# UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
# FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

**TITULO DEL TRABAJO DE TITULACIÓN**

**APLICACIÓN DE TÉCNICAS DE IA Y ACÚSTICA FORENSE PARA IDENTIFICAR AUDIOS SINTÉTICOS**


**AUTOR**

ORTIZ ACOSTA, SULY YULEXY


**MODALIDAD DE TITULACIÓN**

EXAMEN COMPLEXIVO

Previo a la obtención del grado académico en
INGENIERO EN TECNOLOGÍAS DE LA INFORMACIÓN

**TUTOR**

**ING. LÍDICE HAZ LÓPEZ**

**Santa Elena, Ecuador**

**Año 2024**

UPSE

# UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
# FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

**TRIBUNAL DE SUSTENTACIÓN**

Ing. José Sánchez A. Mgt.
**DIRECTOR DE LA CARRERA**

Ing. Lidice Haz L. Mgt
**TUTOR**

Ing. Shendry Rosero V. Msc.
**DOCENTE ESPECIALISTA**

Ing. Mónica Jaramillo I. Mgt.
**DOCENTE GUÍA UIC**

# UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
# FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

## CERTIFICACIÓN

Certifico que luego de haber dirigido científica y técnicamente el desarrollo y estructura final del trabajo, este cumple y se ajusta a los estándares académicos, razón por el cual apruebo en todas sus partes el presente trabajo de titulación que fue realizado en su totalidad por **SULY YULEXY ORTIZ ACOSTA**, como requerimiento para la obtencióndel título de Ingeniero en Tecnologías de la Información.

La Libertad, a los 18 días del mes de junio del año 2024

**TUTOR**

Firmado electrónicamente por:
LIDICE VICTORIA HAZ LOPEZ

**ING. LIDICE VICTORIA HAZ LÓPEZ, Msi.**

# UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
# FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

**DECLARACIÓN DE RESPONSABILIDAD**

Yo, **Suly Yulexy Ortiz Acosta**

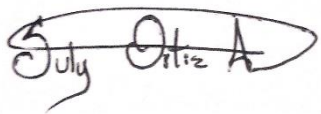
**DECLARO QUE:**

El trabajo de Titulación, **Aplicación de técnicas de IA y acústica forense para identificar audios sintéticos** previo a la obtención del título en Ingeniero en Tecnologías de la Información, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

La Libertad, a los 18 días del mes de junio del año 2024

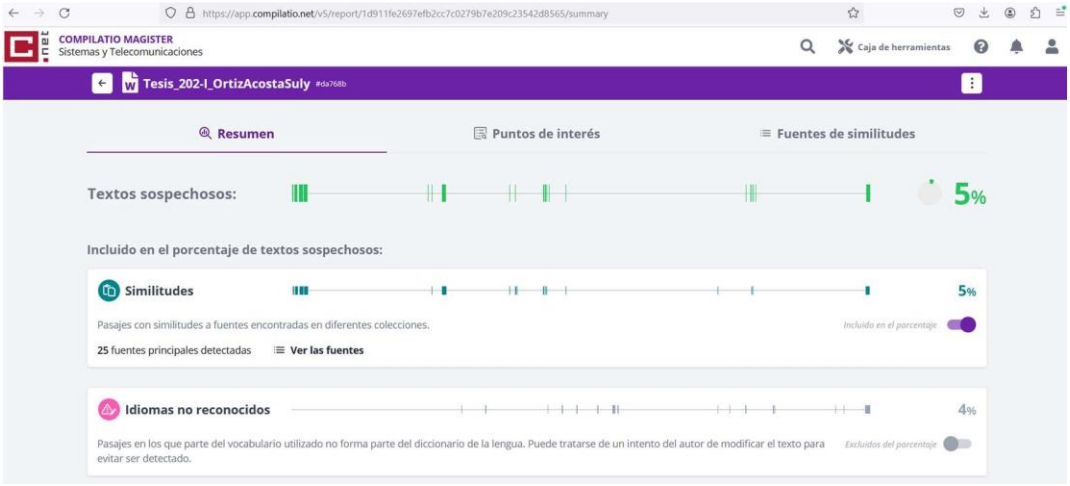
**EL AUTOR**



**Suly Yulexy Ortiz Acosta**

# UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
# FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

## CERTIFICACIÓN DE ANTIPLAGIO




Certifico que después de revisar el documento final del trabajo de titulación denominado "**Aplicación de técnicas de IA y acústica forense para identificar audios sintéticos",** presentado por el estudiante, **SULY YULEXY ORTIZ ACOSTA** fue enviado al Sistema Antiplagio, presentando un porcentaje de similitud correspondiente al 5%, por lo que se aprueba el trabajo para que continúe con el proceso de titulación.

**TUTOR**

Firmado electrónicamente por:
LIDICE VICTORIA HAZ LOPEZ

**ING. LIDICE VICTORIA HAZ LÓPEZ, Msi.**

# UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
# FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

## AUTORIZACIÓN

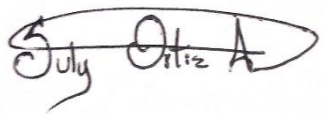
Yo, **Suly Yulexy Ortiz Acosta**

Autorizo a la Universidad Estatal Península de Santa Elena, para que haga de este trabajo de titulación o parte de él, un documento disponible para su lectura consulta y procesos de investigación, según las normas de la Institución.

Cedo los derechos en línea patrimoniales del trabajo de titulación con fines de difusión pública, dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor

La Libertad, a los 18 días del mes de junio del año 2024

## EL AUTOR



**Suly Yulexy Ortiz Acosta**

## AGRADECIMIENTO

Agradezco a Dios, por la salud y por permitirme esforzarme cada dia para finalizar mi trabajo de titulación, agradezco profundamente a mis padres, por todo el apoyo y la confianza que me han brindado, gracias a mis hermanos y a toda mi familia por los valores y conocimientos que han sido fundamental para mi crecimiento personal y profesional.

Tambien quiero expresar mi agradecimiento a mi tutora, Ing. Lídice Haz López por guiarme en este proceso tan importante.

Por último, pero no menos importante, a mi enamorado David Alejandro, por su apoyo incondicional.

*Suly Yulexy Ortiz Acosta*

**DEDICATORIA**

El presente trabajo se lo dedico a nuestro creador, cuya guía y bendiciones han sido mi fortaleza constante. A mi familia, especialmente a mis padres y hermanos, por ser mi inspiración, brindándome su amor y apoyo esencial.

*Suly Yulexy Ortiz Acosta*

INDICE GENERAL

INDICE DE TABLAS

INDICE DE FIGURAS

INDICE DE IMÁGENES

## RESUMEN

La nueva era digital permite la creación y difusión de audios sintéticos o también conocidos como “deepfakes”, este gran ímpetu marca un gran desafío a la seguridad informática al realizar el análisis correspondiente para identificar los audios manipulados. Para dicho enfoque se utilizó una metodología conocida en computación forense denominada UNE 71506:2015 que se conceptualiza como una metodología de pruebas experimentales para el análisis forense de evidencias digitales. Se desarrolló 3 casos de estudio experimental de audios clasificados en original, clonación IA y generación de voz y por ende se seleccionaron técnicas computacionales de forense para enmarcar el estudio y límite de investigación como; el análisis espectral, coeficientes cepstrales, entre otras. El objetivo del proyecto es elaborar un procedimiento aplicando técnicas de acústica forense como de IA para evaluar la autenticidad de los archivos de audios, adicionalmente, ejercer la documentación de los análisis de los audios como a su vez.



Palabras Claves: Computación Forense, Audios Sintéticos, Acústica Forense

## ABSTRACT

The new digital era allows the creation and dissemination of synthetic audios, also known as "deepfakes." This surge presents a significant challenge to computer security in performing the necessary analysis to identify manipulated audios. For this approach, a forensic computing methodology known as UNE 71506:2015 was used, which is conceptualized as a methodology of experimental tests for the forensic analysis of digital evidence. Three experimental study cases were developed with audios classified as original, AI cloning, and voice generation. Consequently, computational forensic techniques such as spectral analysis and cepstral coefficients were selected to frame the study and research limits. The project's objective is to develop a procedure applying forensic acoustics and AI techniques to evaluate the authenticity of audio files, and additionally, to document the audio analysis.



Keywords: Forensic Computing, Synthetic Audios, Forensic Acoustics

# INTRODUCCIÓN

La tecnología ha evolucionado drásticamente a pasos agigantados, permitiendo ejercer la creación de diversas herramientas tecnológicas avanzadas para clonar audios, editar videos, crear imágenes con herramientas de Inteligencia Artificial, una tecnología muy buena, pero a su vez muy negativa al uso que se le otorgue. La creación de audios sintéticos o denominados "deepfakes" permite que el audio este a un nivel de realismo impactante, se entiende que estos audios falsos pueden ser generados por la IA y presentar una gran amenaza a la seguridad información, privacidad y autenticidad de la información que pone en peligro la integridad de los datos que dobleguen la ley de protección de datos personales.

Para el estudio se adopta el uso de tres casos de investigación que permitan el análisis forense en bases a técnicas de acústica forense como lo es el análisis espectral, cepstrales, coeficiente MFCC, forma de onda mediante herramienta que permiten el análisis de los mismos archivos de audios, como también el uso de librerías de Python que ayudan a ejercer una visualización y obtención de datos relevantes de análisis en comparativa de las herramientas Open Source como Audacity y Spek.

Por otra parte, para el proyecto se utilizó la metodología UNE 71506:2015 de computación forense que trata de ejercer pruebas experimentales para el análisis de evidencias digitales. Por lo cual, los tres casos de estudio de selección de audio original, clonación por IA y generación por síntesis de voz permitirán encontrar comparativas y cambios esenciales a través de las técnicas a evaluar. Además, el proyecto también se centró en realizar la recolección de información de justificación estadísticos, temas relacionados a la investigación y literatura de referencias.

El presente proyecta está estructurado de la siguiente forma:

Capítulo I, en esta sección se detalla de manera importante todo lo referente de los antecedentes del proyecto que hace énfasis de la evolución tecnológica y herramientas

para la creación de audios sintéticos, también menciona la descripción del proyecto, los objetivos, la justificación del proyecto y alcance

Capitulo II, en esta sección se describe lo que viene hacer el marco conceptual en donde se exponen los conceptos en relación al tema, también habla del marco teórico que cuenta teorías o estudios relacionado de lo estudiado de técnicas forenses en audios sintéticos, la metodología de investigación y la desarrollo que dirección al proyecto

Capitulo IIII, en esta sección se expone todo lo propuesto del proyecto que viene hacer los componentes como el desarrollo de cada una de las fases de la metodología de UNE 71506:2015 y también lo que viene hacer el proceso de análisis de audios sintéticos. Además, presentar una documentación de los controles que son esenciales para ejercer el análisis de audios sintéticos.

Por ultima, están las conclusiones del proyecto, recomendaciones y los anexos que explican lo desarrollado del estudio investigativo

# CAPÍTULO I

# 1 FUNDAMENTACIÓN

## 1.1 ANTECENDENTES DEL PROYECTO

El rápido progreso de las tecnologías modernas ha ampliado la disponibilidad de herramientas digitales que en el pasado únicamente estaban al alcance de fuerzas militares o compañías cinematográficas, con recursos significativos, tenían el privilegio de utilizarlas a su conveniencia. Este es el caso de los "deepfakes", tambien conocidos como "falsificaciones profundas" en español, que surgen a partir de un reciente software basado en inteligencia artificial en el que, al igual que el Photoshop transformó la fotografía hace más de 30 años, este programa está generando un impacto significativo en el ámbito audiovisual de la actualidad [1].

A pesar de tener a disposición estas herramientas y sus continuas evoluciones a través del tiempo, asi como de los avances en informática, hasta el momento no se ha logrado encontrar una herramienta efectiva en el contexto legal que ayude a determinar la compatibilidad entre un registro de voz en tiempo real (voz viva) y las grabaciones de voz utilizadas como pruebas en un proceso judicial.

El proyecto tiene como principal enfoque brindar apoyo en las investigaciones a profesionales del ámbito judicial, como jueces, fiscales y peritos informáticos, quienes buscan esclarecer procesos legales mediante la utilización de herramientas de inteligencia artificial y técnicas de acústica forense.

Diversas entidades de investigación han querido abordar este problema, inicialmente desarrollando conjuntos de datasets que se puedan gestionar y comparar resultados, como los conjuntos de datos Imagenet o YOLO, que son utilizados en tareas como la Clasificación de Imágenes o la Detección de Objetos y estos intentan proporcionar una amplia variedad de DeepFakes generados con distintos métodos, con el objetivo de brindar una representación completa de los DeepFakes en situaciones reales [2]. El realismo creciente de los deepfakes está haciendo que sea cada vez más complicado

detectar y reconocer estos archivos, lo que a su vez conlleva riesgos de fraude y manipulación en situaciones como grabaciones de audio destinadas a difundir información falsa o como pruebas en investigaciones criminales, esta situación puede obstaculizar la toma de decisiones.

A continuación, se muestra una comparativa de los datasets más utilizados. [2]:

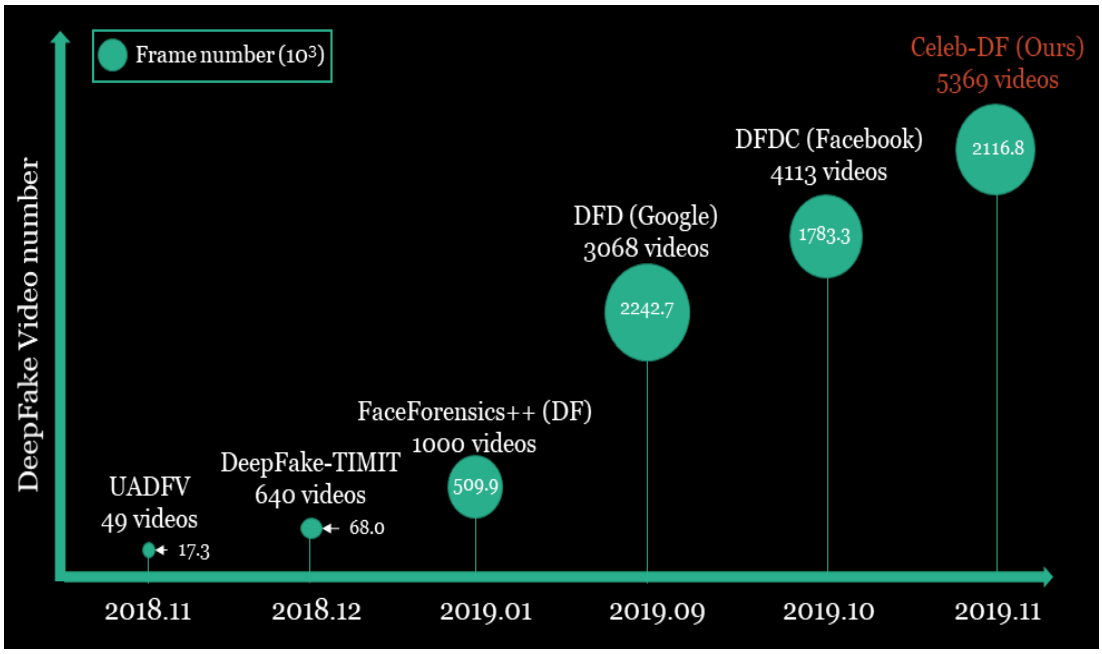



***Figura 1: Comparación Datasets de DeppFakes  Fuente:*** *[3]*

Por ejemplo, en el caso de DFDC (Facebook), se involucraron 4326 voluntarios para generar 25 terabytes de datos sin procesar, y utilizaron software ya existente como DeepFaceLab junto con diversos tipos de modelos de GANs y Autoencoders [2]. El uso de este software mencionado plantea una perspectiva negativa, dado que posibilita la creación de DeepFakes de manera sorprendentemente efectiva. Sin embargo, es importante destacar que no debe emplearse con intenciones criminales o maliciosas.

En un documento judicial que la revista Forbes descubrió se dio a conocer un caso de fraude de voz DeepFake en el que se llevó a cabo un ataque coordinado. Primero, los estafadores tomaron el control del correo electrónico, ya sea mediante el robo de Tokens OAuth o la suplantación de identidad. Luego, utilizaron la inteligencia Artificial para

hacer que la llamada telefónica fuera mucho más convincente [4], lo que desconocía era que había sido víctima de una sofisticada estafa en la que los estafadores emplearon tecnología de “voz profunda” para replicar el discurso del director [5].
El artículo de Changsha, China, se centra en la detección y reconocimiento de suplantación de voz, una preocupación clave para la verificación automática de hablantes. Aborda el desafío de generalizar métodos de detección de suplantación a nuevos algoritmos. Propone un sistema que integra la detección de suplantación de voz con el reconocimiento de algoritmos de suplantación, mejorando así la capacidad de identificar ataques. Este enfoque permite una detección más efectiva y la identificación de nuevas formas de suplantación de voz. La investigación busca fortalecer la seguridad en sistemas de verificación de hablantes frente a ataques fraudulentos. La combinación de detección y reconocimiento potencia la capacidad de adaptación del sistema frente a tácticas de suplantación emergentes. [6].

La investigación realizada en Buenos Aires, Argentina, titulada "Reconocimiento automático de hablantes, empleando técnicas de Deep Learning, en peritajes informáticos", se enfoca en desarrollar un prototipo basado en Inteligencia Artificial (IA) para abordar la problemática de la identificación de la autoría de registros de voz. El estudio lleva a cabo un análisis exhaustivo sobre el reconocimiento de voz, que incluye antecedentes, usos, aplicaciones y diversos enfoques y métodos utilizados en la identificación de hablantes. Su objetivo principal es contribuir al campo de la identificación de hablantes en el ámbito forense mediante el desarrollo de un prototipo basado en técnicas de inteligencia artificial. Este prototipo proporcionaría a los peritos informáticos una herramienta confiable para determinar la compatibilidad de los registros de voz y evitar impugnaciones técnicas en procesos judiciales. [7].

A nivel nacional, en Cuenca – Ecuador el documento titulado "Diseño e Implementación de un Sistema de Síntesis de Voz" presenta un proyecto relacionado con la generación de voz artificial. El sistema desarrollado tiene la capacidad de crear voz artificial a partir de un texto en formato ASCII, se ha diseñado de manera que sea adaptable en diferentes plataformas de hardware incluyendo microcontroladores y dispositivos lógicos

programables (PLD), con el objetivo de complementar tecnologías existentes en sistemas electrónicos [8].
Mediante la revisión bibliográfica se han encontrado artículo, tesis y estudios que abordan este tema en diferente contexto e incluso en procesos judiciales, pero estos se enfocan en la creación de herramientas para la creación de tipos de deepfakes.

Con base a lo expuesto el proyecto se enfoca en la creación de un procedimiento detallado que implica la realización de pruebas experimentales y el seguimiento de pasos específicos, empleando avanzadas herramientas de inteligencia artificial y técnicas de acústica forense. Es fundamental llevar a cabo un minucioso análisis para la identificación de audios sintéticos, con el propósito de ofrecer una guía sólida que permita abordar de manera efectiva la evaluación de la autenticidad e integridad de grabaciones de voz. De esta manera, se busca proporcionar a los investigadores y profesionales una herramienta valiosa para asegurar la veracidad de la información y la identificación precisa de las voces involucradas en diversos contextos.

## 1.2 DESCRIPCIÓN DEL PROYECTO

Una de las modalidades más recientes de desinformación llamadas deepfakes, se han convertido en un verdadero desafío dentro del entorno comunicativo debido a su propagación a través de las noticias en línea y los espacios de las redes sociales. Aunque las noticias falsas han existido hace siglos, su impacto perjudicial se ha acentuado en la actualidad dado a su amplio alcance a herramientas de producción y difusión [9]. Debido a lo expuesto anteriormente, se propone realizar un análisis mediante la investigación de diversas fuentes y además realizar un procedimiento informático forense que describa los pasos a seguir para evaluar la integridad y autenticidad del archivo de audio.

El proyecto se enmarca en la identificación de archivos de audios falsos también, denominados deepfake de voz, o audios sintéticos. El objetivo es utilizar técnicas de inteligencia artificial (IA) y herramientas de acústica forense para evaluar la integridad y veracidad de archivos de audios en formato (.mp3, wav, y .ogg). Para lo cual, se plantea

un marco teórico detallando conceptos relacionados con el funcionamiento de los audios sintéticos, y los principios de la acústica forense, para lograr comprender las métricas que identifican la autenticidad e integridad de las grabaciones de audios. El estudio se desarrolla usando la metodología de investigación exploratoria, descriptiva y experimental; para lo cual, se diseñan pruebas experimentales en un laboratorio virtual donde se automatiza el análisis de las muestras reales de audios; es decir, la comparación entre la voz real y la voz sintética.

En el proyecto, se empleará la metodología UNE 71506/2013 como enfoque para llevar a cabo las pruebas experimentales que está compuesta por 5 fases [10].

- **Fase 1: PRESERVACIÓN**

  - ✓ Verificar la autenticidad del audio mediante el uso del algoritmo hash para los diferentes casos de generación de audios deepfake.
  - ✓ Revisar la información contenida en los metadatos de cada archivo de audio para identificar y comprender los datos asociados.

- **Fase 2: ADQUSICIÓN**

  - ✓ Investigar las diversas técnicas e instrumentos para identificar audios sintéticos.
  - ✓ Agrupar las diferentes técnicas de procesamiento de datos para la identificación de audios sintéticos.
  - ✓ Preparar los ambientes virtuales de trabajo para utilizar las técnicas de procesamiento de datos.

- **Fase 3: ANÁLISIS**

  - ✓ Examinar las grabaciones de audios a través de diferentes herramientas informáticas que utilicen técnicas forenses.

- ✓ Emplear un cuadro descriptivo que detalle la recopilación de información de las pruebas de experimentación.
- ✓ Realizar observaciones técnicas.

- **Fase 4: DOCUMENTACIÓN**
  - ✓ Introducción
  - ✓ Objetivo
  - ✓ Alcance
  - ✓ Descripción de controles de autenticidad de audios
  - ✓ Recomendaciones

En esta etapa, se llevará un registro detallado de todo el proceso, desde el inicio del análisis hasta la entrega del informe pericial al solicitante. Se documentará minuciosamente todos los procedimientos realizados y las herramientas empleadas, siguiendo un orden temporal preciso.

A través de este estudio, contribuirá al campo de investigación en tecnologías y sistemas de información y comunicación, específicamente dentro de la sub-línea de investigación en redes y seguridad de la información, explorando y analizando enfoques alternativos para abordar los desafíos relacionados con la seguridad de la información en el contexto de la ciberseguridad, análisis forense informático, auditoria de redes informáticas, gestión de riesgos tecnológicos, la implementación de plataformas e infraestructura de seguridad informática, asi como la seguridad dentro de las infraestructuras en la nube y evaluación de datos [11].

## 1.3 OBJETIVOS DEL PROYECTO

### 1.3.1 OBJETIVO GENERAL

Diseñar un procedimiento aplicando técnicas de IA y acústica forense para evaluar la veracidad e integridad de los archivos de audios

### 1.3.2 OBJETIVOS ESPECÍFICOS

- Recopilar un conjunto de audios sintéticos para uso del estudio.
- Realizar pruebas experimentales para analizar audios sintéticos utilizando técnicas de IA y acústica forense.
- Evaluar los resultados de la experimentación mediante el análisis de las métricas del estudio forense.

## 1.4 JUSTIFICACIÓN DEL PROYECTO

El crecimiento de las tecnologías de medios sintéticos y deepfakes nos está llevando a comprender algo de gran relevancia y al mismo tiempo preocupante: nuestra larga creencia de que los videos y el audio son testimonios confiables de la realidad ya no es sostenible [12]. Paradójicamente, en la sociedad actual, a pesar de tener un acceso amplio a la información, nos encontramos en un estado de desinformación más pronunciado que nunca. La vulnerabilidad no hace excepciones por edad ni se limita a contextos políticos o mediáticos específicos. Las noticias falsas o fake news, que promueven el engaño y la desinformación, están ejerciendo una influencia deliberada y desconcertante sobre la población [13].

La encuesta presentada por Jumio, una empresa, abordó la detección de deepfakes a través de un estudio global que revela la falta de entendimiento entre la inteligencia artificial generativa y los deepfakes, y el riesgo que estas tecnologías representan para la seguridad en línea. Los resultados de la investigación de este año destacan como los consumidores comprenden que tanto la inteligencia artificial como las tecnologías deepfake pueden incrementar el riesgo de robo de identidad, lo que crea la necesidad de contar con identidades digitales para verificar y autenticar en línea [14].

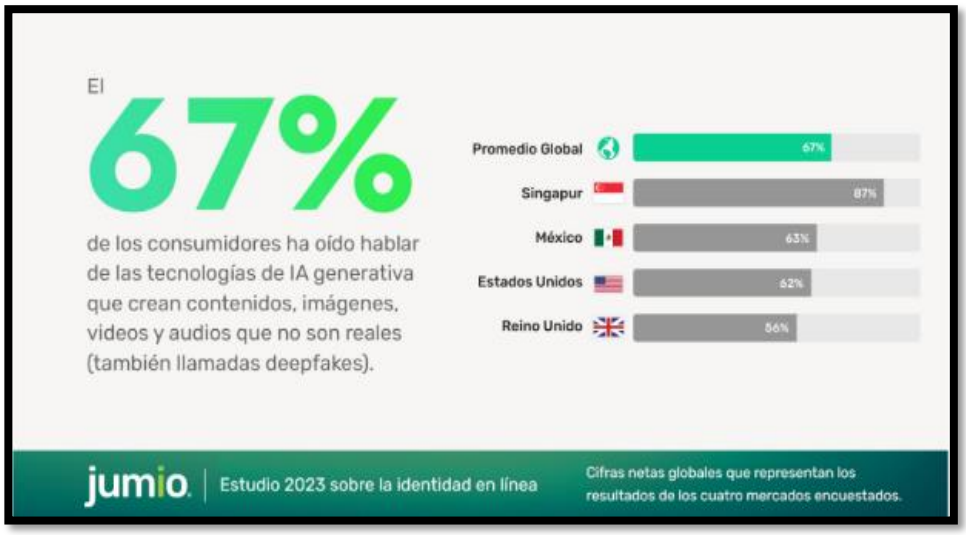



**Figura 2: Estudio 2023 de identidad en línea – Fuente:** *[14]*

Mas del 67% de las personas encuestadas aseguran estar familiarizadas con las herramientas de inteligencia artificial generativa, como ChatGPT, DALL-E y Lensa AI, las cuales tienen la capacidad de crear contenido ficticio, como videos, imágenes y audio. Estos datos tambien muestran un aumento constante en el uso de tecnologías deepfake cada vez más sofisticadas en todo el mundo y en todos los sectores, con una presencia mayor en los sectores de pagos y criptomonedas [14].

El reconocimiento de esta nueva amenaza surgida en el uso de las plataformas de redes sociales digitales causa que el estudio de esta herramienta, conocida como “deepfake” sea relevante para la sociedad a nivel global, ya que tiene el potencial de ejercer un impacto negativo en las opiniones de los usuarios en línea y, por consiguiente, en la sociedad en su conjunto a nivel global. Además, se destaca la necesidad de establecer regulaciones legales para lograr controlar esta práctica, debido a la amenaza que representa para la sociedad, incluso en relación con los delitos derivados de aquellos sujetos que utilizan deepfakes en el entorno digital [15].

En este estudio de investigación, se busca divulgar de manera publica este innovador método de engaño o confusión de personas, al mismo tiempo que se intenta establecer la

capacidad de determinar la autenticidad de un archivo de audio en situaciones donde exista incertidumbre sobre su veracidad. Además, se pretende emplear técnicas forenses e informáticas ampliamente reconocidas en el proceso de análisis, garantizando asi la confiabilidad del estudio.

Al desarrollar un procedimiento forense detallando el paso a paso servirá como una herramienta relevante para investigaciones judiciales y forenses, permitiendo la autenticación de audios y a la vez prevenir el uso malicioso de deepfakes. Además, contribuye al avance de la tecnología de detección de audios falsos, con posibles aplicaciones en otros campos de la ciberseguridad y la inteligencia artificial. Su divulgación aumenta la concienciación del público sobre las amenazas de los deepfakes y previene la propagación de información falsa.

El proyecto a realizar se centra en el Plan De Creación de Oportunidades en los siguientes ejes, objetivos y políticas:

**Eje 1:** Eje económico

**Objetivo 4:** Garantizar la gestión de las finanzas públicas de manera sostenible y transparente [16].

**Política 4.1:** Priorizar el gasto público para la atención en salud, educación, seguridad, con enfoque en los derechos humanos [16].

**Eje 2:** Eje Social

**Objetivo 7**: Potenciar las capacidades de la ciudadanía y promover una educación innovadora, inclusiva y de calidad en todos los niveles [16].

**Política 7.2:** Promover la modernización y eficiencia del modelo educativo por medio de la innovación y el uso de herramientas informáticas [16].

**Eje 3:** Eje Seguridad Integral

**Objetivo 10:** Garantizar la soberanía nacional, integridad territorial y seguridad del Estado [16].

**Política 10.1:** Fortalecer al estado para mantener la confiabilidad, integridad y disponibilidad de la información frente a amenazas provenientes del ciberespacio y proteger su infraestructura critica [16].

## 1.5 ALCANCE

En el mundo digital actual, la capacidad de crear audios sintéticos con alta calidad y realismo ha aumentado considerablemente, lo que plantea desafíos significativos en términos de autenticidad y confiabilidad de la información. Estos audios sintéticos pueden ser utilizados con diversos propósitos, desde la generación de contenido multimedia hasta la manipulación maliciosa de información, lo que podría tener graves implicaciones en campos como la seguridad, la justicia y la reputación de las personas o instituciones.

Por tanto, la investigación busca abordar esta problemática mediante el desarrollo y aplicación de técnicas de IA y acústica forense para identificar audios sintéticos. Para lograrlo, se plantea la necesidad de crear un conjunto de datos de audios sintéticos que representen una variedad de escenarios y condiciones de grabación. Estos datos serán generados utilizando herramientas como el Generador de texto a voz y la Clonación de voz, las cuales permiten crear audios sintéticos con diferentes niveles de realismo y calidad.

