rango repetitivo en que se han producido las erupciones, no existe duda de que se va a suscitar una nueva erupción en el plazo de décadas, motivo por el cual el volcán Cotopaxi es monitoreado constantemente. (IGEPN, Cotopaxi, 2017)

La relevancia del volcán Cotopaxi ha aumentado a nivel mundial debido a su peligrosidad, es por eso, que al igual que otros volcanes activos, se mantiene en observación y estudio constante para prever los efectos desastrosos de su erupción. La monitorización de la actividad volcánica mediante redes de sensores es un proceso que se viene realizando desde hace más de una década en el Ecuador. Como se detalla en (Werner-Allen, Johnson, Ruiz, Lees, & Welsh, 2005), estos sensores son utilizados para determinar el origen y localización de un sismo o una explosión, para estudiar la estructura interior del volcán o para diferenciar en la señal, las verdaderas erupciones, del ruido que provocan otras señales que no son de interés volcanológico.

La erupción del volcán Cotopaxi puede dar lugar a la aparición de una serie de fenómenos como: gases volcánicos que si bien no son letales, en altas cantidades podrían perjudicar la salud de los seres vivos que se expongan a ellos; flujos de lava que descienden lentamente por las faldas del volcán pero no llegan a cubrir largas distancias hasta llegar a un estado de petrificación que podría mantener una temperatura alta incluso por meses; domos de lava formados por magma altamente viscoso que le obliga a estancarse; flujos piroclásticos o nubes ardientes formados por gases, ceniza y fragmentos de roca que descienden por el cono del volcán y tiene una alta probabilidad de ocurrencia debido a la historia eruptiva del volcán; lluvia de ceniza y piroclásticos que por su menor peso son acarreadas por el viento y llegan a cubrir áreas a mayor distancia del cráter; lahares que son flujos de lodo y escombros de material volcánico formado por la fusión del casquete glaciar que facilita el deslizamiento del mismo a velocidades de hasta 100km/h y finalmente se presentarían sismos de baja magnitud producidos en el cono del volcán y que no afectan ni son percibidos por las comunidades aledañas. Estos sismos de baja magnitud serán el objeto de la presente investigación ya que se analiza el espectro tanto en tiempo como frecuencia que generan estas señales sísmicas. (Hall, y otros, 2005)

El Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN) ejecuta un sistema de monitorización y vigilancia permanente durante las 24 horas del día y los 365 días del año. Cuenta con un Sistema Informático para el Procesamiento y Análisis de Señales Sísmicas de Volcanes en el Ecuador (SIPASS), mismo que fue desarrollado en el 2011 con el fin de reemplazar los anteriores programas para el análisis de señales sísmicas que ya estaban obsoletos y requerían de intervención manual. Con esto se busca brindar una respuesta inmediata al identificarse una crisis sísmica o volcánica en el territorio ecuatoriano. En (Viracucha & De la Bastida, 2014) se explica el funcionamiento del sistema, mismo que reconoce los ESV que los volcanes producen, utilizando la transformada rápida de Fourier (FFT) para calcular el espectro de la señal, luego el usuario verifica el rango de frecuencias y determina el tipo de evento de manera manual, posteriormente esta información es almacenada en una hoja de cálculo de Excel.

En este contexto, en la presente investigación se desarrolla un algoritmo automático de detección de los ESV del volcán Cotopaxi. Al tratar las señales sísmicas que fueron obtenidas por los sensores, se verificó la similitud de estas con las señales de voz. Es así como se optó por adaptar las rutinas y procesos utilizados por los algoritmos de Detección de Actividad de Voz (VAD, del inglés *Voice Activity Detection*) de manera que se acoplen a los parámetros de trabajo de las señales sísmicas y consecuentemente reconozca los segmentos de la señal en que se produjeron ESV que pueden servir de alerta temprana para declarar el estado eruptivo del volcán.

Se han realizado ya varias investigaciones relacionadas al reconocimiento del habla utilizando técnicas de Detección de Actividad de Voz. Las principales alternativas se basan en el cálculo y verificación de la energía en cortos periodos, funciones de autocorrelación, tasa de cruce por cero, etc. En la presente investigación se desarrolla un algoritmo adaptativo de Detección de Actividad de Voz (A-VAD) cuyas técnicas han sido adaptadas para obtener un correcto

funcionamiento con las señales sísmicas. En el detector desarrollado se realiza un preprocesamiento de la señal, junto con técnicas como la sustracción espectral de potencia de la señal, el cálculo de los coeficientes cepstrales de la frecuencia de Mel, modelos de entrenamiento de mezcla Gaussiana (GMM) y pruebas de máxima verosimilitud. Se busca mejorar las medidas de rendimiento de detectores de ESV desarrollados anteriormente para el volcán Cotopaxi y así aportar a la ejecución de una alerta temprana que disminuirá las cifras de tragedias al presentarse una eventual erupción.

Las señales sísmicas utilizadas como datos de este trabajo están organizadas en una base de datos que consta de 1190 ESV etiquetados por el IGEPN. Dichas etiquetas muestran información relevante para la presente investigación, como la fecha, hora, tiempo de inicio y finalización del evento, la estación de monitoreo de la cual proviene, la duración del evento y su tipología.

El presente trabajo investigativo se ha desarrollado con el apoyo del IGEPN y bajo la tutela de los docentes anexados al proyecto de investigación “SRASI - Implementación de un Sistema de Reconocimiento Automático de Señales Sísmicas del Volcán Cotopaxi” registrado con número N.-2016-EXT-038 en la Universidad de las Fuerzas Armadas ESPE.

1.2. Justificación e Importancia

Las erupciones volcánicas son incidentes geológicos que ocurren frecuentemente en la zona conocida como Cinturón de Fuego de los Andes. Tanto geólogos como vulcanólogos son profesionales que buscan predecir con precisión cuando un volcán va a entrar en periodo de erupción y cuál será su impacto en las áreas aledañas. Motivo por el cual los sistemas de alerta temprana se han convertido en un tema de relevancia para preservar la vida en general y mitigar el riesgo de pérdidas materiales. Es por eso que se intenta aumentar el nivel de precisión de los

sistemas de detección, ya que estos sistemas forman parte de la monitorización que ejecuta el IGEPN y al presentarse un aumento de la actividad microsísmica en el volcán, el IGEPN debe notificar a las autoridades pertinentes.

En el contexto de un Sistema Nacional de Alerta Temprana de Volcanes (Guffanti, y otros, 2007) afirma que el alcance técnico de la monitorización es un proceso crítico ya que está directamente relacionado a la investigación de los componentes clave que suponen un riesgo en un desastre natural de tipo volcánico, para de esta manera iniciar acciones y tomar decisiones políticas que disminuyan el impacto en las comunidades que corren riesgo. En Ecuador, el IGEPN fue decretado como encargado oficial del diagnóstico y la vigilancia de los peligros sísmicos y volcánicos en todo el territorio nacional (IGEPN, Presentación, 2019). Sus alertas están basadas en la información y datos recolectados por una red de sensores instalados en las inmediaciones del volcán. Debido al ruido de fondo que pueden captar los sensores, es de vital importancia que las señales obtenidas sean primero preprocesadas de manera óptima, reduciendo así el nivel de ruido y filtrando las componentes de frecuencia que no están dentro del rango de interés; para luego ser analizadas por el algoritmo de detección que mostrará la cantidad de eventos existentes en un reducido tiempo de procesamiento, conociendo dicha información los especialistas encargados podrán generar los informes correspondientes a la actividad sismo-volcánica. A continuación, el IGEPN remite estos informes a la Secretaria de Gestión de Riesgos que, junto con el Ministerio Coordinador de Seguridad toman las medidas pertinentes para salvaguardar la integridad de la ciudadanía.

Con la detección de los microsismos del volcán Cotopaxi se busca optimizar el proceso de etiquetado y clasificación que realizan en el IGEPN, ya que se extraería directamente el segmento de la señal en que ocurre el evento, evitando así, desplazarse manualmente en la señal en búsqueda

de la siguiente ocurrencia. La implementación de un algoritmo de VAD en el estudio de reconocimiento de ESV de la actividad del volcán Cotopaxi aportará al rendimiento y precisión del sistema general utilizado por el IGEPN, realizándose también un estudio comparativo con los métodos ya implementados.

1.3. Alcance del Proyecto

La presente investigación busca fortalecer y colaborar con los objetivos del proyecto “SRASI - Implementación de un Sistema de Reconocimiento Automático de Señales Sísmicas del Volcán Cotopaxi”, enmarcándose en el contexto de la detección de los ESV en las señales obtenidas de los sensores situados en el volcán Cotopaxi.

Para la etapa de experimentación se utilizan dos bases de datos que contienen ESV previamente etiquetados por el IGEPN. Las señales pasaran por el A-VAD desarrollado entregando la información sobre el tiempo de inicio y finalización de los ESV encontrados. Se busca comparar y verificar los eventos ya etiquetados por el IGEPN en estas señales para calcular las medidas de rendimiento, que son: exactitud, precisión, especificidad, sensibilidad y tasa de error balanceado (BER, del inglés *Balanced Error Rate*).

Posteriormente, se generan curvas de la Característica Operativa del Receptor (ROC) para los resultados de cada base de datos, con el objetivo de no limitar las estadísticas sobre el rendimiento del detector a un solo punto específico de operación, y así tener una idea generalizada del desempeño de este. Por lo tanto, en el presente proyecto se detectan los eventos sísmicos, aportando el tiempo en que inició y finalizó el evento, lo cual ayudará a los profesionales pertinentes a predecir cuando el volcán Cotopaxi entre en una etapa eruptiva para que así, las entidades correspondientes, tomen medidas tempranas de prevención y alerta.

1.4. Objetivos

1.4.1. General

Desarrollar e implementar un algoritmo de VAD en el estudio de la detección de Eventos Sismo Volcánicos (ESV) del volcán Cotopaxi.

1.4.2. Específicos

- Adquirir información mediante un estudio previo de los diferentes algoritmos de VAD, para analizar la compatibilidad con las características de las señales sísmicas del volcán Cotopaxi.
- Desarrollar el detector A-VAD definiendo los valores de los parámetros que se ajusten a las señales sísmicas.
- Evaluar el desempeño del detector en las bases de datos de señales del volcán Cotopaxi, tanto para los ESV etiquetados individualmente como para las señales de simulación de tiempo real.
- Calcular las métricas de desempeño del detector que son: exactitud, precisión, especificidad, sensibilidad y tasa de error balanceado, asegurándose de cumplir con los requerimientos del IGEPN.
- Comparar el algoritmo A-VAD con trabajos previos realizados sobre la misma línea de investigación.

1.5. Trabajos Relacionados

Se han realizado una serie de trabajos investigativos relacionados con la detección de ESV en las señales del volcán Cotopaxi, para lo cual se han utilizado una variedad de métodos y procesos distintos con el objetivo de experimentar y mejorar las métricas de rendimiento del sistema general.

En el estudio (Lara-Cueva, Moreno, Larco, & Benitez, 2016) el preprocesamiento de las señales sísmicas es el primer paso, luego se realiza la detección de un par de puntos al inicio y al final del ESV, pero adicionalmente se analizan los bordes que se forman por el espectro relacionado a la energía, es decir se encuentra un borde ascendente que precede un punto inicial y un borde descendente relacionado a un punto final; para lo que se utiliza un filtro óptimo basado en técnicas de procesamiento digital de imágenes y la selección fija de umbrales para la lógica de decisión. Este detector tuvo una precisión de 95.2% y BER de 0.005, siendo testeado en una base de datos de 436 ESV etiquetados por el IGEPN.

En el artículo (Lara, Paillacho, & Villalva, 2017) se presenta un sistema de detección basado en la clasificación supervisada de los eventos sismo-volcánicos y no volcánicos registrados durante el año 2010 en el volcán Cotopaxi; para la detección, el algoritmo emplea cuatro características adquiridas por medio de la energía de los coeficientes de aproximación y la descomposición wavelet analizando con las familias Daubechies y Symlet obteniendo una precisión del 98%.

Otro tipo de procesamiento para la detección de ESV del volcán Cotopaxi es propuesto en (Jaramillo Aranha, 2015), donde se realiza un estudio tiempo-frecuencia tomando a la señal sísmica como un proceso localmente estacionario con ventanas deslizantes y por medio de una hipótesis de test binario se maximiza la presencia de eventos y se minimiza la ausencia de los mismos. Se logró diseñar 12 detectores de eventos basándose en técnicas usadas en aplicaciones de radar y que

trabajan incluso en presencia de ruido; con estas técnicas se buscó brindar información para definir el tipo de sismos de las zonas detectadas.

Con el objetivo de identificar el incremento de la actividad sísmica en el volcán Cotopaxi, la investigación detallada en (Lara, Carrera, Benítez, Ruiz, & Rojo, 2016) detecta de manera automática los ESV que se presentan en la señal, misma que pasa primero por un subsistema de preprocesamiento compuesto por un filtro pasabanda de orden 128, para luego ser segmentada en ventanas no solapadas de 15 segundos que forman una matriz de datos que servirá para entrenar la máquina y así distinguir entre las ventanas que contienen un evento y las que no. La etapa de detección del sistema mostro una tasa de reconocimiento del 99% siendo aplicada en un grupo de prueba de 116 señales de tipo Largo Periodo (LP) y Volcano Tectónicos (VT) etiquetadas por el IGEPN.

Dentro de las investigaciones enmarcadas en la implementación de detección de actividad de voz para señales sísmicas, el trabajo de (Rodriguez & Lara-Cueva, 2018) propone mejorar un detector de microsismos basado en VAD, para lo cual se utiliza la segmentación de la señal para trabajar con ventanas fijas y analizar la relación señal a ruido. Tanto el algoritmo de VAD como un filtro óptimo de detección de borde es adaptado al sistema, obteniéndose una precisión del 96% en la implementación en tiempo real de las señales del volcán Cotopaxi.

En cuanto al desarrollo de los algoritmos de VAD, (Qing Ong, Chiat Tan , Vengadasalam, Heng Tan , & Hai Ooi, 2017) reconoce tres grupos: algoritmos basados en reglas, algoritmos de modelamiento estadístico y aquellos que manejan un enfoque al aprendizaje automático conocido como “Machine Learning”. Algoritmos codificadores comerciales como G.729B y el formato Multi-Tasa Adaptativo (AMR, del inglés *Adaptive Multi-Rate*) realizan la clasificación de los segmentos de voz basándose en reglas con umbrales predefinidos. Los modelos estadísticos son

superiores a los mencionados anteriormente cuando el segmento de habla no está claramente definido, entre ellos se tiene al modelo de prueba de razón de verosimilitud (LRT), el modelo oculto de Márkov (HMMs) y los modelos mixtos Gaussianos (GMMs). En los últimos años, se han obtenido resultados impresionantes en los algoritmos de VAD que realizan la clasificación utilizando los coeficientes cepstrales de la frecuencia de Mel (MFCCs) como parámetros de discriminación. (Kinnunen & Padmanabhan, 2013). Además, el algoritmo implementado en (Kinnunen & Padmanabhan, 2013) es de tipo adaptativo, motivo por el cual, la variación del ruido de fondo no sería un problema para el reconocimiento de los ESV.

Tras haber indagado en la bibliografía existente se concluye que, si bien han existido varias investigaciones relacionadas a la detección de los ESV que se presentan en las señales del volcán Cotopaxi, es importante adaptar las nuevas técnicas de VAD que se han ido creando como la sustracción espectral de potencia, el cálculo de coeficientes cepstrales de la frecuencia de Mel, los modelos de entrenamiento de mezcla gaussiana, y, además, tomar en cuentas las mejoras que se implementaron en las últimas investigaciones con relación al preprocesamiento de las señales, con el objetivo de mejorar las medidas de rendimiento.

1.6. Organización del Trabajo

El trabajo presentado consta de cinco capítulos que fueron organizados de manera que se facilite su lectura. El primer capítulo engloba las razones que motivaron el desarrollo de este trabajo investigativo. Con la introducción se brinda una idea general del contexto en el cual está inmerso el trabajo, mientras que la justificación e importancia aclaran los argumentos que se tuvo para realizarla. Tanto en el alcance como en los objetivos del proyecto se definen las metas a lograr y

que resultados se busca presentar; finalmente, los trabajos relacionados hacen de guía y sustento de partida para la presente investigación.

En el segundo capítulo se desarrolla el marco teórico de la investigación, abordando los temas principales que deben ser conocidos para lograr entender el proceso del algoritmo propuesto. Se profundiza en las técnicas utilizadas, dando el trasfondo necesario para la comprensión del siguiente capítulo.

La metodología utilizada para la adaptación del algoritmo y su correcta implementación para las señales sísmicas es explicada en el tercer capítulo. Es necesario explicar cómo fueron utilizadas las técnicas explicadas en el segundo capítulo para entender los procesos que se llevaron a cabo con el fin de lograr los objetivos propuestos.

El capítulo cuatro describe los resultados que se generaron tras implementar el algoritmo del detector de ESV en las bases de datos disponibles. Son expuestas las medidas de rendimiento calculadas para un punto de operación fijo. También se presentan las gráficas de las curvas ROC que brindan una idea general del desempeño del detector en cualquier punto de operación.

Finalmente, en el capítulo cinco se realiza la discusión, misma que es un compendio de los resultados obtenidos al probar el algoritmo A-VAD en contraste con trabajos similares realizados anteriormente. Se generan las conclusiones y recomendaciones del trabajo investigativo en relación con los objetivos planteados inicialmente y, por último, se habla sobre los posibles trabajos futuros que se pueden realizar en la línea de este proyecto.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Vulcanología

2.1.1. Generalidades

En el marco del pasado geológico de nuestro planeta se sabe que más de 1300 volcanes han pasado por un estado eruptivo en los últimos 10000 años y en promedio 50 volcanes entran anualmente en erupción. Además, tomando en cuenta que aproximadamente el 10% de la población mundial vive sobre o cerca de volcanes potencialmente eruptivos, es de crucial importancia elaborar planes de emergencia y planificación territorial que salvaguarden la vida de millones de personas. En este contexto opera la Vulcanología, que, como tal, es una ciencia que se dedica al estudio y análisis de los fenómenos eruptivos magmáticos y volcánicos que se producen en el manto y la corteza terrestre y están relacionados directamente con la evaluación y vigilancia de los peligros que suponen las erupciones volcánicas.

Para realizar una correcta clasificación generalizada de los tipos de volcanes en relación con la composición de lava y su estilo eruptivo se presenta la Tabla 1 donde constan las diferentes características que se pueden presentar. Es importante señalar que el volcán de tipo Escudo está ubicado generalmente en el interior de las placas tectónicas y se presentan en la superficie puntos calientes, mientras que el tipo Estrato – volcán está ubicado a lo largo o en los márgenes de las placas tectónicas convergentes.

Tabla 1

Relaciones entre tipos de volcanes, estilos y características eruptivas.