Posteriormente, se llevarán a cabo pruebas exhaustivas utilizando este conjunto de datos, utilizando técnicas de IA y acústica forense para analizar y comparar los audios sintéticos con grabaciones reales. Esto permitirá desarrollar y validar algoritmos y modelos que puedan identificar de manera efectiva la presencia de audios sintéticos y distinguirlos de los audios genuinos.

En cuanto al contexto de escenarios a simular, es importante considerar una amplia gama de situaciones en las que los audios sintéticos podrían ser utilizados, como llamadas telefónicas, mensajes de voz, entrevistas, discursos políticos, entre otros. Cada escenario podría presentar desafíos únicos en términos de ruido de fondo, calidad del audio y características lingüísticas, por lo que es crucial tener en cuenta esta diversidad al generar los datos y diseñar las pruebas.

Para llevar a cabo este estudio, se pueden utilizar máquinas virtuales que permitan simular diferentes configuraciones de hardware y entornos de grabación. Esto proporcionaría un entorno controlado y reproducible para realizar las pruebas, además de facilitar la escalabilidad y la distribución del trabajo en el desarrollo y análisis de los algoritmos.

Para garantizar la coherencia y el rigor en el desarrollo de la investigación, se ha seleccionado cuidadosamente una metodología que se ajuste a las necesidades del estudio. En este sentido, se ha optado por la metodología UNE 71506:2013, la cual proporciona un marco estructurado y detallado para la realización de análisis forenses en el ámbito digital, se ajusta la metodología hasta la fase 1 de preservación y fase 4 de documentación.

**Fase 1: Preservación -** Durante esta etapa inicial, se verificará la autenticidad de los audios utilizando el algoritmo hash, una técnica que proporciona una huella digital única para cada archivo. Además, se revisarán meticulosamente los metadatos de cada audio para comprender mejor los datos asociados, lo que garantizará la integridad y confiabilidad de la evidencia durante todo el proceso.

**Fase 2: Adquisición -** Una vez que se ha preservado la integridad de la evidencia, se procederá a adquirirla. Esto incluye investigar diversas técnicas e instrumentos para identificar audios sintéticos y seleccionar las más adecuadas para el procesamiento de datos. Se recolectará un conjunto de datos representativo y se almacenará en un repositorio local para su uso posterior en las pruebas y análisis.

**Fase 3: Análisis -** En esta fase, se definirán los ambientes de trabajo necesarios y se combinarán herramientas informáticas y forenses para examinar detalladamente las grabaciones de los audios. Se realizarán experimentos exhaustivos y se describirán en detalle cada caso en una tabla, lo que permitirá una evaluación completa de los resultados obtenidos.

**Fase 4: Documentación -** Finalmente, se documentarán todas las herramientas utilizadas en el proceso judicial, asegurando así la transparencia y la reproducibilidad del procedimiento. Se elaborará un cuadro detallando las vulnerabilidades relacionadas con la seguridad informática identificadas durante el análisis, lo que proporcionará información valiosa para futuras investigaciones y acciones de seguridad.

# CAPÍTULO II

# 2 MARCO TEÓRICO Y METODOLOGÍA DEL PROYECTO

## 2.1 MARCO CONCEPTUAL

### 2.1.1 ACUSTICA FORENSE

Se conoce como acústica forense como una parte criminalística que enmarca toda la índole de aplicaciones de técnicas para el desarrollo de la ingeniería acústica para elucidar a los delitos emitidos y hallar a los causantes de la misma. Esta área combina principios de física acústica como la ingeniería de sonido, psico acústica y análisis de señales para examinar y evaluar las grabaciones de audio en contextos legales [17]. En la Acústica Forense, se aplican técnicas avanzadas de análisis de audio para examinar y evaluar evidencia sonora con el objetivo de proporcionar pruebas sólidas en investigaciones criminales y procesos legales. Esta disciplina combina el rigor científico de la acústica con la meticulosidad de la investigación criminal, desempeñando un papel crucial en la búsqueda de la verdad y la justicia [18].

#### 2.1.1.1 FONACIÓN

La voz humana se caracteriza por ser una serie de variaciones rápidas en la presión del aire, producidas por el complejo sistema del aparato fonador humano. Este aparato, compuesto por diversos órganos, desempeña un papel fundamental en la emisión y modulación de sonidos, lo que permite la comunicación verbal y la expresión de emociones de manera efectiva. El estudio de estos procesos es fundamental en campos como la lingüística, la fonética y la comunicación humana [19]


**Imagen 1:TRACTO VOCAL– Fuente** *[20]*

#### 2.1.1.2 AUDICION

La audición se refiere a la habilidad de detectar sonidos, lo que implica la capacidad de percibir vibraciones en el ambiente y convertirlas en información interpretada por nuestro sistema auditivo. Esta capacidad nos permite interactuar con nuestro entorno y comunicarnos con otros individuos de manera efectiva. A través de un proceso complejo de decodificación, nuestro cerebro interpreta estas señales sonoras, lo que nos permite comprender el mundo que nos rodea y participar en interacciones sociales significativas [21].

- **Oído Externo:** La parte externa del oído, conocida como aurícula u oreja, es la porción visible y reconocible que recoge las ondas sonoras del entorno. Además de su función de captar el sonido, esta estructura también actúa como una barrera protectora para el oído medio, resguardándolo de influencias externas. De esta manera, la aurícula no solo recoge las ondas de sonido, sino que también cumple un papel crucial en la preservación y seguridad del sistema auditivo interno [21].
- **Oído Medio:** El tímpano, también llamado membrana timpánica, es una estructura que amplifica y transmite los sonidos desde la aurícula hasta el oído interno. Este componente esencial del sistema auditivo está compuesto por tres diminutos huesos conocidos como los "huesecillos del oído": el martillo, el yunque y el estribo. Estos huesos trabajan en conjunto para llevar las vibraciones del tímpano al oído interno, donde se procesa la información auditiva antes de ser enviada al cerebro para su interpretación [21].
- **Oído Interno:** La cóclea, también llamada caracol debido a su forma característica, despliega la importante función de convertir las ondas sonoras en señales eléctricas, las cuales son luego enviadas al cerebro para su interpretación. Este órgano esencial del oído interno juega un papel fundamental en la percepción auditiva, siendo responsable de transformar la energía acústica en impulsos nerviosos que son procesados y comprendidos por el cerebro [21].

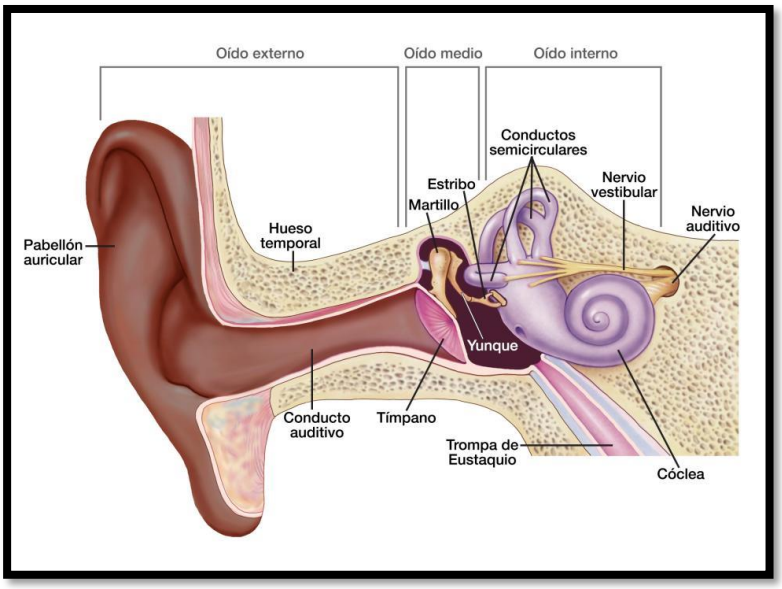



**Imagen 2: LA AUDICIÓN - Fuente:** *[22]*

### 2.1.2 INFORMÁTICA FORENSE

La informática forense comprende un conjunto de métodos y técnicas sistemáticas destinados a identificar, recopilar, preservar, extraer, interpretar, documentar y presentar evidencias digitales de equipos informáticos. Estas evidencias deben ser admisibles en procedimientos legales o administrativos en tribunales. Este campo es crucial para investigar delitos digitales, centrándose en crímenes cometidos a través de dispositivos informáticos como redes, ordenadores y medios de almacenamiento digital, donde la tecnología actúa como fuente o víctima del delito [23].

Con la creciente digitalización del mundo, la informática forense se vuelve cada vez más importante. La resolución de delitos cibernéticos y la recuperación de datos importantes o comprometidos dependen de la gestión de evidencias digitales. El trabajo de un investigador de informática forense consiste en recopilar, analizar y proteger tales pruebas [24].

#### 2.1.2.1 TIPOS DE INFORMATICA FORENSE

- **SISTEMAS OPERATIVOS**

  El proceso implica recopilar datos importantes del sistema operativo de un dispositivo. La informática forense en sistemas operativos tiene como objetivo obtener evidencia que se pueda utilizar contra el ciberdelincuente [25].

- **REDES**

  Su trabajo consiste en analizar, recopilar y rastrear la actividad en una red para identificar y prevenir tráfico sospechoso, así como para determinar las fuentes de ataques o virus [26].

- **BASE DE DATOS**

  Recuperación y análisis de datos y metadatos presentes en bases de datos. Este proceso es esencial para extraer información crítica y realizar investigaciones precisas en entornos de datos estructurados [27].

- **CORREO ELECTRÓNICO**

  El encabezado, que contiene información crucial sobre el mensaje, es la principal evidencia en las investigaciones por correo electrónico. Dado que la información del remitente está en la parte inferior y la del receptor en la parte superior, el análisis del encabezado debe hacerse de abajo hacia arriba. El encabezado también muestra el viaje del correo a través de varios MTA [28].

- **DISPOSITIVOS MOVILES**

  El análisis de teléfonos celulares puede realizarse tanto de forma activa como remota, lo que es útil en casos de sospecha de mala conducta o interferencia de los empleados. En situaciones críticas, es esencial contar con un experto que pueda actuar como testigo y presentar las interpretaciones de los datos ante las

autoridades. Nuestros expertos cuentan con la formación, certificaciones, experiencia y conocimientos especializados necesarios para llevar a cabo estas investigaciones [29].

- **NUBE O CLOUD**

  La computación en la nube, también conocida como computación en la nube, tiene el potencial de convertirse en una de las tecnologías informáticas más transformadoras, junto con otras tecnologías como computadoras centrales, portátiles, internet y teléfonos inteligentes. La forma en que se crean, entregan, acceden y gestionan los servicios de TI está cambiando drásticamente debido a esta tecnología [30].

- **MEMORIA**

  Los datos del disco duro pueden borrarse completamente o no dejar rastro alguno durante ataques sofisticados, lo que dificulta la obtención de pruebas para una investigación forense. El análisis forense de la memoria busca artefactos potenciales en la memoria RAM de una computadora [31].

- **MALWARE**

  Incluye el análisis de programas maliciosos para determinar su comportamiento, origen y método de propagación. Para crear estrategias de defensa y mitigación efectivas contra amenazas cibernéticas, este análisis es esencial [32].

#### 2.1.2.2 INFORMATICA FORENSE VS INFORMATICA ANTI-FORENSE

**INFORMATICA FORENSE**

La ciberseguridad está estrechamente relacionada con la informática forense. La informática forense puede ayudar a los equipos de ciberseguridad a detectar y resolver amenazas cibernéticas rápidamente y prevenir futuros ataques. La Forense Digital y Respuesta ante Incidentes (DFIR), una disciplina emergente en ciberseguridad, combina

actividades forenses informáticas con la respuesta ante incidentes para acelerar la mitigación de amenazas y garantizar la integridad de la evidencia digital [33].

**INFORMATICA ANTI-FORENSE**

Debido a la sofisticación de las técnicas de evasión, los analistas forenses enfrentan desafíos cada vez mayores en el rastreo y detección de atacantes. Las técnicas antiforense son métodos destinados a limitar la disponibilidad de evidencia durante un proceso forense después de que se descubrieron fallas en las herramientas forenses. Utilizando la destrucción, el ocultamiento, la eliminación o la falsificación, estas técnicas buscan manipular los datos [34].

### 2.1.3 TECNICAS TRADICIONALES DE ANALISIS DE AUDIO

**Análisis Espectral:** El análisis espectral implica descomponer las señales en componentes sinusoidales, donde cada componente representa una oscilación a una frecuencia específica con una amplitud determinada. Este enfoque proporciona una representación detallada de la información en el dominio de la frecuencia, permitiendo visualizar la contribución de cada componente de frecuencia a una señal. Es una herramienta fundamental en el análisis de audio, ya que permite identificar características importantes de la señal, como picos de frecuencia, patrones armónicos y componentes de ruido. Además, el análisis espectral es ampliamente utilizado en diversas áreas, incluyendo la ingeniería de sonido, la música, la comunicación y la acústica forense [35].

**Comparación de forma de onda:** Los métodos de comparación de formas de onda son esenciales para verificar la funcionalidad, el tiempo y la cobertura del diseño RTL. Sin embargo, la tarea puede volverse desafiante cuando las formas de onda son complejas y dinámicas, lo que implica la presencia de múltiples señales, estados y transiciones que dependen de diversas entradas y condiciones. En este artículo, exploraremos estrategias y técnicas para abordar este desafío, garantizando resultados de comparación precisos y eficientes en entornos de diseño RTL con formas de onda complejas y cambiantes [36].

**Análisis envolvente de datos:** Según estudios como el titulado "Método de análisis envolvente de datos y redes neuronales en la evaluación y predicción de la eficiencia técnica de pequeñas empresas exportadoras", el análisis envolvente de datos ha demostrado ser una herramienta efectiva en la industria. En este artículo, los autores han propuesto un método para evaluar y predecir la eficiencia utilizando el análisis envolvente de datos en el contexto de pequeñas empresas exportadoras. Este método enfatiza el uso y la flexibilidad del DEA en la evaluación y optimización de procesos comerciales, brindando una herramienta útil para la mejora continua y la toma de decisiones basada en datos [37].

**Análisis de Fase:** El análisis de fase, que se basa en la observación de las vibraciones de las máquinas, es una herramienta fundamental para monitorear la condición de las máquinas. Es responsable de evaluar la sincronización temporal de las señales de vibración, que normalmente provienen de diferentes partes de la máquina. Este método es esencial para identificar problemas como desalineaciones, desequilibrios, holguras en las piezas o fallas en los engranajes. La presente guía discutirá los fundamentos del análisis de fase, cómo es importante para el mantenimiento predictivo, las técnicas de medición relacionadas y cómo ayuda a identificar fallas tempranas en equipos industriales [38].

**Análisis de espectrograma:** El espectrograma es una representación gráfica que visualiza las variaciones en la frecuencia y la intensidad del sonido a lo largo del tiempo. Para analizarlo, se emplea el análisis espectral, el cual se centra en descomponer sonidos o fragmentos musicales en múltiples componentes. Este análisis requiere la capacidad de distinguir y localizar sonidos con frecuencias cercanas entre sí. Es importante tener en cuenta que los parámetros de los componentes de los sonidos reales son variables y cambian según su duración. Además, cada nota musical, ya sea alta o baja, tiene una frecuencia específica, medida en Hertz (Hz) o kilohertzio (1.000 Hz) [39].

### 2.1.4 INTELIGENCIA ARTIFICIAL

#### 2.1.4.1 AUDIO DIGITAL

Un archivo de audio digital es una representación digital de una señal eléctrica en forma de onda sonora. En la actualidad, hay una amplia variedad de programas de edición de

audio disponibles, tanto gratuitos como de pago, que hacen más fácil editar grabaciones de audio en formato digital. Estas herramientas se utilizan con frecuencia para realizar ajustes, agregar efectos, mejorar la calidad o convertir un archivo a varios formatos. El uso de técnicas de narración y el cumplimiento de ciertas normas son necesarios para producir contenido auditivo de calidad [40].

Dado que el sonido se propaga en forma de onda longitudinal a través de cualquiera de estos estados de la materia, las ondas sonoras son el resultado de la vibración de las moléculas en un medio, que puede ser sólido, líquido o gaseoso. Puede propagarse como una onda transversal en medios sólidos. Una vibración que se mueve es lo que se conoce como sonido. Estas vibraciones se propagan en el espacio a través de un medio elástico después de que un cuerpo vibrante se active [41].

#### 2.1.4.2 CLONACIÓN DE VOZ

La tecnología de síntesis del habla ha permitido la creación de réplicas digitales de voz con un alto nivel de realismo en el campo de la inteligencia artificial. La capacidad de replicar tu propia voz a través de la IA es una aplicación destacada de esta tecnología, lo que abre un amplio abanico de posibilidades tanto en el ámbito personal como profesional. Esta guía detallada analizará las ventajas y desventajas de clonar tu voz con inteligencia artificial [42].

#### 2.1.4.3 PELIGROS ASOCIADOS CON LA CLONACIÓN DE VOZ POR IA

**Fraude y extorsiones:** Clonación de voz para cometer fraudes y extorsiones. Por ejemplo, pueden realizar llamadas telefónicas simulando ser familiares o colegas y solicitar dinero o información confidencial de manera fraudulenta [43].

**Suplantación de identidad:** La clonación de voz permite a los agresores imitar la voz de una persona real en llamadas telefónicas, mensajes de voz o grabaciones de video [43].

**Manipulación de contenido multimedia:** Los delincuentes cibernéticos pueden usar fragmentos de voz de videos de dominio público, como los de TikTok, YouTube e Instagram, para crear mensajes falsificados [43].

**Impacto en la seguridad nacional:** La utilización de inteligencia artificial para la clonación de voces también genera inquietudes en el ámbito de la seguridad nacional. Esta tecnología podría ser empleada para fabricar discursos falsos de líderes políticos o figuras de autoridad, lo que podría desencadenar confusiones, conflictos o incluso crisis diplomáticas. Además, podría ser utilizada para generar información engañosa en mensajes de audio que podrían afectar la estabilidad y la seguridad de un país o una región [43].

## 2.2 HERRAMIENTAS

- **AUDACITY**

  Audacity es una aplicación de grabación y edición de sonido que destaca por su facilidad de uso. Es un programa de software gratuito y de código abierto. Es capaz de grabar, reproducir, importar y exportar archivos de sonido en una variedad de formatos, incluidos WAV y MP3. Tiene herramientas de edición como cortar, copiar y pegar, así como la capacidad de trabajar simultáneamente con varias pistas de audio, mezclarlas y aplicar una variedad de efectos sonoros [44].

- **SPEK**

  Spek es una herramienta de análisis de audio que, gracias a las bibliotecas de FFmpeg, soporta todos los formatos de archivos de audio populares, tanto con pérdidas como sin pérdidas. Para acelerar el análisis, proporciona procesamiento de señal ultrarrápido con múltiples subprocesos. El nombre del códec y los parámetros de la señal de audio se muestran en Spek, que también permite guardar el espectrograma como un archivo de imagen. Permite el arrastrar y soltar y funciona con formatos de archivo de audio comunes. Además, se ajusta automáticamente a las reglas de frecuencia, densidad espectral y tiempo, lo que le

permite cambiar el rango de densidad espectral. Varios idiomas lo han traducido [45].

- **PYTHON**

  Python es un lenguaje de programación de alto nivel conocido por su simplicidad y versatilidad. Ampliamente utilizado en el desarrollo de software y aplicaciones, ofrece una sintaxis clara y legible que lo hace ideal tanto para principiantes como para expertos. Su amplia gama de bibliotecas y su comunidad activa lo convierten en una opción popular para una variedad de proyectos, desde aplicaciones web hasta análisis de datos y aprendizaje automático [46].

- **CAINE**

  CAINE (Computer Aided INvestigative Environment): CAINE es una distribución de Linux basada en Ubuntu que se centra en proporcionar un entorno completo para realizar investigaciones forenses digitales. Viene preinstalada con una amplia gama de herramientas forenses, incluyendo herramientas para análisis de audio [47].

- **KALI LINUX**

  Kali Linux es una distribución de Linux basada en Debian que se utiliza ampliamente en seguridad informática y pruebas de penetración. Viene preinstalada con una amplia gama de herramientas forenses y de seguridad, algunas de las cuales podrían ser útiles para el análisis de audio forense [48].

- **NUMPY**

  NumerPy es una biblioteca de Python esencial para la informática científica. Proporciona soporte para arrays y matrices multidimensionales, así como una colección de funciones matemáticas para operar de manera efectiva con estos arrays. Gracias a su implementación en C, NumPy es ampliamente utilizado por su capacidad para realizar operaciones matemáticas y estadísticas de alto

rendimiento. Además, es la base sobre la cual se construyen muchas otras bibliotecas científicas de Python, como SciPy, Pandas y scikit-learn [49].

- **MATPLOTLIB**

Matplotlib es una biblioteca de Python que se puede usar para crear visualizaciones interactivas, animadas y estáticas. Se utiliza ampliamente en la comunidad científica y de ingeniería porque puede producir gráficos en 2D de alta calidad. Matplotlib puede crear una variedad de gráficos, incluidas barras, dispersiones, líneas, histogramas y gráficos de sectores. La flexibilidad y la extensibilidad de su diseño le permiten ajustar prácticamente todos los aspectos de un gráfico. Además, se integra bien con otras bibliotecas científicas de Python como NumPy, SciPy y Pandas, lo que facilita la creación de visualizaciones a partir de datos complejos [50].

- **LIBROSA**

La biblioteca Python Librosa se especializa en el análisis y procesamiento de señales de audio. Es ampliamente utilizado en aplicaciones de ingeniería de audio y música, así como en investigación de audio y música. Para tareas como la carga y visualización de archivos de audio y la extracción de características como espectrograma, coeficientes de mel-frecuencia cepstral (MFCC), tiempo, tono y ritmo, Librosa ofrece una amplia gama de herramientas. Además, facilita la manipulación del sonido, como la velocidad y el tono. La biblioteca es una excelente opción para el análisis de datos de audio porque funciona bien con otras herramientas del ecosistema científico de Python, como NumPy y SciPy [51].

- **EXIFTOOL**

ExifTool es un conjunto personalizable de módulos Perl, junto con una aplicación completa de línea de comandos llamada exiftool, utilizada para leer y escribir metadatos en una amplia variedad de archivos. Esto incluye la información específica de muchos fabricantes de cámaras digitales, como Canon, Casio, DJI,

FLIR, FujiFilm, GE, HP, JVC/Victor, Kodak, Leaf, Minolta/Konica-Minolta, Nikon, Nintendo, Olympus/Epson, Panasonic/Leica, Pentax/Asahi, Phase One, Reconyx, Ricoh, Samsung, Sanyo, Sigma/Foveon y Sony [52].

- **GTKHASH**

  La extensión GtkHash para caja permite a los usuarios usar la biblioteca mhash para calcular resúmenes de mensajes o sumas de verificación. MD5, MD6, SHA1, SHA256, SHA512, RIPEMD, TIGER y WHIRLPOOL son funciones hash actualmente soportadas [53].

- **PRATT**

  Es una herramienta gratuita y de código abierto para el análisis, síntesis y manipulación del habla, desarrollada por Paul Boersma y David Weenink en la Universidad de Ámsterdam. Es utilizada ampliamente en los campos de la lingüística y la fonética para el análisis detallado de las características acústicas del habla [54].

## 2.3 MARCO TEÓRICO

### 2.3.1 UN ESTUDIO DE LINGÜÍSTICA APLICADA: EL RECONOCIMIENTO DE VOCES EN EL ÁMBITO FORENSE

El análisis lingüístico aplicado ha emergido como un campo de estudio dinámico y relevante en diversos ámbitos, incluyendo la Lingüística Forense. En este contexto, el reconocimiento de voces se ha posicionado como una herramienta esencial para la investigación y resolución de crímenes. Este estudio se propone indagar en los procesos y metodologías empleadas en el ámbito forense para la identificación de voces, explorando tanto las técnicas tradicionales como las innovaciones impulsadas por la tecnología y la lingüística computacional. A través del análisis de casos de estudio y la

revisión de literatura especializada, se examinarán los desafíos y las perspectivas en esta área, así como se propondrán recomendaciones para mejorar la efectividad y precisión del reconocimiento de voces en contextos forenses [55].

### 2.3.2 LA EFICACIA PROBATORIA DE LA PERICIA DE ANÁLISIS COMPARATIVO DE HABLANTES EN EL PROCESO PENAL COLOMBIANO: EL COTEJO DE VOZ

El propósito de este trabajo es captar la atención del juez encargado del proceso penal en curso respecto a la fiabilidad de la prueba pericial científica, específicamente en relación con el análisis comparativo de la voz, siguiendo los estándares establecidos por el caso Daubert. Estos estándares, derivados de la jurisprudencia estadounidense y adoptados en el Código de Procedimiento Penal, se orientan a la admisión en juicio de pruebas científicamente relevantes. Se procede a una conceptualización de la pericia de cotejo de voz, resaltando su posición tanto a nivel nacional como internacional, especialmente en contextos jurídicos como España y Estados Unidos. Se pone énfasis en el desarrollo de esta técnica en el ámbito criminalístico local, argumentando que la credibilidad de esta prueba científica se fundamenta en demandas sociales que trascienden el proceso tradicional de valoración probatoria, siendo influenciadas por aspectos de índole social que actualizan la perspectiva del juez [56]

### 2.3.3 INGENIERIA ACUSTICA APLICADA A LA CRIMINALISTICA "ACUSTICA FORENSE"

La ingeniería acústica, a lo largo de su historia, ha demostrado una versatilidad destacada en sus diferentes áreas profesionales, con la criminalística emergiendo como una especialización incipiente en este campo. La acústica forense, aún en su relativa novedad con apenas cinco años de desarrollo, ha sido mayormente impulsada por la sección de sonido del laboratorio de criminalística. El propósito fundamental de esta tesis es divulgar las diversas áreas que abarca esta especialidad, incentivando así el interés por futuras investigaciones que contribuyan al perfeccionamiento de las técnicas empleadas en la

acústica forense en nuestro país. La comprensión de la generación de la voz humana y los procesos auditivos permite a un experto en audio colaborar de manera efectiva con otros profesionales, especialmente en el ámbito del reconocimiento de voz en contextos forenses. Además, el dominio de los sistemas de grabación y microfonía resulta esencial para respaldar la credibilidad de un peritaje de audio, ya que cualquier falta de conocimiento en estas áreas podría poner en entredicho la competencia del experto al defender sus análisis [57].

## 2.4 METODOLOGÍA DE INVESTIGACIÓN

La investigación exploratoria tiene como objetivo proporcionar una perspectiva general sobre algún tema determinado, que suele ser desconocida, que esta presente en el estudio que se llevara a cabo [58]. En el Ecuador se dispone de poca información acerca del estudio propuesto, existe un conocimiento limitado en las personas sobre la tecnología asociada a la suplantación de identidad mediante herramientas que posibilitan la creación de audios sintéticos.

El objetivo principal de la investigación descriptiva es proporcionar una descripción detallada de lo que está sucediendo en el momento. Se centra en aspectos como el porcentaje del grupo objetivo que utiliza la marca en una ubicación concreta o las características de las personas que utilizan un servicio concreto [59].

Se basa en el método científico y utiliza la inducción y la deducción como procesos lógicos. Incluye actividades destinadas a autenticar, probar o reproducir determinados fenómenos naturales o provocados por el hombre. Esto se hace de una manera que facilita la creación de experiencias, que a su vez ayudan a generar hipótesis. A través del proceso científico, estas hipótesis conducen a generalizaciones científicas que pueden ser confirmadas por hechos concretos de la vida cotidiana [60].

### 2.4.1 BENEFICIARIOS

Para la población de estudio estaría ligado de manera general para aquellas personas involucradas en temas de investigación académica, profesionales de la materia y propios usuarios para emitir la concientización de la evolución tecnológica tanto para bien y mal.

| **BENEFICIARIOS** | **CANTIDAD** |
|---|---|
| INVESTIGADORES ACADEMICOS | 5 |
| PROFESIONALES DE SEGURIDAD Y FORENSES | 3 |
| USUARIOS DE PLATAFORMAS – REDES SOCIALES Y COMUNICACIÓN | 10 |
| TOTAL | 18 |

**Tabla 1: DETALLE DE BENEFICIARIOS**

### 2.4.2 VARIABLE

Para medir el impacto de un antes y un después, se considerará el tiempo de análisis de las muestras mediante herramientas forenses. Este proceso integral busca evaluar la complejidad del audio y los procesamientos eficaces en un lapso reducido. De esta manera, se determinará qué herramientas son más óptimas para diversos casos prácticos, investigativos y teóricos, sentando así las bases para futuras investigaciones en el tema.

### 2.4.3 TÉCNICAS E INTRUSMENTOS DE RECOLECCIÓN DE DATOS

**TÉCNICAS**

- Estado del arte y fuentes bibliográficas

**INSTRUMENTO**

Para garantizar la calidad y relevancia de los datos recopilados, se llevarán a cabo entrevistas con el personal docente de la Facultad de Sistemas y Telecomunicaciones de

la Universidad Estatal Península de Santa Elena. Estas entrevistas serán esenciales para comprender a fondo el panorama actual en relación con el tema investigativo, asegurando así que la investigación se base en una comprensión sólida y actualizada de las prácticas y perspectivas en el uso de tecnologías relevantes.

## 2.5 METODOLOGÍA DE DESARROLLO

En el presente proyecto se utiliza la metodología UNE 71506:2013 que será base esencial para la realización de las pruebas experimentales para el análisis forense de evidencias digitales. Este marco referencial proporciona un marco estructurado para llevar a cabo investigaciones forenses sobre dispositivos digitales, audios, computadores, teléfonos móviles, tables, dispositivos de almacenamiento USB, entre otros [61].

A continuación, se presenta las fases fundamentales del marco referencial seleccionado para el estudio del proyecto.

**FASE I: PERVERSIÓN -** Durante esta etapa, se asegura la preservación de todas las pruebas, tanto digitales como físicas, evitando cualquier pérdida. Se lleva a cabo la verificación de los dispositivos o elementos sujetos a análisis y se garantiza la validez legal de las pruebas recopiladas. Se toman medidas para prevenir la alteración, daño o manipulación de la información, tanto por acciones humanas como por eventos naturales, en caso de ser necesario.