Tipo de volcán	Lava predominante		Estilo eruptivo	Características eruptivas más comunes
	Composición	Viscosidad relativa		
Escudo – Volcán	Basáltica	Fluidal	Generalmente no explosivo a débilmente explosivo	Fuentes de lava, flujos de lava (extensos), lagos y pozas de lava
Estrato – Volcán	Andesítica	Menos fluidal	Generalmente explosivo, pero a veces no explosivo	Flujos de lava (medianos), productos balísticos explosivos, caídas de tefra, flujos y oleadas piroclásticas.
	Dacítica o Riolítica	Viscosa a muy viscosa	Típicamente muy explosivo, pero puede ser no explosivo, especialmente después una explosión grande	Productos balísticos explosivos, caídas de tefra, flujos y oleadas piroclásticas, flujos de lava (cortos) y domos de lava.

Fuente: (Tilling & Beate, 1993)

Independientemente del tipo de volcán, los procesos geológicos son peligrosos por los efectos que conllevan, por lo que en la Tabla 2 serán clasificados de acuerdo con sus características.

Tabla 2

Tipos de eventos volcánicos peligrosos

Tipo de evento volcánico	Características
Flujos de lava	
Corrientes de densidad piroclástica	Son flujos impulsados por la gravedad,
<ul style="list-style-type: none"> • Explosiones laterales dirigidas • Flujos y oleadas piroclásticas 	típicamente impactan aquellas áreas ubicadas dentro de valles cuyas cabeceras están en los volcanes.
Avalanchas de escombros	
Flujos de escombros volcánicos o lahares	

CONTINÚA 

Crecientes o aluviones	
Caídas de tefra y proyectiles balísticos	Son lanzados a la atmósfera por las explosiones volcánicas y dispersadas por los vientos hasta cubrir áreas extensas.
Gases volcánicos	Escapan durante y entre erupciones y pueden afectar grandes áreas.
Terremotos	
Tsunamis	Se presentan en menor probabilidad, pero al generarse, representan el mayor peligro.
Ondas atmosféricas de choque	

Fuente: Modificado de (Blong, 1984)

Para llevar a cabo la mitigación del peligro y la reducción del riesgo volcánico las instituciones académicas y los programas de investigación científica realizan un sustento investigativo de: procesos eruptivos, productos volcánicos, fenomenología volcánica, modelos eruptivos, intervalos de recurrencia y redes de vigilancia, que, basándose en el comportamiento pasado y presente de cada volcán ayudan a predecir el posible comportamiento a futuro del este. Posteriormente, la entidad científica responsable crea mapas de zonificación de la peligrosidad del volcán, junto con programas de concientización sobre los peligros y pronósticos de erupciones. Finalmente, las entidades gubernamentales junto con las autoridades ejecutoras crean planes de aviso y contingencia sobre desastres naturales para salvaguardar la seguridad de la población. (Tilling & Beate, 1993)

2.1.2. Vulcanología en el Ecuador

Ecuador está ubicado en un margen continental activo ya que está en la subducción de la Placa Oceánica Nazca por debajo de la Placa Continental Sudamericana, mismas que forman un ángulo entre 25° y 30° que ha influenciado directamente el comportamiento de los volcanes a lo

largo del Ecuador. Junto con Colombia, el país forma parte de los Andes septentrionales (NVZ, del inglés *Northern Volcanic Zone*) presentando elevaciones que varían entre 3500 y 6000 msnm. Como se explica en (Vallejo Vargas, 2011), en el país se diferencian tres regiones fisiográficas: el Oriente o Llanura Amazónica al este de los Andes, la Sierra que comprende dos cordilleras paralelas, la Cordillera Occidental y Cordillera Real separadas por una depresión estructural denominada Valle Interandino y la región Costa que se refiere a las zonas de baja elevación al occidente de los Andes.

En Ecuador se han producido varias erupciones volcánicas que han alcanzado altos índices de magnitud y explosividad volcánica (VEI), siendo las más destructivas las generadas en el volcán Chalupas, la Isla de la Plata y el volcán Guagua Pichincha. Ecuador acoge en su territorio un total de 68 volcanes, de los cuales 33 son considerados activos (Corominas, 2011) y dentro de los cuales, el volcán Cotopaxi es uno de los seis volcanes que más afectación ha causado a lo largo de la historia del país.

El Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN) en Ecuador, es la organización encargada de mantener en vigilancia permanente las elevaciones que podrían representar amenazas sísmicas y volcánicas, con el objetivo de reducir el impacto negativo y promover una cultura de prevención en el país reduciendo la vulnerabilidad frente a los fenómenos sísmicos y volcánicos, llevando a cabo monitoreos geoquímicos, de deformación, cambios físicos y vigilancia volcánica para emitir boletines semanales y durante estados de emergencia, diarios.

2.1.3. Volcán Cotopaxi

Ubicado en las coordenadas 0°38' S, 78°26' W, el volcán Cotopaxi (ver Figura 1) se levanta hasta una altura de 5897m sobre el nivel del mar, está ubicado en la Cordillera Real de los Andes

Ecuatorianos en la provincia que lleva el mismo nombre. Con 20 km de diámetro basal, sobre su forma cónica y simétrica reposa un casquete glacial cuya pendiente está entre los 30° y 35°.



Figura 1. Volcán Cotopaxi visto desde el Norte.

Fuente: S. Hidalgo (IGEPN)

Este volcán es considerado de carácter bimodal debido a que ha generado dos tipos de erupciones: andesíticas y riolíticas, que, como se explicó en la Tabla 1, difieren en su estilo eruptivo y en los efectos adversos que generan, siendo la erupción riolítica la que genera una mayor distribución de los desechos volcánicos. Según (Hall, y otros, 2005), el primer periodo prehistórico eruptivo del volcán Cotopaxi data de hace ~500 ka, este antiguo complejo remanente volcánico yace en el flanco sur y sur occidente del cono actual, cuya explosión produjo principalmente magmas y lavas riolíticas y de obsidiana. Después de una época de reposo de 350 ka, el volcán tuvo una vigorosa reactivación produciendo al menos cinco erupciones riolíticas y frecuentes erupciones andesíticas que por su menor alcance contribuyeron a la formación del cono del volcán. Debido a la fundición de gran porción del glaciar se formaron lahares de tamaño gigantesco conocido como “Lahar del Valle de Los Chillos” que flujo por los ríos Pita, San Pedro y

Guayllabamba hasta llegar al océano Pacífico por el río Esmeraldas. En el periodo histórico, las erupciones del volcán azotaron a las culturas prehispánicas durante los últimos 500 años con un promedio de ocurrencia de una vez por siglo. La última gran erupción se dio en 1877, con presencia de flujos piroclásticos, pequeñas corrientes de lava, ceniza y lahares que provocaron destrucción a su paso por haber rebosado fácilmente los cauces naturales de los ríos provocando inundaciones de lodo y arrasando todo a su paso.

A continuación, se describen los tipos de fenómenos volcánicos ocurridos en el Cotopaxi durante sus erupciones, se los presenta desde el menos peligroso hasta el más devastador. (Hall, y otros, 2005)

- Sismos volcánicos: señales detectadas por una red de sismógrafos instalada en el volcán, aumentan representativamente en las semanas o meses precedentes a la erupción y se localizan cerca o bajo el cono del volcán.
- Gases volcánicos: estos gases de origen magmático se producen antes, durante y después de la erupción y llevan consigo partículas de dióxido de azufre, ácido clorhídrico, dióxido de carbono y monóxido de carbono que pueden ser dañinas para los seres vivos de las cercanías y en el peor panorama provocar lluvias ácidas.
- Flujos de lava: formados por roca fundida que se derrama desde el cráter y fluye por flancos y quebradas aledañas. Es común que avance a bajas velocidades y se estanque a los pocos kilómetros, cubriendo solamente el cono del volcán.
- Domos de lava: se forman cuando el magma es de gran viscosidad y no logra fluir, de manera que se estancan sobre el cráter formando una estructura inestable que se derrumba con facilidad.

- Flujos piroclásticos: mezcla de gases, ceniza y fragmentos de roca a altas temperaturas que se desplazan a manera de avalancha a grandes velocidades alcanzando los 200km/h. Su componente más denso viaja por el suelo y está formado por fragmentos de roca y bloques de escoria, mientras que su componente gaseosa se expande más y prolifera ceniza y gases calientes.
- Lluvia de ceniza y piroclásticos: las partículas más pequeñas producto de la erupción llegan a mayores alturas para ser dispersadas por el viento, llegando a distancias más lejanas del volcán, cubriendo con una capa de varios centímetros o decímetros de espesor y generando pérdidas en agricultura y ganadería.
- Flujos de lodo y escombros: conocidos comúnmente como lahares, son una mezcla de rocas, piedra pómez y arena que, junto con el agua proveniente del deshielo del glaciar por las altas temperaturas, se movilizan a velocidades de hasta 100km/h por los drenajes naturales como ríos y quebradas, llegando a desembocar en el Océano Pacífico.
- Avalanchas de escombros: el ascenso de gran cantidad de magma o un sismo de gran magnitud puede provocar el derrumbe de rocas por los flancos del volcán, mismos que también pueden ser inestables porque la pendiente del cono supera los 30° de inclinación, llevando así a su paso lo que encuentre.

Además, el IGEPN, mantiene un monitoreo por vigilancia instrumental en el volcán

Cotopaxi que se enfoca principalmente en los siguientes parámetros:

- Sismicidad
- Deformación
- Fumarolas y aguas termales

- Temperatura superficial
- Lahares
- Glaciar del Cotopaxi

Es de suma importancia tener una clara percepción de los riesgos asociados a una eventual erupción del volcán Cotopaxi. Según (Salazar & D'Ercole , 2009) los vulcanólogos consideran que son posibles varios tipos de crisis volcánicas, con fases largas de actividad previa a la erupción (al menos varias semanas), aun así el tiempo de las crisis volcánicas no puede ser previsto todavía, por lo que se elaboran varios estudios relativos a los microsismos como indicador de la cercanía de un proceso eruptivo.

2.2. Sismicidad volcánica

Los microsismos volcánicos se producen como parte del proceso de reactivación del volcán al entrar en un proceso eruptivo, se relacionan con el ascenso del magma que debe abrirse paso entre las formaciones rocosas provocando fracturas. Estos microsismos generan ondas sísmicas que viajan hacia la superficie del volcán llegando a los sensores.

Las ondas sísmicas están compuestas de varios segmentos de sub señales cuya ocurrencia sucede consecutivamente, comenzando con la onda P, seguida de la onda S, la onda L y finalmente la coda como se visualiza en la Figura 2.

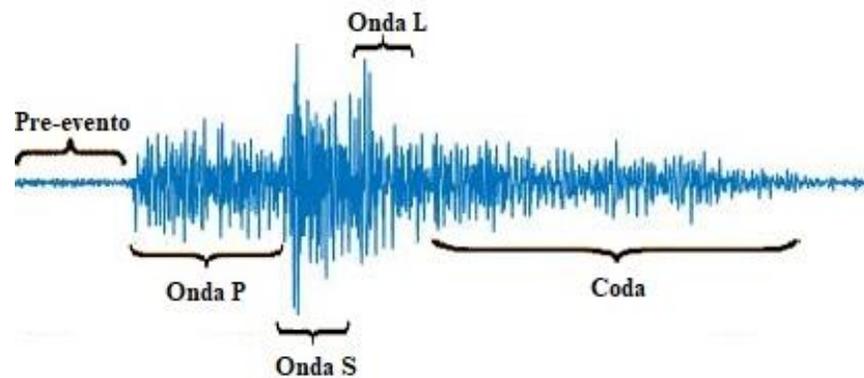


Figura 2. Partes de un evento sísmico

A continuación, se explican las características de cada tipo de ondas sísmicas. Las principales son las ondas de volumen: P y S. Las ondas primarias (P) son ondas longitudinales de compresión que se propagan a gran velocidad a través de cualquier material y oscilan en la dirección de propagación de la onda, se asimilan a las típicas ondas sonoras.

Las ondas secundarias (S) son transversales o de cizalla, se mueven en dirección perpendicular a la dirección de propagación de la perturbación (Figura 3), con menor velocidad de las ondas primarias y sin poder propagarse por masas líquidas (Jiménez Tintaya, 2007). Se subdividen en ondas sh cuyas partículas oscilan en el plano horizontal y las ondas sv que lo hacen en el plano vertical. En el caso de los sismos regionales, el tiempo de diferencia causado por la velocidad de propagación entre la onda P y la onda S es de suma importancia para localizar el foco donde se originó el temblor o terremoto.

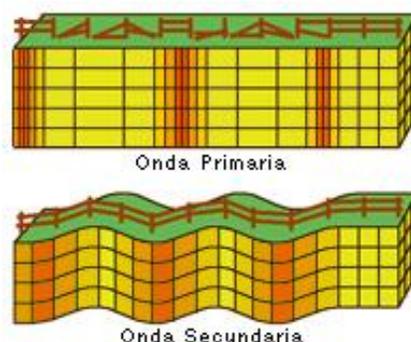


Figura 3. Propagación de las ondas S y P
Fuente: (EducarChile, 2012)

Existen también las llamadas ondas superficiales que se propagan por las capas superiores de la tierra y pierden abruptamente su amplitud dependiendo de la profundidad donde se originen, en su clasificación constan las ondas Rayleigh que causan un movimiento similar a las ondas del mar, con sus partículas moviéndose en forma elipsoidal en el plano vertical y ondas Love que oscilan solamente en el plano horizontal, su velocidad es un 90% de la velocidad de las ondas S y es ligeramente superior a la velocidad de las ondas Rayleigh. En la Figura 4 se ilustra la propagación de las ondas Rayleigh y Love.

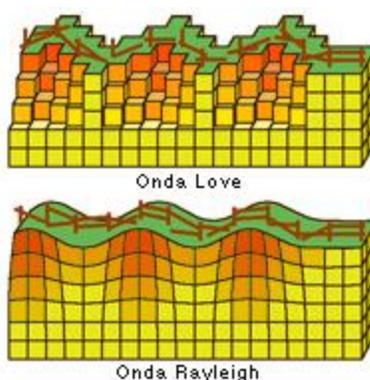


Figura 4. Propagación de las ondas L y R
Fuente: (EducarChile, 2012)

Finalmente, la Coda es considerada una onda de retrodispersión causada por las numerosas heterogeneidades distribuidas uniformemente sobre la corteza terrestre, representa la pérdida de energía del evento sísmico y finalmente su dispersión. (Aki & Chouet, 1975)

2.2.1. Monitoreo de la sismicidad

Existe una amplia variedad de instrumentos de medición que se utilizan para captar las señales sísmicas de un volcán, entre los más avanzados tecnológicamente están las estaciones sísmicas de banda ancha, las cámaras termográficas y los satélites. (Hall, y otros, 2005)

Según el Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN), Ecuador cuenta con una Red Nacional de Sismógrafos (RENSIG) desde 1970, habiendo sido instalados primero sensores de periodo corto de una componente y con el tiempo ser renovados por sensores de tres componentes de periodo corto. Además, se tiene colaboraciones con varios centros investigativos a nivel internacional como: Geological Survey of Canada (GSC), University of Hawaii (UH), University of Mississippi (UM), Incorporated Research Institutions for Seismology (IRIS); quienes han contribuido con la instalación de estaciones con sensores banda ancha, acelerógrafos, equipos de infrasonido y un GPS de alta resolución para el monitoreo de la deformación cortical, formando parte de una red mundial de monitoreo volcánico.

2.2.2. Tipos de Eventos Sismo Volcánicos

Conocer el tipo de evento sísmico que está ocurriendo en un volcán es de suma importancia para el descubrimiento de posibles interacciones entre estos ligeros movimientos telúricos volcánicos y futuros procesos volcánicos eruptivos (Alvarez, Henao, & Duque, 2007). A continuación se describen los principales tipos de señales sísmicas cuya ocurrencia es común en un

panorama de posible erupción volcánica y son sismos: Volcano Tectónicos (VT), de Largo Período (LP), Híbridos (HB) y Tremores Volcánicos (TRE).

a. Eventos Volcano Tectónicos

Los sismos Volcano Tectónicos (VT) o de alta frecuencia se deben a la formación o propagación de fracturas al interior del volcán (Hall, y otros, 2005) por lo que su duración es variable. Es común que el comienzo de la señal sea de manera abrupta, es decir se puede visualizar la onda P y a continuación la onda S.

La forma de onda que se observa en los sismos VT es característica ya que se puede diferenciar la llegada de las partículas de las ondas P y S seguida de las ondas superficiales L y R y finalizando con un decaimiento en forma exponencial que forma la Coda (Figura 5). En el dominio de la frecuencia se tiene un amplio espectro que en singulares ocasiones hasta podría superar los 10 Hz.

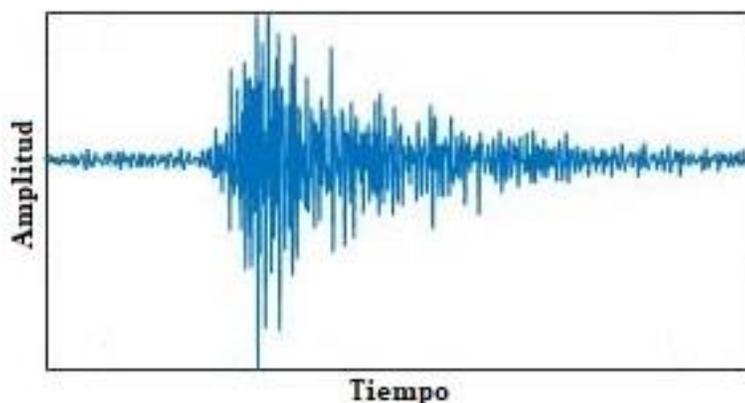


Figura 5. Ejemplo de evento tipo VT

b. Eventos de Largo Período

Los eventos de Largo Período (LP) se originan por movimientos de corta duración de fluidos como gases o magma que se producen en el interior del volcán. También son conocidos

como eventos de baja frecuencia debido a que generalmente mantienen sus componentes en un estrecho umbral de 0.5 Hz a 5 Hz. (Ibáñez & Carmona, 1997)

En cuanto a su forma de onda se debe señalar que los eventos se destacan por tener un comienzo emergente en el cual no se puede distinguir la onda P ni la onda S (Figura 6), su espectro se asemeja a un “huso de tejer” (Ibáñez & Carmona, 1997) por lo que es complicado determinar el momento exacto de su comienzo, en cuanto a su duración puede ir desde pocos segundos hasta más de un minuto.

La ocurrencia consecutiva de varios eventos de un mismo tipo es conocido como un enjambre sísmico, según (Chouet, 1996) los eventos de tipo LP suelen presentarse de esta manera cuando sucederá una erupción volcánica.