**FASE II: ADQUISICIÓN -** Durante esta etapa, se realizará la recopilación de las pruebas, ya sean físicas o digitales, verificando la autenticidad de los dispositivos sujetos a análisis. Se adquirirán las herramientas necesarias, tanto de hardware como de software, para la recolección de datos. Asimismo, se llevará a cabo una copia de seguridad de la evidencia encontrada para preservar la integridad de la evidencia original.

**FASE III: ANÁLISIS -** Durante esta fase, se utilizan herramientas especializadas para llevar a cabo el análisis forense. Se selecciona el método de investigación más adecuado para el caso en cuestión. Se toman medidas para garantizar que la integridad de la información extraída de las pruebas no se vea comprometida en ningún momento del proceso.

**FASE IV: DOCUMENTACIÓN -** En esta etapa, se lleva a cabo la documentación mediante un informe pericial que detalla el proceso de extracción de las pruebas. Se evidencian los datos obtenidos utilizando fotografías o capturas de pantalla. Además, se elabora un informe técnico que resume las conclusiones del proceso realizado.

**FASE V: PRESENTACIÓN -** Durante esta fase, se exponen los resultados obtenidos del análisis forense. Se realiza una presentación clara y convincente, donde se detallan las herramientas y metodologías empleadas durante todo el proceso. Se explican con detalle las conclusiones previamente expuestas en la fase de documentación, garantizando que sean comprendidas de manera accesible para el público objetivo.


**Figura 3: Metodología UNE 71506:2013**

# CAPÍTULO III

# 3 PROPUESTA

## 3.1 DESARROLLO

### 3.1.1 FASE 1: PRESERVACIÓN

En la siguiente fase denominada preservación se desarrolló lo que es la verificación de los audios sintéticos como estudio de caso tanto por generación y por IA. Por lo tanto, son orientado a herramientas para capturar el código hash y el metadato correspondiente del formato

#### AUTENTICIDAD DE AUDIOS - HASHES

La herramienta gtkhash y el desarrollo de un algoritmo dio la apertura de encontrar el codigo hash de los archivos en diferentes formas de algoritmos de proveedor criptográfico como se muestra a detalle el siguiente grafico

| *Herramienta* | *Audios* | | |
|---|---|---|---|
| *Algoritmo* | **AudioIA.mp3** | **SHA256** | **18ec37d205700d6196bf896e596878f8ed3369fa8610f33aa8ad6b584708b8e9** |
| | **AudioOri** | **SHA256** | **731ab1325d862fe5f16c207907b104dcc6a01f0be8d15b6057db88721e877667** |
| *GtkHash* | **AudioIA** | **MD5** | **751469d4380965e4bea83a3e2a001313** |
| | | **SHA1** | **56107a6d8eda0ecf9c7a94e73845dd106cbfc752** |
| | | **SHA256** | **18ec37d205700d6196bf896e596878f8ed3369fa8610f33aa8ad6b584708b8e9** |
| | | **CRC32** | **213a0a4a** |
| | **AudioOri.mp3** | **MD5** | **b97fd947ab7592738f61ed8bbe814fe6** |
| | | **SHA1** | **2902933475cc538af5bf4306f752770034250e2c** |
| | | **SHA256** | **731ab1325d862fe5f16c207907b104dcc6a01f0be8d15b6057db88721e877667** |
| | | **CRC32** | **986e7b4f** |

**Tabla 2: DETALLE DE AUTENTICIDAD DE AUDIOS – HASES**

**METADATA DE AUDIOS POR HERRAMIENTA EXIFTOOL – KALI LINUX**

La herramienta de comando Exiftool permite analizar los metadatos de los archivos de caso de estudio, presentando a detalle información relevante de su creación, modificación, permisos de compatibilidad, entre otros. Como se detalle en el siguiente cuadro

***EXIFTOOL v12.16***

| | | |
|---|---|---|
| ***FILE NAME*** | **AudioIA.mp3** | **AudioOri.mp3** |
| ***DIRECTORY*** | . | . |
| ***FILE SIZE*** | 14 KB | 56 KB |
| ***FILE MODIFICATION DATE/TIME*** | 2024:04:13 15:46:10-04:00 | 2024:04:13 15:46:06-04:00 |
| ***FILE ACCESS DATE/TIME*** | 2024:04:13 15:46:10-04:00 | 2024:04:13 15:46:06-04:00 |

| | | |
|---|---|---|
| ***FILE INODE CHANGE DATE/TIME*** | 2024:04:13 15:46:11-04:00 | 2024:04:13 15:46:07-04:00 |
| ***FILE PERMISSIONS*** | -rw-r—r-- | -rw-r—r-- |
| ***FILE TYPE*** | MP3 | Mp3 |
| ***FILE TYPE EXTENSION*** | Mp3 | Mp3 |
| ***MINE TYPE*** | Audio/mpeg | Audio/mpeg |
| ***MPEG AUDIO VERSION*** | 1 | 1 |
| ***AUDIO LAYER*** | 3 | 3 |
| ***AUDIO BITRATE*** | 64 kbps | 113 kbps |
| ***SAMPLE RATE*** | 48000 | 48000 |
| ***CHANNEL MODE*** | Single Channel | Single Channel |
| ***MS STEREO*** | Off | Off |
| ***INTENSITY STEREO*** | Off | Off |
| ***COPYRIGHT FLAG*** | False | False |
| ***ORIGINAL MEDIA*** | False | False |
| ***EMPHASIS*** | None | None |
| ***VBR FRAMES*** | None | 167 |
| ***VBR BYTES*** | None | 56448 |
| ***VBR SCALE*** | None | 0 |
| ***ID3 SIZE*** | 45 | 45 |
| ***ENCODER SETTINGS*** | Lavf58.29.100 | Lavf58.76.100 |
| ***DURATION*** | 1.70 s (approx) | 4.01 s (approx) |

**Tabla 3: METADATOS EXTRAÍDOS POR EXIFTOOL**

### 3.1.2 FASE 2: ADQUISICIÓN

En esta fase denominada adquisición se desarrolla la investigación exhaustiva de las diversas técnicas e instrumentos necesarios para identificar los audios sintéticos. Se recolecta toda la información necesaria para ejercer la agrupación por parámetros base para la mejor clasificación posible. A continuación, se detalla la información de técnicas para identificar audios sintéticos

| CUADRO COMPARATIVO DE TECNICAS PARA IDENTIFICAR AUDIOS S | | | | |
|---|---|---|---|---|
| **TÉCNICA** | **INSTRUMENTO/HERRAMIENTA** | **DESCRIPCIÓN** | **VENTAJAS** | **USO** |
| **Análisis Espectral** | Software de Análisis Espectral (por ejemplo: Audacity) | Examinar la distribución de frecuencias en el espectro de un audio. | - Facilidad de visualización de características del audio. | - Análisis de calidad de au detección de cambios en e espectro. |
| **Inspección Visual de Formas de Onda** | Software de Edición de Audio (por ejemplo: Adobe Audition) | Herramientas de software que proporcionan representaciones visuales de las formas de onda de los sonidos. | - Intuitivo para detectar anomalías visuales. | - Verificació manual de ca de audio, identificació irregularidad |
| **Análisis de Coeficientes Cepstrales** | Librerías de Procesamiento de Audio en Python (por ejemplo: librosa) | Características extraídas de la transformada cepstral de una señal de audio. | - Captura de características relevantes para reconocimiento de voz. | - Reconocimi de voz, identificació patrones en a |
| **Extracción de Características MFCC** | Librerías de Procesamiento de Audio en Python (por ejemplo: librosa) | Características derivadas de la transformación de Mel de una señal de audio y la transformada de Fourier. | - Representación compacta y discriminativa de características. | - Reconocimi de habla, identificació locutores. |

| | | | | |
|---|---|---|---|---|
| **Comparación con Base de Datos** | Base de Datos de Sonidos Sintéticos | Mantener una base de datos de sonidos sintéticos conocidos y comparar audios sospechosos con esta base. | - Referencia para comparar audios sospechosos. | - Detección d deepfakes, autenticación voz, análisis forense de au |
| **Preprocesamiento de Señales** | Software de Procesamiento de Señales (por ejemplo: MATLAB) | Filtrado para eliminar ruido y artefactos que podrían interferir con la identificación precisa de sonidos. | - Mejora la calidad de la señal y facilita el análisis posterior. | - Mejora de calidad de au eliminación d ruido, preprocesami para análisis audio. |
| **Extracción de Características** | Librerías de Procesamiento de Audio en Python (por ejemplo: librosa) | Obtener características relevantes de la señal de audio, como los MFCC o los coeficientes cepstrales. | - Permite representación de datos de manera compacta y discriminativa. | - Clasificació audio, reconocimien habla, detecc eventos de so |
| **Clasificación de Patrones** | Librerías de Aprendizaje Automático en Python (por ejemplo: scikit-learn) | Utilización de algoritmos de clasificación, como SVM o árboles de decisión, para distinguir sonidos. | - Capacidad para automatizar la identificación de audios sintéticos. | - Detección d deepfakes, autenticación voz, análisis forense de au |
| **Validación y Evaluación del Modelo** | Librerías de Aprendizaje Automático en Python (por ejemplo: scikit-learn) | Evaluación de la precisión y eficacia del modelo utilizando técnicas de validación cruzada y métricas. | - Permite evaluar la calidad del modelo y ajustar hiperparámetros. | - Evaluación modelos de detección de deepfakes, autenticación voz, análisis forense de au |

**Tabla 4: DETALLE DE TÉCNICAS DE AUDIOS SINTÉTICOS**

**Preparación de Ambiente Virtuales**

Aquí se prepara los ambientes virtuales para usar herramientas esenciales para el análisis que viene preinstaladas con las dependencias para trabajar

- **Descargar Virtual Box y configurar**

  - **Descargar VirtualBox V 7.0.14 del siguiente link** https://download.virtualbox.org/virtualbox/7.0.14/VirtualBox-7.0.14-161095-Win.exe
  - **Descargar Extensión Pack del siguiente link** https://download.virtualbox.org/virtualbox/7.0.14/Oracle_VM_VirtualBox_Extension_Pack-7.0.14.vbox-extpack

- **Preparación de Máquina Virtual - Caine**

  Para la preparación del entorno virtual Caine se debe establecer la correcta configuración de la mismo, como descargar la Imagen Iso del Sitio Oficial Caine, establecer la preparación del entorno en virtualBox, Configurar opciones de usuarios, particiones, entre otros factores esenciales para el uso pertinente de Proyecto Digital Forensics. Esta máquina virtual cuenta con herramientas esenciales para uso de computación forenses que ayudaran al proceso de investigación y análisis de los comportamientos digitales como es el caso de Audios sintéticos. Para más detalle (**VER ANEXO 2: INSTALACIÓN DE CAINE**)

- **Preparación de Máquina Virtual - Kali Linux**

  Para la preparación del entorno virtual Kali linux una distribución basada en Debian que cuenta con una amplia gama de herramientas de seguridad informática con una interfaz de usuario amigable y entendible. La instalación correspondiente del mismo es descargar la imagen iso o el archivo rar que cuenta con el ova de instalación preconfigurada para el uso. Tanto el proceso manual de instalación o preconfiguración permitirá establecer las reglas de uso y que ventajas útiles para

el análisis e investigación del caso de Audios Sintéticos. Para más detalle (**VER ANEXO 2: INSTALACIÓN DE KALI LINUX**

### 3.1.3 FASE 3: ANÁLISIS

En la fase de análisis del proyecto se desarrolla lo que es la generación de los escenarios de pruebas para ejercer el análisis de los audios sintéticos de muestra. Se crea respetivamente el cuadro de resumen en donde se detalla cada parámetro de información del caso/escenario

| | Descripción | Tiempo de Audio | Formato de Archivo | Ambiente de Grabación | Dispositivo de Grabación | Configuración de Grabación | Ubicación d Dispositivo Durante la Grabación |
|---|---|---|---|---|---|---|---|
| **1** | Generación de audio mediante la aplicación grabadora – Celular | 2 minutos | MP3 | Entorno no controlado, posible presencia de ruido ambiental | Xiaomi Redmi Note 10 | Configuración predeterminada de la aplicación grabadora | En la mano, cerca del hablante o fuente de sonido |
| **2** | Generación de audio texto a voz | 2 minutos | MP3 | Sitio de grabación ElevenLabs | Laptop | Configuración predeterminada de la tecnología | Auriculares cerca del hablante o fuente de sonido |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **3** | Generación del Audio a través de la tecnología ElevenLabs IA – Clonación | 6 segundos | MP3 | Laptop a través del soporte de grabación de ElevenLabs | Laptop | Configuración de similitud/semejanza de audio 100% | Auriculares/p rlante de laptop |

**Tabla 5: RESUMEN DE ESCENARIOS DE AUDIOS SELECCIONADOS**

En esta sección una vez desarrollado los escenarios se comienza a ejercer el análisis mediante las técnicas seleccionada para el estudio y se crea respectivamente el cuadro descriptivo con los parámetros se muestra a continuación

| ANALISIS DE CASO 1 – GENERACION TEXTO A VOZ | | | | |
|---|---|---|---|---|
| **CASO 1** | **TÉCNICA** | **HERRRAMIENTA** | **PARAMETROS ANALIZADOS** | **RESULT** |
| Generación de audio Original Obtenido mediante la herramienta ElevenLabs | **Análisis de Coeficientes MFCC** | Python usando biblioteca de Librosa, matplotlib y numpy | Estructura temporal clara, con picos y valles que corresponden a los diferentes sonidos del habla.<br>Típico del habla humana, compuesta por fonemas con características espectrales distintivas. | Se presenta una del audio tras el |
| | **Análisis Espectral** | Audacity | El diagrama muestra que la señal de audio tiene una gran cantidad de energía en las frecuencias bajas, entre 50 Hz y 500 Hz. Esto podría indicar que la señal contiene graves o voces. También hay una cantidad significativa de energía en las frecuencias altas, entre 2000 Hz y 10000 Hz. Esto podría indicar que la señal contiene platillos o sibilancia. | El pico del espe encuentra en 12 significa que est frecuencia domi señal. Es posibl frecuencia corre nota fundamenta instrumento mu humana. |
| | | SPEK | En la representación del grafico se muestra una distribución de energía de señal de audio en diversos parámetros de frecuencias. El área más brillante del espectro representa las frecuencias donde hay más energía, se observa energía de bajas frecuencias que indica que la señal de audio es rica en graves. | Se presentó una variación de la mostrando valor de frecuencias |
| | **Análisis de Coeficientes Cepstrales** | Python usando biblioteca de Librosa, matplotlib y numpy | **Estructura regular y definida:** Los coeficientes cepstrales se encuentran ordenados de manera predecible, formando una secuencia con picos y valles bien definidos.<br>**Picos y valles marcados:** La amplitud de los picos y valles es considerable, indicando una fuerte presencia de frecuencias específicas en la señal original.<br>**Frecuencias discretas:** La distribución de los coeficientes sugiere que la señal de audio está compuesta por un conjunto finito de frecuencias, típicas de una grabación limpia y libre de distorsiones. | Los coeficientes Audio 1 muestra sinusoidal, sugi frecuencia domi señal. El pico m alrededor de los indica una frecu dominante de aproximadamen |

| | | | | |
|---|---|---|---|---|
| | **Análisis de Forma de Onda** | Python usando biblioteca de Librosa, matplotlib | La forma de onda tiene como variedad:<br>• **Picos:** Los picos representan los puntos de máxima amplitud de la señal de audio. Estos picos corresponden a los sonidos más fuertes de la grabación.<br>• **Valles:** Los valles representan los puntos de mínima amplitud de la señal de audio. Estos valles corresponden a los sonidos más suaves de la grabación. | La forma de onc<br>muestra una señ<br>con múltiples co<br>frecuencia. Los<br>indican alta frec<br>mientras que las<br>en la amplitud r<br>cambios en la in<br>sonido. |

**Tabla 6: DETALLE DE ANALISIS DE CASO 1 – AUDIO TEXTO A VOZ - ELEVENLABS**

| **ANALISIS DE CASO 2 – AUDIO CLONACIÓN IA - ELEVENLABS** | | | | |
|---|---|---|---|---|
| **CASO 2** | **TÉCNICA** | **HERRRAMIENTA** | **PARAMETROS ANALIZADOS** | **RESULT** |
| Grabación de audio Por ElevenLabs | **Análisis de Coeficientes MFCC** | Python usando biblioteca de Librosa, matplotlib y numpy | □ Coeficientes MFCC más ruidosos y menos estructurados que los del Audio 1.<br>□ Distorsiones introducidas por la clonación por IA:<br>• Mayor variabilidad en los valores de los coeficientes.<br>Pérdida de información temporal. | Posible uso de<br>acústico entren<br>datos de habla |
| | **Análisis Espectral** | Audacity | El rango de frecuencia de la señal de audio se extiende aproximadamente entre 50 Hz y 20.000 Hz. Este es el rango de frecuencias que los humanos podemos escuchar. Las frecuencias más bajas (50-500 Hz) se asocian típicamente con sonidos graves, mientras que las frecuencias más altas (2.000-20.000 Hz) se asocian con sonidos agudos. | El gráfico tam<br>varios picos es<br>que son puntos<br>amplitud de la<br>particularment<br>picos correspo<br>frecuencias do<br>la señal de auc<br>caso, el pico m<br>prominente se<br>alrededor de 2<br>está dentro del<br>de tenor. |

| | | | | |
|---|---|---|---|---|
| | | SPEK | la frecuencia fundamental de la voz clonada es ligeramente más alta que la de la voz original. Esto se debe a que la técnica de clonación de voz suele modificar la voz original para que suene más similar a la voz de destino.<br><br>Además, los armónicos de la voz clonada están menos definidos que los de la voz original. Esto se debe a que la técnica de clonación de voz suele distorsionar la señal original. | La variación d<br>original con el<br>un cambio rad<br>presentado una<br>adaptación sin<br>de factores de |
| | **Análisis de Coeficientes Cepstrales** | Python usando biblioteca de Librosa, matplotlib y numpy | □ **Estructura irregular y difusa:** Los coeficientes cepstrales carecen de un ordenamiento claro, presentando una distribución aleatoria y sin patrones definidos.<br>□ **Picos y valles menos marcados:** La amplitud de los picos y valles es menor, indicando una menor presencia de frecuencias específicas y una mayor presencia de ruido.<br>□ **Frecuencias no discretas:** La distribución de los coeficientes sugiere que la señal de audio contiene un espectro más amplio y menos definido, propio de una señal con ruido y distorsión. | e esperaría obs<br>espectrograma<br>coeficientes ce<br>exhiba caracte<br>irregularidad,<br>picos y valles<br>marcados, y fr<br>discretas, lo qu<br>la presencia de<br>distorsión en l<br>audio analizad |

| | | | | |
|---|---|---|---|---|
| | **Análisis de Forma de Onda** | Python usando biblioteca de Librosa, matplotlib | • **Picos y valles:** Los picos y valles de la forma de onda del audio clonado son menos pronunciados que los del audio original. Esto significa que el audio clonado tiene un rango dinámico más pequeño.<br>• **Frecuencia:** La frecuencia de la señal de audio clonado es ligeramente diferente de la del audio original. Esta diferencia es más notable en las frecuencias altas.<br>• **Timbre:** El timbre del audio clonado es similar al del audio original, pero no es exactamente el mismo. Esto significa que el audio clonado no suena exactamente como la voz original. | se esperaría ob<br>diferencias en<br>dinámico, la fi<br>timbre entre el<br>original y el au<br>lo que indicarí<br>audio clonado<br>réplica perfect<br>original |

| | | | | |
|---|---|---|---|---|
| | **Análisis Acústico** | PRAAT | • Espectrograma<br>• Frecuencia máxima<br>• Ancho de Banda<br>• Periodo más largo – corto<br>• Factor de periodo máximo | Se genera com<br>una banda baja<br>banda alta, ade<br>intensidad del<br>audio presento<br>negativo dand<br>que el audio es<br>ese resultado s<br>es todo lo cont<br>original |

**Tabla 7: DETALLE DE ANÁLISIS DE CASO 2 – AUDIO CLONACIÓN IA - ELEVENLABS**

### 3.1.4 FASE 4: DOCUMENTACIÓN

En este aparto de documentación se describe de manera detalla todos los puntos necesarios desarrollado en el análisis de los casos – audios sintéticos generador de forma general como proporcionado por IA. Cuenta con una estructura amigable que dispone de una introducción, alcance, objetivo de estudio, descripción de controles de las técnicas, observaciones técnicas y recomendaciones. La finalidad de esta documentación permite que tanto expertos como o profesionales cuenten con una guía de desarrollo para ejercer los controles pertinentes del análisis forense en el contexto de audios Para más detalle ver (ANEXO 5: DOCUMENTACIÓN)

## 3.2 PROCEDIMIENTO WORKFLOW

En esta sección se desarrolló el procedimiento de workflow que permite describir de manera lógica como el proceso de análisis de audios sintéticos se proporciona mediante algunos filtros para ejercer la verificación autentica del caso a través de técnicas forenses en audios sintéticos

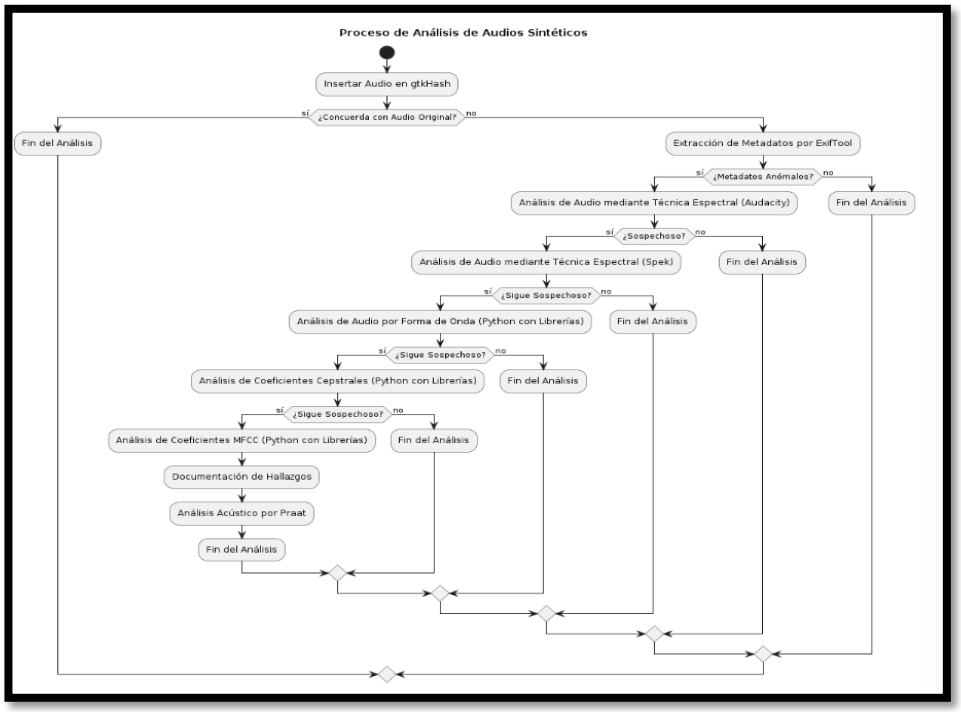


**Figura 4: PROCEDIMIENTO WORKFLOW DE VERACIDAD DE AUDIOS**

# CONCLUSIONES

- Se determinó que la recopilación de un conjunto amplio y diverso de audios sintéticos fue esencial para el estudio. Esta diversidad permitió realizar un análisis más exhaustivo y preciso, facilitando la identificación de patrones característicos de los audios manipulados mediante técnicas de IA.
- Se concluyó que las pruebas experimentales con técnicas de IA y acústica forense fueron efectivas para diferenciar audios sintéticos de los originales. El uso de análisis espectral y coeficientes cepstrales proporcionó una base sólida para identificar discrepancias en los audios. La precisión de los resultados mejoró significativamente al aplicar técnicas avanzadas y algoritmos optimizados, reduciendo el margen de error.
- Se determinó que evaluar los resultados con métricas forenses fue crucial para validar la eficacia del procedimiento. Las métricas demostraron que las técnicas aplicadas fueron efectivas para identificar audios sintéticos, proporcionando resultados consistentes y fiables. Detallar la ejecución de las actividades, especificando herramientas, algoritmos y técnicas, facilitó la comprensión de los procesos y la replicabilidad del estudio en futuros trabajos.

# RECOMENDACIONES

- Se recomienda diseñar y desarrollar nuevos algoritmos y herramientas específicas para la identificación de audios sintéticos. Estos deben ser capaces de adaptarse a las nuevas técnicas de creación de deepfakes, garantizando así una detección precisa y actualizada.
- Es fundamental realizar pruebas y evaluaciones continuas de las nuevas herramientas y técnicas desarrolladas. Este proceso de validación constante permitirá mejorar y ajustar las metodologías forenses, asegurando su eficacia y estabilidad
- Se recomienda aplicar las herramientas y metodologías desarrolladas en futuros estudios y contextos prácticos, ampliando así el alcance y la utilidad de la

investigación. Además, es importante documentar detalladamente cada análisis para contribuir a la base de conocimientos en acústica forense y IA

# BIBLIOGRAFÍAS


[1] M. L. Oliva, «Deepfake: Cuando la inteligencia artificial amenaza el Derecho y la Democracia,» Revista de Derecho y Tecnología, p. 85.

[2] C. Alonso, «Un informático en el lado del mal,» 5 Mayo 2021. [En línea]. Available: https://www.elladodelmal.com/2021/05/dfaas-deepfakes-as-service-en-la.html. [Último acceso: 30 Septiembre 2023].

[3] Yuezun Li1, Xin Yang1, Pu Sun2, Honggang Qi2 and Siwei Lyu1, «Celeb-DF (v2): A New Dataset for DeepFake Forensics,» 2020. [En línea]. Available: https://cse.buffalo.edu/~siweilyu/celeb-deepfakeforensics.html. [Último acceso: 26 Mayo 2024].

[4] C. Alonso, «Un informatico en el lado del mal,» 16 Octubre 2021. [En línea]. Available: https://www.elladodelmal.com/2021/10/robo-de-400000-usd-clonando-la-voz-del.html. [Último acceso: 2 Octubre 2023].

[5] T. Brewster, «Estafadores clonaron la voz del director de una empresa en un atraco de 35 millones de dólares, según descubre la policía,» pp. 2-7, 15 Enero 2020.

[6] Y. Z. y. H. W. Jinlin Guo, «Detección de suplantación generalizada y algoritmo incremental. Reconocimiento de suplantación de voz.,» Universidad Nacional de Tecnología de Defensa, Changsha, 2023.

[7] J. M. Miguez, «Reconocimiento automático de hablante, empleando técnicas deep learning en peritaje informáticos,» Universidad de Palermo, Buenos Aires, 2023.

[8] J. A. y. E. C. K. Palacio, «Diseño e implementación de un sistema de síntesis de voz,» Revista Tecnológica ESPOL, Cuenca, 2008.

[9] M. V. Á. y. X. L. G. Ángel Vizoso, «Luchando contra los deepfakes: los medios y los gigantes de Internet convergen y divergen Estrategias contra la desinformación de alta tecnología,» Facultad de Ciencias de la Comunicación, Universidad de Santiago de Compostela, 2021.

[10] G. Lab, «peritos informáticos,» [En línea]. Available: https://peritosinformaticos.es/iso-71506-2013-perito-informatico/#:~:text=La%20norma%20UNE%2071506%2F2013,de%20la%20norma%20UNE%2071505.. [Último acceso: 2 Octubre 2023].

[11] U. E. P. d. S. Elena, «Resolución RCF-FST-SO-09 No. 03-2021,» La Libertad, 2021.

[12] G. P. F. C. y. L. C. Henry Ajder, «El estado de los deepfakes Paisaje, amenazas e impacto.,» ©2019 Deeptrace, 2019.

[13] B. R. Rebaque, «La comunicación científica contra la desinformación,» Universidad Rey Juan Carlos (URJC), 2020.

[14] Jumio, «América digital México,» 12 Junio 2023. [En línea]. Available: https://mx.america-digital.com/jumio-presenta-encuesta-sobre-deteccion-de-deepfakes/. [Último acceso: 16 Octubre 2023].

[15] H. L. T. B. C. y. G. F. S. Moyana Mariano Robles Lessa, «Deepfake: inteligencia artificial y algoritmos que provocan riesgos a la sociedad en el ciberespacio,» Derecho y cambio social, Lima, 2020.

[16] Secretaria Nacional de Planificación , «Plan de Creación de Oportunidades 2021 2025,» p. 122, 2021.

[17] Ing. Edurdo García L., «Acústica Forense: Identificación de locutores automatizada,» Universidad Nacional Autónoma de México, México, 2014.

[18] Josefina Riva P., «Suspensión de ruido en audio mediante programas informáticos,» Universidad FASTA, Buenos Aires - Argentina, 2023.

[19] Pablo Félix C., «Las emisiones sonoras del habla son producidad por el aparato fonatorio compuesto por el sistemas respiratorio, la laringe, las cuerdad vocales y la cavidad bucal - aparato fonador,» [En línea]. Available: https://sisbib.unmsm.edu.pe/bibvirtualdata/libros/linguistica/leng_nino/pdf/explor_producc.pdf. [Último acceso: 15 Mayo 2024].

[20] Nieto R. F., «Estudio de la percepción de la identidad del locutor utilizando técnicas de análisis-síntesis de voz,» EUITT-UPM, Madrid - España, 2010.

[21] Sancho Ramírez Centro Auditivo, «¿Qué es la audición?,» LinkedIn, 24 Marzo 2022. [En línea]. Available: https://es.linkedin.com/pulse/qu%C3%A9-es-la-audici%C3%B3n-sancho-ram%C3%ADrez-centro-auditivo. [Último acceso: 15 Mayo 2024].

[22] National Institute on Deafness and Other Communication Disorders, «¿Cómo oímos?,» 14 Junio 2022. [En línea]. Available: https://www.nidcd.nih.gov/es/espanol/como-oimos. [Último acceso: 15 Mayo 2024].

[23] Unir - Universidad en Internet, «Informática forense: en qué consiste, ámbitos de aplicación y perfiles profesionales,» 24 Junio 2021. [En línea]. Available: https://www.unir.net/ingenieria/revista/informatica-forense/. [Último acceso: 26 Mayo 2024].