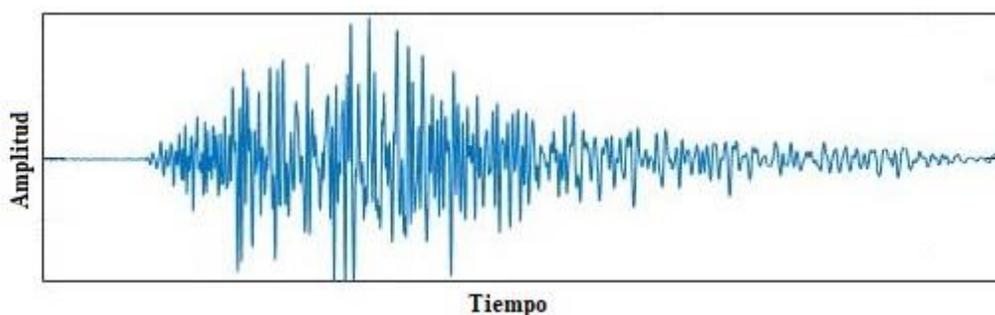


Figura 6. Ejemplo de evento tipo LP.

c. Eventos Híbridos

Los Eventos Híbridos (HYB) son causados por la formación de una fractura seguida de movimientos de fluidos al interior del volcán. En el dominio frecuencial, presenta altas frecuencias al inicio en una amplia banda superior a 10Hz y luego presenta bajas frecuencias típicas de un evento LP, por eso recibió su denominación ya que se entiende como la fusión de dos eventos, uno de tipo VT seguido de un LP.

En el dominio temporal se puede diferenciar claramente la llegada de las fases de las ondas P y S, por lo que daría la idea que comienza como un evento VT, pero a continuación se observa una señal similar a una de tipo LP en forma de onda, duración y contenido espectral. Según (Ibáñez & Carmona, 1997) la aparición de este tipo de eventos en las señales de un volcán está directamente asociada a episodios pre eruptivos inminentes.

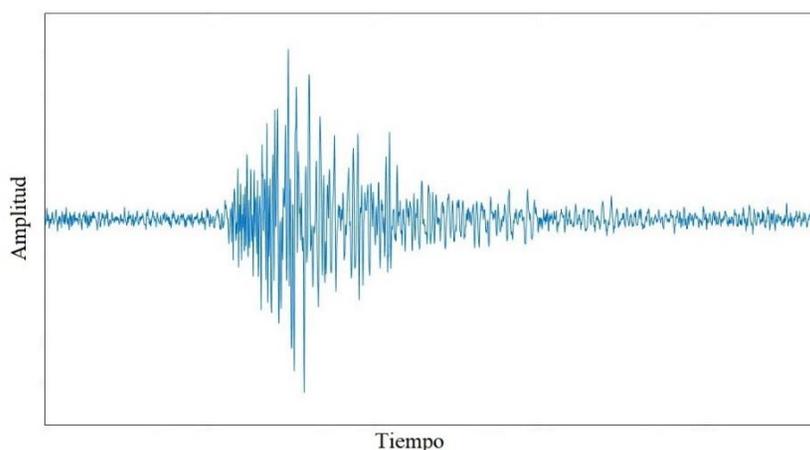


Figura 7. Ejemplo de evento tipo híbrido.

d. Tremores volcánicos

Los Tremores Volcánicos (TRE) son vibraciones volcánicas prolongadas causadas por la circulación de fluidos dentro del cono del volcán. La sustancia que está en movimiento puede ser: magma, gases o la combinación de agua, gases y cenizas. Su duración es variable ya que puede durar desde minutos y prolongarse incluso a días. (Mora & Alvarado , 2001)

Este tipo de señales sísmicas se caracterizan por mantener una duración constante durante un largo periodo de tiempo (Figura 8). Su contenido espectral se limita a una banda estrecha en la que se identifica un estable pico dominante con algunos picos subdominantes, pudiendo presentarse en altas frecuencias con valores mayores a 25 Hz. Además, no se localizan las fases de las ondas tipo P y S.

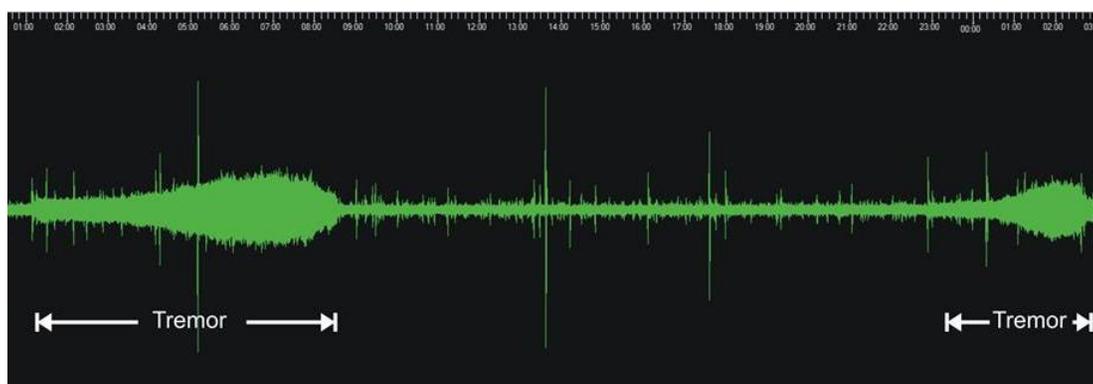


Figura 8. Tremor de la estación CVTR del volcán Turrialba.

Fuente: (Mora & Alvarado , 2001)

Es preciso haber puesto en conocimiento los conceptos relacionados con las señales sísmicas ya que el algoritmo A-VAD que se desarrolla en el presente trabajo va a detectar los ESV que se presenten en las señales provenientes de los sensores de monitoreo. A continuación, se tratan los conceptos básicos relacionados a las técnicas y herramientas que utiliza el A-VAD para cumplir con los objetivos de la presente investigación.

2.3. Conceptos Básicos

En este apartado se presentan los conceptos básicos necesarios para comprender la metodología que se seguirá en el desarrollo del presente trabajo. Exponiendo primero una idea general sobre los algoritmos de detección de actividad de voz y sus tipos, seguido de las técnicas utilizadas como la sustracción espectral de potencia, la creación de modelos Gaussianos mixtos utilizando los Coeficientes Cepstrales de la Frecuencia de Mel. Seguido también de los métodos de decisión empleados como el método de Agrupamiento de k-Medias y finalmente se explican también los parámetros de las medidas de rendimiento en base a los que se presentan los resultados de la investigación.

2.3.1. Detección de actividad de voz

La Detección de Actividad de Voz (VAD, del inglés *Voice Activity Detection*) es una importante técnica utilizada para detectar la presencia y ausencia de la voz humana en una señal, distinguiendo la voz de los sonidos de fondo. Es utilizada en un contexto en que las señales serán procesadas posteriormente para ser utilizadas en aplicaciones de comunicación móvil, automóviles, motores de búsqueda de internet, audífonos, etc.

Los detectores de actividad de voz pueden ser clasificados en tres tipos según (Ong & Tan, 2016):

- De aprendizaje supervisado
- De aprendizaje semi supervisado
- De aprendizaje no supervisado

En los detectores de aprendizaje supervisado, el sistema es entrenado con señales que fueron previamente etiquetadas de manera que se conoce las regiones de voz y de silencio. Mientras que en el aprendizaje semi supervisado, el sistema de VAD es entrenado con una serie de señales de ruido que se asemejan al ruido que va a ser encontrado en las aplicaciones donde será utilizado finalmente. Utilizando un aprendizaje semi supervisado, el sistema asume que algunas de las primeras tramas de la señal son tramas sin habla y las utiliza para estimar los parámetros de ruido y así realizar la discriminación entre habla y ruido en las tramas venideras.

Finalmente, el aprendizaje no supervisado de un sistema de VAD puede funcionar sin la necesidad de entrenar los datos, utilizándolos como se propone en (Ying, Yan, Dang, & Soong,

2011) para la construcción de un modelo estadístico de mezcla gaussiana. Muchos modelos estadísticos buscan construir clasificadores para discriminar la detección de voz o silencio.

A continuación, se muestra la fórmula utilizada por el detector basado en VAD que se adapta a la energía y fue desarrollado en (Kinnunen & Padmanabhan, 2013). Siendo $x[n]$ la señal a ser analizada, donde n denota la n -ésima muestra de voz en la señal. Se calcula la energía E para cada trama en la Ecuación (17):

$$E = 10 \log_{10} \left(\frac{1}{N-1} \sum_{n=1}^N (x[n] - \mu)^2 + \epsilon \right) \quad (1)$$

Donde, $\mu = (1/N) \sum_{n=1}^N x[n]$ es la Energía Media de la señal, N es la longitud de la señal y $\epsilon = 10^{-16}$ es una constante arbitraria para evitar el logaritmo de cero. El detector de voz propuesto en el presente trabajo investigativo utiliza herramientas como Sustracción Espectral de Potencia, Coeficientes Cepstrales de la Frecuencia de Mel, entre otros, que serán explicados a continuación.

2.3.2. Sustracción espectral de potencia

La Sustracción Espectral es uno de los métodos con mejores resultados en procesamiento digital de señales, ya que solo requiere de una señal y puede lidiar con ruido no estacionario. Se busca que tras el proceso la señal tenga una mejora en la Relación Señal a Ruido (SNR, del inglés *Signal-to-Noise Ratio*), que guarde naturalidad con la señal original y que el proceso no tenga un alto costo computacional.

Como se explica en (Borrás, 2006), para esta técnica se asume que los segmentos de voz no están correlacionados con los de ruido y son aditivos en el dominio del tiempo, para lo cual el espectro en potencia de la señal ruidosa es la suma de los espectros en potencia de la voz y el ruido.

Se asume también que el ruido es estacionario porque varía lentamente con relación a la voz, por lo que tiene varianza cero. Se define la señal ruidosa en el dominio del tiempo como muestra la Ecuación (2):

$$y[n] = x[n] + r[n] \quad (2)$$

Donde $y[n]$ representa la señal ruidosa, $x[n]$ es la señal original sin ruido y finalmente, $r[n]$ es el ruido. Expresando la ecuación (2) en el dominio de la frecuencia f se tiene:

$$Y(f) = X(f) + R(f) \quad (3)$$

Se divide la señal de entrada en ventanas, obteniendo segmentos de muestras de N' puntos a los que se denomina tramas. Posteriormente se aplica una función de ventana Hanning o Hamming para transformar al dominio de la frecuencia utilizando la transformada de Fourier, como resultado se obtienen nuevas tramas con N' muestras espectrales.

De una manera general, el proceso de resta espectral se realiza como se muestra en la Ecuación (4):

$$|\hat{X}(f)|^b = |Y(f)|^b - g|\overline{R(f)}|^b \quad (4)$$

Siendo $|X(f)|^b$ la señal original y $|\hat{X}(f)|^b$ la estimación de su espectro, $|\overline{R(f)}|^b$ el espectro del ruido promedio en función del tiempo, $|Y(f)|^b$ la señal ruidosa en el dominio frecuencial, g es la variable que controla la cantidad de ruido que se va a sustraer, tal que al tomar el valor de 1 se realiza una resta completa y para una resta de menor índice se utilizan valores mayores a 1.

La variable b representa dos opciones de resta espectral que son:

- Resta espectral de magnitud: donde $b = 1$

- Resta espectral de potencia: donde $b = 2$

Se calcula la estimación promediada en el tiempo correspondiente al espectro del ruido $|\overline{R(f)}|^b$, en las tramas que no contienen voz, como indica la ecuación (5):

$$|\overline{R(f)}|^b = \frac{1}{N''} \sum_{i=0}^{N''-1} |R_i(f)|^b \quad (5)$$

Donde, el espectro promedio del ruido de la correspondiente trama $R_i(f)$ será $|\overline{R(f)}|^b$ tomando a N'' como el número de tramas en el periodo analizado.

El espectro promedio del ruido $|\overline{R_i(f)}|^b$ pasa también por un filtro paso bajo digital de primer orden, de la forma:

$$|\overline{R_i(f)}|^b = \rho |\overline{R_{i-1}(f)}|^b + (1 - \rho) |R_i(f)|^b \quad (6)$$

Donde la constante ρ típicamente toma valores entre 0.85 y 0.99 para filtros paso bajo. $|\overline{R_i(f)}|^b$ es la trama actual de la señal de ruido y $|\overline{R_{i-1}(f)}|^b$ es la trama anterior a la actual. Finalmente se requiere restaurar la señal al dominio del tiempo, para lo que se combina la estimación del espectro de magnitud $|\hat{X}(f)|$ con la fase de la señal con ruido, utilizando la transformada inversa de Fourier, obteniendo finalmente la Ecuación (7):

$$\hat{x}[n] = \sum_{h=0}^{N-1} |\hat{X}(h)| e^{j\theta_y(h)} - e^{-j\frac{2\pi}{N}hn} \quad (7)$$

Donde $\theta_y(h)$ es la fase de la señal con ruido $Y(f)$ calculada en la Ecuación (3).

Obtener resultados óptimos al aplicar la sustracción espectral depende de la precisión del algoritmo al reducir las variaciones del ruido y eliminando las distorsiones, ya que se busca que el

ruido que no pudo ser eliminado este en una banda estrecha de frecuencias aisladas que no afecte al espectro de voz.

2.3.3. Frecuencia de Mel

La escala de Mel se basa en un mapeo entre la frecuencia real y el tono percibido, ya que aparentemente el sistema auditivo humano no percibe el tono de manera lineal. el mapeo es aproximadamente lineal por debajo de 1KHz y logarítmico por encima.

El éxito de los Coeficientes Cepstrales de la Frecuencia del Mel (MFCCs, del inglés *Mel Frequency Cepstral Coefficients*) se debe a su capacidad para representar el espectro de amplitud de voz en una forma compacta. Cada paso en el proceso de creación de las funciones MFCCs está motivado por consideraciones perceptivas o computacionales. La Figura 9 muestra el proceso de creación de estas. El primer paso es dividir la señal de voz en tramas, generalmente aplicando una función de ventana a intervalos fijos. El objetivo aquí es modelar pequeñas secciones (típicamente 20 ms) de la señal que son estadísticamente estacionarias. La función de ventana, típicamente una ventana de Hamming, elimina los efectos de borde, para luego generar un vector de características cepstrales para cada trama.

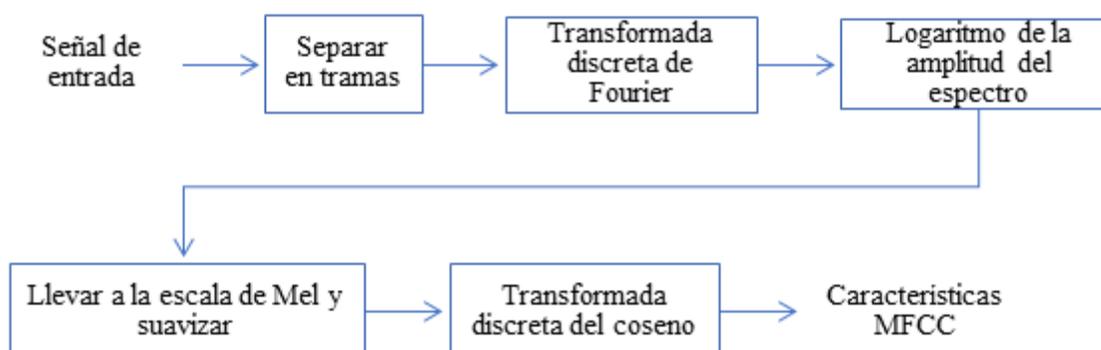


Figura 9. Proceso para obtener las características de los MFCCs

El siguiente paso es tomar la transformada discreta de Fourier de cada trama, entonces se conserva solo el logaritmo del espectro de amplitud, se descarta la información de la fase porque los estudios perceptivos han demostrado que la amplitud del espectro es mucho más importante que la fase (Patel & Patil, 2015). Se toma el logaritmo del espectro de amplitud porque el volumen percibido de una señal es aproximadamente logarítmico. El siguiente paso es suavizar el espectro y enfatizar las frecuencias perceptivamente significativas. Esto se logra mediante la recopilación de los 256 componentes espectrales en 40 intervalos de frecuencia como se representa en la Figura 10. Se ha encontrado que, para el habla, las frecuencias más bajas son perceptualmente más importantes que las frecuencias más altas. por lo tanto, el espacio entre las frecuencias sigue la llamada escala de frecuencia de Mel.

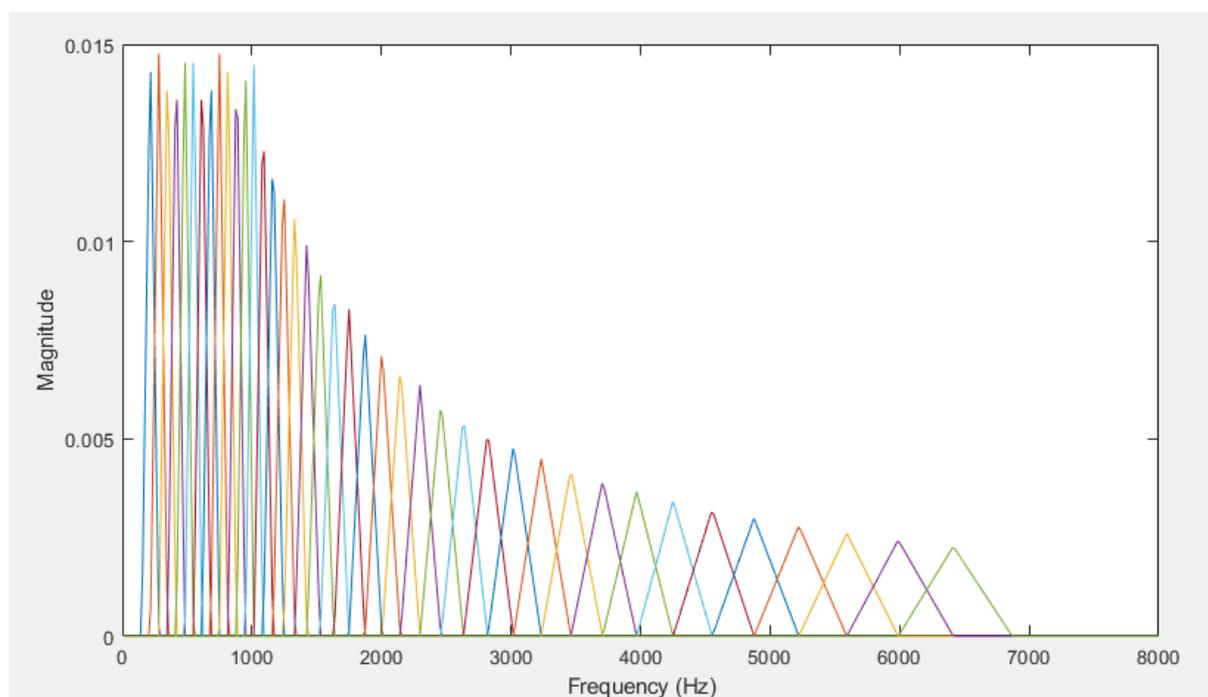


Figura 10. Banco de filtros de la frecuencia de Mel

Fuente: (MathWorks, 2018)

Los componentes de los vectores espectrales de Mel calculados para cada trama de la señal están altamente correlacionados. Las características del habla suelen ser modeladas por mezclas de densidades gaussianas. Por lo tanto, para reducir el número de parámetros en el sistema, el último paso de la construcción de la función MFCC es aplicar una transformación a los vectores espectrales de Mel que descorrelacionen sus componentes. Teóricamente, la transformada Karhunen-Loeve¹ logra esto, ya que se aproxima con la transformada discreta del coseno. Usando esta transformación, se obtienen 13 características cepstrales para cada cuadro o trama (Rubin, y otros, 2016).