[24] Coursera Staff, «¿Qué es la informática forense? Tipos, técnicas y carreras,» Coursera.Org, 22 Abril 2024. [En línea]. Available: https://www.coursera.org/mx/articles/computer-forensics. [Último acceso: 26 Mayo 2024].

[25] Kassandra Ortega, «¿Qué es la informática forense?,» Saint Leo University, 07 Septiembre 2022. [En línea]. Available: https://worldcampus.saintleo.edu/noticias/que-es-la-informatica-forense-analisis-forense-informatico. [Último acceso: 26 Mayo 2024].

[26] Equipo de Expertos en Ciencia y Tecnología, «Ciencia y Tecnología - La Labor del informático Forense a detalle,» Universidad Internacional de Valencia, 08 Septiembre 2022. [En línea]. Available: https://www.universidadviu.com/ec/actualidad/nuestros-expertos/la-labor-del-informatico-forense-detalle. [Último acceso: 26 Mayo 2024].

[27] Martín Marco, «Seguridad informática e informática forense,» 26 Febrero 2023. [En línea]. Available: https://www.economiadehoy.es/seguridad-informatica-e-informatica-forense. [Último acceso: 26 Mayo 2024].

[28] Moises Díaz, «Análisis forense del correo electrónico: técnicas de investigación,» 17 Enero 2021. [En línea]. Available: https://moisesdiaz.com.mx/analisis-forense-del-correo-electronico-tecnicas-de-investigacion/. [Último acceso: 26 Mayo 2024].

[29] En Vista Forensics, «Análisis Forense de Telefónos Celulares y Dispositivos Móviles,» 2016 2022. [En línea]. Available: https://www.envistaforensics.com/es-mx/servicios/servicios-forenses-digitales/analisis-forense-de-telefonos-celulares-y-dispositivos-moviles/. [Último acceso: 26 Mayo 2024].

[30] Keyun R. , Ibrahim B. , Joe C. & Tahar K. , «urvey on Cloud F y on Cloud Forensics and Critical Criteria for Cloud F ensics and Critical Criteria for Cloud Forensic,» Journal of Network Forensics, 2011.

[31] Ciberseguridad.com, «Análisis Forense - Tipos,» [En línea]. Available: https://ciberseguridad.com/servicios/analisis-forense/. [Último acceso: 26 Mayo 2024].

[32] Valeria Ara, «Tipos de informática forense,» 22 Noviembre 2023. [En línea]. Available: https://msmk.university/ciberseguridad/que-es-la-informatica-forense-msmk-university. [Último acceso: 26 Mayo 2024].

[33] IBM, «¿Qué es la informática forense?,» IBM, [En línea]. Available: https://www.ibm.com/es-es/topics/computer-forensics. [Último acceso: 26 Mayo 2024].

[34] Lucas Paus, «¿Qué son las técnicas anti forenses?,» WeliveSecurity ESET, 02 Julio 2015. [En línea]. Available: https://www.welivesecurity.com/la-es/2015/07/02/tecnicas-anti-forenses/. [Último acceso: 26 Mayo 2024].

[35] Microlab IoT, «Entendiendo el Análisis Espectral: FFT, FS, DTFS, FT, DTFT, DFT,» 03 Mayo 2022. [En línea]. Available: https://microlab.ec/blog/entendiendo-el-analisis-espectral/. [Último acceso: 15 Mayo 2024].

[36] LinkedIn, «¿Cómo se manejan las formas de onda RTL complejas y dinámicas en los métodos de comparación de formas de onda?,» [En línea]. Available: https://es.linkedin.com/advice/3/how-do-you-handle-complex-dynamic-rtl-waveforms-waveform?lang=es. [Último acceso: 15 Mayo 2024].

[37] T. Fontalvo & E. de la Hoz, «Método análisis envolvente de datos y redes neuronales en la evaluación y predicción de la eficiencia técnica de pequeñas empresas exportadoras",» Información Técnologica, 2018.

[38] Thierry Erbessd, «Análisis de Fase,» 13 Mayo 2022. [En línea]. Available: https://www.erbessd-instruments.com/es/tutoriales/analisis-de-fase/#:~:text=El%20an%C3%A1lisis%20de%20fase%20es,diferentes%20partes%20de%20una%20m%C3%A1quina.. [Último acceso: 15 Mayo 2024].

[39] Unir - La Universidad en Internet, «¿Qué es el espectrograma y cuáles son sus usos en el análisis musical?,» 26 Noviembre 2021. [En línea]. Available: https://ecuador.unir.net/actualidad-unir/espectrograma/. [Último acceso: 15 Mayo 2024].

[40] Francisco Ángeles A., «Mundos Digitales - Creación de Audio Digital,» [En línea]. Available: https://www.uaeh.edu.mx/docencia/P_Presentaciones/prepa3/2020/mundos-digitales.pdf. [Último acceso: 15 Mayo 2024].

[41] Agustín Jamele, «Audio digital: qué es, cómo funciona y ventajas,» 06 Febrero 2023. [En línea]. Available: https://www.innovaciondigital360.com/i-a/audio-digital-que-es-como-funciona-y-ventajas/. [Último acceso: 15 Mayo 2024].

[42] Cliff Weitzman, «Cómo clonar tu voz con IA: la guía definitiva,» 29 Enero 2024. [En línea]. Available: https://speechify.com/es/blog/como-clonar-la-voz-con-ai/#:~:text=La%20clonaci%C3%B3n%20de%20voz%20es,igual%20que%20una%20voz%20humana.. [Último acceso: 15 Mayo 2024].

[43] Adam Kails, «CIBERATAQUES CON CLONACIÓN DE VOZ POR INTELIGENCIA ARTIFICIAL,» LinkedIn, 06 Febrero 2024. [En línea]. Available: https://es.linkedin.com/pulse/ciberataques-con-clonaci%C3%B3n-de-voz-por-inteligencia-artificial-kalis-taawf. [Último acceso: 15 Mayo 2024].

[44] Pedagogía Virtual, «Qué es Audacity?,» [En línea]. Available: https://pedagogiavirtual.org/cursowebhm/documentos/talleres/AUDACITY.pdf. [Último acceso: 26 Mayo 2024].

[45] Spek, «Spek — Analizador de espectro,» 2024. [En línea]. Available: https://www.spek.cc/about. [Último acceso: 26 Mayo 2024].

[46] Amazon Web Services, «¿Qué es Python?,» 2023. [En línea]. Available: https://aws.amazon.com/es/what-is/python/#:~:text=Python%20es%20un%20lenguaje%20de,ejecutar%20en%20muchas%20plataformas%20diferentes.. [Último acceso: 11 11 2023].

[47] Course Hero, «Course Hero - Que es Caine,» 07 10 2023. [En línea]. Available: https://coursehero.com/file/207906933/CAINE-FORENSIC-PPTpptx/. [Último acceso: 27 03 2024].

[48] Sebastián Bortnik & Martina López, «WeliveSecurity - Pruebas de penetración para principiantes: 5 Herramientas para empezar,» 24 11 2023. [En línea]. Available: https://www.welivesecurity.com/es/recursos-herramientas/herramientas-pentesting-para-principiantes/. [Último acceso: 27 03 2024].

[49] Alfredo Sánchez A., «La librería Numpy,» 12 Mayo 2022. [En línea]. Available: https://aprendeconalf.es/docencia/python/manual/numpy/. [Último acceso: 26 Mayo 2024].

[50] Alfredo Sánchez A., «La librería Matplotlib,» 04 Octubre 2020. [En línea]. Available: https://aprendeconalf.es/docencia/python/manual/matplotlib/. [Último acceso: 26 Mayo 2024].

[51] Julieta J., «Librosa para análisis de audio,» 25 Junio 2021. [En línea]. Available: https://l52mas.gitlab.io/posts/librosa-report/. [Último acceso: 26 Mayo 2024].

[52] Kali , «libimage-exiftool-perl,» 16 Febrero 2024. [En línea]. Available: https://www.kali.org/tools/libimage-exiftool-perl/. [Último acceso: 26 Mayo 2024].

[53] Kali, «Gtkhash,» 23 Mayo 2024. [En línea]. Available: https://www.kali.org/tools/gtkhash/. [Último acceso: 26 Mayo 2024].

[54] Paul B. & David W., «Praat: Doing Phonetics by computer,» 2019. [En línea]. Available: https://www.fon.hum.uva.nl/praat/. [Último acceso: 10 Junio 2024].

[55] Beatriz Salmón T., «Un estudio de lingüística aplicada: El reconocimiento de voces en el ámbito forense,» Universidad de Cantabria, Cantabria - España, 2018-2019.

[56] Irina Duarte R., «La eficacia probatoria de la pericia de análisis comparativo de hablantes en el proceso penal colombiano: El cotejo Voz,» Universidad Sergio Arboleda, Bogotá - Colombia, 2019.

[57] Eduardo Pérez B., «Ingeniería acústica aplicada a la criminalística : "Acústica Forense",» Universidad Austral de Chile, Valdivia - Chile, 2008.

[58] N. Morales, «Investigación Exploratoria: Tipos, Metodología y Ejemplos».

[59] M. J. M. Sanz, Introducción a la investigación de mercados, Madrid : ESIC Editorial, 2015.

[60] D. M. P. ROBERTO, «DocPlayer,» 2013. [En línea]. Available: https://docplayer.es/81811126-Metodologia-de-la-investigacion.html. [Último acceso: 19 Octubre 2023].

[61] G., Lab, «Peritos Informáticos,» 2022. [En línea]. Available: https://peritosinformaticos.es/iso-71506-2013-perito-informatico/#:~:text=La%20norma%20UNE%2071506%2F2013,de%20la%20norma%20U NE%2071505. [Último acceso: 02 Octubre 2023].

# ANEXOS

## Anexo 1. Árbol de problemas

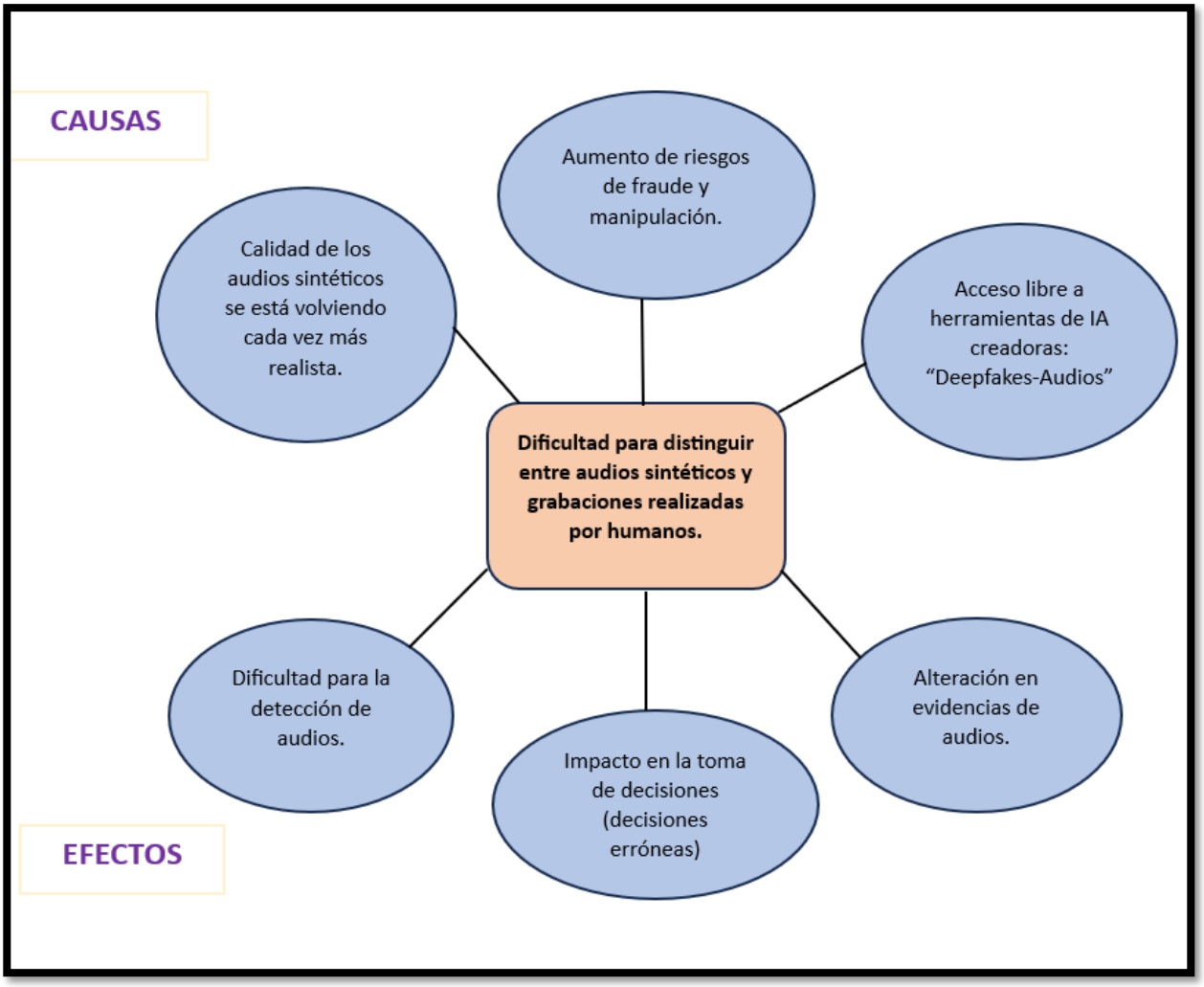



**Figura 5: Causa y Efectos de Dificultad de detectar audios sintéticos**

# ANEXO 2: MANUAL DE INSTALACIÓN

# INSTALACION DE CAINE

1. **Primero se descarga la Imagen Iso Caine del siguiente Link**
   https://deb.parrot.sh/direct/parrot/iso/caine/caine13.0.iso

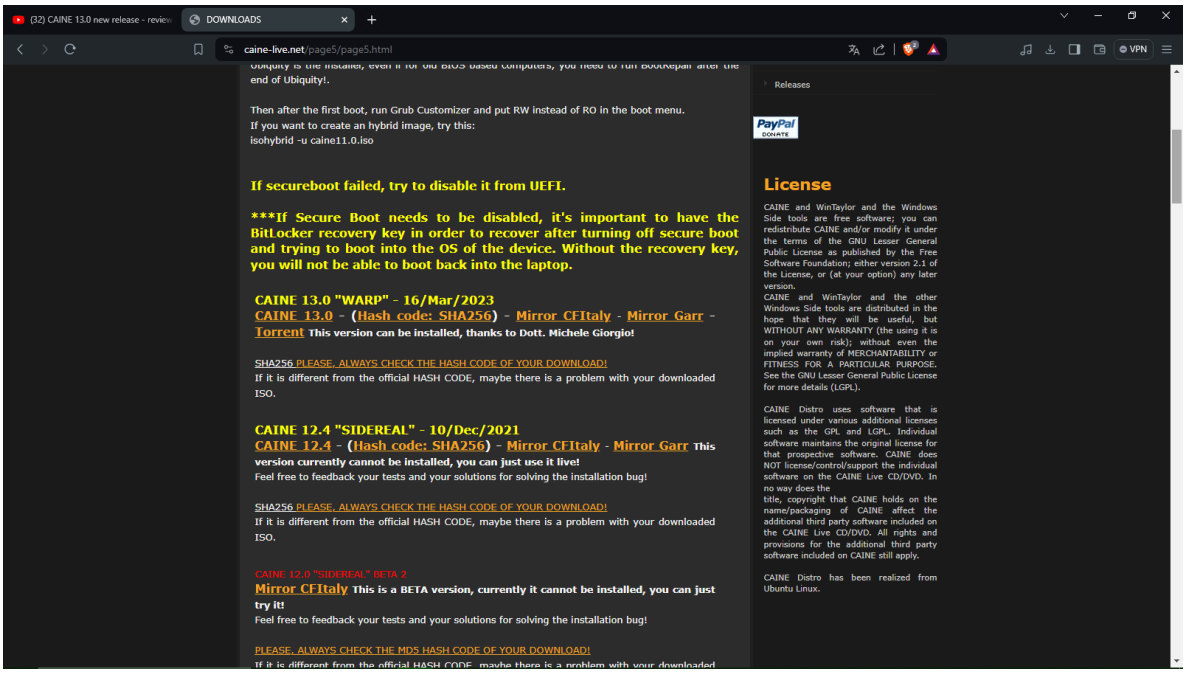



**Imagen 3: Sitio de Caine**

2. **Archivo Descargado.**

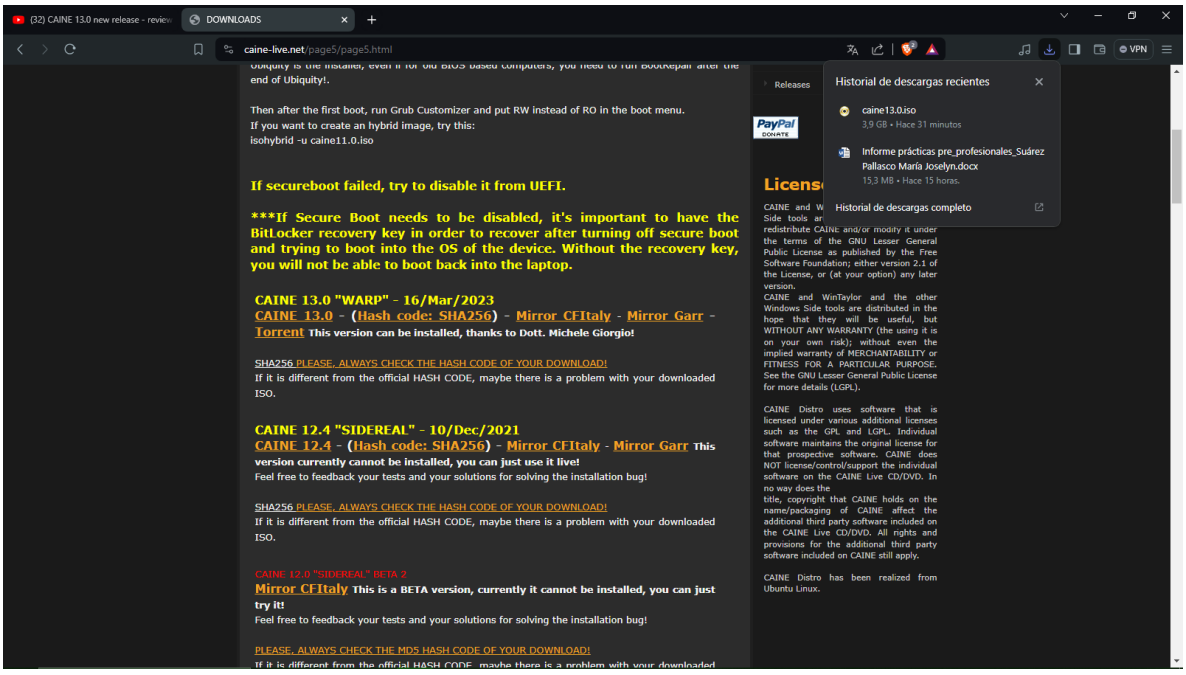



**Imagen 4: Archivo Caine Descargar**

3. **Abrir VirtualBox.**

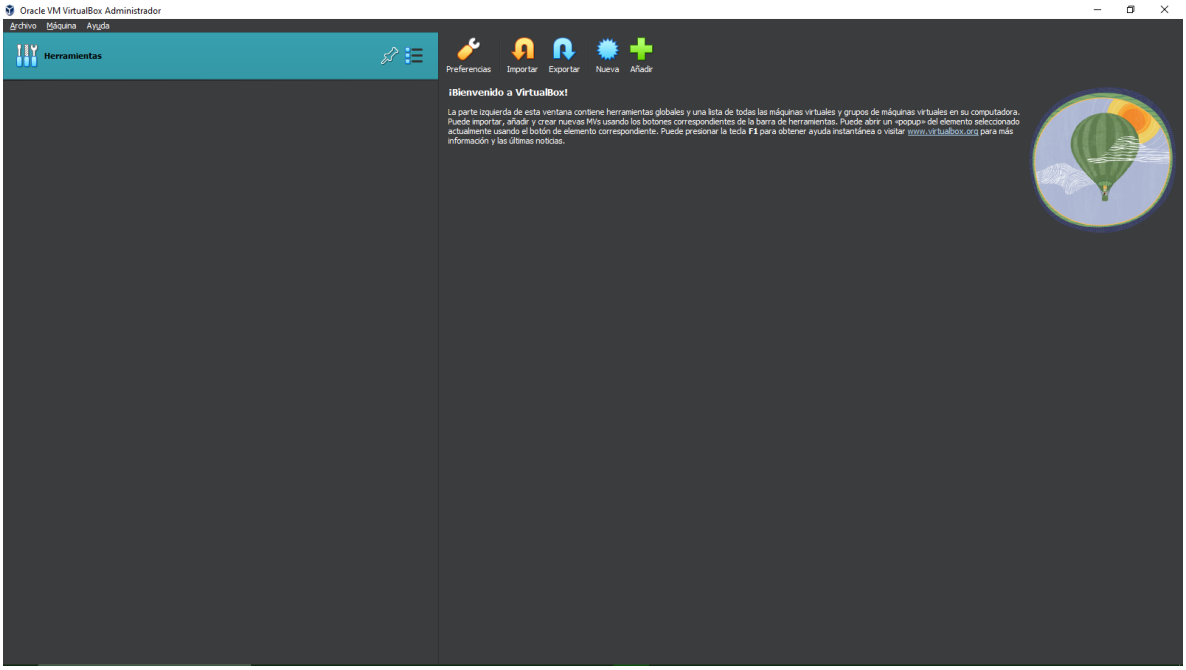



**Imagen 5: Inicio de VirtualBox**

4. **Seleccionar en la opción nueva para establecer la creación de la máquina Virtual.**