2.3.4. Prueba de razón de máxima verosimilitud

La prueba de Razón de Probabilidad Estadística es una herramienta ampliamente utilizada en los métodos de VAD, en el que la razón de probabilidad de la trama actual se compara con un umbral. Este umbral representa tanto el límite superior de la razón de probabilidad para las tramas que no son de voz, como el umbral inferior que representa la razón de probabilidad para las tramas de voz. Como resultado, se logra una alta tasa de aciertos en las tramas sin voz y se mantiene una alta tasa de reconocimiento de tramas de voz (Li, Horaud, & Girin, 2016).

Se considera una señal que contiene voz y ruido mezclados en el dominio de la transformada de Fourier de tiempo reducido (STFT, del inglés *Short-time Fourier Transform*). Se denotan como los vectores de coeficientes de la señal con ruido $Y = [Y_1, \dots, Y_W]$, la señal de voz $X = [X_1, \dots, X_W]$ y la señal de ruido $R = [R_1, \dots, R_W]$, donde W representa el intervalo de frecuencias. Para cada trama el algoritmo de VAD toma una decisión entre dos hipótesis:

¹ Según (Altuve & Warrick, 2015), la descomposición de Karhunen-Loeve (KL) es utilizada para reducir la dimensionalidad de los datos y capturar las variaciones más importantes en los primeros componentes de la descomposición.

$$H_0: Y = R \quad \text{ausencia de habla}$$

$$H_1: Y = X + R \quad \text{presencia de habla}$$

Dado que $\lambda_R = E\{|R|^2\}$ y $\lambda_X = E\{|X|^2\}$ denotan la densidad espectral de potencia (PSD, del inglés *Power Spectral Density*) de las señales de ruido y de voz respectivamente. La función de densidad de probabilidad es medida mediante el espectrograma de potencia $|X|^2$ que sigue una distribución exponencial con media λ_R y $\lambda_R + \lambda_X$ para cada hipótesis respectivamente. La razón de verosimilitud Λ para el intervalo de frecuencias W se muestra en la Ecuación (8):

$$\Lambda \triangleq \log \left\{ \frac{p(|X_W|^2|H_1)}{p(|X_W|^2|H_0)} \right\} = \frac{\gamma_W \xi_W}{1 + \xi_W} - \log\{1 + \xi_W\} \quad (8)$$

Donde $\gamma_W \triangleq |X_W|^2/\lambda_R$ y $\xi_W \triangleq \lambda_X/\lambda_R$ son el SNR a posteriori y a priori respectivamente. Finalmente se asume que la PSD del ruido λ_R puede ser estimada utilizando el ruido existente.

2.3.5. Modelos de entrenamiento de mezcla Gaussiana

Para describir el modelo de mezcla Gaussiana como se lo realizó en (Jin, Zhang, Balakrishnan, Wainwright, & Jordan, 2016) , es preciso señalar que se utiliza M para denotar el conjunto $\{1, 2, \dots, M\}$ y $F(\mu, \Sigma)$ para referirse a la distribución Gaussiana d -dimensional con vector de la media μ y matriz de covarianza Σ . Se utiliza $\phi(\cdot | \mu, \Sigma)$ para denotar la función de densidad de probabilidad de la distribución Gaussiana, de la siguiente manera:

$$\phi \left(x | \mu, \Sigma \right) := \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \quad (9)$$

Un modelo de mezcla Gaussiana (GMM, del inglés *Gaussian Mixture Model*), con d dimensiones y M componentes puede ser especificado por una colección de $\mu^* = \{\mu_1^*, \dots, \mu_M^*\}$ vectores de la media d -dimensionales, un vector $\lambda^* = (\lambda_1^*, \dots, \lambda_M^*)$ de los pesos de la mezcla positivos que suman uno, y una colección $\Sigma^* = \{\Sigma_1^*, \dots, \Sigma_M^*\}$ de matrices de covarianza. Dados

estos parámetros, la función de densidad de un modelo de mezcla Gaussiana $p(x|\lambda^*, \mu^*, \Sigma^*)$ toma la forma de la Ecuación (10):

$$p(x|\lambda^*, \mu^*, \Sigma^*) = \sum_{i=1}^M \lambda_i^* \phi \left(x | \mu_i^*, \Sigma_i^* \right) \quad (10)$$

Donde la función de densidad Gaussiana ϕ fue previamente definida en la Ecuación (9). Se asume el caso en que todos quienes componen la mezcla tienen pesos iguales, y la covarianza de cada componente de la mezcla es idéntico. Esto lleva a un modelo mixto que se denomina GMM(μ^*) de la forma:

$$p(x|\mu^*) := \frac{1}{M} \sum_{i=1}^M \phi \left(x | \mu_i^*, I \right) \quad (11)$$

En este caso, los únicos parámetros a ser estimados son los vectores de la media $\mu^* = \{\mu_i^*\}_{i=1}^M$ de los M componentes. La dificultad de estimar una distribución de mezcla Gaussiana depende de la cantidad de separación existente entre los vectores de la media. Es decir, para un parámetro dado $\xi > 0$, se dice que el modelo GMM(μ^*) está ξ -separado si:

$$\left\| \mu_i^* - \mu_j^* \right\|^2 \geq \xi \text{ para todos los pares } i, j \in [M]$$

Suponiendo que se tiene una secuencia de variables aleatorias independientes e idénticamente distribuidas (i.i.d) $\{x_l\}_{l=1}^n$ que siguen una distribución GMM(μ^*), si se busca estimar el grupo de vectores de medias desconocido (μ^*), la función de máxima verosimilitud logarítmica $\zeta_n(\mu)$ viene dada por la Ecuación (12) :

$$\zeta_n(\mu) := \frac{1}{n} \sum_{l=1}^n \log \left(\frac{1}{M} \sum_{i=1}^M \phi(x_l | \mu_i, I) \right) \quad (12)$$

Dado que el tamaño de la muestra n tiende al infinito, la probabilidad de la muestra converge con la probabilidad de verosimilitud de la función total ζ dada por la Ecuación (13):

$$\zeta(\mu) := \mathbb{E}_{\mu^*} * \log \left(\frac{1}{M} \sum_{i=1}^M \phi(X|\mu_i, I) \right) \quad (13)$$

Donde \mathbb{E}_{μ^*} denota la esperanza del vector aleatorio X modelado acorde el modelo GMM(μ^*).

2.3.6. Método de agrupamiento de k medias

El agrupamiento por clústeres facilita trabajar con grandes volúmenes de información. Este proceso consiste en agrupar un determinado conjunto de datos en una cantidad predeterminada de conjuntos disjuntos, llamados grupos. De tal manera que la similitud de los elementos dentro del clúster sea alta y la similitud entre grupo sea baja.

En base al estudio descrito en (Capó, Pérez, & Lozano, 2017), el método de agrupamiento de k-medias se puede explicar de la siguiente manera. Dado un conjunto de n datos (instancias) $D = \{x_1, \dots, x_n\}$ y un entero K , el objetivo de K-medias es determinar un conjunto de K centroides $C = \{c_1, \dots, c_K\}$ de manera que se minimice la siguiente función de error $E(C)$:

$$E(C) = \sum_{x \in D} \sum_{k=1, \dots, K} \|x - c_k\|^2 \quad (14)$$

Este es un problema de optimización combinatoria ya que es equivalente a encontrar una partición de las n instancias en K grupos, cuyos conjuntos asociados de centro de masa minimiza

la función de error. En ese caso, el número de posibles particiones es un número de Stirling del segundo tipo².

El algoritmo de k-medias tiene dos etapas: la inicialización, en la cual se establece el conjunto inicial de centroides y la etapa iterativa llamada algoritmo de Lloyd. Este algoritmo consta de dos pasos: un primer paso en el que cada instancia se asigna a su centroide más cercano conocido como paso de asignación y luego se actualiza el conjunto de centroides (paso de actualización). Finalmente, se verifica un criterio de detención, siendo el más común aquel que implica el cálculo de la función de error presentada anteriormente. Si el error no disminuye significativamente con respecto a la iteración anterior, el algoritmo se detiene. Siendo d las dimensiones en la que está definido el conjunto, el tiempo requerido para el paso de asignación depende de n, K, d mientras que el paso de actualización para el conjunto de centroides y el proceso del criterio de detención dependen de n, d . Siendo el paso de asignación el más extenuante computacionalmente hablando (Capó, Pérez, & Lozano, 2017).

² Un número de Stirling de segunda clase representa el número de particiones distintas de subconjuntos que pueden definirse en un conjunto de n elementos.

CAPÍTULO III

3. METODOLOGÍA

Para el desarrollo del presente trabajo de investigación, la metodología utilizada consta de tres etapas como se muestra en la Figura 11. A continuación, se presenta una breve descripción de las etapas que conforman el proceso metodológico utilizado y que más adelante en este capítulo serán detalladas a profundidad.

En la primera etapa se lleva a cabo el preprocesamiento de las señales sísmicas, eliminando los picos de frecuencia que no son de interés para el presente estudio y eliminando también la tendencia lineal de la señal. Como segunda etapa se tiene el algoritmo adaptativo de VAD (A-VAD) que consta de tres subetapas: la extracción de los Coeficientes Cepstrales de la Frecuencia de Mel, la Sustracción Espectral y el Modelo de Entrenamiento que utiliza una mezcla de modelos Gaussianos. Finalmente, en la denominada etapa de decisión se eliminan los eventos detectados que son considerados falsos positivos, basándose en la duración de estos, y se preparan los datos para dar a conocer los puntos de inicio y fin del ESV tanto en muestras como en tiempo. El proceso antes mencionado se realiza para cada una de las señales analizadas, mientras que el proceso de obtención de las Medidas de rendimiento para las bases de datos será explicado más adelante.

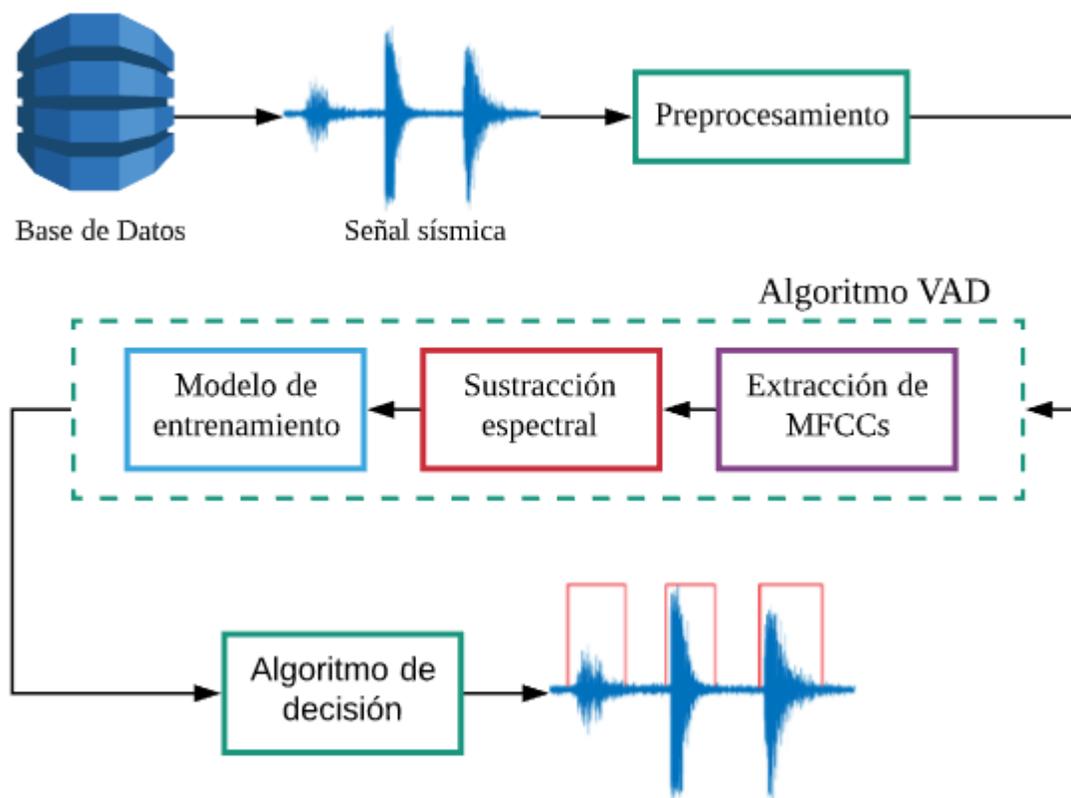


Figura 11. Diagrama de bloques de la metodología utilizada en la investigación.

3.1. Bases de Datos

El IGEPN cuenta con seis estaciones sismológicas de corto periodo (SP, del inglés *Short Period*) cuya respuesta en frecuencia está en el rango de 1 a 50 Hz, y seis estaciones de banda ancha (BB, del inglés *BroadBand*) con respuesta en frecuencia en el rango de 0.1 a 50 Hz (IGEPN, Cotopaxi, 2017). Cada sismómetro recolecta información en tres componentes espaciales, siendo la más utilizada la componente vertical debido a su alto SNR. Las estaciones, cuyas señales van a ser utilizadas son las denominadas BREF, BVC2 y VC1. Estas estaciones fueron digitalizadas con una frecuencia de muestreo de 100 Hz en el caso de VC1 y 50Hz para BREF y BVC2. Posteriormente la información es almacenada en el formato denominado *The Standard for the*

Exchange of Earthquake Data (SEED), guardando información relevante como: nombre de la estación, componente espacial, red de sismómetros a la que pertenece, tiempo de inicio de grabación, número de muestras, frecuencia de muestreo, tamaño de la grabación, y datos de la señal. Posteriormente, el IGEPN detecta visualmente los ESV, define el punto de inicio y punto de fin del evento y genera las etiquetas correspondientes a la fecha de detección, tipo de evento, amplitud, periodo y duración del evento.

Las Bases de Datos a utilizarse son extraídas de información brindada por el IGEPN en formato SEED, para luego ser transformadas a la extensión de archivo .mat que permite su utilización en Software Matlab R2018a ®. En el capítulo de Experimentación y Resultados se explicará con mayor profundidad las Bases de Datos utilizadas.

3.2. Preprocesamiento

El objetivo de la etapa de preprocesamiento es remover el valor medio de la tendencia lineal de la señal sísmica de entrada, para lo cual se utiliza la técnica de Ajuste por Mínimos Cuadrados que determina los valores de los parámetros a y b de la recta r (Ecuación (17)) que mejor se ajuste a los valores de la señal sísmica. Los parámetros a y b se obtienen mediante las fórmulas (18) y (19) donde n es el número de puntos que van a representar la recta r .

$$r = ah + b \quad (15)$$

$$a = \frac{n(\sum h_i r_i) - (\sum h_i)(\sum r_i)}{n(\sum h_i)^2 - (\sum h_i)^2} \quad (16)$$

$$b = \frac{(\sum r_i) - a(\sum h_i)}{n} \quad (17)$$

Para su implementación se utilizó la función `detrend` del Software Matlab R2018a ® como se observa en el Segmento de código 1.

Segmento de código 1. Etapa de preprocesamiento

```

% Etapa de Preprocesamiento
Fmin=2;
Fmax=Fs-2;
Signal=detrend(Si); % Remueve la tendencia lineal
B = fir1(128,[Fmin/Fs Fmax/Fs]); % Filtro FIR pasabanda de 1 Hz a 50 Hz.
[H,f] = freqz(B,1,1024,Fs); % Respuesta en frecuencia del filtro digital
Sp = filtfilt(B,1,Signal); % Preprocesado de la señal

```

Luego la señal es filtrada aplicando un filtro FIR pasabanda de orden 128 que permite el paso de frecuencias entre 1 y 50 Hz. El motivo por el cual el valor de frecuencia mínimo utilizado es de 1 Hz, es la existencia de un gran pico frecuencial en el rango de 0.1 Hz a 0.3 Hz que según (Gubbins, 2004) es producido por movimientos de las placas oceánicas y, al no estar relacionado con la presente investigación se ha decidido eliminarlo. Se utilizó la función `fir1` del Software Matlab R2018a® como se indica en el Segmento de código 1 para la implementación del filtro indicado, mientras que se utiliza la función `filtfilt` para el filtrado digital de fase cero que procesa los datos de entrada hacia delante, luego invierte la secuencia filtrada y lo realiza hacia atrás obteniendo una distorsión de fase cero y haciendo que el orden del filtro `fir1` se duplique a 256. En la Figura 12 se muestra una señal sísmica original (a) y después de haber sido preprocesada (b).

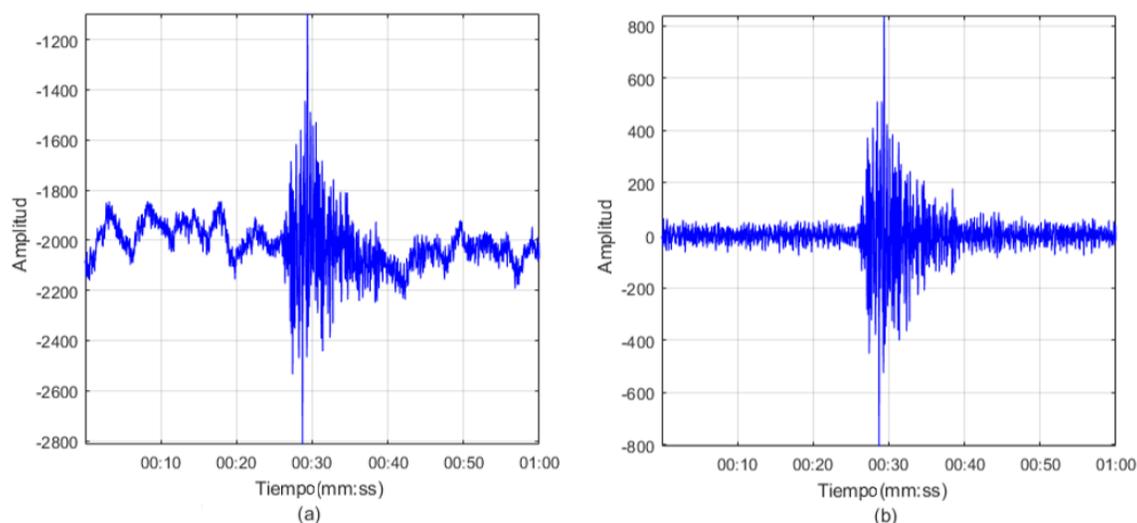


Figura 12. Ejemplo de una señal microsísmica antes y después de ser preprocesada.
(a) Señal original. (b) Señal preprocesada.

Fuente: Modificado de (Rodríguez & Lara-Cueva, 2018)

3.3. Algoritmo A-VAD

El algoritmo A-VAD desarrollado tiene el objetivo de identificar los ESV que se presenten en las señales sísmicas del volcán Cotopaxi. Por lo que su principal función es determinar el punto de inicio y de finalización del evento. El algoritmo está conformado como tal, por varios bloques de funciones que ejecutan operaciones específicas, como la extracción de los Coeficientes Cepstrales de la Frecuencia de Mel, el proceso de Sustracción Espectral y el Modelo de Entrenamiento utilizado. En el proceso se utilizan un total de 13 parámetros de control que fueron definidos específicamente para obtener el mejor desempeño posible al tratar con señales sísmicas de las bases de datos del IGEPN, mismos que se irán describiendo en cada etapa y se muestran recopilados al finalizar la sección en la Tabla 3.

3.3.1. Extracción de MFCCs

El proceso de extracción de los Coeficientes Cepstrales de la Frecuencia de Mel se observa en la Figura 13. La señal sísmica de entrada S_i es aquella que ya fue preprocesada en pasos

anteriores, F_s representa la frecuencia de muestreo, $Num_Filtros$ representa al parámetro `params.num_filters` que es el número de bandas de frecuencia del Filtro de Mel, F_{min} y F_{max} son frecuencias mínima y máxima de la señal respectivamente y Tam_FFT representa al parámetro `param.NFFT` que indica el tamaño de la Transformada Rápida de Fourier.

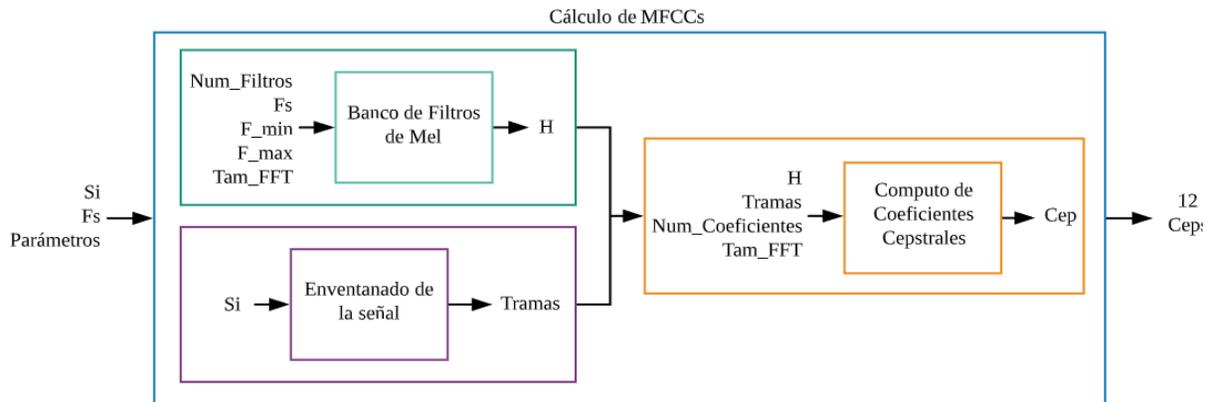


Figura 13. Diagrama de bloques del proceso de extracción de MFCCs.

El primer bloque que se ejecuta se encarga de la creación del Banco de Filtros de Mel que cuenta con 27 filtros triangulares que permiten obtener información de una banda de frecuencia. Estas bandas se basan en la percepción del oído humano ya que las frecuencias bajas son aquellas que se perciben con mayor peso, perdiendo así relevancia las componentes de altas frecuencias. Es por eso posible su adaptabilidad a las señales sísmicas ya que se encuentran principalmente en bajas frecuencias. La Figura 14 muestra las 26 bandas de frecuencias utilizadas en la presente investigación. Cada filtro se utilizará para concentrar las energías en la banda en un solo valor, de manera que, tras aplicar el filtro en cada ventana, se obtendrán 26 valores. La salida de este bloque es la función de transferencia H que será utilizada para el computo de los MFCCs en el último bloque.

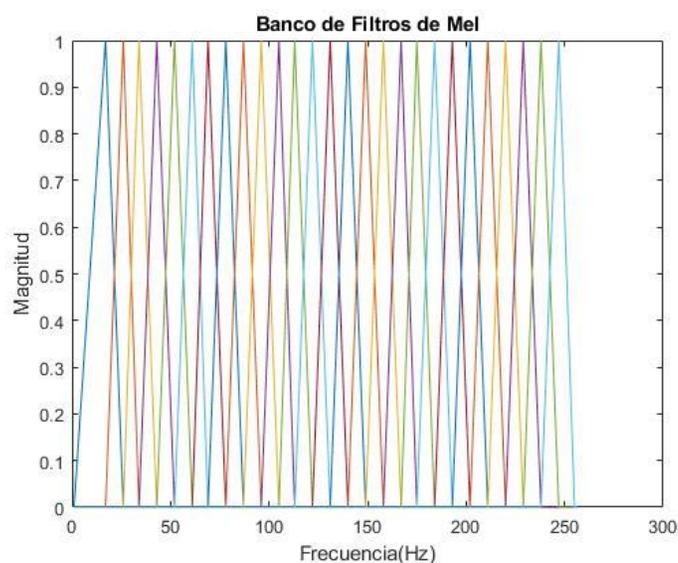


Figura 14. Banco de Filtros de Mel utilizado para señales sísmicas

En el segundo bloque de la Figura 13 se lleva a cabo el inventariado de la señal que define una ventana cada dos segundos, con un segundo de solapamiento. Separar la señal en ventanas es necesario ya que no se habla de una señal periódica por lo que no es posible calcular las componentes en frecuencia para toda la señal. Entonces al separar en ventanas se puede asumir que cada una de ellas representa una función periódica y se pueden extraer sus componentes frecuenciales. De manera que se obtendrá una secuencia de componentes que evolucionan en el tiempo (Meza, 2013). Cada ventana pasa por una ventana Hamming³ de longitud igual a la de cada trama de señal (100 o 200 muestras dependiendo de la f_s) que ayuda a suavizar la señal, filtrando frecuencias espurias que se generan por el inventariado. La salida del segundo bloque de funciones son las tramas que fueron inventariadas y serán utilizadas como parámetro de entrada en el tercer bloque.

³ La ventana de Hamming es una función matemática utilizada en el procesamiento de señales que evita las discontinuidades al inicio y al final de las tramas de señal cuando el procesamiento sobrepasa las capacidades del equipo de cómputo. Es también conocida como la ventana de coseno elevado debido a la forma de su función.

El tercer bloque de la Figura 13 realiza el computo de los Coeficientes Cepstrales de la frecuencia de Mel, utilizando las tramas de la señal (Frames), NumCoeffs que representa al parámetro `params.num_cep` que es el número de Coeficientes Cepstrales de la Frecuencia de Mel. La Transformada Discreta de Fourier de tamaño NFFT (`param.NFFT`) permite calcular los componentes en el dominio de la frecuencia, pero al necesitar la energía de estas, se eleva al cuadrado la información para extraer la magnitud de la energía como se muestra en el Segmento de código 2. Finalmente se utiliza la Transformada del Coseno Discreto (DCT, del inglés *Discrete Cosine Transform*) para encontrar los valores MFCCs finales Cep que son la salida de este bloque de funciones.

Segmento de código 2. Cómputo de MFCCs

```
%%Computo de MFCCs
function Cep = ComputeFFTCepstrum(Frames, NumCoeffs, NFFT)
ESpec      = abs(fft(Frames',NFFT)).^2;
LogSpec    = log(ESpec(1:NFFT/2+1, :)) + 1e-11;
Cep        = dct(LogSpec);
```

La extracción de los Coeficientes Cepstrales de la Frecuencia de Mel es un proceso necesario para el Modelo de Entrenamiento que se realizará posteriormente.

3.3.2. Sustracción espectral

El algoritmo de Sustracción Espectral busca mejorar la señal sísmica de entrada. Consta como una herramienta o *toolbox* en Matlab conocido como “Voicebox1⁴” que debe ser añadido al

⁴ Voicebox es una caja de herramientas de procesamiento de voz que consta de rutinas desarrolladas en Matlab que son escritas y actualizadas por Mike Brookes del Departamento de Ingeniería Eléctrica y Electrónica del Imperial College en Reino Unido. Las rutinas están disponibles como un repositorio de GitHub bajo los términos de la Licencia Pública GNU en el sitio web: <https://github.com/ImperialCollegeLondon/sap-voicebox>

directorío para su funcionamiento. Trabaja en base a parámetros definidos que serán explicados en esta sección.

En presencia de una baja SNR en la señal, los algoritmos de VAD tradicionales tienden a marcar la mayoría de las tramas como poseedoras de eventos. Para lo cual es necesario utilizar una estrategia que permita aumentar la SNR. Como se mostró en la Ecuación (4), g es la variable que controla la cantidad de ruido que se va a sustraer, por lo que es conocido como factor de sustracción y su fórmula se muestra en la Ecuación (18):

$$g = \max \left\{ \left(1 - \left(\alpha \frac{|R|^2}{|Y|^2} \right)^{\frac{\gamma}{2}} \right)^{\frac{e}{\gamma}}, \min \left(g_h, \left(\beta \frac{|R|^2}{|Y|^2} \right)^{\frac{e}{2}} \right) \right\} \quad (18)$$

Donde α es el factor de sobre sustracción, γ determina el dominio de sustracción, e es el exponente de ganancia, g_h es la ganancia máxima para el ruido de fondo y β determina la atenuación de ruido máxima en el dominio de la potencia. Como se sugiere en (Kinnunen & Padmanabhan, 2013) se fija el valor de $g_h = 1$ y $\beta = 0.01$, para ahora enfocarse en: a) el dominio de sustracción, b) la cantidad de sobre extracción y c) el estimador de ruido. Con respecto al dominio de sustracción, la sustracción del dominio de magnitud se obtiene al elegir $(\gamma, e) = (1,1)$, la sustracción espectral del dominio de energía por $(\gamma, e) = (2,1)$ y el filtro de Wiener⁵ por $(\gamma, e) = (2,2)$. Para esta investigación se define los parámetros para aplicar una sustracción espectral utilizando el filtro de Wiener como se indica en el Segmento de código 3 donde las variables `pp.g` y `pp.e` toman el valor de 2. Con respecto a la cantidad de sustracción, el parámetro `pp.am = 10` permite llevar a cabo una sobre sustracción máxima sugerida en (Gerkman & Hendriks, 2012) para

⁵ El filtro de Wiener es de tipo lineal y busca determinar la respuesta al impulso $h(n)$ de forma que el error sea lo más pequeño posible utilizando la minimización del valor cuadrático medio del error.

señales con baja SNR. En cuanto al estimador de ruido α toma el valor de 1 para usar el estimador de error cuadrático medio mínimo (MMSE) definido por la variable `pp.ne`.

Segmento de código 3. Etapa de sustracción espectral

```
% Limpieza de la señal para mejorar el nivel de energía
if params.clean_energy
    % Se definen los parámetros del filtro de Wiener
    pp.g      = 2;
    pp.e      = 2;
    pp.ne     = 1;
    pp.am     = 10;
    s_limpia = specsus(Sp, Fs, pp);
```

Finalmente, en el Segmento de código 3 se muestra la función `specsus` cuyas entradas son: a) la señal sísmica original, b) la frecuencia de muestreo f_s y c) los parámetros de los que se trató anteriormente, generando como salida una señal que mantiene solo los principales picos de frecuencia (ver Figura 15).

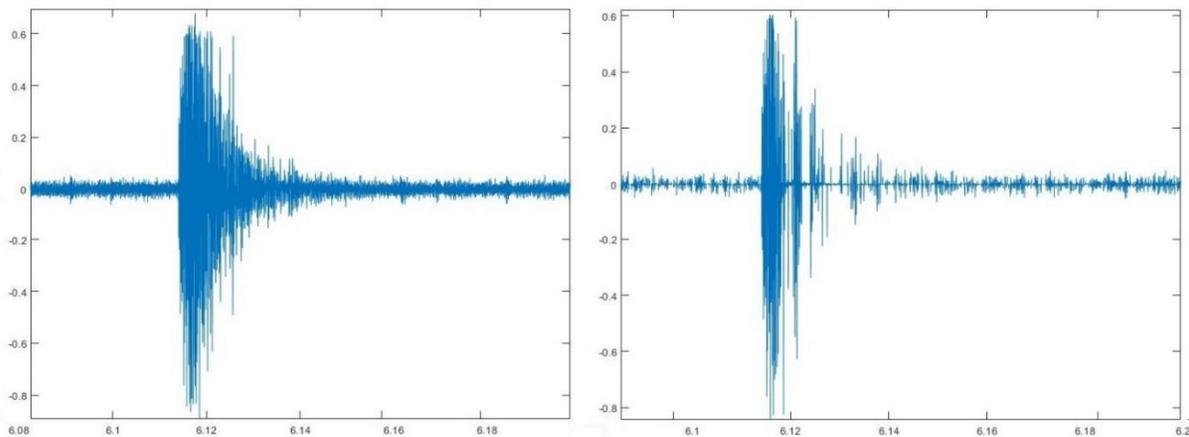


Figura 15. Ejemplo de señal sísmica que pasó por el proceso de Sustracción Espectral.

3.3.3. Modelo de entrenamiento

El Modelo de entrenamiento tiene el objetivo de crear dos vectores de entrenamiento, uno que esté modelado con las tramas que contienen ESV y otro vector cuyo modelamiento responda a las tramas que no presentan un evento. De manera que dichos vectores sirvan como referente para poder discriminar de manera adaptativa la presencia de un ESV en una señal.

Para el Modelo de Entrenamiento utilizado se comienza ordenando los valores de la energía de cada ventana, con el objetivo de determinar las de menor y mayor valor, extrayendo un porcentaje fijo de 30% de los cuadros de energía más bajos y más altos que representan los más confiables para ser utilizados en el modelo de entrenamiento. Los modelos se entrenan utilizando los MFCCs correspondientes a los índices de las tramas de confianza seleccionadas. Ambos Modelos de Entrenamiento de mezcla Gaussiana (GMMs) siguen la forma de la Ecuación (10) presentada anteriormente. Se utiliza el método de agrupamiento por vectores de cuantización conocido como k-media y la Prueba de Razón de Máxima Verosimilitud Estadística (LLR).

El algoritmo k-media opera con un tamaño de vector de cuantificación de tamaño 16 por sugerencia y experimentación de (Kinnunen & Padmanabhan, 2013) y tras realizar su proceso de agrupamiento por “clúster” devuelve para cada modelo, 16 vectores de entrenamiento, cada uno relacionado con 12 MFCCs.

A continuación, se calcula la distancia euclidiana en parejas entre dos conjuntos de observaciones como se visualiza en el Segmento de código 4, utilizando la función `pdist2` que compara par a par el vector de MFCCs denotado con la variable `Cep`, con el modelo de ESV obtenido con el algoritmo de k-media.

Segmento de código 4. Cálculo de la distancia euclidiana entre MFCCs y GMMs

```
% Calculo de distancias entre los modelos y los MFCCs
D_speech    = pdist2(Cep, speech_model, 'euclidean').^2;
D_nonspeech = pdist2(Cep, nonspeech_model, 'euclidean').^2;
LLR         = min(D_nonspeech) - min(D_speech');
```

Finalmente, se obtiene el vector LLR como la resta entre los valores mínimos del modelo de entrenamiento correspondiente a las tramas que no contiene ESV, menos el valor mínimo del modelo que contiene un evento. En la Figura 16 se observa una señal de ejemplo del vector LLR obtenido tras el modelamiento realizado, se resalta la coincidencia temporal de los picos de amplitud de la señal original Figura 16a, con respecto al vector LLR graficado Figura 16b.

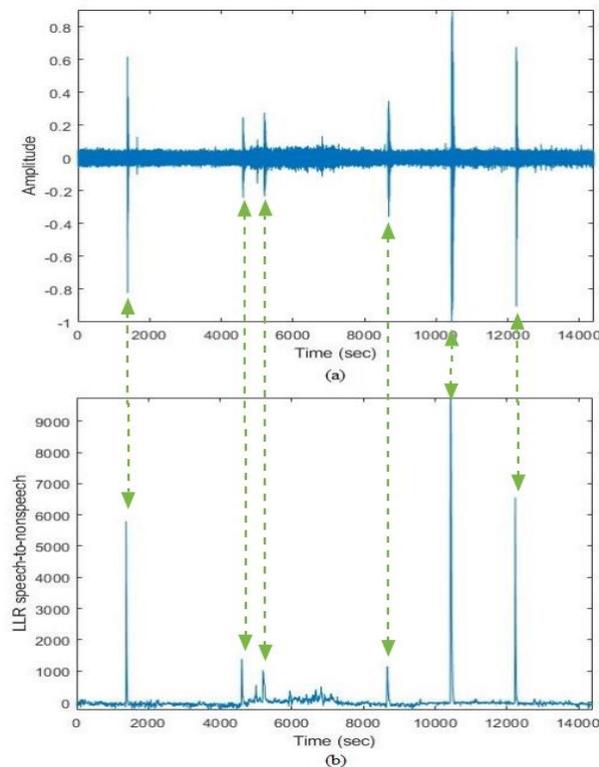


Figura 16. Coincidencia picos de frecuencia con vector LLR calculado.
a) Señal original b) Vector LLR graficado

3.1. Algoritmo de decisión

Una vez obtenidos los puntos de inicio y finalización de los ESV encontrados por el algoritmo A-VAD, es necesario discriminar aquellos picos de energía que debido a su corta duración no pueden ser considerados ESV, pero fueron reconocidos por el detector debido a su alta energía momentánea, por ese motivo este algoritmo de decisión busca eliminarlos para que su presencia no genere alertas falsas a los expertos del IGEPN.

Para definir que tramas de la señal contienen ESV, cada ventana debe cumplir dos parámetros: estar en el vector LLR obtenido del GMM y cumplir con un valor de energía mínima (Segmento de código 5), que, tras experimentación, se definió en -20dB. Al cumplir ambos parámetros el número de trama pasa a formar parte del vector `speechInd` que será parámetro de salida del A-VAD.

Segmento de código 5. Proceso de decisión del A-VAD.

```
%Proceso de decisión
VQ_speechInd = (LLR >= 0);
speechInd = VQ_speechInd;
speechInd = VQ_speechInd & (energy >= params.min_energy);
```

Posteriormente se lleva a cabo la eliminación de falsos positivos, para lo cual es necesario tomar en cuenta que la duración promedio de un ESV según (Rodriguez & Lara-Cueva, 2018) es de 20 segundos. El vector `speechInd` está formado por valores binarios, donde un '1' representa la presencia de un ESV en la trama y un '0' la ausencia de este. Es posible encontrar casos en que se tienen picos de energía que no cumplen con la condición de duración mínima para considerarse un ESV, en ese caso las tramas etiquetadas con un '1' pasan a ser '0'.

Finalmente, para determinar si un evento fue correctamente detectado se comparan el tiempo de inicio y finalización del evento según el detector A-VAD con respecto al tiempo definido en las etiquetas dadas por el IGEPN. Dichos valores serán expuestos en la sección de Resultados junto con los valores de las Medidas de rendimiento para ambas Bases de Datos.

En la Tabla 3 se muestra un compendio de los parámetros utilizados en el algoritmo A-VAD desarrollado, describiendo brevemente su objetivo y el valor fijado para su correcto funcionamiento con las señales sísmicas.