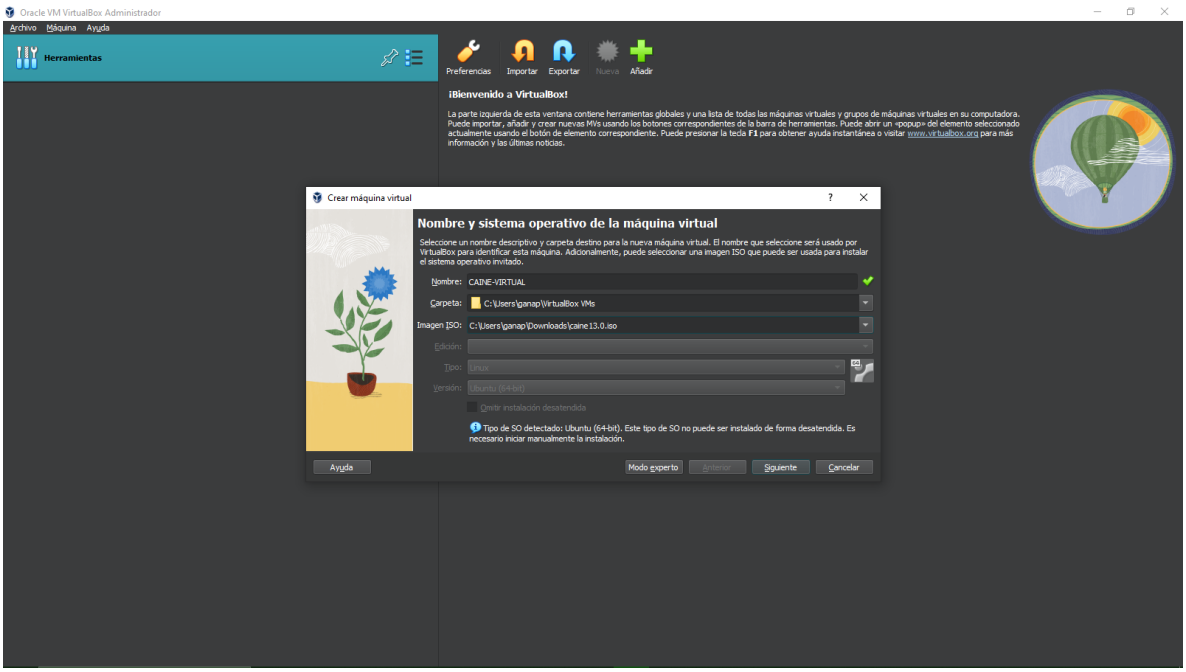



**Imagen 6: Creación de máquina virtual Caine**

5. **Se establece el valor de memoria Ram a usar en la máquina virtual.**

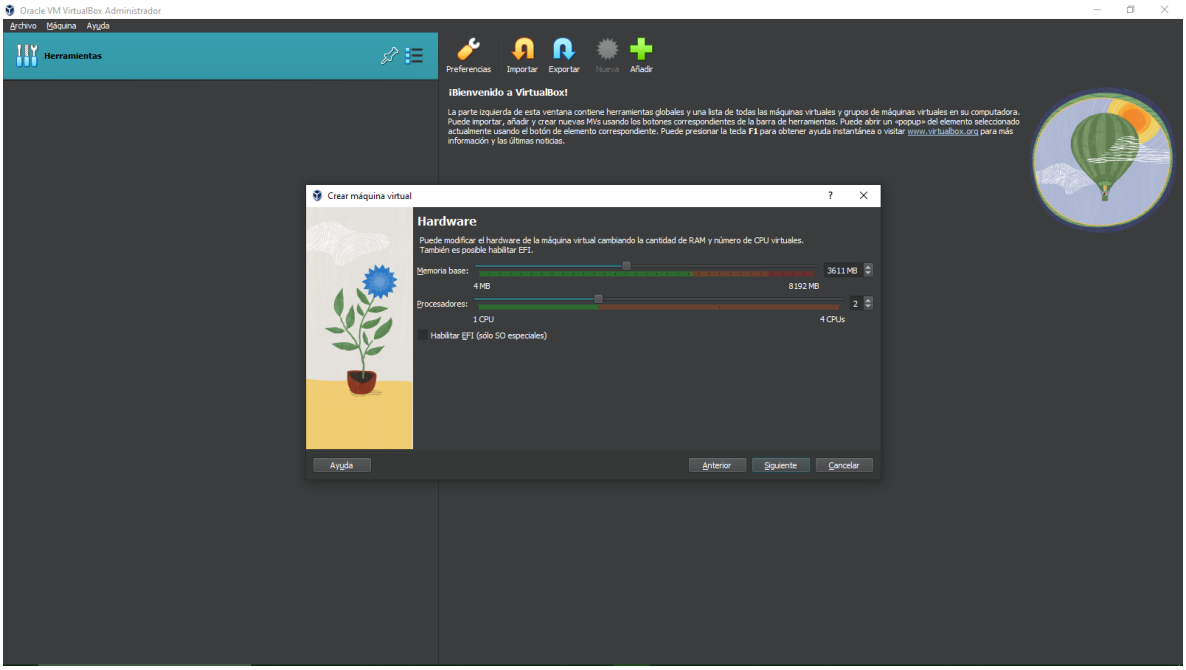



**Imagen 7: Seleccionar memoria ram – Caine**

6. **Capacidad del disco duro Virtual.**

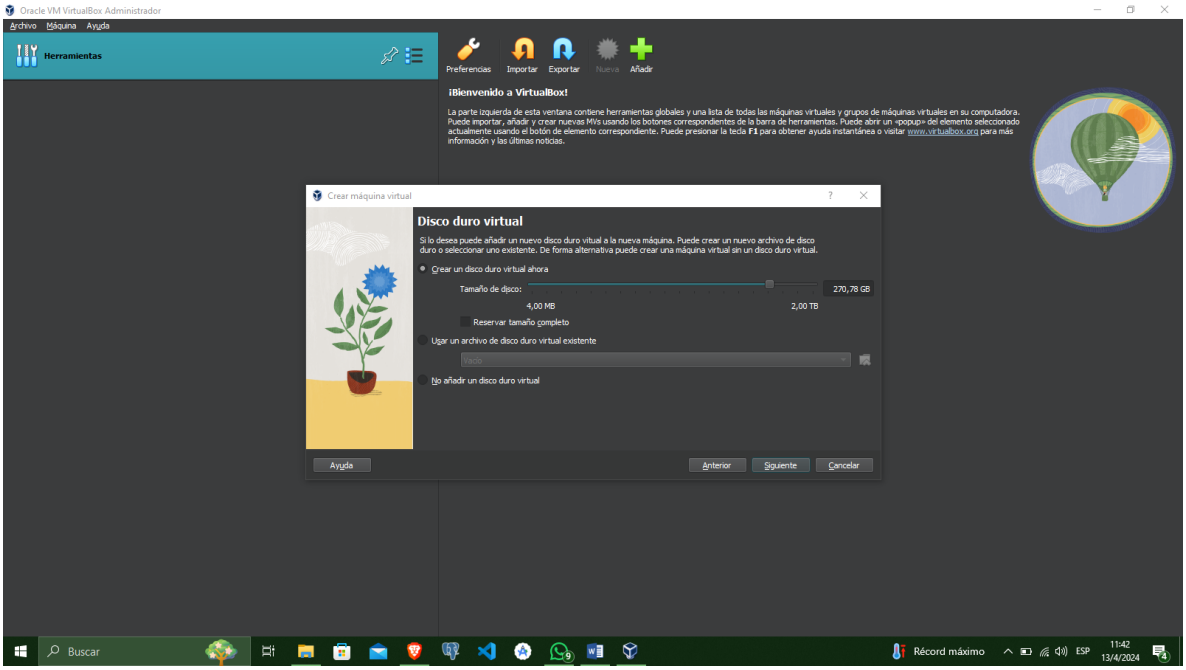



**Imagen 8: Disco Duro Virtual – Caine**

7. **Dar click en terminar la configuración.**

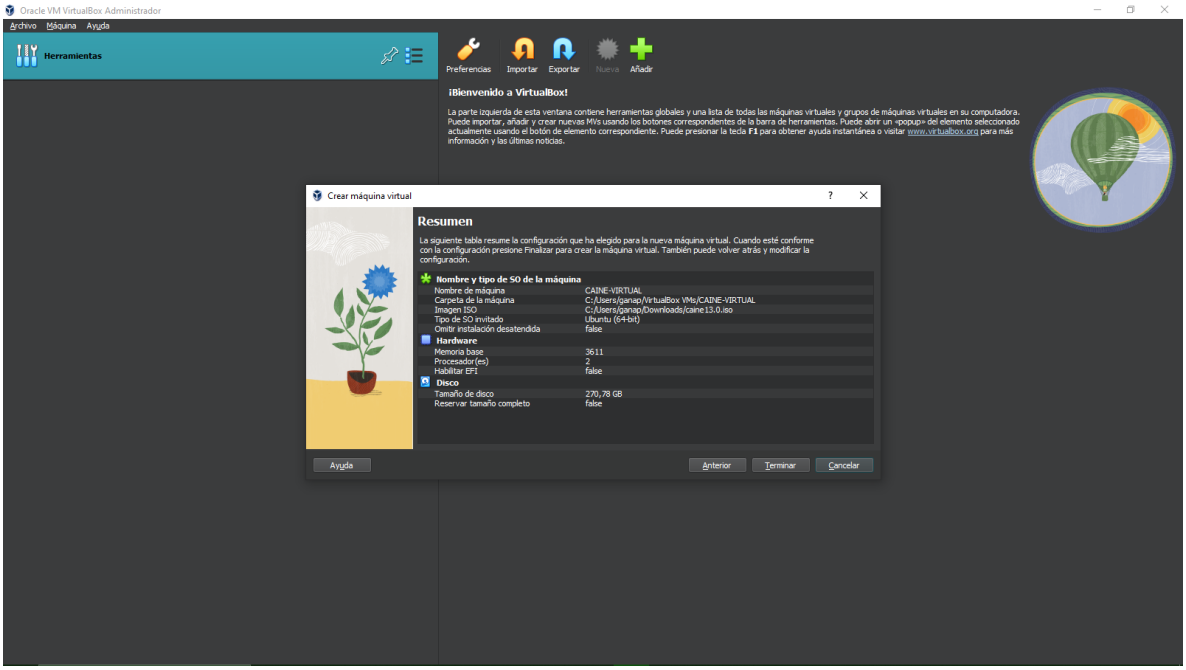



**Imagen 9: Configuración – Caine**

8. **Máquina Virtual creada perfectamente.**


**Imagen 10: Maquina Caine Creada Perfectamente**

9. **Dar click en iniciar para que la maquina encienda y comience a trabajar.**

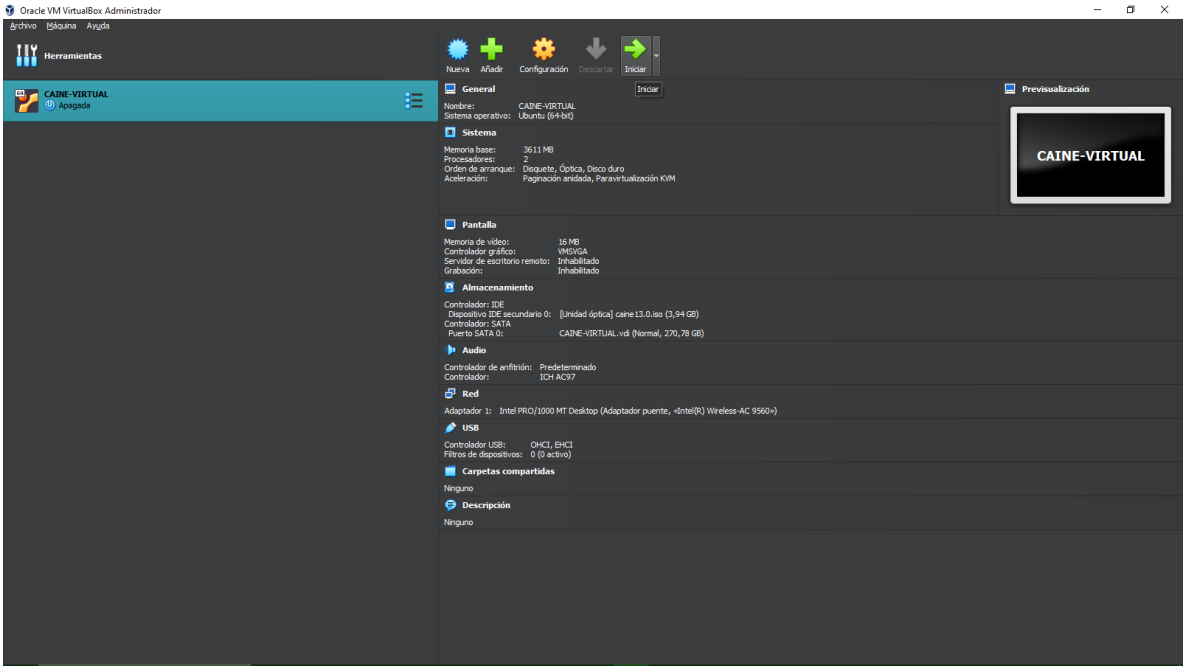



**Imagen 11: Inicio de maquina Caine**

10. **Proceso de inicio de CAINE WARP.**


**Imagen 12: Portal de desarrollo Caine**

**CAINE iniciado perfectamente.**

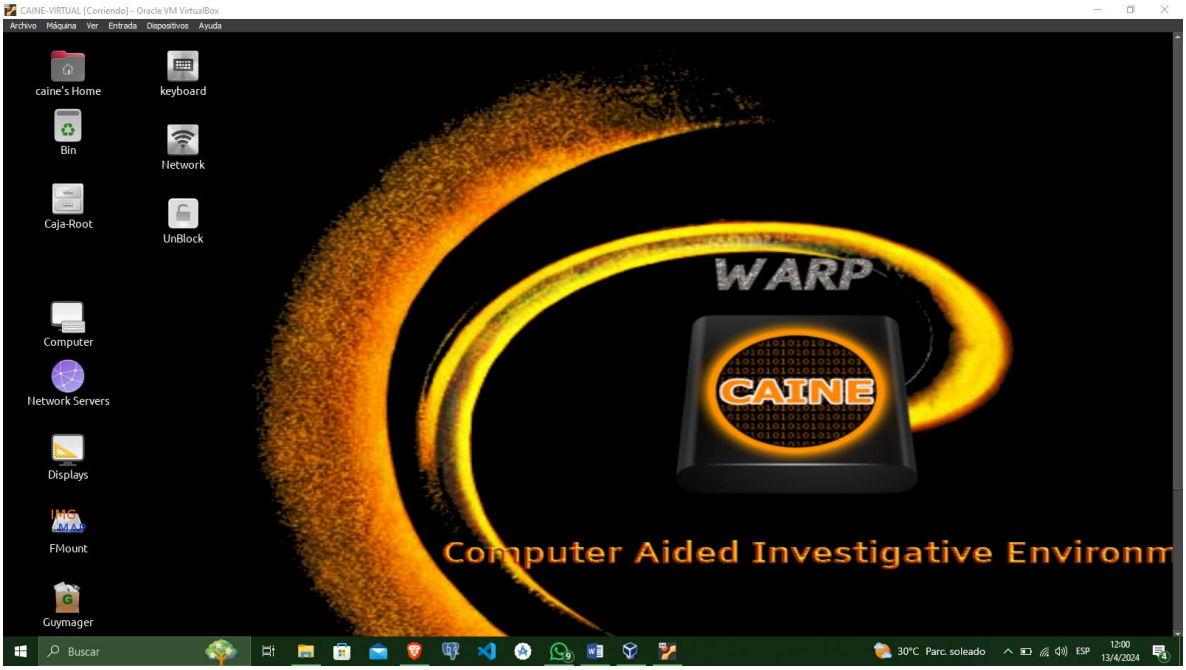



**Imagen 13: Inicio de Caine**

## 11. Herramientas de Forensics Tools para el trabajo de investigación


**Imagen 14: Herramientas de Caine**

# INSTALACIÓN DE KALI LINUX

1. **Visitar la página oficial de Kali Linux en el siguiente Link** https://www.kali.org/

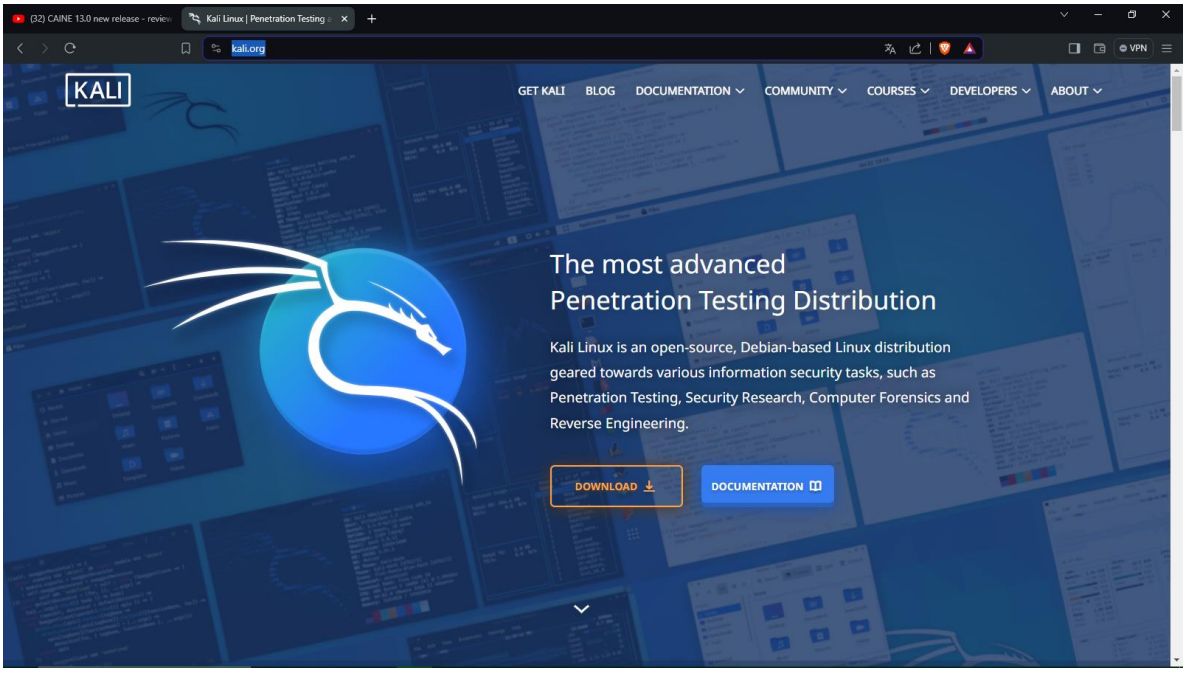



**Imagen 15: Sitio Oficial Kali Linux**

2. **Dar click en Download**

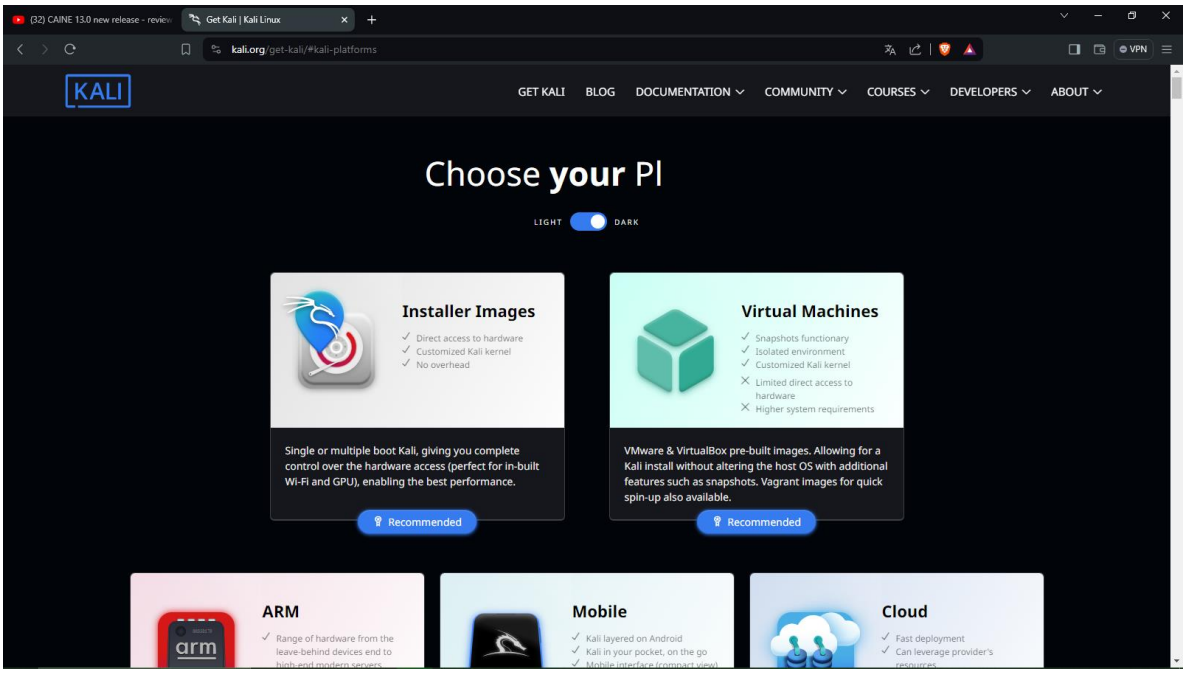



**Imagen 16: Dowloand del Ova de Kali Linux**

3. **Dar click en la sección virtual Machines**

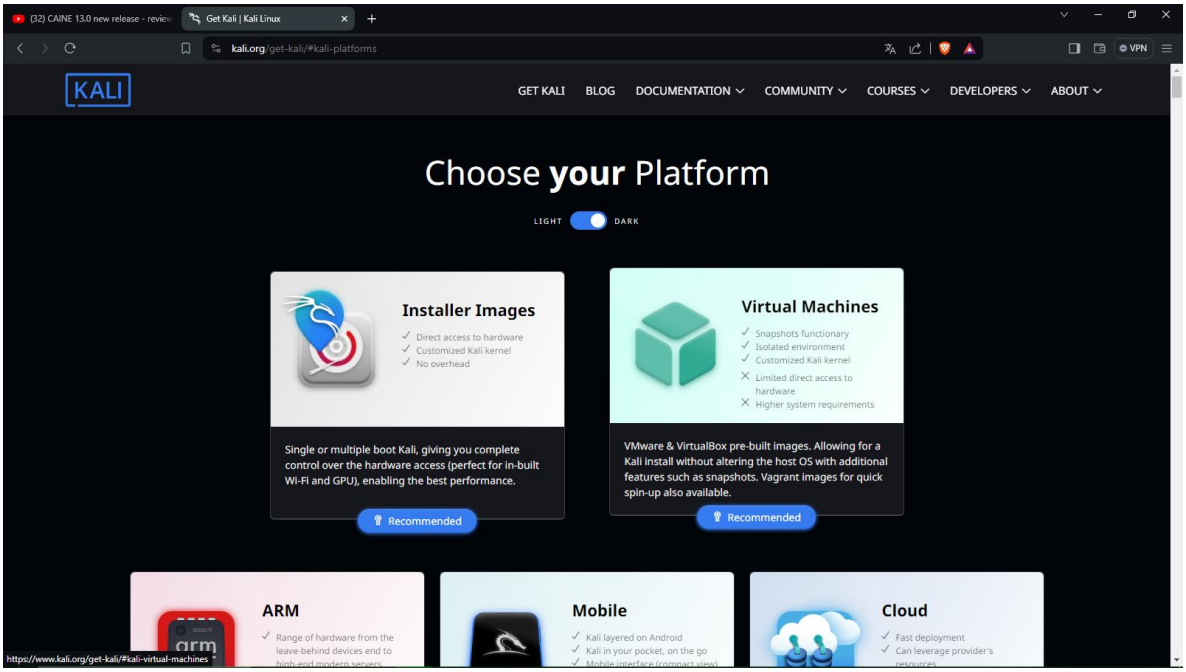



**Imagen 17: Sección Máquina Virtual – Kali Linux**

4. **Aparecen máquinas de soporte, en el caso de estudio es en Windows por ende dar click en la opción VirtualBox en el apartado de la capacidad 2.9 G**

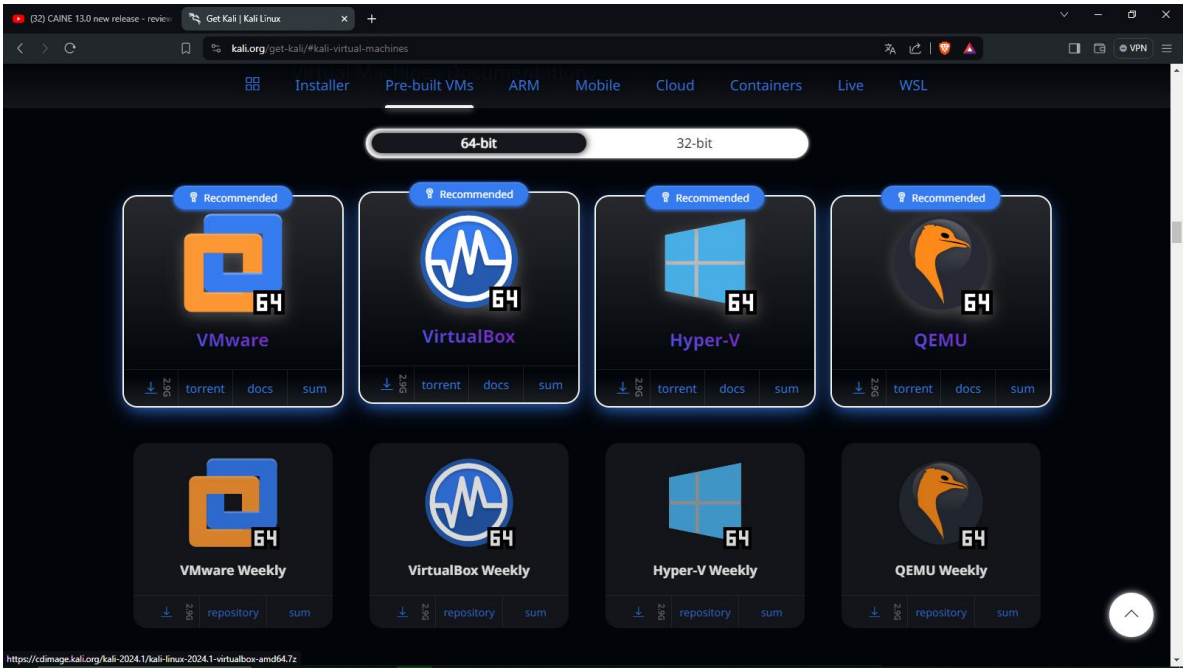



**Imagen 18: OVA Kali Linux para VirtualBox**

5. **Guardar el archivo de descarga que es formato Rar**

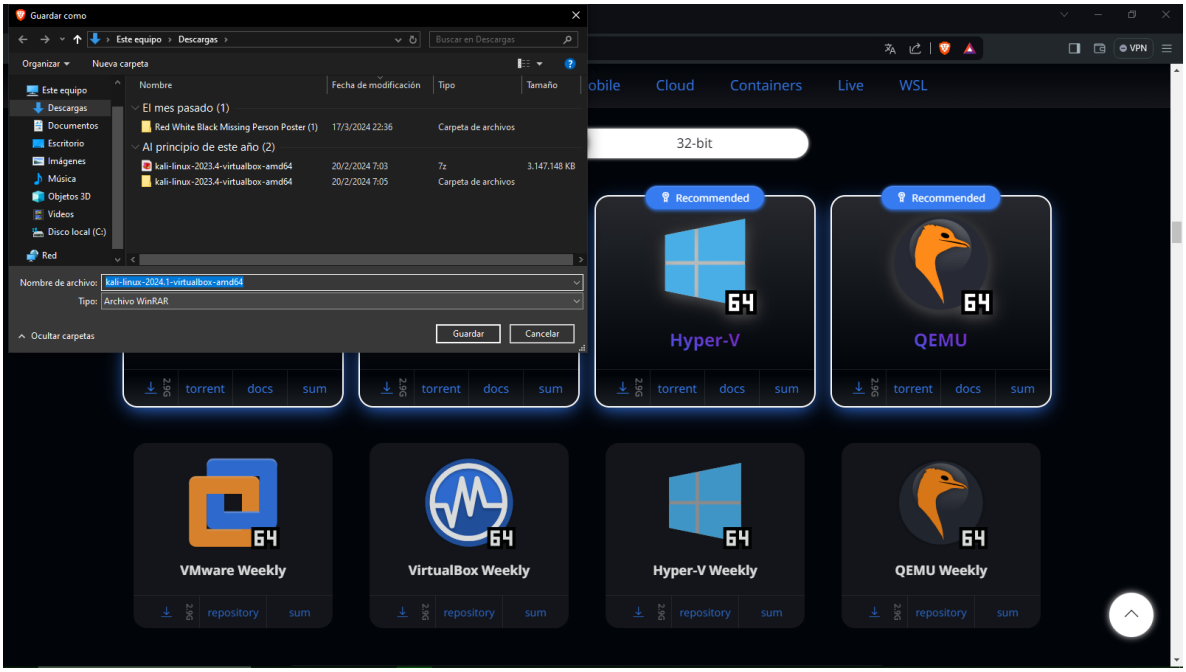



*Imagen 19: Descargar Ova*

6. **Finaliza la descarga extraer la carpeta con los formatos de la máquina que en nuestro caso será el ova**

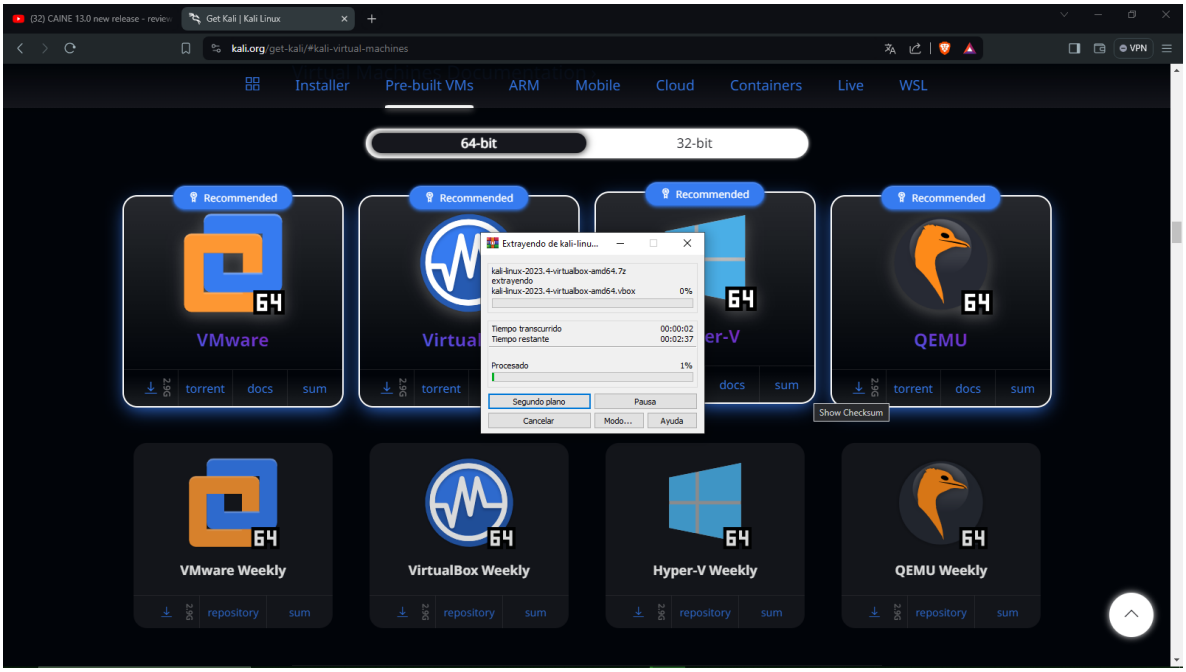



**Imagen 20: Descomprimir archivo Zip para el Ova Kali Linux**

7. **Seleccionar el formato virtual machine para la instalación en virtual box.**

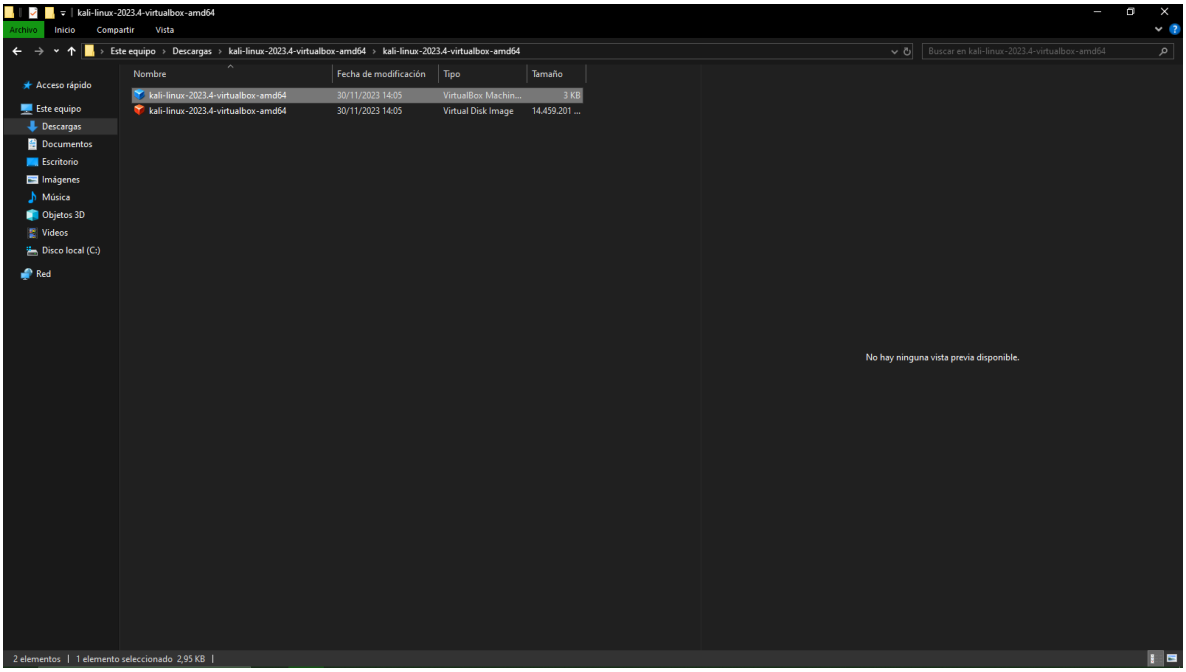



**Imagen 21: Seleccionar el formato de instalación**

8. **Dar click en Añadir.**

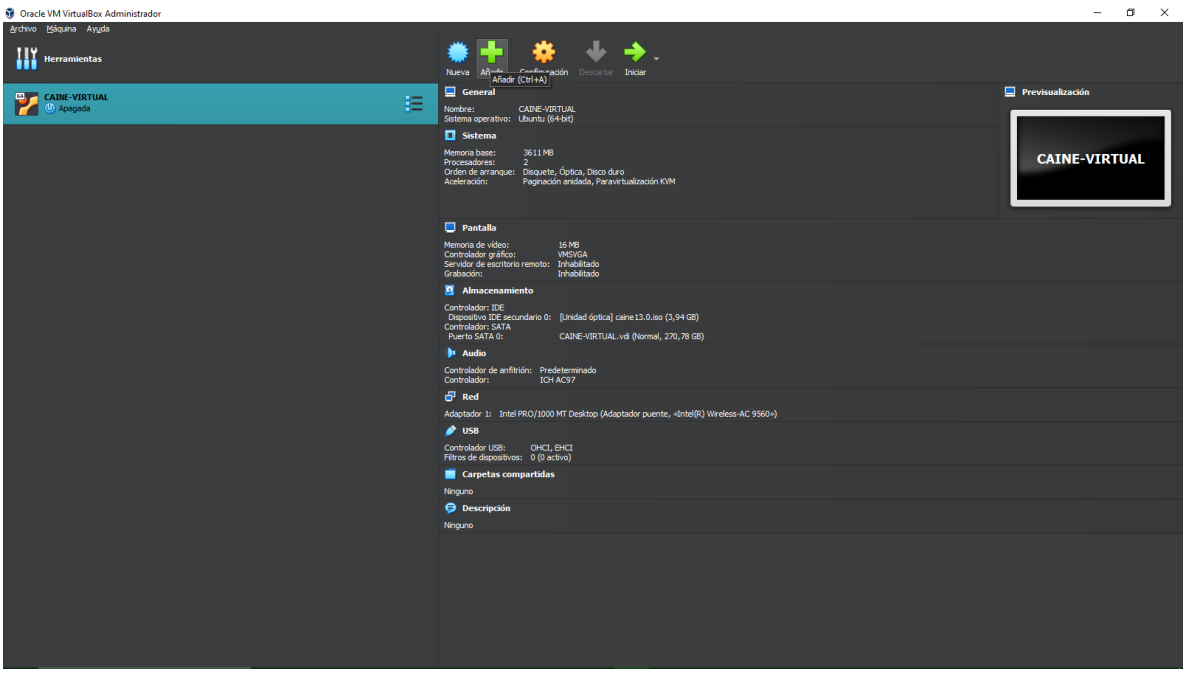



**Imagen 22: Dar en Anadir para la creación de la máquina virtual**

9. **Seleccionar el formato Virtual Machine para crear la maquina virtual de manera rápida con sus preconfiguraciones.**