Tabla 3
Parámetros de control en el algoritmo A-VAD

Parámetro	Valor	Descripción
parms.frame_len	2 seg	La señal sísmica de entrada será dividida en ventanas de duración definida por este parámetro.
parms.frame_shift	1 seg	Parámetro que define el solapamiento entre ventanas.
parms.dither	True	Se agrega ruido Gaussiano despreciable de pequeña amplitud para evitar vectores MFCC duplicados.
parms.clean_energy	True	Se acepta aplicar sustracción espectral en los valores de la energía de la señal original.
parms.clean_MFCCs	False	No se aplica sustracción espectral para el cálculo de los MFCCs ya que el proceso usa características de la señal original.
parms.energy_fraction	0.30	Fracción de las muestras de energía seleccionadas para el modelo de entrenamiento.
parms.min_energy	- 20dB	Constante que define la energía mínima necesaria para que una muestra de energía pueda ser considerada como un ESV.
parms.vq_size	16	Tamaño del vector de cuantificación para el modelo de entrenamiento, sugerido por el algoritmo desarrollado en (Kinnunen & Padmanabhan, 2013).
parms.max_kmeans_iter	20	Número máximo de iteraciones en el algoritmo de k-media.

CONTINÚA 

params.num_filters	26	Número de bandas de frecuencia sugerido para los Filtros de la Frecuencia del Mel.
params.num_cep	12	Número de Coeficientes Cepstrales de la Frecuencia de Mel incluido el coeficiente C_0 .
params.include_C0	True	Se incluye el coeficiente C_0 dentro del análisis.
params.NFFT	512	Tamaño de la Transformada Rápida de Fourier (FFT, del inglés <i>Fast Fourier Transformation</i>).

CAPÍTULO IV

4. RESULTADOS

El proceso que se llevó a cabo para la obtención de las Medidas de rendimiento se muestra en la Figura 17, esto permitirá evaluar el algoritmo desarrollado. Para obtener los resultados de esta investigación se utilizan dos Bases de Datos cuya realización fue supervisada por el IGEPN y sus señales sísmicas no fueron consolidadas en una sola debido a las distintas conclusiones que se pueden sacar analizándolas por separado. Una explicación detallada de cada una de ellas se presenta más adelante en este capítulo.

Se comienza con las Bases de Datos que van a ser preprocesadas, luego pasarán por el algoritmo A-VAD y de decisión, posteriormente se compara con los valores definidos en las etiquetas del IGEPN para poder obtener las Medidas de rendimiento para cada Base de Datos por separado.

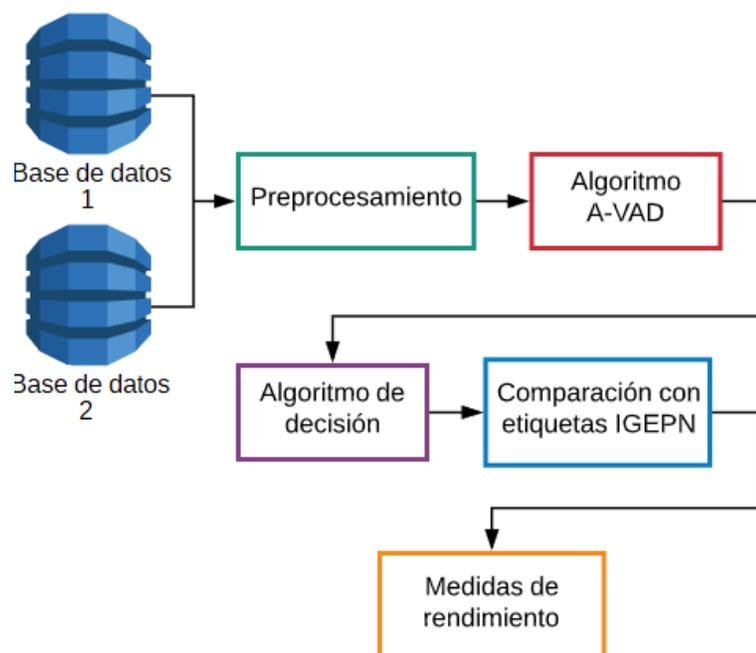


Figura 17. Diagrama de bloques de los resultados a presentarse

A continuación, se explica detalladamente las Bases de Datos a utilizarse, como también los experimentos llevados a cabo con los parámetros del algoritmo A-VAD y finalmente los valores de las Medidas de Rendimiento.

4.1. Bases de Datos

Se utilizan dos tipos de Bases de Datos para la fase de experimentación y resultados. La Base de Datos 1 corresponde a señales sísmicas de tres días consecutivos, en las cuales se puede encontrar más de un evento en cada tramo de señal de 20 minutos, o bien no detectar ninguno. Las señales corresponden a los tres primeros días del año 2018 y fueron tomadas de las estaciones BREF, BVC2 y VC1 en la componente vertical. El IGEPN etiquetó 77 ESV encontrados, de lo cual se obtuvo las estadísticas que se muestran en la Tabla 4. En la Figura 18 se muestra la ventana de almacenamiento de los ESV donde se visualizan los campos que se incluyen. En la Figura 19 se muestra la estructura de la nominación de las señales sísmicas en su formato .mat en la Base de Datos 1.

Trama	'DIA'	'HORA INICIO'	'MIN INICIO'	'SEG INICIO'	'HORA FINAL'	'MIN FINAL'	'SEG FINAL'	'PUNTO INICIO'	'PUNTO FINAL'	'DURACION'	'TIPO'
3	1	0	41	5	0	42	15	123250	126750	70	'REGIONAL'
7	1	2	11	30	2	12	30	394500	397500	60	'LP'
12	1	3	57	0	3	59	10	711000	717500	130	'REGIONAL'
14	1	4	24	0	4	24	30	792000	793500	30	'LP'
15	1	4	47	25	4	48	0	862250	864000	35	'LP'
17	1	5	38	10	5	38	45	1014500	1016250	35	'LP'

Figura 18. Ejemplo de almacenamiento de ESV de la Base de Datos 1

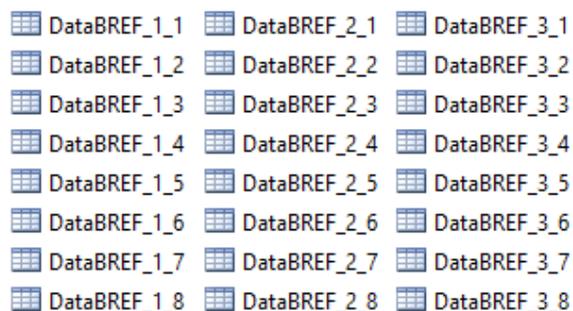


Figura 19. Señales sísmicas en formato .mat

Tabla 4*Número de eventos por tipo en la Base de datos 1*

Estación de referencia	Frecuencia de muestreo (Hz)	Tipo de evento	Número de eventos
BREF	50	LP	46
		VT	13
		Regional	18
Total eventos			77

Al someter la Base de Datos 1 al proceso detallado en la Figura 17, el algoritmo A-VAD desarrollado identificó 55 nuevos eventos que fueron etiquetados individualmente por personal del IGEPN definiendo la hora de inicio y finalización del EVS y su tipología, encontrándose 2 eventos regionales, 45 eventos de tipo LP, 3 eventos de tipo VT, 1 de tipo HB y 4 eventos no recibieron clasificación (S/C) debido a una baja SNR. De manera que la Base de Datos 1 se muestra modificada y actualizada en la Tabla 5 y será utilizada para el cálculo de las medidas de rendimiento.

Tabla 5*Número de eventos por tipo en la Base de datos 1 extendida*

Estación de referencia	Frecuencia de muestreo (Hz)	Tipo de evento	Número de eventos
BREF	50	LP	86
		VT	16
		Regional	20
		HB	1
		S/C	4
Total eventos			127

La Base de Datos 2 o Base de Datos Discreta consta de 1190 eventos discretos etiquetados por el algoritmo detector desarrollado en (Rodriguez & Lara-Cueva, 2018) en cooperación y con la supervisión del IGEPN. Esta base de datos es un compendio de ESV que fueron extraídos de la señal sísmica original y recortados tomando 10 segundos de señal adicional tanto al inicio como al final del evento desde el punto definido por los expertos como se indica en la Figura 20. Las señales pertenecen al periodo desde enero del 2012 hasta marzo del 2019, algunas de sus características relevantes son:

- Corresponden a la red de sismómetros “EC” de Ecuador.
- Se extrajeron 481 eventos de la estación VC1 con una frecuencia de muestreo de 100Hz y 709 eventos de la estación BREF con una frecuencia de muestreo de 50Hz.
- Los eventos de las estaciones se obtienen de la componente espacial vertical HHZ.
- El número de eventos clasificados por su tipo se muestran en la Tabla 6.

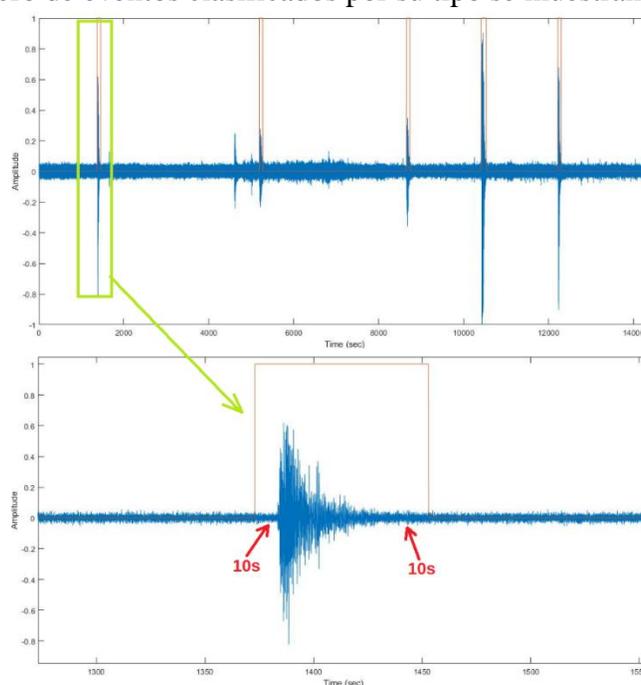


Figura 20. Ejemplo de EVS recortado de la señal sísmica original.

Tabla 6
Número de eventos por tipo en la Base de datos Discreta

Estación de referencia	Frecuencia de muestreo (Hz)	Tipo de evento	Número de eventos
BREF	50	LP	485
		VT	119
		Regional	26
		HB	72
		Icequake	7
Total Eventos BREF			709
VC1	100	LP	404
		VT	51
		Regional	26
		HB	0
		Icequake	0
Total Eventos VC1			481
Total eventos Base de datos 2			1190

4.2. Medidas de Rendimiento

Las Medidas de rendimiento utilizadas para evaluar el desempeño del algoritmo A-VAD son: exactitud (A, del inglés *accuracy*), precisión (P, del inglés *precision*), sensibilidad (R, del inglés *recall* o *sensitivity*), especificidad (S, del inglés *specificity*) y tasa de error balanceado (BER, del inglés *Balanced Error Rate*). Estas medidas de rendimiento fueron utilizadas para evaluar el algoritmo desarrollado en (Rodriguez & Lara-Cueva, 2018), con el cual será comparado el presente trabajo más adelante.

Los términos utilizados para describir los conceptos de las medidas de rendimiento son: los verdaderos positivos (VP), verdaderos negativos (VN), falsos negativos (FN) y falsos positivos

(FP). Se entiende como VP a los ESV que son detectados correctamente, los verdaderos negativos representan aquellas señales que no contienen ningún ESV y el detector tampoco reconoció un evento en la misma, los falsos negativos son aquellos eventos que están en la señal y no fueron reconocidos por el detector y finalmente los falsos positivos son tramas de ruido que el detector reconoció erróneamente como ESV. Tanto los falsos positivos como los falsos negativos indican que el detector está generando un resultado contrario al correcto. Las medidas de rendimiento antes mencionadas vienen dadas en función de los términos ya definidos, como se muestra en las siguientes ecuaciones:

$$A(\%) = \frac{\text{Número de detecciones correctas}}{\text{Número total de eventos}} * 100 \quad (19)$$

$$P(\%) = \frac{VP}{VP + FP} * 100 \quad (20)$$

$$R(\%) = \frac{VP}{VP + FN} * 100 \quad (21)$$

$$S(\%) = \frac{VN}{VN + FP} * 100 \quad (22)$$

$$BER = 1 - \frac{R + S}{2 * 100} \quad (23)$$

Como Medida de Rendimiento también se utilizará la Curva Característica Operativa del Receptor (ROC, del inglés *Receiver Operating Characteristic*), que es una representación gráfica que muestra la sensibilidad en relación con la especificidad en un sistema clasificador que maneja dos clases. Esta curva nació en la Segunda Guerra Mundial para evaluar radares de detección y actualmente es ampliamente utilizado en la Teoría de detección de señales (Zou, O'Malley, & MauriMD, 2007). La curva ROC es obtenida al variar un umbral de discriminación y así, obtener para cada valor de umbral, un valor de VP y de FP. El indicador más utilizado para interpretar una

curva ROC es el área bajo la curva (AUC, del inglés *Area Under Curve*), mismo que será obtenido para ambas Bases de Datos en la presente investigación.

En la siguiente sección se detallan los experimentos realizados y los resultados obtenidos en cuanto a las medidas de rendimiento propuestas.

4.3. Etapa de experimentación

Se llevaron a cabo experimentos para optimizar los parámetros de operación del algoritmo A-VAD de manera que mejoren las medidas de rendimiento y se ejecute en un tiempo de procesamiento mínimo. Se experimentó con todas las señales sísmicas de ambas Bases de Datos ya que sus parámetros de funcionamiento deben ser los mismos, teniendo 216 señales sísmicas de la Base de Datos 1 y 1190 ESV de la Base de Datos 2 que pasaron por el algoritmo A-VAD y cuyos resultados sirvieron para seleccionar los mejores parámetros en esta etapa de experimentación.

El primer parámetro sometido a pruebas es el tamaño de la ventana y su solapamiento, ya que de esto dependerá el proceso de inventariado y cálculo de energía por tramas, que son las bases del algoritmo. La sugerencia inicial de (Kinnunen & Padmanabhan, 2013) para señales de voz, es utilizar ventanas de 0.03 segundos con 0.01 segundos de solapamiento, y si bien dichas señales guardan similitud con las señales sísmicas, se demostró por método experimental en (Rodríguez & Lara-Cueva, 2018) que, utilizando ventanas del 10% de la longitud promedio de un microsismo, que es 20 segundos, se puede considerar las tramas como un proceso estrictamente estacionario, donde la distribución de probabilidad en un instante de tiempo fijo es la misma para todos los instantes de tiempo.

En la Tabla 7 se muestran los resultados obtenidos para los diferentes valores sugeridos y se concluye que la utilización de una ventana de 2 segundos con solapamiento de 1 segundo permite que las medidas de rendimiento sean las más altas en ambas Bases de Datos.

Tabla 7
Resultados experimentación con el tamaño de ventana

		Base de Datos 1	Base de Datos 2
Ventana	Solapamiento	A [%]	A [%]
[seg]	[seg]		
0.03	0.01	81.2653	77.4464
0.5	0.25	84.2563	90.6744
1	0.5	97.11563	91.4307
1.5	0.75	97.4261	92.8918
2	1	99.1039	96.6849
3	1.5	95.3651	90.436
5	2.5	90.2563	86.153
10	5	87.1567	43.256

A continuación, en la Figura 21 se muestra un ejemplo del cálculo realizado para obtener la precisión correspondiente en cada uno de los valores experimentales que se muestran en la Tabla 7. El proceso para la Base de Datos 1 consiste en obtener las etiquetas generadas para cada estación sismológica y verificar si cumplen con la Opción 1 (que el ESV conste en la estación BREF y BVC2) o la Opción 2 (que el ESV conste en la estación BREF y VC1), posteriormente verificar si cumple con el umbral fijo de 5 segundos en los puntos de inicio y finalización, y como último paso realizar el cálculo de la precisión utilizando el número de detecciones correctas obtenidas. Mientras que en la Base de Datos 2 se obtienen las 1190 etiquetas que corresponden a cada una de las señales

sísmicas y continúan con el proceso de verificación del umbral fijo, de donde resultan 1147 etiquetas de ESV detectados, que permiten calcular la precisión que se muestra.

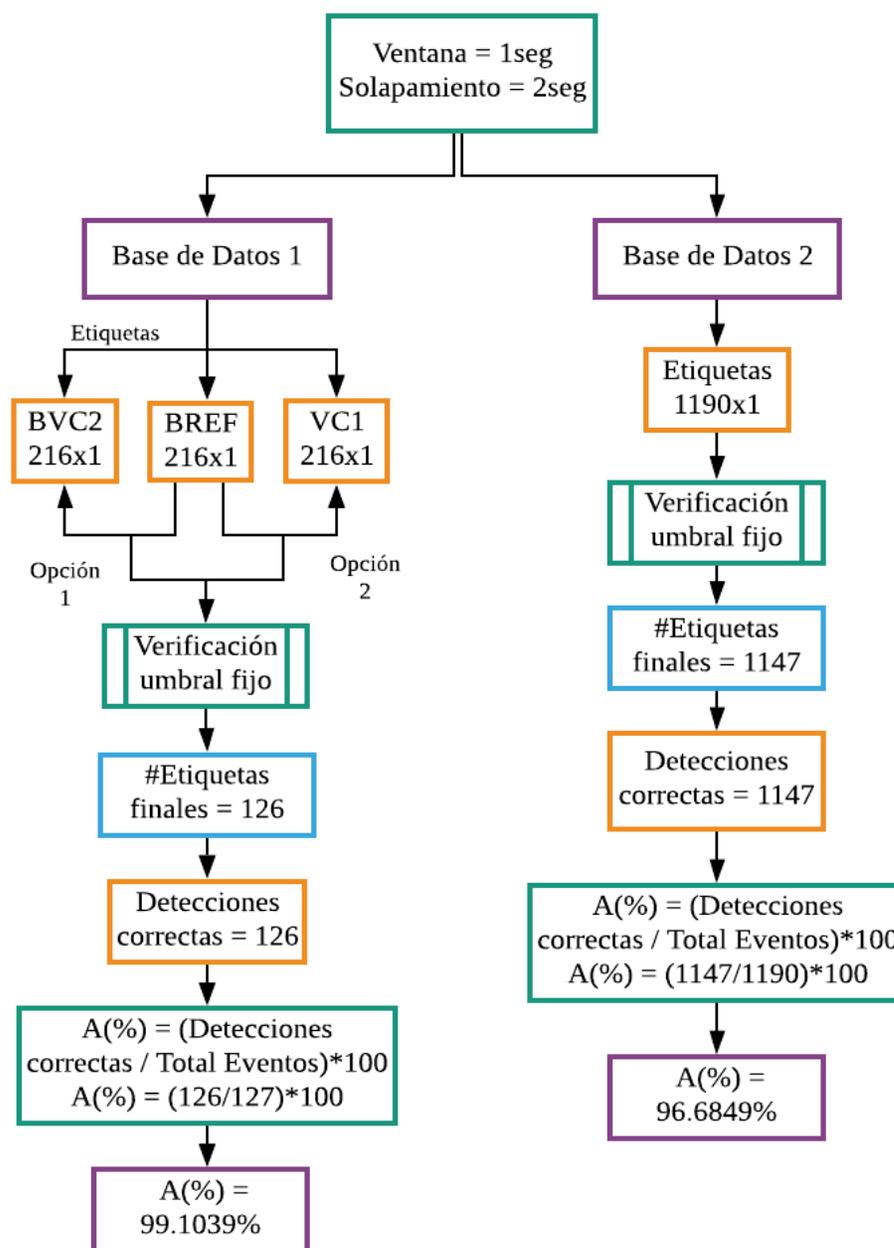


Figura 21. Ejemplo de cálculo de precisión en la etapa de experimentación.

Una vez seleccionado el tamaño de ventana a utilizar, se requiere decidir la energía mínima necesaria que una ventana debe tener para ser considerada como poseedora de un ESV, la energía

es calculada con la Ecuación (17). Para lo cual, se analizaron 10 ESV etiquetados por el IGEPN, tomados de la Base de Datos Discreta, fijándose principalmente en el valor de la energía en su trama de inicio y trama de finalización como se muestra en la Tabla 8. Se muestran también el valor promedio de las tramas analizadas para tener un intervalo de valores de energía con los cuales realizar las pruebas experimentales.

Tabla 8

Experimentación con valores de energía de los ESV.

	Energía [dB]										Promedio
Valor Máximo	-10.35	-21.07	-20.04	-16.93	-5.19	-9.21	-11.26	-13.24	-19.27	-20.16	-14.672
Valor Mínimo	-32.5	-36.22	-30.24	-33.2	-38.39	-30.11	-33.14	-30.16	-31.25	.30.76	-32.597

Se realizaron pruebas con valores dentro del rango de -14.672dB y -32.597dB. Se obtuvo los resultados mostrados en la Tabla 9 y se concluye que la energía óptima a ser utilizada en ambas Bases de Datos es de -20dB ya que presenta el mejor resultado en cuanto a medidas de rendimiento.

Tabla 9

Resultados experimentación con valor mínimo de energía

Energía [dB]	Base de Datos 1 A [%]	Base de Datos 2 A [%]
-17	91.2563	84.1563
-19	96.2533	91.1575
-20	99.1039	96.6849
-22	98.1658	92.5116
-24	97.2678	92.5231
-28	97.1685	92.4643

Como se explicó en secciones anteriores, el Modelo de Entrenamiento a utilizarse requiere de fracciones de las muestras de energía que fueron ordenadas de manera ascendente. Tal que un cierto porcentaje de los valores más bajos de energía del vector de tramas que no presentan ESV y los valores más altos de energía de las tramas que presentan ESV, sean utilizadas para el modelo de entrenamiento. Se considera que dichas tramas de energía mínima y máxima brindan un mayor nivel de confiabilidad para la generación de los vectores de entrenamiento. El porcentaje de tramas seleccionadas fue modificado como se muestra en la Tabla 10 y se observa que el mejor resultado en ambas Bases de Datos se obtiene al utilizar el 30% de las tramas.

Tabla 10
Resultados experimentación con porcentaje de tramas

Número de tramas	Base de Datos 1	Base de Datos 2
[%]	A [%]	A [%]
10	96.1240,	92.3918
20	97.6744	93.2489
30	99.1039	96.6849
40	91.2562	93.0935
50	89.5478	92.5231

A continuación, se presentan los resultados obtenidos con los parámetros modificados, que fueron: tamaño de ventana y solapamiento, valor de energía mínima y porcentaje de tramas para el modelo de entrenamiento. Estos valores fueron definidos tras experimentación en el algoritmo A-VAD para las dos Bases de Datos utilizadas en la presente investigación.

4.4. Evaluación del desempeño del algoritmo

La evaluación de algoritmo A-VAD se ejecutó utilizando Matlab R2018a ® en una computadora con procesador Core i7 @2.40GHz y 8GB de RAM. Los puntos a verificarse son: a)

la detección del ESV y la similitud de sus puntos de inicio y finalización con respecto a las etiquetas del IGEPN y b) el tiempo de procesamiento que le toma al detector, llevar a cabo el análisis de cada señal y de las Bases de Datos enteras.

El desempeño del algoritmo se va a medir utilizando 216 señales sísmicas pertenecientes a la Base de Datos 1 y 1190 señales sísmicas discretas contenidas en la Base de Datos 2, mismas que van a ser analizadas por el algoritmo A-VAD desarrollado, obteniendo así las etiquetas de los puntos de inicio y finalización de los ESV detectados.

4.4.1. Resultados Base de Datos 1

La Base de Datos 1 consta de 72 señales diarias por cada estación de monitorización (se cuenta con tres estaciones sismológicas de monitorización denominadas BREF, BVC2, VC1 cuyas gráficas se muestran en la Figura 22), se analizan los tres primeros días consecutivos de enero de 2018, por lo tanto, se tienen 216 señales sísmicas por cada estación que pueden o no contener ESV. Las estaciones utilizadas son BREF, BVC2 y VC1, siendo BREF la estación de referencia de la que se almacenarán los puntos de inicio y fin del ESV.

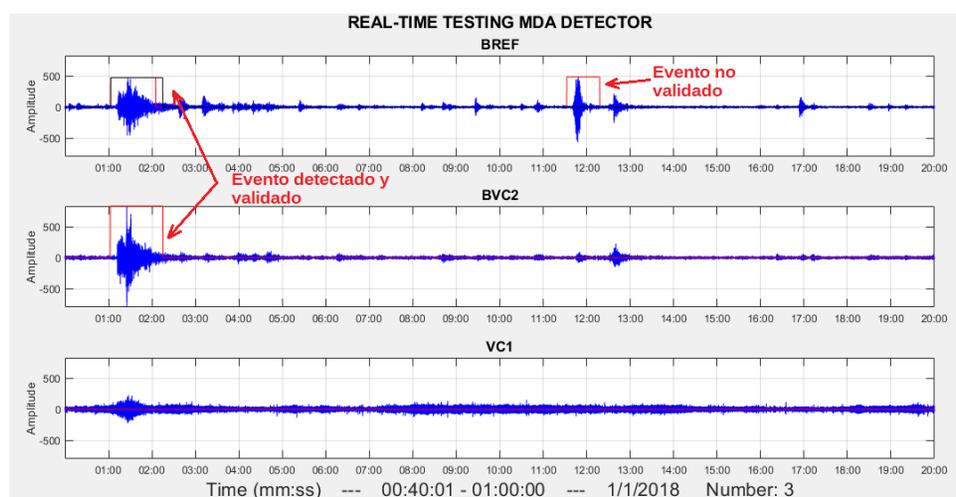


Figura 22. Validación de un ESV en la Base de Datos 1

Por indicación del IGEPN, se realizó un sistema de votación entre las tres estaciones sísmicas que como se observa en la Figura 22, de manera que se considera la ocurrencia de un ESV cuando se verifica su presencia en dos de las tres estaciones, siendo siempre necesaria su aparición en la estación BREF. Entonces pueden existir dos casos, que el ESV sea detectado en BREF y BVC2 y que el ESV sea detectado en BREF y VC1.

Además, para que el ESV sea almacenado también debe cumplir la condición de que sus valores de inicio y finalización coincidan con los valores dados por el IGEPN, este valor debe coincidir dentro de un límite específico de 5 segundos, a lo cual, en adelante se le conocerá como umbral.

En la Figura 22 se muestra un ejemplo de detección de ESV. En el minuto 01:00 se verifica un evento en la estación BREF y BVC2, que, al cumplir con el umbral de 5 segundos en sus puntos de inicio y fin, fue validado. Mientras que en el minuto 11:00 se detecta un evento solamente en la estación BREF, lo cual no cumple con el sistema de votación y no se valida dicho evento.

La Base de Datos 1 está conformada por 127 ESV que al ser analizados por el algoritmo A-VAD propuesto se obtuvieron los siguientes resultados:

- Verdadero Positivos: Se detectaron 124 ESV que coincidieron con las etiquetas del IGEPN.
- Falsos Positivos: No se detectó ningún evento adicional a los etiquetados y verificados por el IGEPN.
- Falsos Negativos: Se presentaron 3 ESV que no fueron detectados debido al tamaño reducido de su amplitud y energía con respecto a las demás señales que se ha estado analizando.

- Verdaderos Negativos: Tomando en cuenta que se tienen 216 señales sísmicas, se consideran VN a todas aquellas señales que no poseen una etiqueta de ESV y que tampoco fueron catalogadas como eventos por el algoritmo A-VAD, por lo tanto, se entiende como la resta de 216 señales menos 124 VP y menos 3 FN, obteniendo 89 VN.

Con estos valores de VP, FP, FN, VN, las Medidas de rendimiento para el punto específico de operación, fijando el valor del umbral en 5 segundos, son las que se muestran en la Tabla 11, donde se muestra el valor de BER obtenido que satisface el requerimiento del IGEPN de tener un valor de BER de 0.01 según (Rodriguez & Lara-Cueva, 2018).

Tabla 11
Resultados Medidas de rendimiento Base de Datos 1

Medida de Rendimiento	Valor
A	98.4496 %
P	100 %
R	97.7273 %
S	100 %
BER	0.0114

Como se mencionó anteriormente, se presenta también la Curva ROC la cual se genera a partir de variar el valor del umbral desde 0 hasta 40 segundos obteniendo un punto para la gráfica cada 0.5 segundos, ya que esto permite tener una idea generalizada del desempeño del algoritmo A-VAD con diferentes valores de umbral. La Figura 23 muestra la Curva ROC resultante, en el eje X se tiene el número de Falsos Positivos y en el eje Y el número de Verdaderos positivos detectados. Se observa que el AUC obtenida es de 99.1039%.

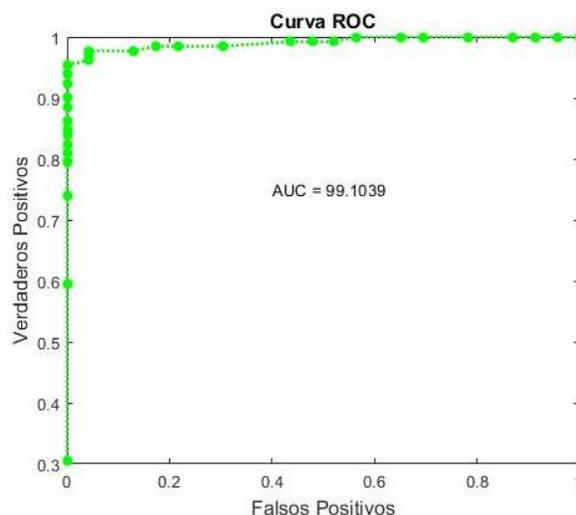


Figura 23. Curva ROC resultante Base de Datos 1

El tiempo de procesamiento necesario para analizar una sola señal de duración 20 minutos es de 5.57 segundos, ahora bien, tomando en cuenta que para analizar toda la Base de Datos 1 se requieren analizar 216 señales por cada estación de monitoreo, el tiempo de procesamiento junto con el sistema de decisión y de votación es de 7.14 minutos.

4.4.2. Resultados Base de Datos 2

Esta Base de Datos Discreta contiene 1190 señales sísmicas con un ESV cada una. A diferencia de la Base de Datos 1, no se comparan las señales entre las tres estaciones, ya que el evento fue almacenado y recortado extrayéndolo de la estación que permita su mejor visualización. De este modo, el valor del umbral seleccionado toma vital importancia en la validación de la detección del ESV en comparación con las etiquetas ya establecidas.

En la Figura 24 se muestra un ejemplo de ESV perteneciente a la Base de Datos 2. Se indica también el cuadro de detección: a) en rojo, el etiquetado por el IGEPN, b) en lila, el generado por el algoritmo A-VAD. Se pueden presentar los siguientes casos que validan a la señal como poseedora de un ESV:

- Punto de inicio IGEPN = Punto de inicio A-VAD \pm 5 segundos
- Punto de fin IGEPN = Punto de fin A-VAD \pm 5 segundos

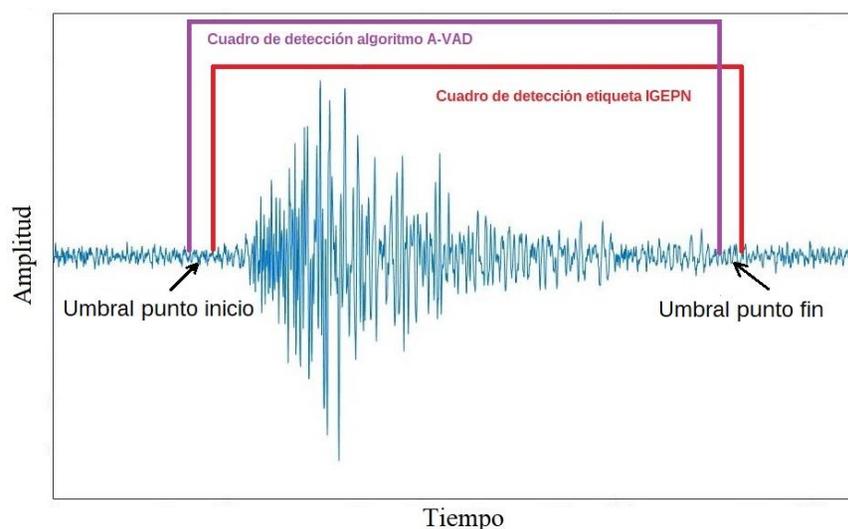


Figura 24. Validación de ESV Base de Datos 2

La evaluación de esta Base de Datos discreta se realiza ingresando las 1190 señales sísmicas en el algoritmo A-VAD desarrollado para obtener las etiquetas de los puntos de inicio y finalización de los ESV. Esto se lo realiza trabajando en un punto fijo de operación y se obtuvieron los siguientes resultados:

- Verdaderos Positivos: fueron detectadas correctamente y dentro del umbral de 5 segundos 1147 señales de 1190.
- Verdaderos Negativos: se considera que cada señal sísmica que fue correctamente detectada consta también como VN ya que los tramos de señal que no son considerados parte del ESV tanto al inicio como al final de la señal (denominados pre y post ESV), no fueron catalogados como un ESV por el algoritmo A-VAD, lo cual es correcto y lo verifica como VN.

- Falsos Positivos: no se detectaron eventos adicionales a los ya presentados por la Base de Datos.
- Falsos Negativos: se registran 43 señales cuyos puntos de inicio y fin no coincidieron con las etiquetas del IGEPN guardando el límite de umbral de 5 segundos.

Las Medidas de rendimiento para el punto específico de operación, fijando el valor del umbral en 5 segundos, son las que se muestran en la Tabla 12, y al igual que el valor de BER obtenido para la Base de Datos 1, el BER alcanzado en la Base de Datos 2 satisface el requerimiento del IGEPN de no superar el 0.01.

Tabla 12

Resultados Medidas de rendimiento Base de Datos 2

Medida de Rendimiento	Valor
A	96.3866 %
P	100 %
R	96.3866 %
S	100 %
BER	0.0118

La Curva ROC obtenida toma la forma que se visualiza en la Figura 25, generada al variar el valor de umbral desde 0 hasta 40 segundos en los resultados de la Base de Datos 2. Permite tener una idea generalizada del desempeño del A-VAD para no limitar los resultados al punto de operación específico. El AUC obtenido es de 96.6849%, valor que cumple con las expectativas de la presente investigación.

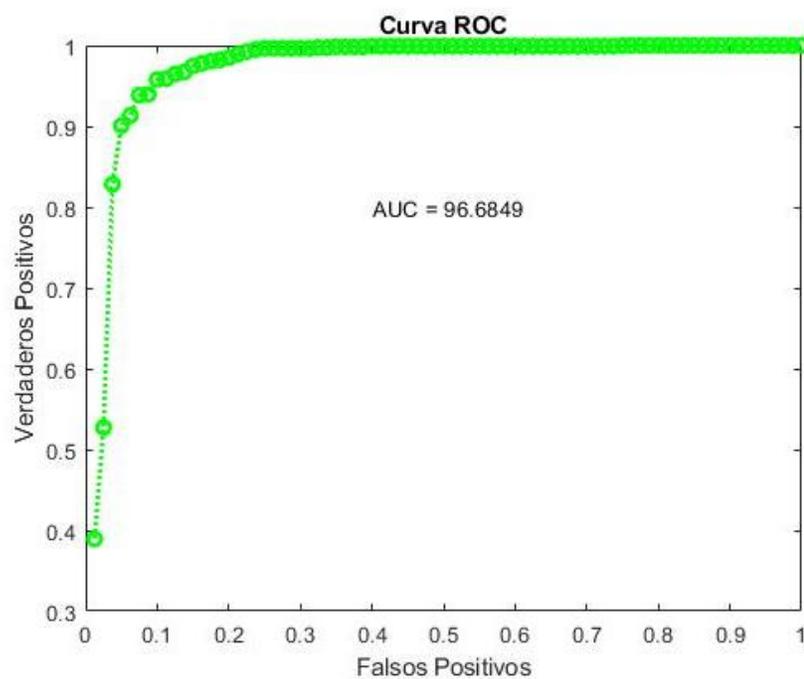


Figura 25. Curva ROC resultante Base de Datos 2

Finalmente, el tiempo de procesamiento necesario para analizar cada una de las 1190 señales sísmicas pertenecientes a esta Base de Datos discreta es en promedio de 0.3354 segundos, mientras que para analizar toda la Base de Datos 2 junto con el algoritmo de decisión y obtención de las Medidas de rendimiento el tiempo necesario es de 8.3448 minutos.

CAPITULO V

5. DISCUSIÓN

En la presente investigación se desarrolló un algoritmo adaptativo que utiliza técnicas de VAD para la detección de ESV. El método utilizado comenzó con el preprocesamiento de la señal sísmica, filtrándola y eliminando la tendencia lineal de la misma como se realizó también en (Rodriguez & Lara-Cueva, 2018). A continuación, el algoritmo A-VAD como tal, demostró que la utilización de los MFCCs se adapta a las señales sísmicas brindando características relevantes de los tramos de la señal que presentan un ESV. Posteriormente, la Sustracción Espectral aplicada a la señal permitió aumentar la diferencia entre la energía de las tramas que presentan un ESV con respecto a las que no. El paso final del algoritmo A-VAD fue el Modelo de Entrenamiento, que utilizando el Método de Agrupamiento de k-medias y la Prueba de Razón de Máxima Verosimilitud permitió obtener dos modelos que discriminan entre los tramos de señal que contienen ESV y aquellos que no. El algoritmo de decisión fue desarrollado específicamente para señales sísmicas, ya que se tomaron en cuenta las características propias del ESV como su duración, para la eliminación de Falsos Positivos. Hasta este punto del proceso, nada de lo llevado a cabo difiere para las dos Bases de Datos utilizadas.