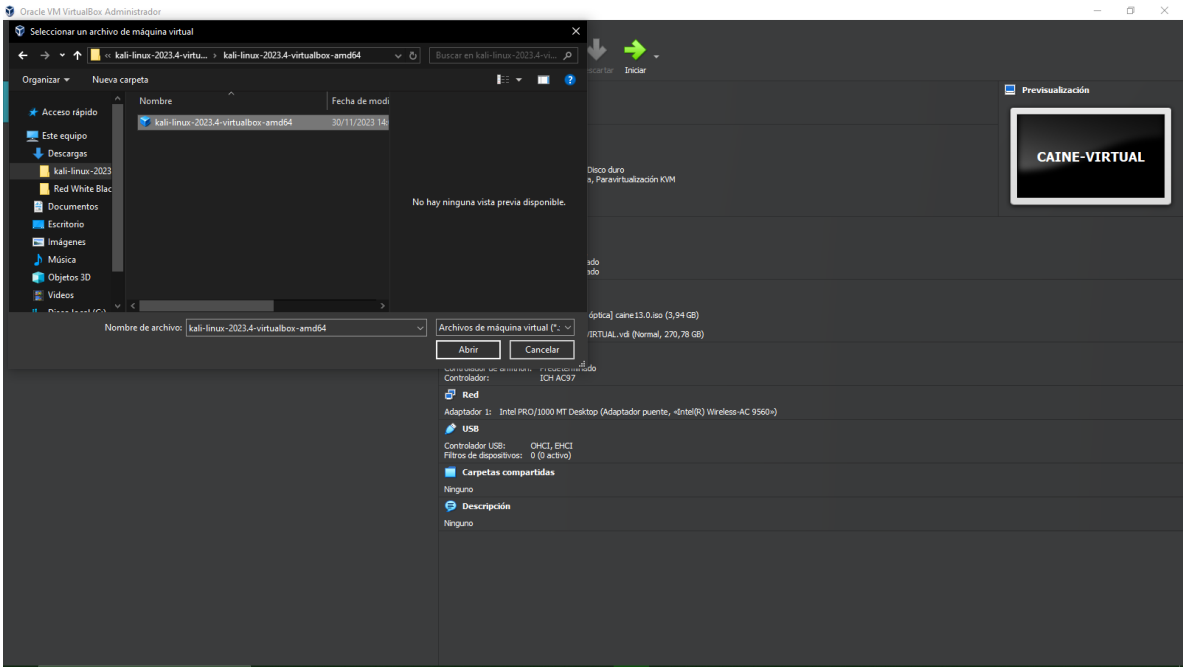



*Imagen 23: Selección Kali Linux*

10. **Máquina Virtual creada exitosamente.**


*Imagen 24: Maquina creada exitosamente*

11. **Dar click en Iniciar para empezar a correr la máquina virtual Kali Linux.**


**Imagen 25: Inicio de Kali Linux**

12. **Presentación de credenciales, como la maquina viene con preconfiguraciones el usuario y contraseña es kali - kali**

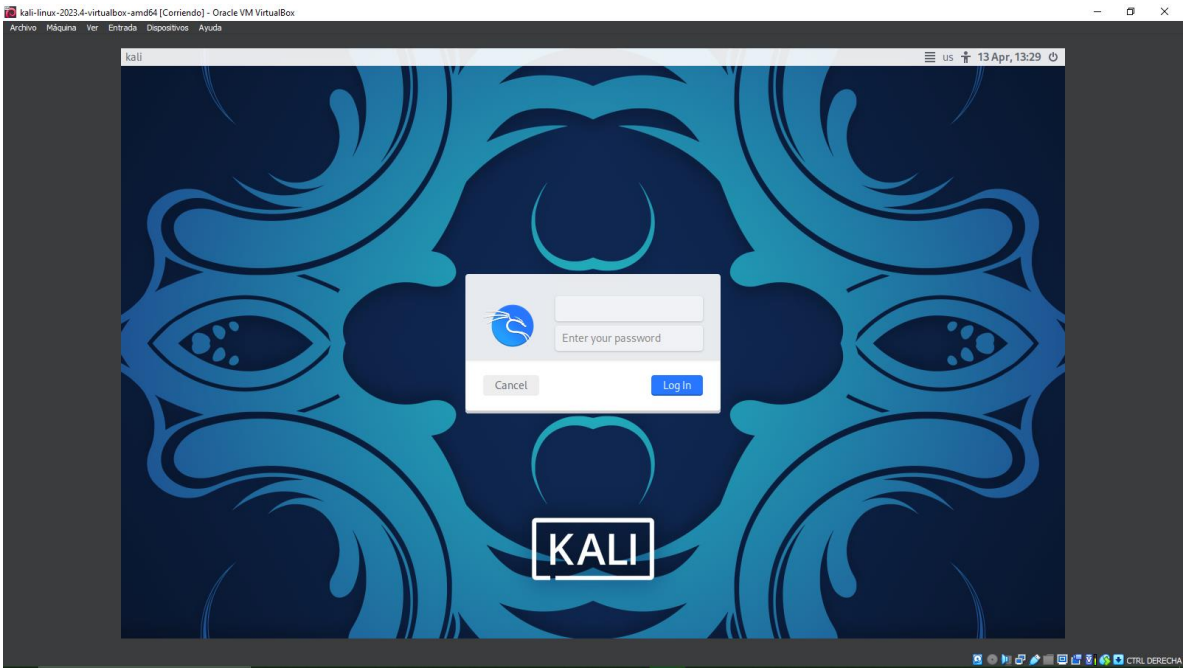



**Imagen 26: Sesión de Kali Linux**

13. **Inicio de la máquina Kali Linux**

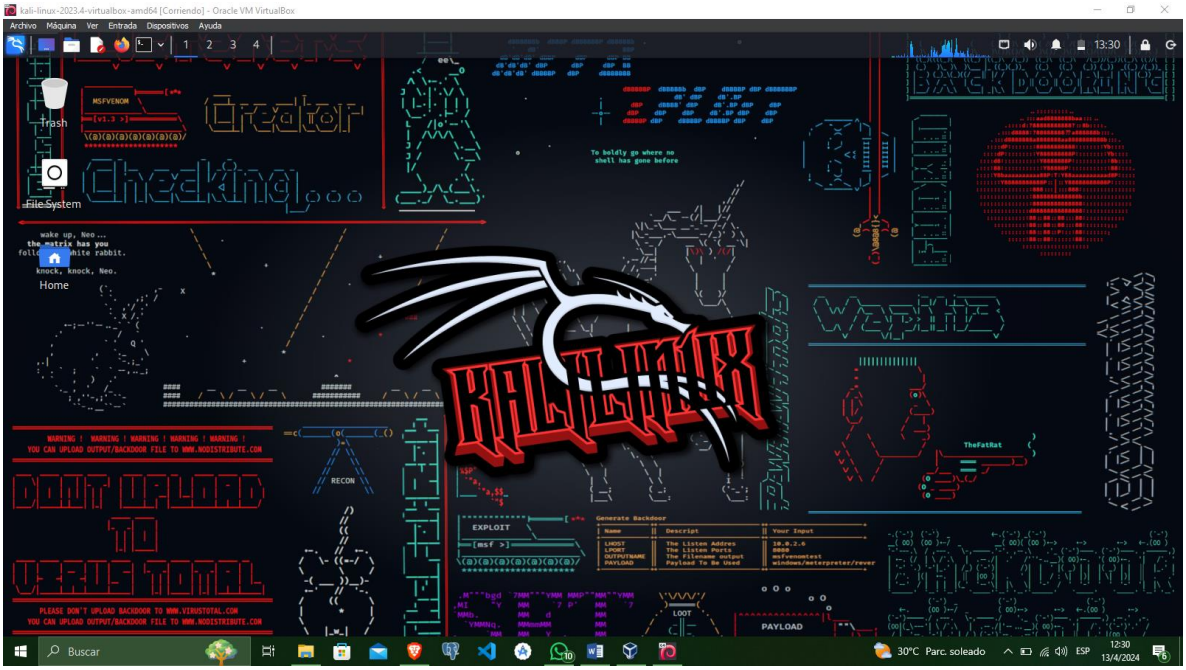


**Imagen 27: Kali Linux**

14. **Herramientas para Forensics**

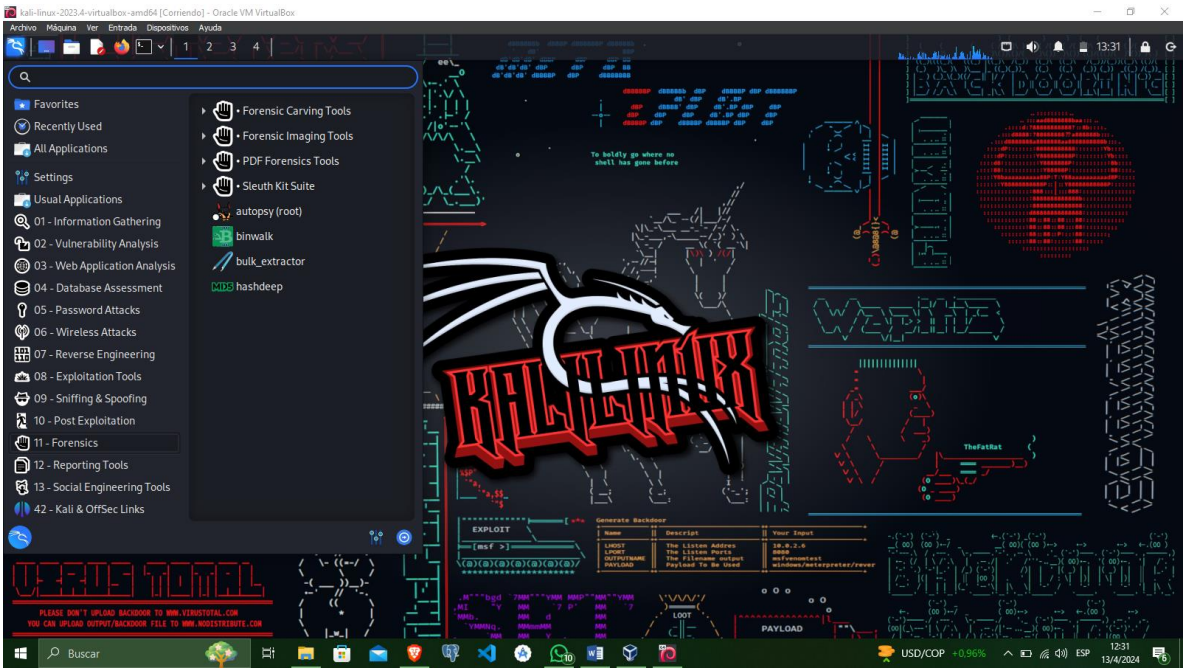



**Imagen 28: Herramientas Kali Linux**

# ANEXO 3:

# GENERACIÓN AUDIO CASOS

## AUDIO ORIGINAL POR CELULAR EMITIDO POR LA GRABADORA

1. **Se realiza establecer a buscar la aplicación grabadora en el dispositivo seleccionado**

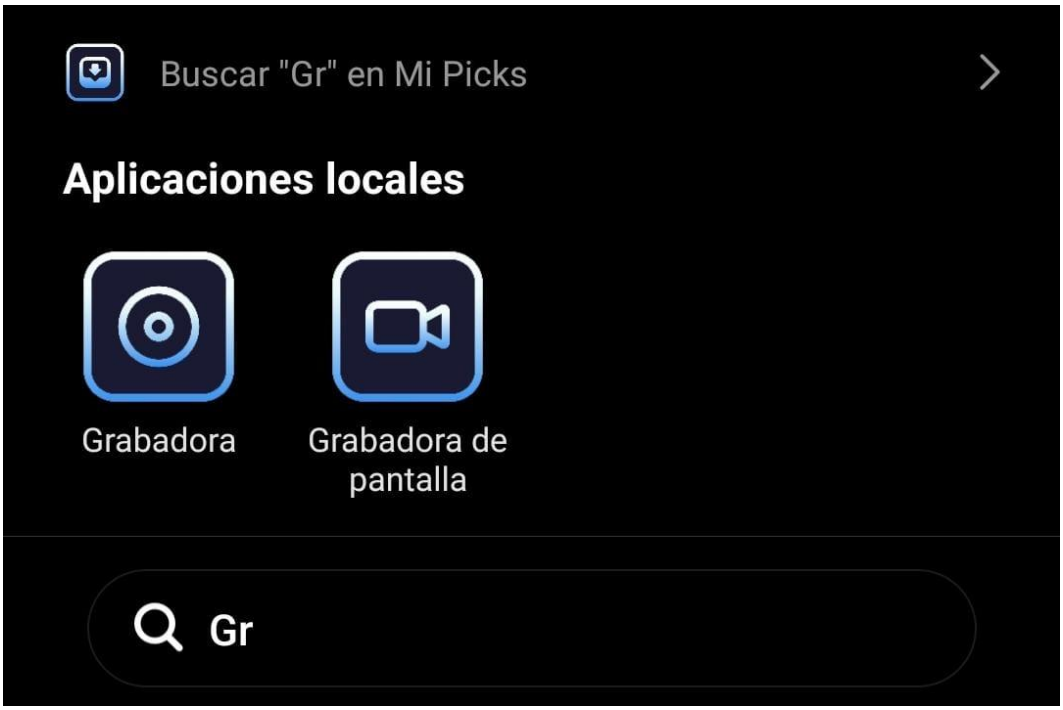



**Imagen 29: Aplicación grabadora - Xiomy**

2. **Se empieza a grabar el audio de manera fluida**

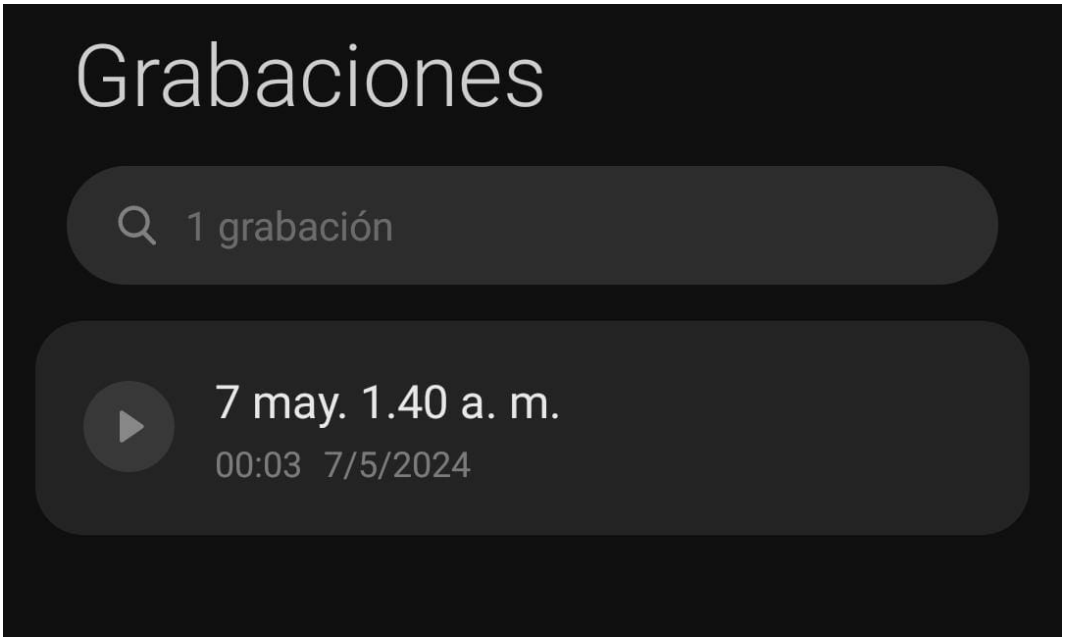



**Imagen 30: Grabación de Audio cumplida**

3. **Se envía el audio por WhatsApp a un chat predeterminado para descargar**


**Imagen 31: Audio enviado por WhatsApp**

4. **Formato descargado en formado Opus**

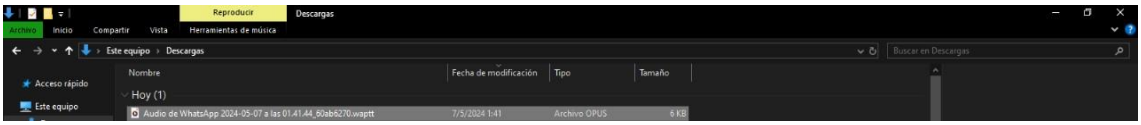



**Imagen 32: Archivo de audio descargado**

5. **Convertir formato Opus a MP3**

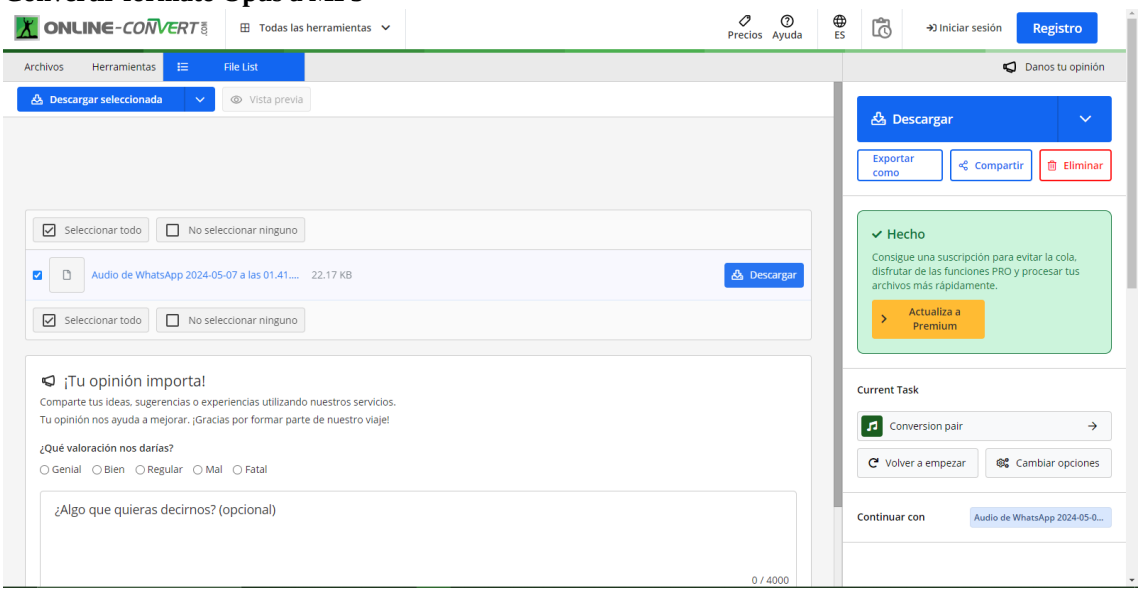



**Imagen 33: Conversión de Opus a Mp3 de la grabación**

6. **Y se establece la evidencia en una carpeta de Análisis para el estudio del proyecto**

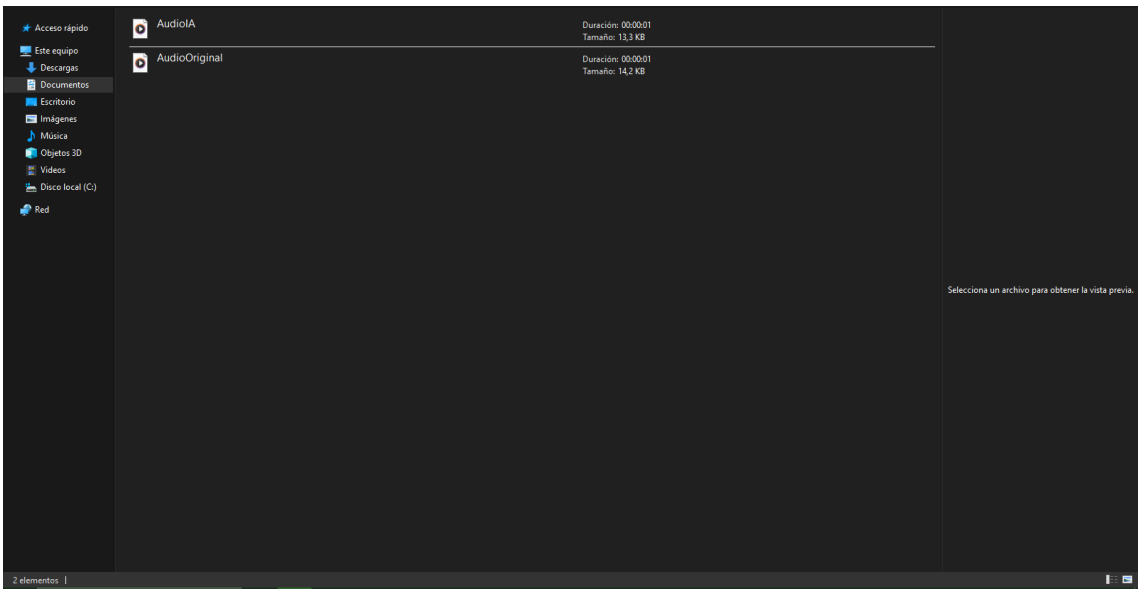



**Imagen 34: Evidencia almacenada**

# CASO 1 - GENERACION DE VOZ (TEXTO A VOZ)

1. Agregar las muestras de audio a utilizar

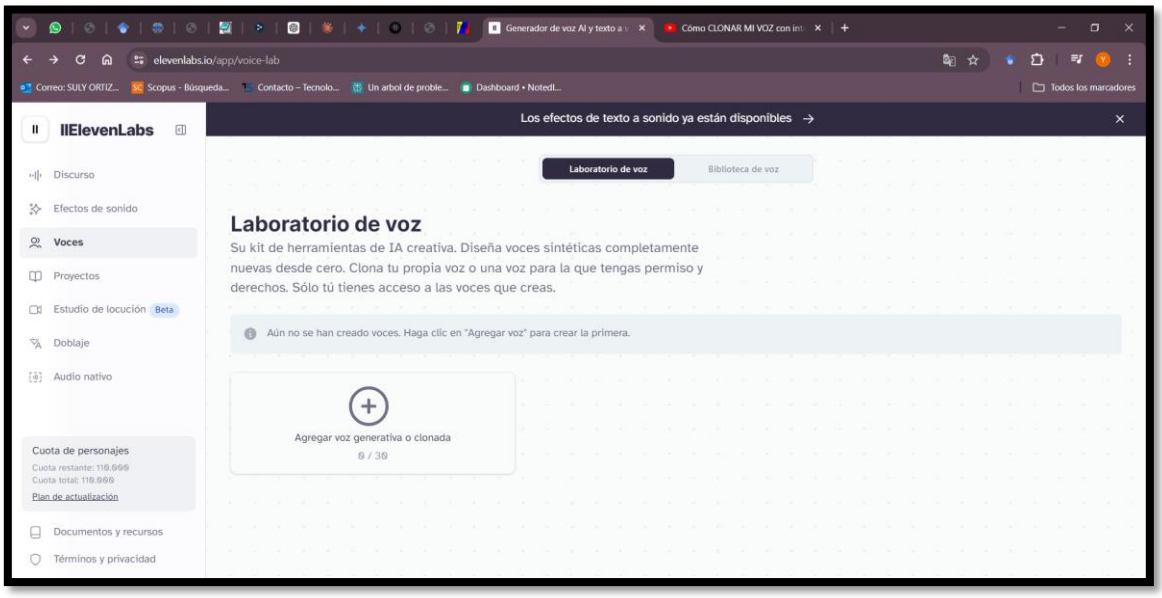



**Imagen 35: Muestras de audio**

2. Seleccionar CLONACION DE VOZ PROFESIONAL, así el audio tendrá una salida con mejor calidad.

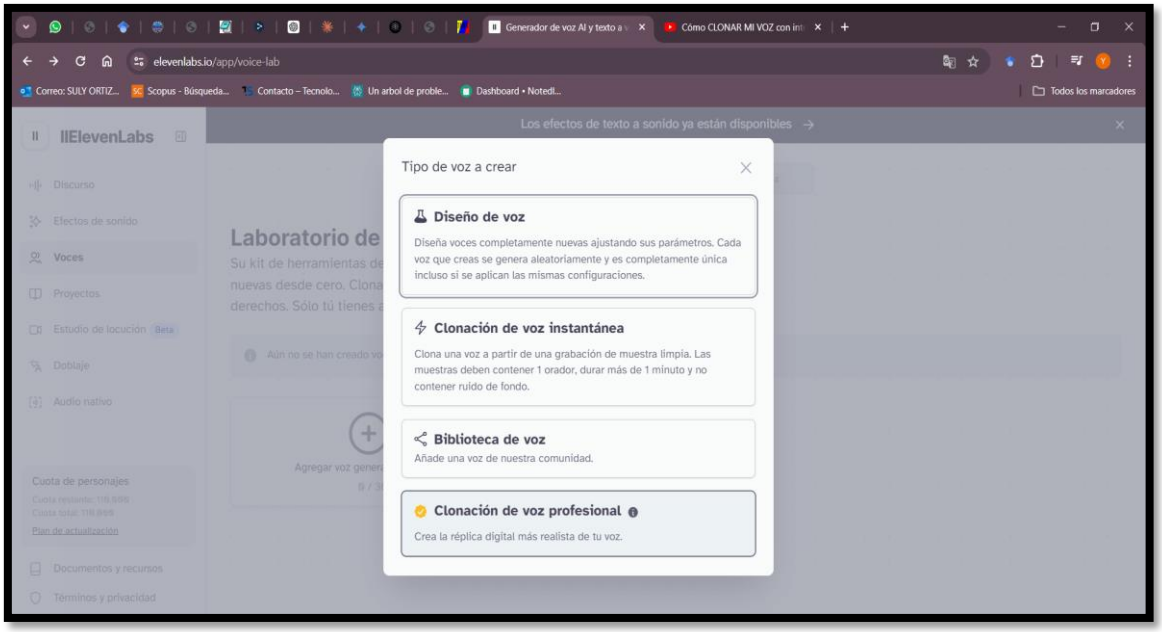



**Imagen 36: Aceptar Clonación Profesional**

3. Tendremos la siguiente información sobre el uso de las muestras de audio a ingresar, si tiene en cuenta que se debe ser muy cuidadoso ya que solo se podrá generar una sola vez nuestra voz profesional.

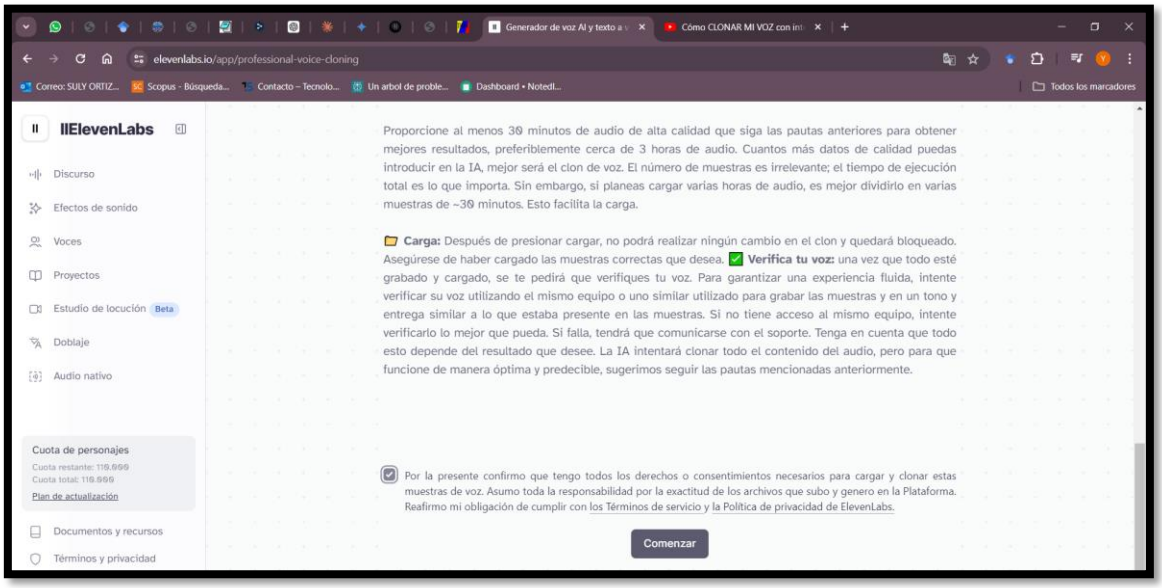



**Imagen 37: Generar una sola muestra profesional**

4. Para la creación de nuestra voz, nos pide ingresar algunos datos => se coloca cualquier nombre en este caso “Mi Voz” => buscamos el idioma correspondiente (Español) => Cargar los archivos de audio

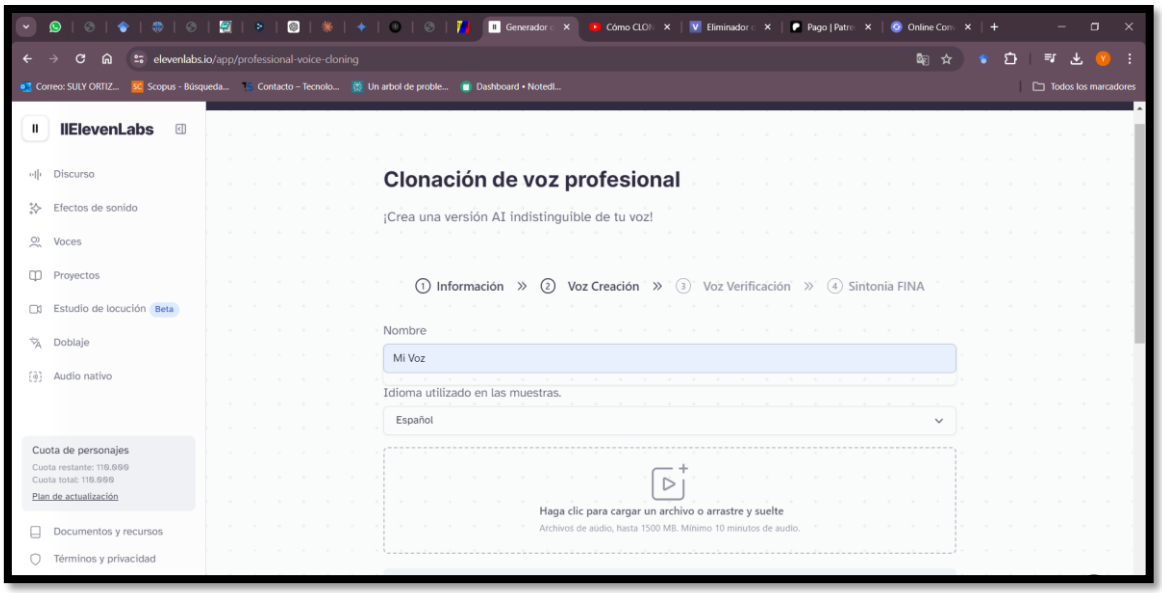



**Imagen 38: Configuración para la creación del audio profesional**

5. Se ven reflejadas las muestras de audios y además agregar una descripción de la voz.

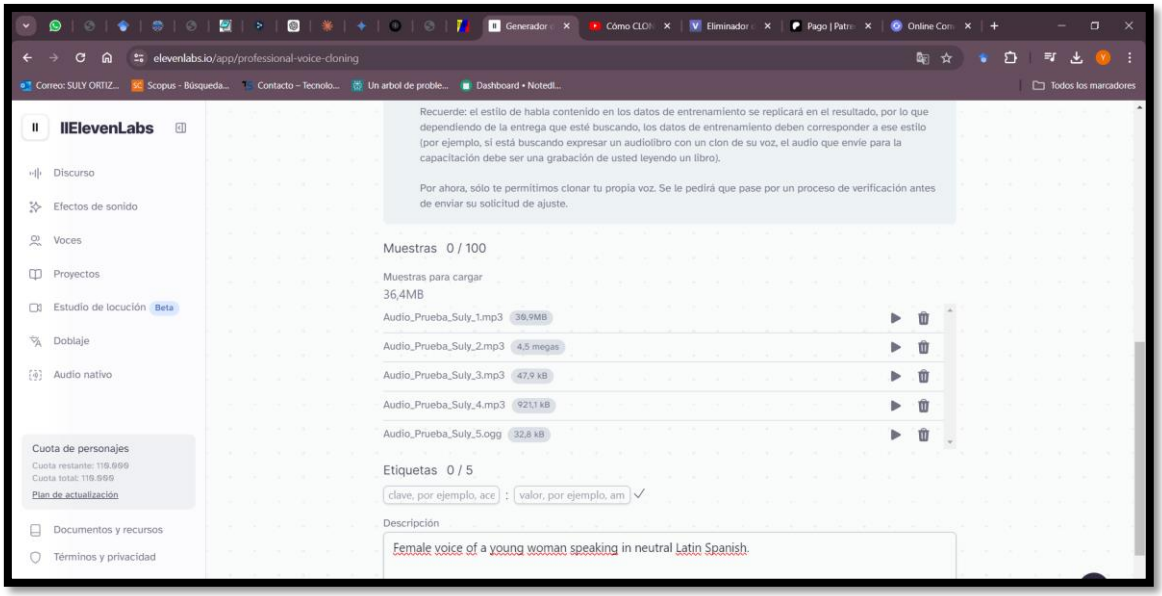



**Imagen 39: Muestras de audios**

6. Se verifica que la voz ingresada coincide con los archivos de audio ingresados.

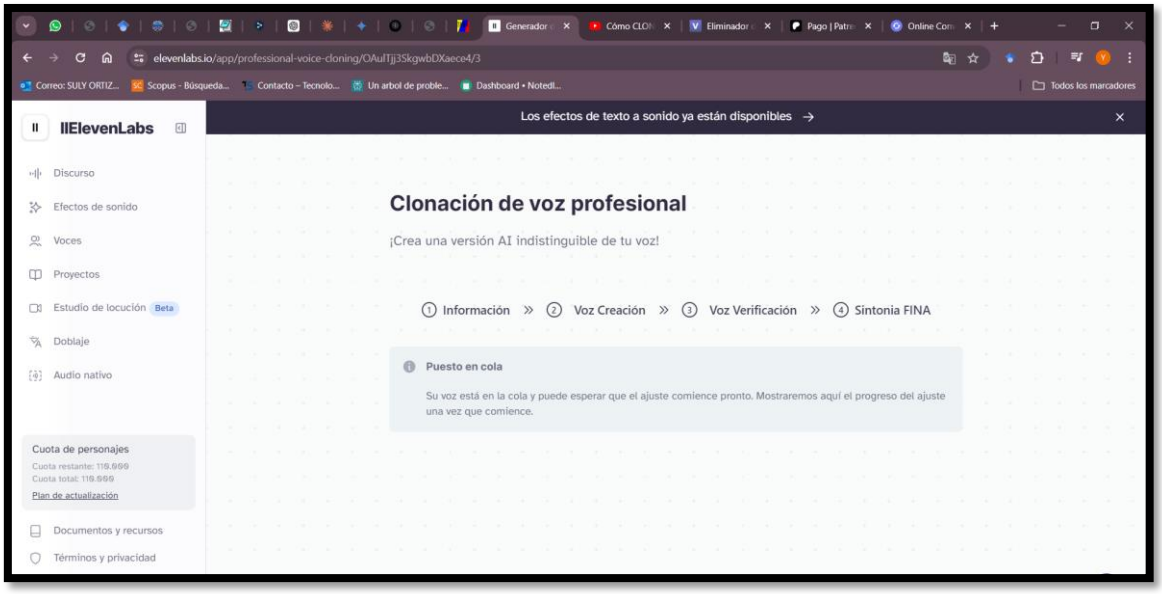



**Imagen 40: Verificación del archivo insertado**

7. Una vez realizado el proceso se espera de 3 a 5 minutos que la voz profesional esta lista para ser usada

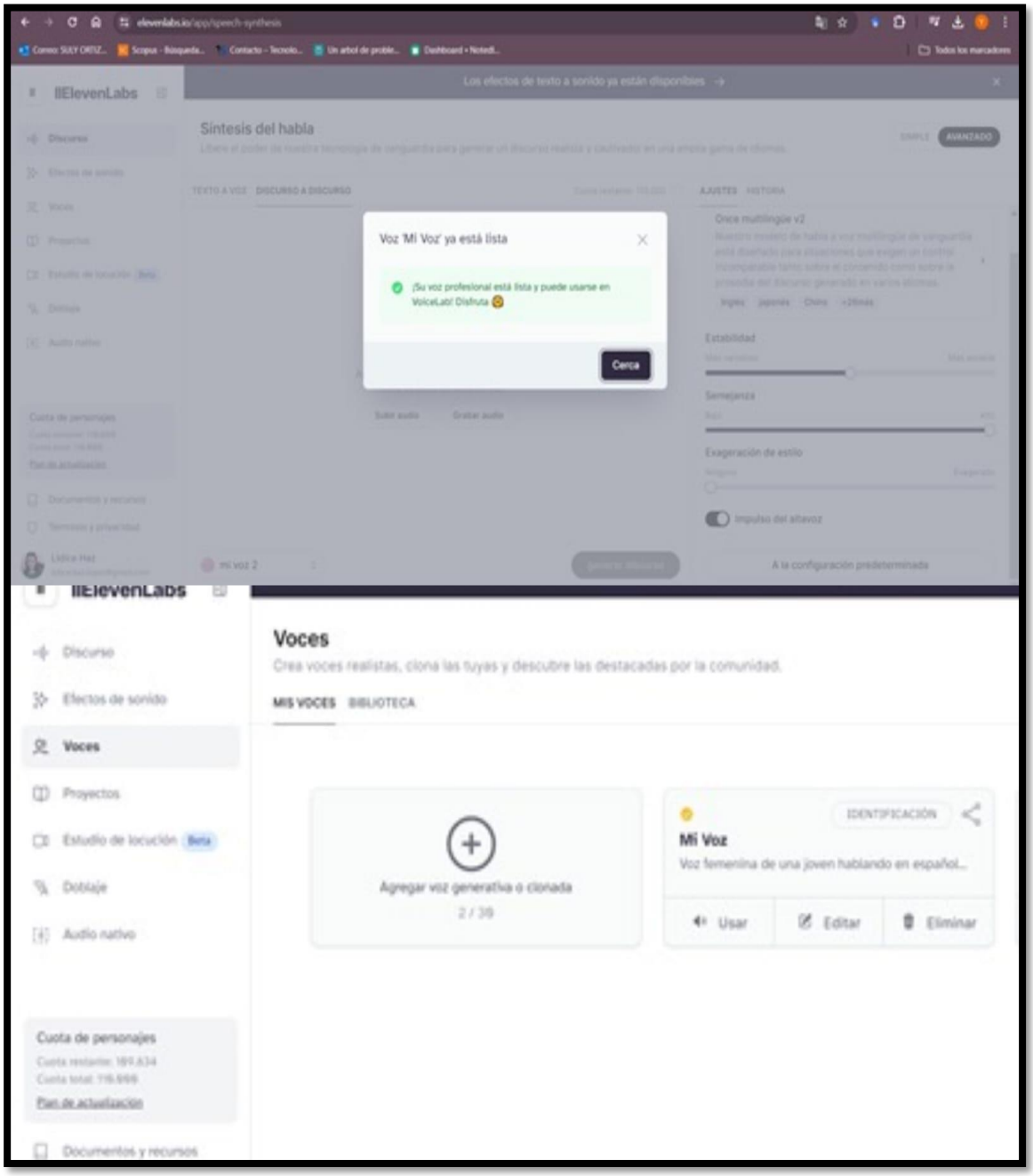



**Imagen 41: Proceso de creación del audio Profesional**

8. Agregar el texto que deseamos que se genere en el audio mediante las muestras anteriormente entrenadas -> Presionamos generar discurso, realizamos ajustes si es necesario y descargamos el audio creado.

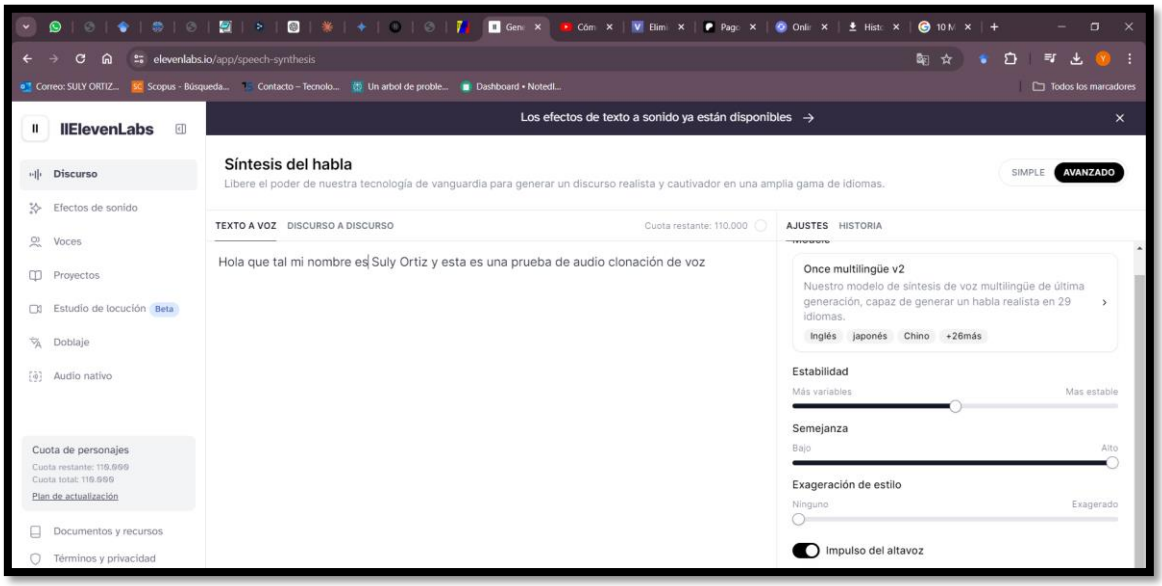



**Imagen 42: Insertar texto a generar por el audio**

# CASO 2 - VOZ A VOZ (CLONACION DE VOZ)

1. Para realizar la clonación de voz seleccionamos USAR

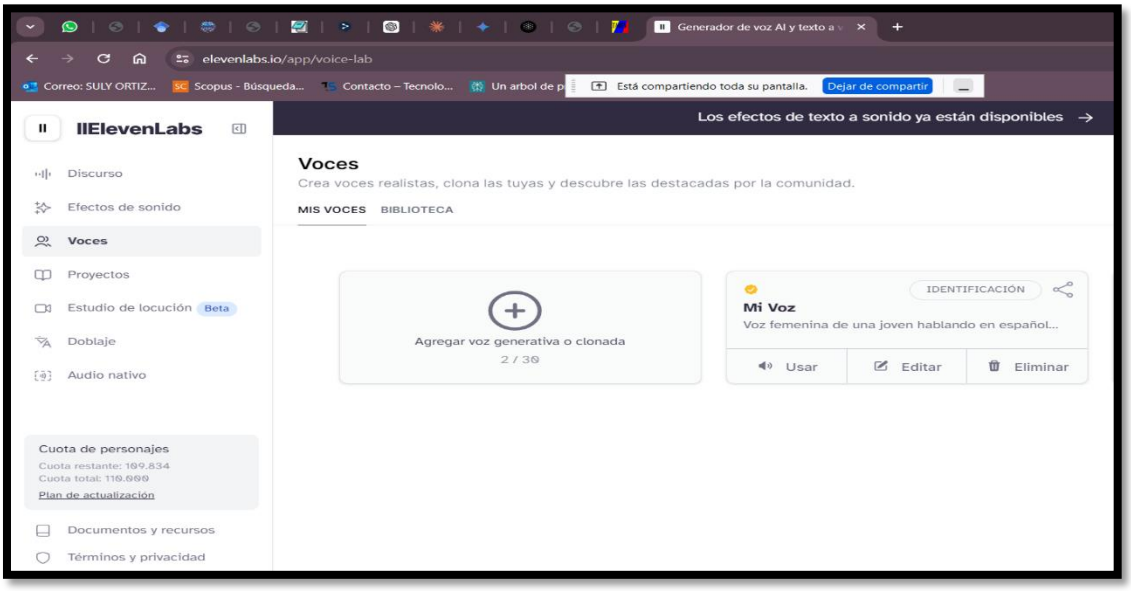



**Imagen 43: Seleccionar el audio a usar para clonar**

2. Nos dirigimos a la opción de DISCURSO A DISCURSO, podemos elegir subir audio o grabar audio.

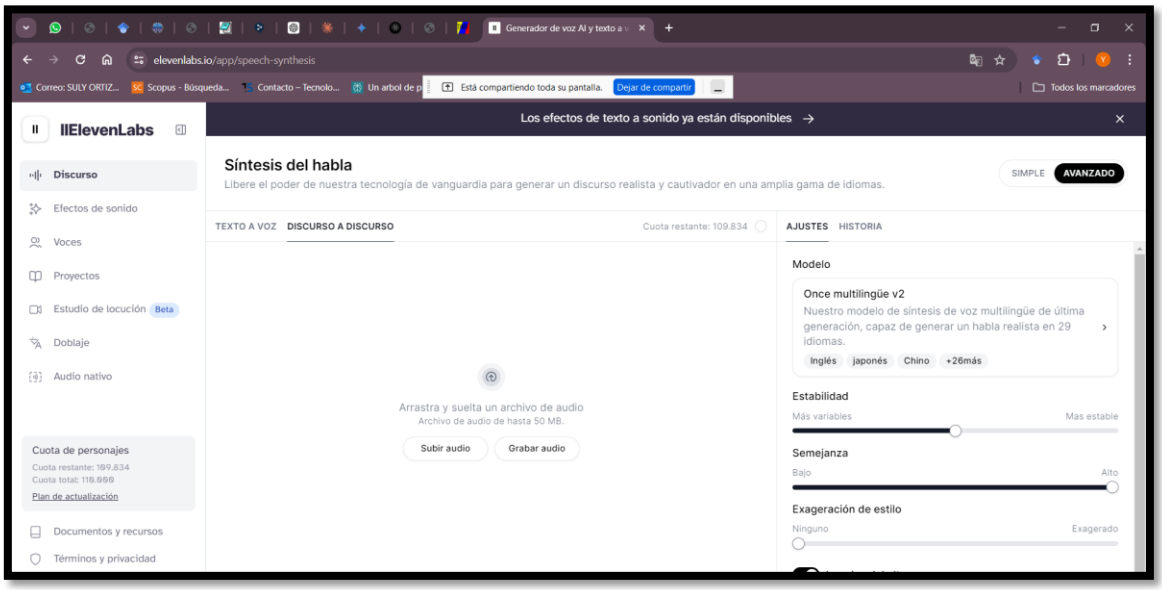



**Imagen 44: Discurso a Discurso**

3. En este caso se utilizó grabar audio, procedemos a generar el discurso y descargamos el audio generado por la IA.

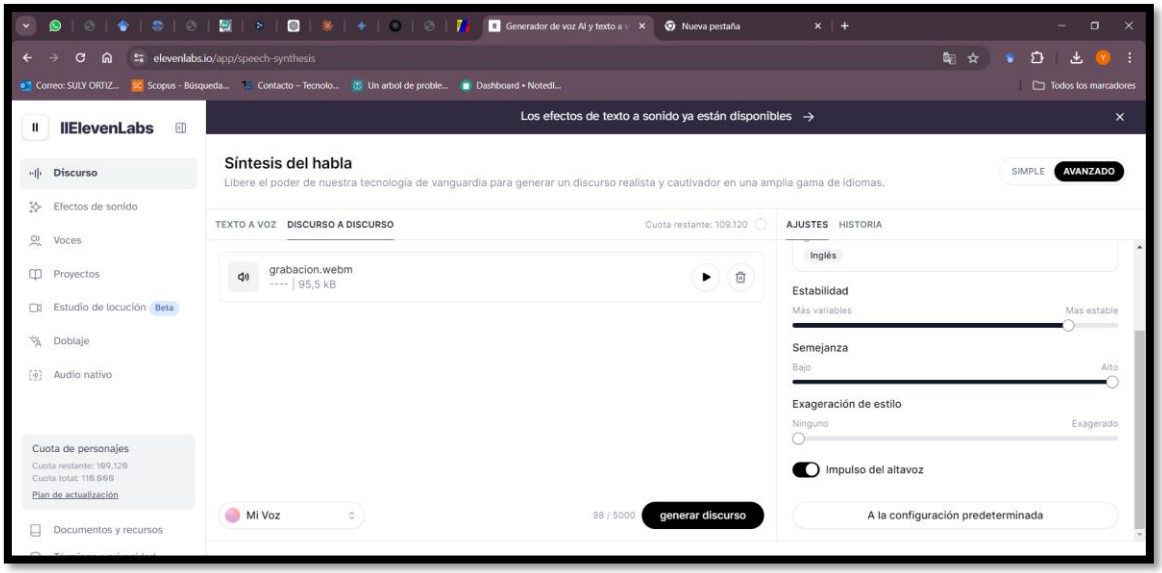



**Imagen 45: Descargar audio generado por IA**

# ANEXO 4: FASE ANÁLISIS

## ANÁLISIS COEFICIENTE MFCC (AUDIO ORI – AUDIO IA DEEPFAKE) POR PYTHON

1. **Importaciones de Bibliotecas**

```
import librosa
import librosa.display
import matplotlib.pyplot as plt
import numpy as np
```

**Imagen 46: Importación de Bibliotecas en Python**

2. **Definición de la Función "calcular_mfcc" , la función toma la ruta del archivo como entrada y calcula los coeficientes MFCC utilizando "librosa .freature.mfcc"**

```
# Función para calcular los coeficientes MFCC
def calcular_mfcc(ruta_audio):
    audio, sr = librosa.load(ruta_audio, sr=None)
    mfcc = librosa.feature.mfcc(y=audio, sr=sr, n_mfcc=13)
    return mfcc
```

**Imagen 47: Creación de la Función Mfcc para calcular los coeficientes de los audios muestras**

3. **Se define las rutas de los archivos de audio que se van a procesar**

```
# Rutas de los archivos de audio
ruta_audio1 = 'Audios/AudioOriginal.mp3'
ruta_audio2 = 'Audios/AudioIA.mp3'
```

**Imagen 48: Definición de las rutas de las muestras**

4. **Se calculan los coeficientes MFCC para ambos archivos de audio utilizando la función calcular_mfcc.**

```
# Calcular coeficientes MFCC para ambos audios
mfcc_audio1 = calcular_mfcc(ruta_audio1)
mfcc_audio2 = calcular_mfcc(ruta_audio2)
```

**Imagen 49: Definición de cálculo de los archivos muestras**

5. **Graficar de los coeficientes MFCC para ambos archivos de audio**

```
# Graficar los coeficientes MFCC para ambos audios
plt.figure(figsize=(12, 6))

plt.subplot(2, 1, 1)
librosa.display.specshow(mfcc_audio1, x_axis='time', cmap='viridis')
plt.colorbar()
plt.title('Coeficientes MFCC - Audio 1')
plt.xlabel('Tiempo (s)')
plt.ylabel('Coeficiente MFCC')

plt.subplot(2, 1, 2)
librosa.display.specshow(mfcc_audio2, x_axis='time', cmap='plasma')
plt.colorbar()
plt.title('Coeficientes MFCC - Audio 2')
plt.xlabel('Tiempo (s)')
plt.ylabel('Coeficiente MFCC')

plt.tight_layout()
plt.show()
```

**Imagen 50: Proceso de graficar los coeficientes MFCC para los archivos muestras**

6. **Grafico resultante del codigo**

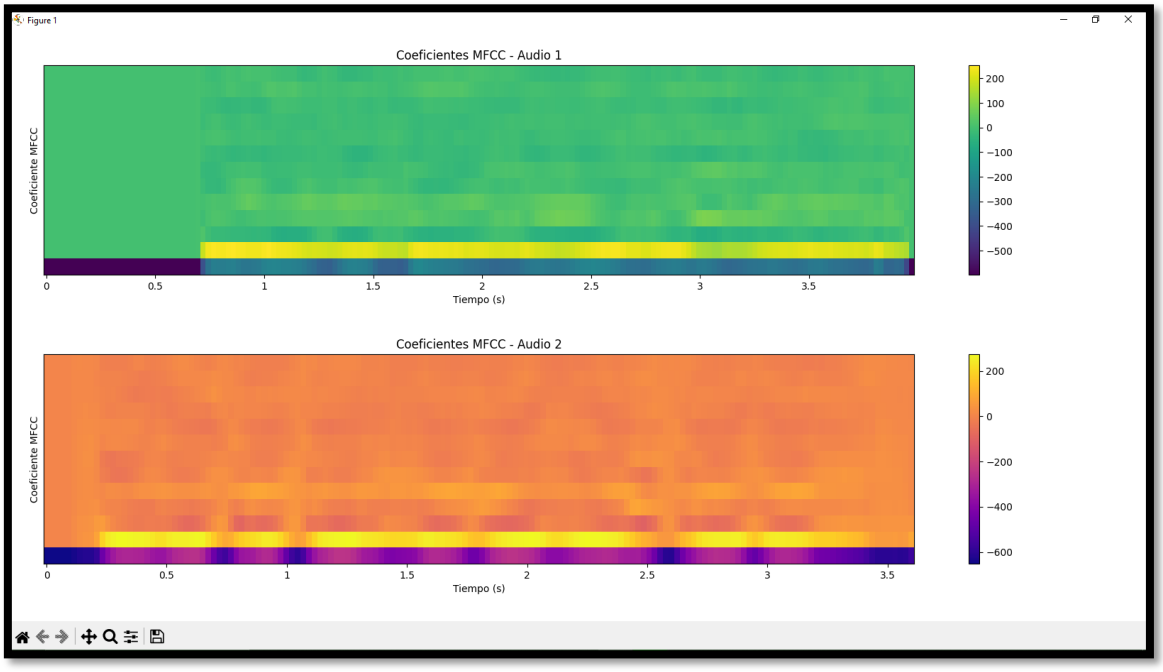



**Imagen 51: Grafico Comparativo de los archivos muestras**

# ANÁLISIS ESPECTRAL (AUDIO ORI)

## HERRAMIENTA: AUDACITY

1. **Dar clic en la opción Archivo y seleccionar la opción importar en la opción audio**