Fue de vital importancia la etapa de experimentación ya que permitió definir el valor de algunos parámetros relevantes de los cuales dependieron las Medidas de rendimiento resultantes. Al modificar los parámetros en el proceso de encontrar los mejores resultados se analizó la precisión que se obtenía, ya que es la Medida de Rendimiento que más variaba por depender directamente del número de detecciones correctas con respecto al número total de eventos etiquetados por el IGEPN. En resumen, se modificaron los parámetros que se muestran en la Tabla

13, indicando el porcentaje de mejora obtenido con respecto a los valores iniciales sugeridos por (Kinnunen & Padmanabhan, 2013), que constan en la columna denominada Valor Sugerido. Se observa que el mayor porcentaje de mejora obtenido depende de la longitud de ventana y su solapamiento, ya que al tomar secciones muy pequeñas de la señal se aumentaba tiempo de procesamiento y se producían cortes en los eventos, es decir, existían eventos que siendo uno solo, era catalogado como dos o tres ESV. Mientras que, al aumentar el tamaño de la ventana, se pierde exactitud al comparar los puntos de inicio y fin con las etiquetas del IGEPN, lo que se traduce disminución de las Medidas de Rendimiento, es así que con longitud de ventana de 2 seg y 1 seg de solapamiento se obtuvo una mejora del 17.8386% en la Base de Datos 1 y del 19.2385% en la Base de Datos 2. Es importante analizar que el mayor porcentaje de mejora se obtuvo en la Base de Datos 2 con respecto a todos los parámetros modificados, ya que en esta Base de Datos Discreta se cuenta con un mayor número de ESV (1190) en comparación con la Base de Datos 1 que posee 216 señales sísmicas de 20 minutos de duración.

Tabla 13
Mejora de Medidas de rendimiento por experimentación

Parámetro modificado	Valor	Valor	Mejora Base de Datos 1	Mejora Base de Datos 2
	sugerido	implementado		
Longitud de ventana y solapamiento	0.03 seg	2 seg	17.8386%	19.2385%
	0.01 seg	1 seg		
Valor de energía mínima	-17 dB	-20 dB	7.8476%	12.5286%
Porcentaje de tramas para Modelo de Entrenamiento	10%	30%	2.9799%	4.2931%

Para el cálculo de las Medidas de rendimiento es necesario tratar a cada Base de Datos por separado debido a la manera en que se presentan las señales que van a ser analizadas, ya que esto

va a permitir saber en qué tipo de Base de Datos el algoritmo A-VAD desarrollado trabaja de mejor manera. En la Tabla 14 se indican los resultados obtenidos, apreciándose que la Base de Datos 1 brindó mejores resultados en todos los parámetros de las Medidas de Rendimiento, esto en gran parte se debe a la validación de los ESV extras que pudieron ser verificados por personal del IGEPN, ya que de no haber sido así, hubieran sido tratados como eventos Falsos Positivos, disminuyendo las Medidas de Rendimiento.

Tabla 14

Contraste de resultados entre las Bases de Datos

Medida de Rendimiento	Base de Datos 1	Base de Datos 2
A	98.4496 %	96.3866 %
P	100 %	100 %
R	97.7273 %	96.3866 %
S	100 %	100 %
BER	0.0114	0.0118
AUC	99.1039%.	96.6849%,

La Base de Datos 2, por su parte, también presentó valores que satisfacen el requerimiento del IGEPN pero la disminución en promedio del 1.94% en sus Medidas de rendimiento (obtenida de la resta entre la Medida de Rendimiento de la Base de Datos 1 menos la Base de Datos 2) es atribuida a que cada una de las 1190 señales sísmicas fueron recortadas de la señal original, y fueron almacenadas con solamente un tramo de señal de 10 segundos antes y después del ESV. Entonces, dado que el algoritmo A-VAD presentado utiliza un Modelo de Entrenamiento Adaptativo, existe un menor número de tramas de cada tipo, para los vectores de entrenamiento. A diferencia de las señales de la Base de Datos 1, donde la señal tenía una duración de 20 minutos, tiempo que se evidenció mejora los resultados del Modelo de Entrenamiento Adaptativo.

5.1. Comparación con trabajos similares

El presente trabajo de investigación va a ser comparado con el algoritmo desarrollado en (Rodríguez & Lara-Cueva, 2018) denominado MDA (del inglés, *Microseisms Detector Algorithm*), mismo que utiliza una combinación de técnicas para cumplir con los requerimientos del IGEPN. Al igual que en el algoritmo A-VAD desarrollado, el preprocesamiento y la segmentación de la señal en ventanas solapadas son los primeros pasos del método usado. Se presenta una diferencia con la utilización de un filtro óptimo utilizado para la detección del ESV y en cuyo diagrama de estados se realizaron varias modificaciones para adaptar su funcionamiento con las señales sísmicas del volcán Cotopaxi.

El algoritmo MDA de (Rodríguez & Lara-Cueva, 2018) fue evaluado en lo que en esta investigación se denominó Base de Datos 1, que consta de 216 señales sísmicas con una duración de 20 minutos cada una, obteniéndose los resultados que se muestran en la Tabla 15 en comparación con los valores de las Medidas de rendimiento generadas con el algoritmo A-VAD desarrollado. Se puede observar también que el único parámetro de las Medidas de rendimiento que tiene un valor mayor en el algoritmo MDA es la sensibilidad (R), mientras que los demás parámetros fueron mejorados en la presente investigación gracias a las técnicas y herramientas utilizadas.

Tabla 15
Comparación de resultados con un trabajo previo

Medida de Rendimiento	Algoritmo	Algoritmo MDA
A	98.4496 %	98 %
P	100 %	96 %
R	97.7273 %	98 %
S	100 %	98 %
BER	0.0114	0.020

5.2. Conclusiones

El estudio del estado del arte realizado con respecto a las técnicas de VAD permitió conocer las principales falencias de las técnicas que ya no son muy utilizadas en la actualidad, debido a que varias investigaciones demostraron que se podían presentar mejoras en la implementación. De esta manera se evitó seleccionar erróneamente técnicas que disminuirían las Medidas de rendimiento esperadas.

Las técnicas de VAD utilizadas en la presente investigación pudieron ser adaptadas para que funcionen correctamente con las señales sísmicas del volcán Cotopaxi, comprobándose así la similitud que existe entre el espectro del habla y el de los ESV volcánicos, que, si bien no están en el mismo amplio rango de frecuencias que la voz humana, pueden adaptarse mediante la modificación de algunos parámetros de funcionamiento.

Durante la etapa de experimentación se evidenció que la utilización de los valores para los parámetros sugeridos por (Kinnunen & Padmanabhan, 2013) arrojaban resultados aceptables, pero no cumplían con los requerimientos del IGEPN, por lo que, al modificar el tamaño de ventana y su solapamiento, el valor mínimo de energía y el porcentaje de tramas a utilizarse en el Modelo de Entrenamiento, se mejoraron las Medidas de rendimiento y se logró satisfacer el requerimiento de que el BER no supere el valor de 0.01 en ambas Bases de Datos.

La aplicación del algoritmo A-VAD en las Bases de Datos por separado, respetando como cada una presentaba los ESV de diferente manera, permitió concluir que el detector desarrollado operaría satisfactoriamente en el ambiente de prueba que utiliza el IGEPN para etiquetar en tiempo real los ESV, analizando señales con duración de 20 minutos, donde se pueden suscitar ESV o no, presentando una precisión del 98.4496 %.

En base a las Medidas de rendimiento obtenidas al analizar las señales de la Base de Datos Discreta con el algoritmo A-VAD, se concluye que se podría mejorar la precisión obtenida del 96.3866%, incluyendo más tiempo de la señal pre y post ESV, ya que específicamente en la presente investigación se requiere de un número mayor de muestras de ruido para generar un mejor Modelo de Entrenamiento, y así tomar una decisión más acertada en cuanto a la definición de si una trama representa un ESV o no. La precisión obtenida es alta, pero hubiera sido mayor, de haber contado con un tramo de la señal ruidosa más extenso que 10 segundos.

El desarrollo del presente algoritmo A-VAD contribuye con la línea de investigación a favor de la prevención y alerta temprana de catástrofes producidas por una eventual erupción volcánica del Cotopaxi, que permitirá que los profesionales del IGEPN aumenten su rendimiento en el etiquetado de ESV. De esta manera se podrá emitir informes de mayor confiabilidad hacia las entidades encargadas de generar una alerta temprana hacia la ciudadanía, salvaguardando así decenas de miles de vidas.

5.3. Trabajos Futuros

La presente investigación que se basó en la utilización de técnicas de VAD podría ser mejorada aplicando nuevas técnicas de VAD que están surgiendo y, con el progreso de la tecnología investigativa están siendo utilizadas actualmente en la parte comercial del reconocimiento automático de voz. La técnica publicada en (Qing Ong, Chiat Tan , Vengadasalam, Heng Tan , & Hai Ooi, 2017) utiliza nuevas técnicas como la medida de la Entropía Ponderada de la Envolvente Superior (UEWE, del inglés *Upper Envelope Weighted Entropy*) como un medio para permitir una mejor diferenciación entre los segmentos de voz y silencio en una señal de audio, que junto con un Filtro Asimétrico no Lineal (ANF, del inglés *Asymmetric Nonlinear Filter*)

extraen el factor de peso adaptativo de la señal que compensará el efecto del ruido. Adaptando estas técnicas a las señales sísmicas del volcán Cotopaxi es posible generara mejores Medidas de Rendimiento.

Existe una amplia variedad de filtros que son utilizados para la Detección de Actividad de Voz, en la presente investigación se utilizó el Banco de Filtros de la Frecuencia de Mel que permitió obtener los MFCCs utilizados para el Modelo de Entrenamiento. En el trabajo de (Ong & Tan, 2016) se utiliza el Banco de Filtros de Gammatone que se inspira en el funcionamiento de la cóclea del oído humano, por lo que sería interesante poner a prueba las señales sísmicas de las Bases de Datos bajo este nuevo concepto de filtrado ya que para las señales de voz arroja resultados óptimos especialmente cuando no se tiene un alto nivel de SNR.

Finalmente, como trabajo futuro es recomendable ampliar las Bases de Datos existentes ya que no tienen la suficiente cantidad de señales sísmicas con ESV etiquetados que permitan experimentar con otras técnicas como Machine Learning o Deep Learning que necesitan que las señales sean separadas en un grupo de aprendizaje y otro de prueba.

REFERENCIAS

- Ying, D., Yan, Y., Dang, J., & Soong, F. (2011). Voice Activity Detection Based on an Unsupervised Learning Framework. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(8), 2624-2633. doi:Ying, D., Yan, Y., Dang, J., & Soong, F. K. (2011). Voice Activity Detection Based o10.1109/tasl.2011.2125953
- Aki, K., & Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research*, 80(23), 1896-1977. doi:<https://doi.org/10.1029/JB080i023p03322>
- Altuve, M., & Warrick, P. (2015). Descomposición De Karhunen-loève De Registros Electrocardiográficos Abdominal Materno. *Memorias de V Congreso Venezolano de Bioingeniería*. Caracas: Researchgate.
- Alvarez, M., Henao, R., & Duque, E. (Agosto de 2007). Clasificación de eventos sísmicos empleando procesos Gaussianos. (U. T. Pereira, Ed.) *Scientia et Technica Año XIII*, 35(0122-1701). Recuperado el 9 de Noviembre de 2019
- Blong, J. R. (1984). *Volcanic Hazards*. Estados Unidos: Academic Press.
- Borrás, O. (2006). Reductor de ruido mediante resta espectral en entorno Matlab. En E. TELECOMUNICACIÓN.
- Capó, M., Pérez, A., & Lozano, J. (2017). An efficient approximation to the K-means clustering for massive data. *Elsevier Knowledge-Based Systems*, 56-69.
- Chouet, B. A. (1996). Long-period volcano seismicity: its source and use in eruption forecasting. *Nature*, 309-316.

- Corominas, M. (Noviembre de 2011). Estudio Comparativo De Los Planes De Actuación Frente Al Riesgo Volcánico (Chile, Costa Rica, El Salvador, Ecuador, España, México Y Nicaragua). *Revista Geológica de América Central*, 33-56. doi:: 10.15517/rgac.v0i52.18980
- EducarChile. (2012). *Centro de recursos digitales*. Obtenido de Ondas sísmicas : <http://centroderecursos.educarchile.cl/handle/20.500.12246/52350>
- Gerkman, T., & Hendriks, R. (2012). Unbiased MMSE-Based Noise Power Estimation With Low Complexity and Low Tracking Delay. *IEEE Trans Audio, Speech, Language Processing*, 20, 1383-1393.
- Gubbins, D. (2004). The seismic wavefield: Introduction and theoretical development. *Cambridge University Press*, 143(3), 541-547. doi:<https://doi.org/10.1016/j.pepi.2004.06.002>
- Guffanti, M., Brantley, S., Cervelli, P., Nye, C., Serafino, G., Siebert, L., . . . Wald, L. (2007). *Technical-Information Products for a National Volcano Early Warning System*. U. S Geological Survey. Virginia: U. S. Department of the Interior.
- Hall, M., & Mothes, P. (2008). The rhyolitic-adesitic eruptive history of Cotopaxi volcano, Ecuador. *Springer Bulletin of Volcanology*, 70, 675-702. doi:<https://doi.org/10.1007/s00445-007-0161-2>
- Hall, M., Andrade, D., Mothes, P., Troncoso, P., Eissen, J.-P., Samaniego, P., . . . Yepes, H. (2005). Los peligros volcánicos asociados con el Cotopaxi. En Institut de Recherche Pour Le Développement (IRD), & IGEPN. Corporación Editorial Nacional.
- Ibáñez, J., & Carmona, E. (1997). *Sismicidad Volcánica*. Granada: Instituto Andaluz de Geofísica. Universidad de Granada.
- IGEPN. (2017). *Cotopaxi*. Obtenido de <http://www.igepn.edu.ec/>

- IGEPN. (2019). *Presentación*. Obtenido de Instituto Geofísico de la Escuela Politécnica Nacional:
<https://www.igepn.edu.ec/nosotros>
- Jaramillo Aranha, C. (2015). *Caracterización de señales sísmicas del volcán Cotopaxi utilizando estimadores espectrales clásicos y de máxima entropía*. Universidad de las Fuerzas Armadas ESPE, Departamento de Eléctrica y Electrónica, Sangolquí.
- Jiménez Tintaya, C. (2007). Procesamiento Digital de Señales con Matlab. *Revista de Investigación de Física*, 10(1605-7744), 45-50. Recuperado el 8 de Noviembre de 2019
- Jin, C., Zhang, Y., Balakrishnan, S., Wainwright, M., & Jordan, M. (2016). Local Maxima in the Likelihood of Gaussian Mixture Models: Structural Results and Algorithmic Consequences. *30th Conference on Neural Information Processing Systems (NIPS)*. Barcelona .
- Kinnunen, T., & Padmanabhan, R. (2013). A Practical, Self-adaptive Voice Activity Detector For Speaker Verification With Noisy Telephone And Microphone Data. *ICASSP*.
- Lara, R. C., Paillacho, V., & Villalva, M. (2017). Towards an automatic detection system of signals at cotopaxi volcano. *DYNA*, 4, 176-184.
- Lara, R., Carrera, V., Benítez, D., Ruiz, M., & Rojo, J. (2016). Automatic Recognition of Long Period Events From Volcano Tectonic Earthquakes at Cotopaxi Volcano. *IEEE Transactions on Geoscience and Remote Sensing*.
- Lara-Cueva, R., Moreno, A., Larco, J., & Benitez, D. (2016). Real-Time Seismic Event Detection Using Voice Activity Detection Techniques. *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, 1939-1404. Obtenido de http://www.ieee.org/publications_standards/publications/rights/index.html

- Li, X., Horaud, R., & Girin, L. (2016). Voice activity detection based on statistical likelihood ratio with adaptive thresholding. *IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, 1-5. doi:10.1109/IWAENC.2016.7602911
- MathWorks. (2018). *Speaker identification using pitch and MFCC*. Obtenido de MathWorks: <https://la.mathworks.com/help/audio/examples/speaker-identification-using-pitch-and-mfcc.html>
- Meza, I. V. (Marzo de 2013). MFCCs. (UNAM, Ed.) México. Obtenido de <https://turing.iimas.unam.mx/~ivanvladimir/posts/mfcc/>
- Mora , M., & Alvarado , G. E. (2001). Primer taller en actualización de Sismología Volcánica. *Red Sismológica Nacional (UCR-ICE)*. Costa Rica: Red Sismológica Nacional. Recuperado el 11 de Noviembre de 2019, de <https://rsn.ucr.ac.cr/documentos/educativos/vulcanologia/5099-que-es-un-tremor>
- Ong, W. Q., & Tan, W. C. (2016). Robust Voice Activity Detection Using Gammatone Filtering and Entropy. *Robotics, Automation and Sciences (ICORAS)* (págs. 25-36). IEEE.
- Patel, T., & Patil, H. (2015). Combining Evidences from Mel Cepstral, Cochlear Filter Cepstral and Instantaneous Frequency Features for Detection of Natural vs. Spoofed Speech. *Interspeech 2015*. Gujarat: Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT).
- Qing Ong, W., Chiat Tan , A. W., Vengadasalam, V., Heng Tan , C., & Hai Ooi, T. (2017). Real-time robust voice activity detection using the upper envelope weighted entropy measure and the dual-rate adaptive nonlinear filter. *Entropy*, 19, 487. doi:10.3390/e19110487
- Rodriguez, M., & Lara-Cueva, R. (2018). A Real-Time Microseisms-Detector based on Voice Activity Detection and Endpoint Detection: An Approach to Cotopaxi Volcano. (E. y.

- Departamento de Eléctrica, Ed.) *Maskay*(1390-6712). Recuperado el 18 de Noviembre de 2019, de <http://repositorio.espe.edu.ec/handle/21000/15565>
- Rosenberg, M. (22 de Diciembre de 2018). *Ring of Fire*. Obtenido de ThoughtCo: <https://www.thoughtco.com/ring-of-fire-1433460>
- Rubin, J., Abreu, R., Ganguli, A., Nelaturi, S., Matei, I., & Sricharan, K. (2016). Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients. *Computing in Cardiology Conference (CinC)*, (págs. 813-816). Vancouver. Obtenido de <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7868867&isnumber=7868653>
- Salazar, D., & D'Ercole , R. (2009). Percepción del riesgo asociado al volcán Cotopaxi y vulnerabilidad en el Valle de Los Chillos (Ecuador). *Bulletin de l'Institut Français d'Études Andines*, 38(3), 849-871. doi:10.4000/bifea.2522
- Tilling, R., & Beate, B. (1993). Los peligros volcánicos. Apuntes breves sobre un curso breve. *Organización Mundial de Observatorios Vulcanológicos*, 2-3.
- Vallejo Vargas, S. (Febrero de 2011). Distribución de cenizas volcánicas holocénicas - tardías en la costa del Ecuador. Quito, Pichincha, Ecuador: Escuela Politécnica Nacional.
- Viracucha, E., & De la Bastida, J. (2014). *Sistema Informático para el Procesamiento y Análisis de Señales Sísmicas de Volcanes en el Ecuador*. Quito: EPN.
- Werner-Allen, G., Johnson, J., Ruiz, M., Lees, J., & Welsh, M. (2005). Monitoring Volcanic Eruptions with a Wireless Sensor Network. *Proceedings of the Second European Workshop on Wireless Sensor Networks*. doi:Werner-Allen, G., Johnson, J., Ruiz, M., Lees, J., & Welsh, M. (n.d.). Monitoring volcanic eruptions with a wireless sensor network. Pro10.1109/ewsn.2005.1462003

Zou, K. H., O'Malley, J., & MauriMD, L. (2007). Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models. (H. M. School, Ed.) *Harvard Clinical Research Institute*, 654-657.
doi:<https://doi.org/10.1161/CIRCULATIONAHA.105.594929>