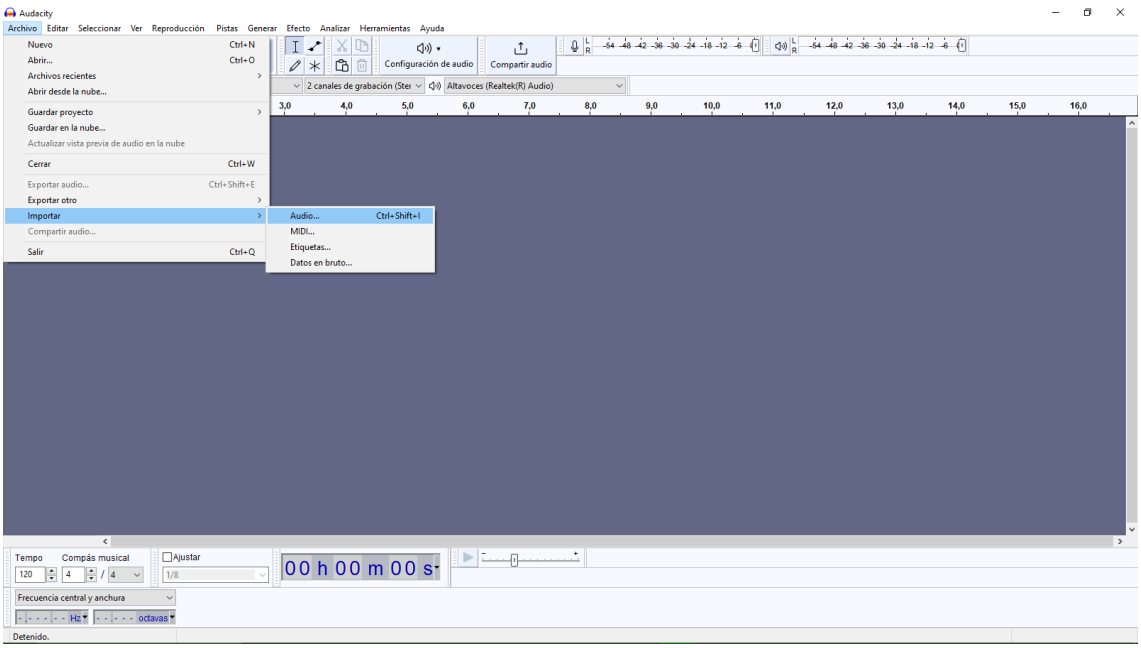



**Imagen 52: Herramienta Audacity**

2. **Seleccionar el audio correspondiente**

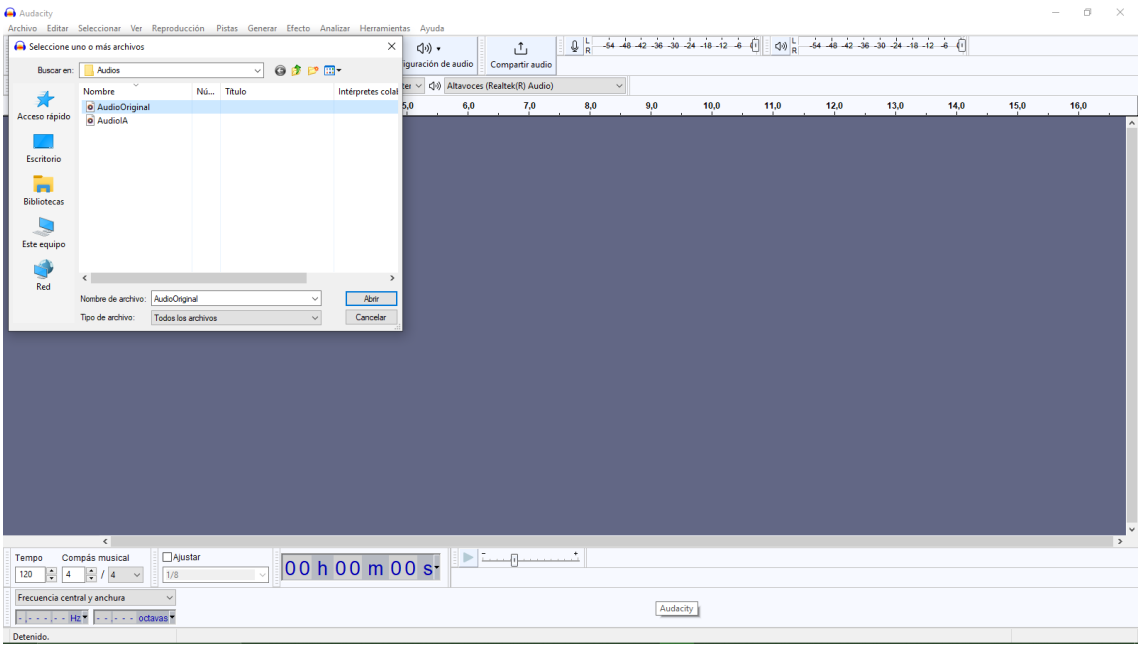



**Imagen 53: Seleccionar el archivo muestra**

3. **Visualizar el archivo seleccionador**

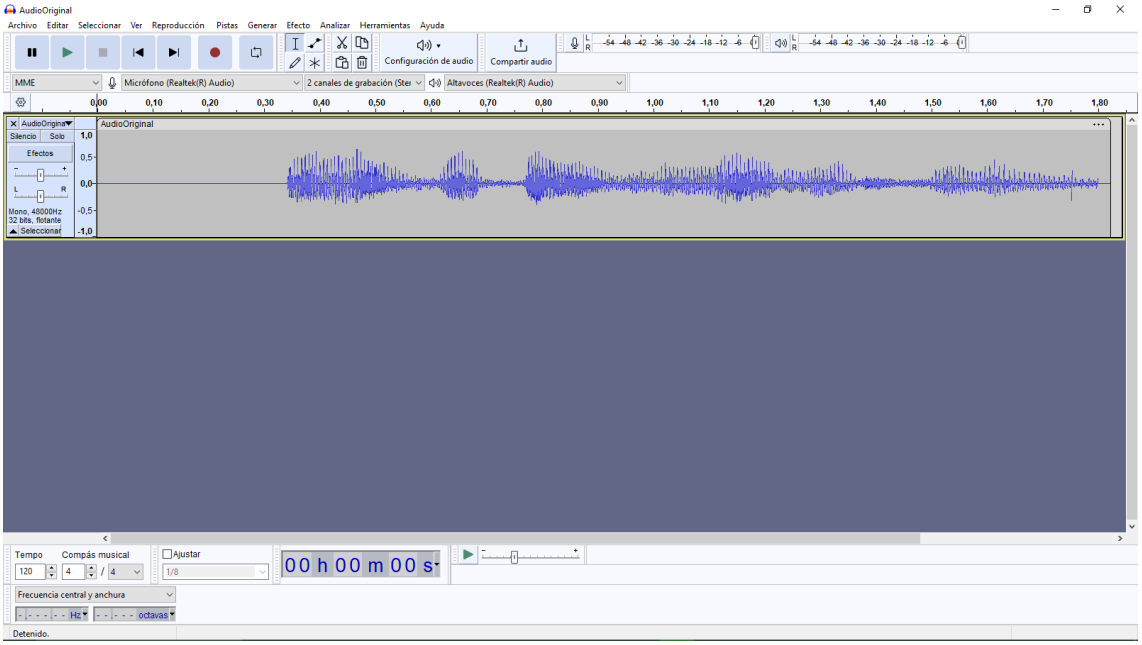



**Imagen 54: Vista Previa de la muestra**

4. **Crtl+A permite seleccionar el audio para realizar el proceso de análisis**

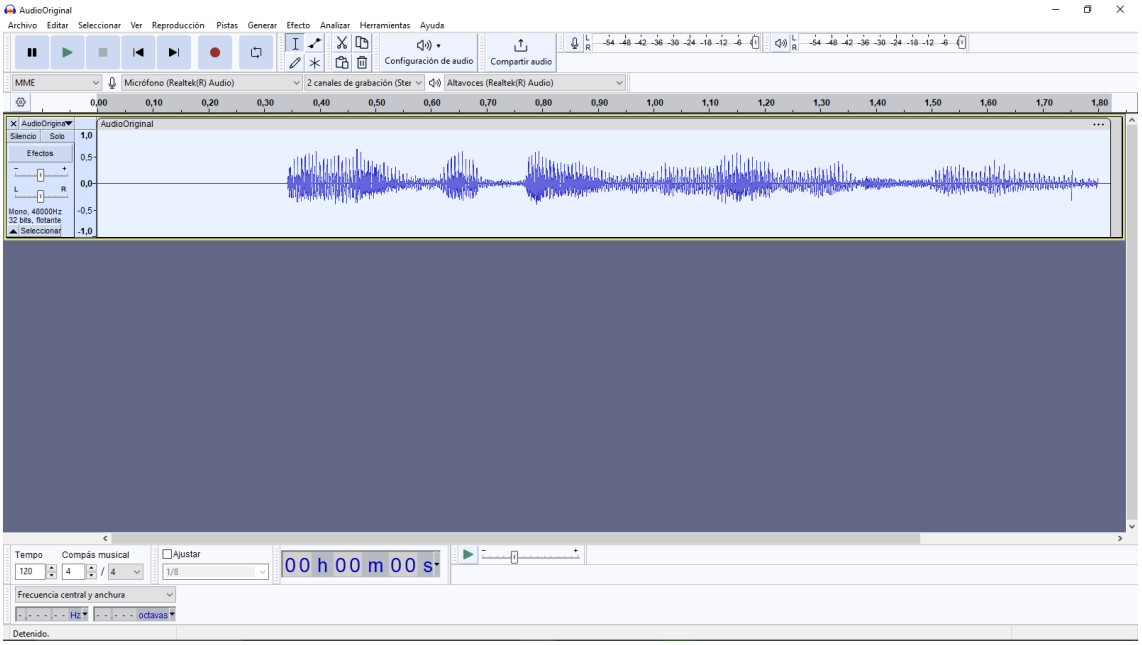



**Imagen 55: Seleccionar la muestra para el análisis**

5. **En el apartado Analizar dar click en la opción trazar espectro**

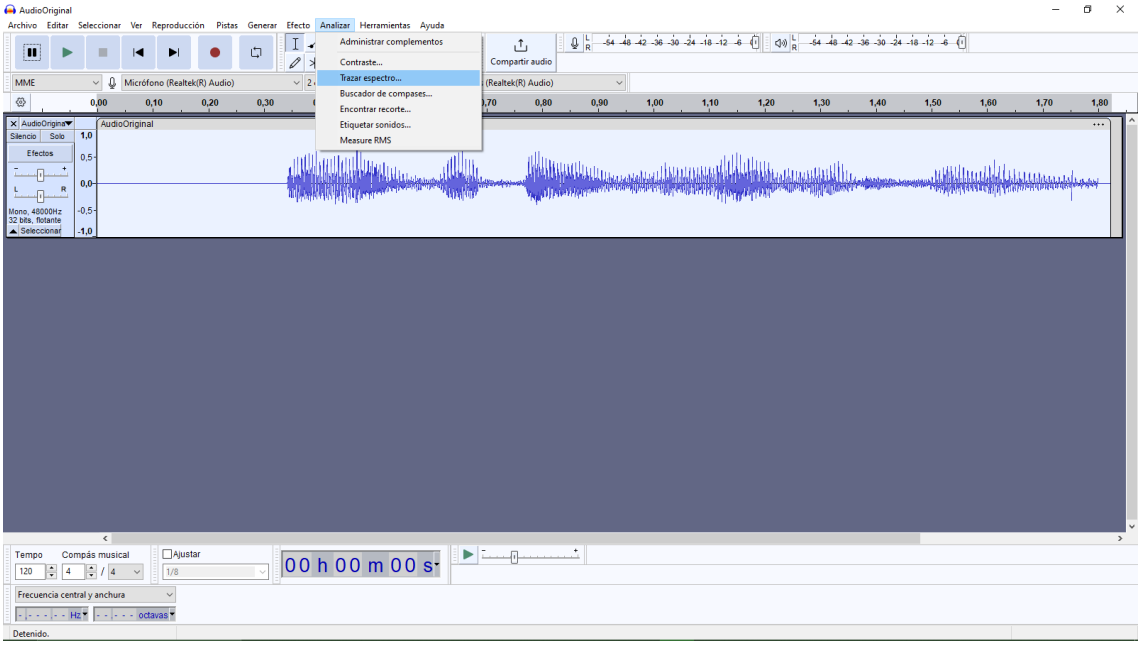



**Imagen 56: Trazar espectro como enfoque de análisis**

6. **Resultado del análisis**

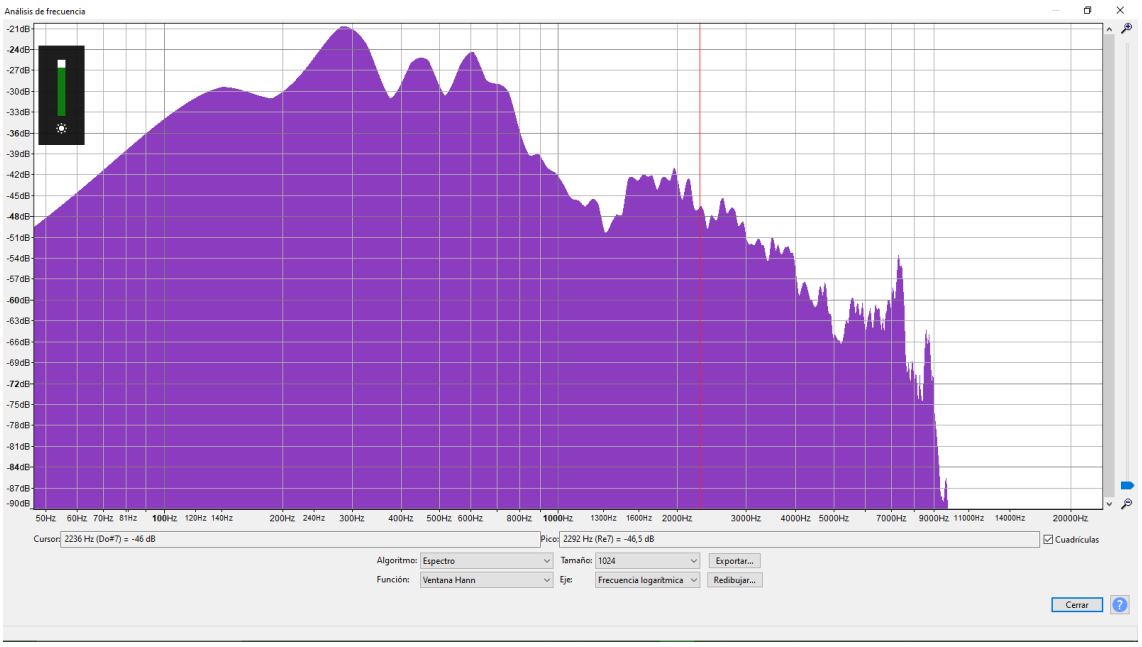



**Imagen 57: Resultado de análisis**

# HERRAMIENTA: SPEK

## 1. Abrir la herramienta Spek


**Imagen 58: Herramienta de Spek – Portal Principal**

## 2. Seleccionar el audio

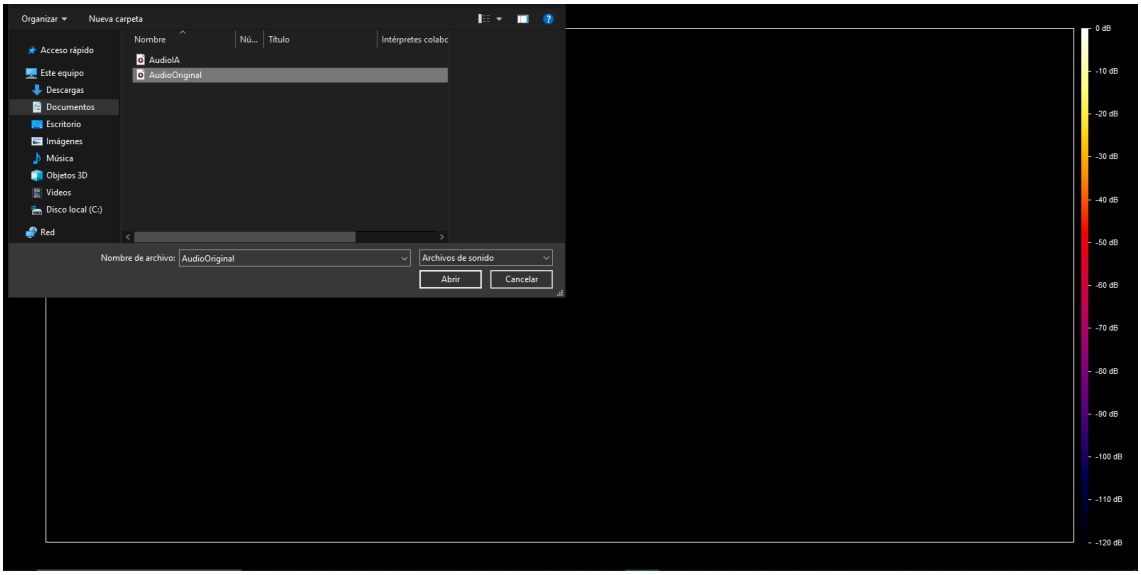



**Imagen 59: Seleccionar el Audio**

### 3. Grafica resultante con valores esenciales

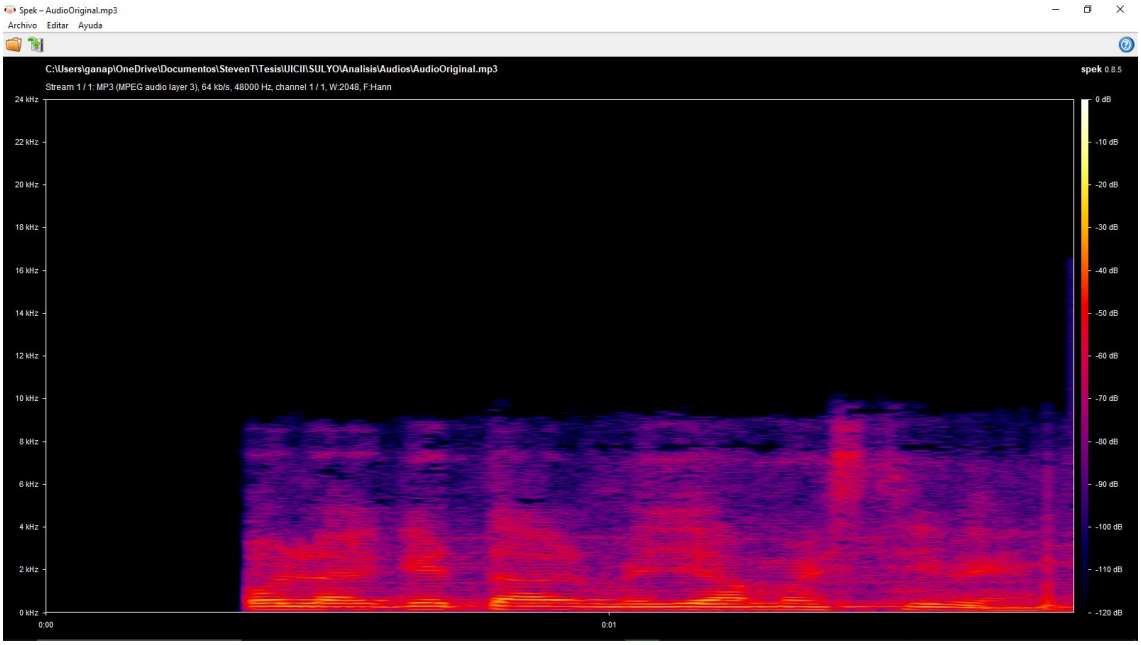



**Imagen 60: Valores Esenciales de la muestra**

## HERAMIENTA AUDACITY – AUIDO IA

### 1. Realizar los mismo paso con el proceso del audio anterior

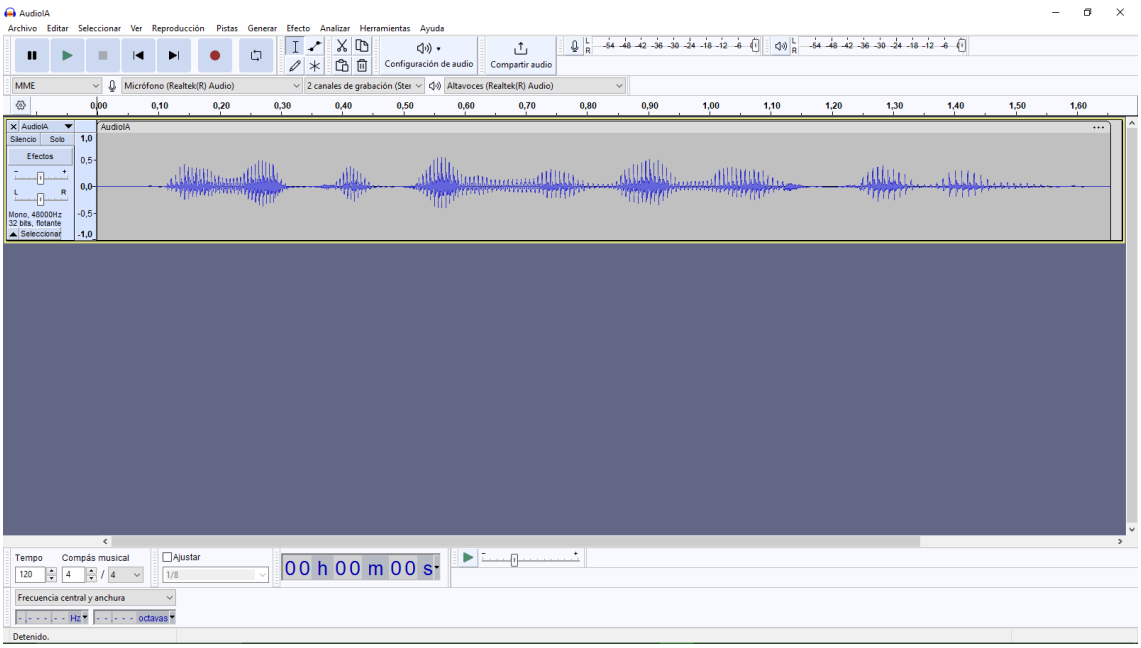



**Imagen 61: Audio IA – Vista Previa**

2. **La opción analizar en trazar espectro para el proceso del audio**

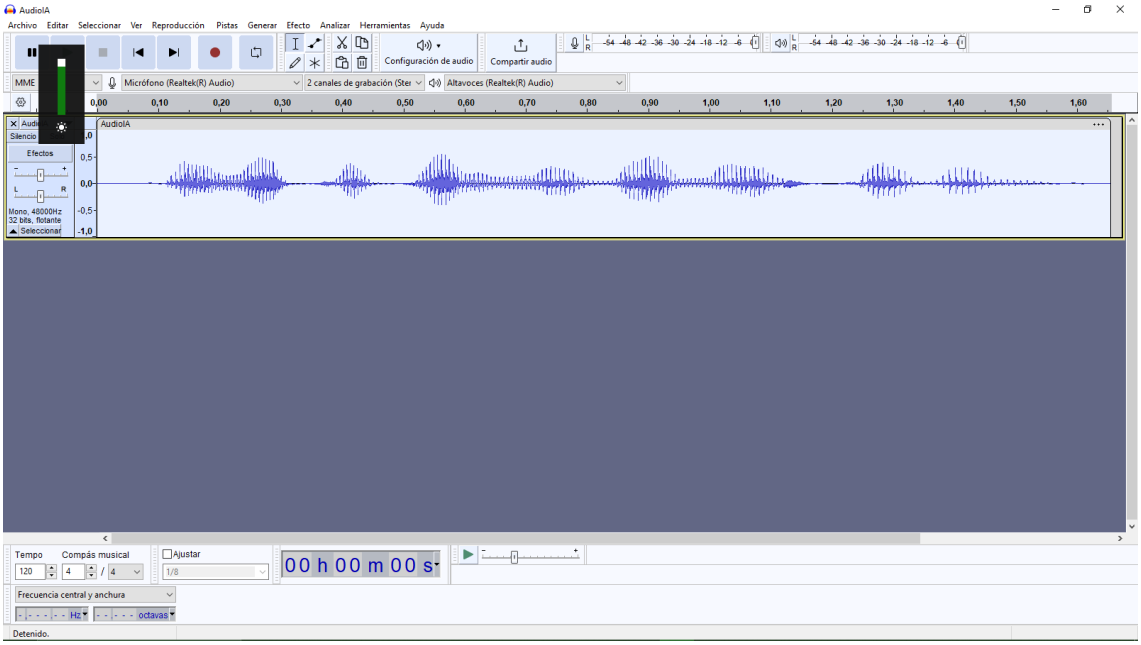



**Imagen 62: Audio en proceso de análisis por Trazar Espectro**

3. **Resultado Final**

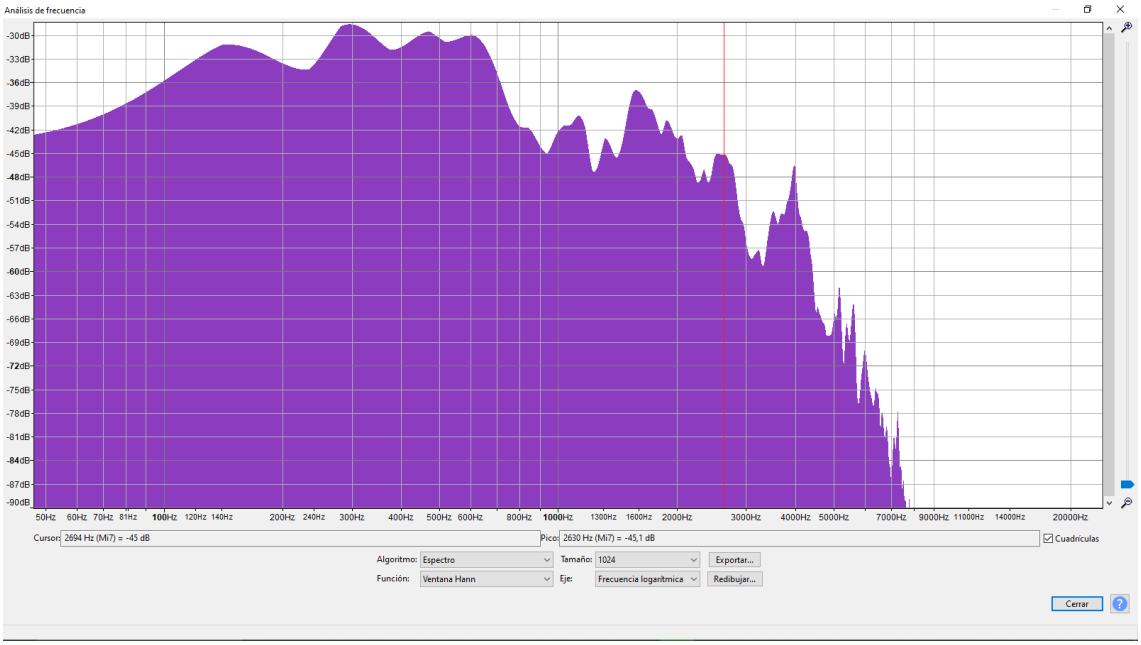



**Imagen 63: Resultado de Audio IA**

# HERRAMIENTA SPEK

1. **Se realiza los pasos anteriores y se captura el resultado**

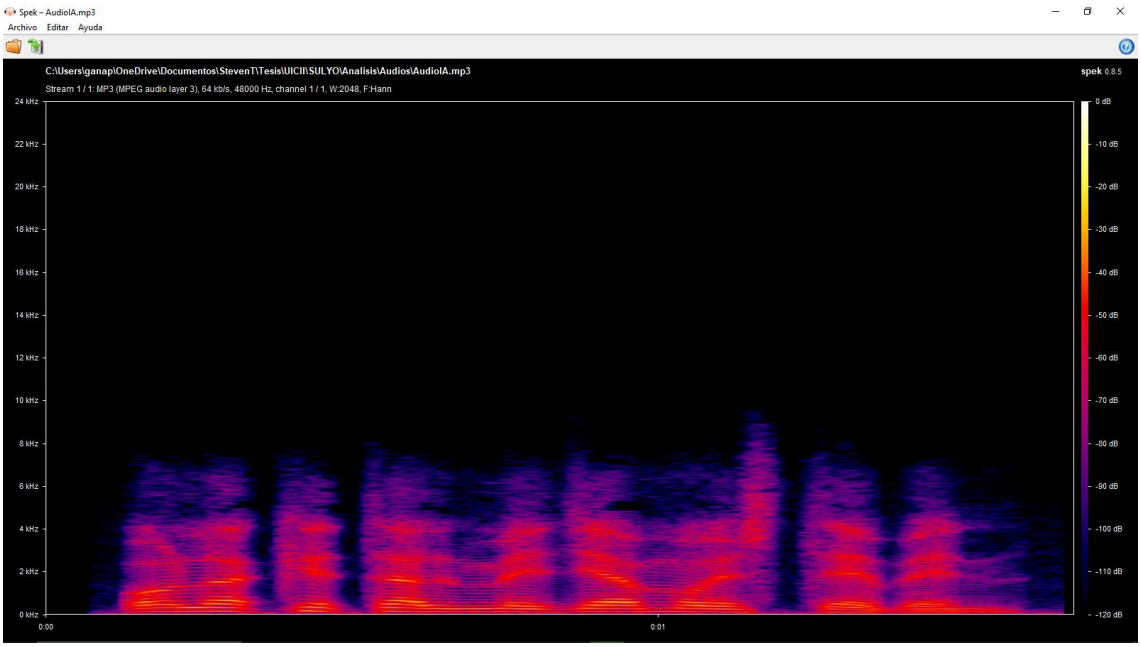



**Imagen 64: Resultado Final del Audio IA por la Herramienta SPEK**

# ANALISIS COEFICIENTES CEPSTRALES - AUDIO ORI – DEEPFAKE POR PYTHON

1. **Importación de bibliotecas: Se importan las bibliotecas necesarias: librosa para el análisis de audio, matplotlib.pyplot para la visualización de datos y numpy para operaciones matemáticas.**

```
import librosa
import librosa.display
import matplotlib.pyplot as plt
import numpy as np
```

**Imagen 65: Importación de bibliotecas en Python**

2. **Definición de las rutas de los archivos de audio: Se especifican las rutas de los archivos de audio que se van a analizar.**

```
# Rutas de los archivos de audio
ruta_audio1 = 'Audios/AudioOriginal.mp3'
ruta_audio2 = 'Audios/AudioIA.mp3'
```

**Imagen 66: Rutas de los archivos muestras**

3. **Función para calcular los coeficientes cepstrales: Se define la función calcular_cepstrales que toma la ruta de un archivo de audio como entrada.**

```
# Función para calcular los coeficientes cepstrales
def calcular_cepstrales(ruta_audio):
    audio, sr = librosa.load(ruta_audio, sr=None)
    espectrograma = np.abs(librosa.stft(audio))
    cepstral_coeficientes = librosa.feature.mfcc(S=librosa.amplitude_to_db(espectrograma), sr=sr, n_mfcc=13)
    return cepstral_coeficientes
```

**Imagen 67: Función para calcular Cepstrales**

4. **Cálculo de los coeficientes cepstrales para ambos audios: Se calculan los coeficientes cepstrales para ambos archivos de audio utilizando la función calcular_cepstrales().**

```
# Calcular coeficientes cepstrales para ambos audios
cepstrales_audio1 = calcular_cepstrales(ruta_audio1)
cepstrales_audio2 = calcular_cepstrales(ruta_audio2)
```

**Imagen 68: Calcular los coeficientes Cepstrales de las muestras**

5. **Graficación de los coeficientes cepstrales para ambos audios**

```
# Graficar los coeficientes cepstrales para ambos audios
plt.figure(figsize=(12, 6))
plt.subplot(2, 1, 1)
librosa.display.specshow(cepstrales_audio1, x_axis='time')
plt.colorbar()
plt.title('Coeficientes Cepstrales - Audio 1')
plt.xlabel('Tiempo (s)')
plt.ylabel('Coeficiente Cepstral')

plt.subplot(2, 1, 2)
librosa.display.specshow(cepstrales_audio2, x_axis='time')
plt.colorbar()
plt.title('Coeficientes Cepstrales - Audio 2')
plt.xlabel('Tiempo (s)')
plt.ylabel('Coeficiente Cepstral')

plt.tight_layout()
plt.show()
```

**Imagen 69: Graficar los coeficientes Cepstrales de los archivos muestras**

6. **Resultado final**

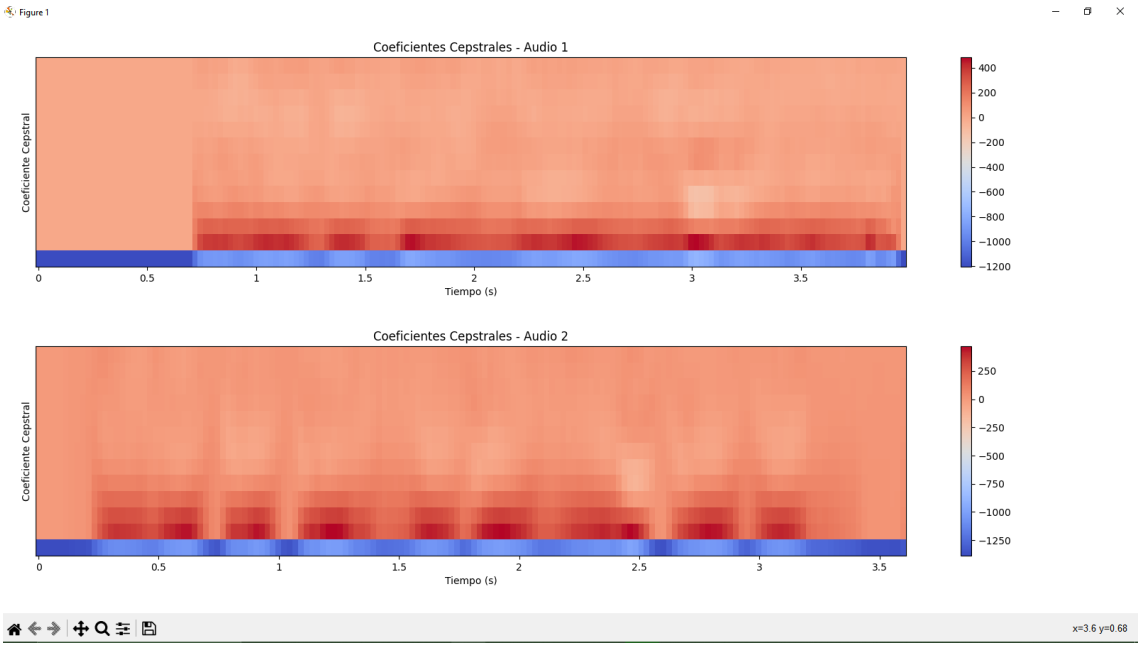



**Imagen 70: Grafico comparativo de los audios**

## ANALISIS DE FORMA DE ONDA - AUDIO ORI – DEEPFAKE POR PYTHON

1. **Importación de bibliotecas: Se importan las bibliotecas necesarias: librosa para el análisis de audio y matplotlib.pyplot para la visualización de datos.**

```
import librosa
import matplotlib.pyplot as plt
```

**Imagen 71: Importación de bibliotecas en Python**

2. **Definición de las rutas de los archivos de audio: Se especifican las rutas de los archivos de audio que se van a cargar.**

```
# Ruta de los archivos de audio
ruta_audio1 = 'Audios/AudioOriginal.mp3'
ruta_audio2 = 'Audios/AudioIA.mp3'
```

**Imagen 72: Ruta de los archivos muestras**

3. **Cargar los archivos de audio: Utilizando librosa.load(), se cargan los archivos de audio especificados en las rutas. El argumento sr=None asegura que se carguen los archivos utilizando la frecuencia de muestreo original.**

```
# Cargar los archivos de audio
audio1, sr1 = librosa.load(ruta_audio1, sr=None)
audio2, sr2 = librosa.load(ruta_audio2, sr=None)
```

**Imagen 73: Carga de los archivos muestra**

4. **Crear el tiempo para cada muestra: Utilizando librosa.times_like(), se crea un arreglo de tiempo correspondiente a cada muestra de los archivos de audio. Esto se hace para asegurar que el eje de tiempo esté correctamente escalado para cada archivo, considerando la frecuencia de muestreo de cada uno**

```
# Crear el tiempo para cada muestra
tiempo_audio1 = librosa.times_like(audio1, sr=sr1)
tiempo_audio2 = librosa.times_like(audio2, sr=sr2)
```

**Imagen 74: Crear el tiempo de cada muestra**

5. **Graficar las formas de onda**

```
# Graficar las formas de onda
plt.figure(figsize=(10, 6))

plt.subplot(2, 1, 1)
plt.plot(tiempo_audio1, audio1)
plt.title('Forma de Onda - Audio 1')
plt.xlabel('Tiempo (s)')
plt.ylabel('Amplitud')

plt.subplot(2, 1, 2)
plt.plot(tiempo_audio2, audio2)
plt.title('Forma de Onda - Audio 2')
plt.xlabel('Tiempo (s)')
plt.ylabel('Amplitud')

plt.tight_layout()
plt.show()
```

**Imagen 75: Graficar las formas de las muestras en onda**

6. **Resultado Final**

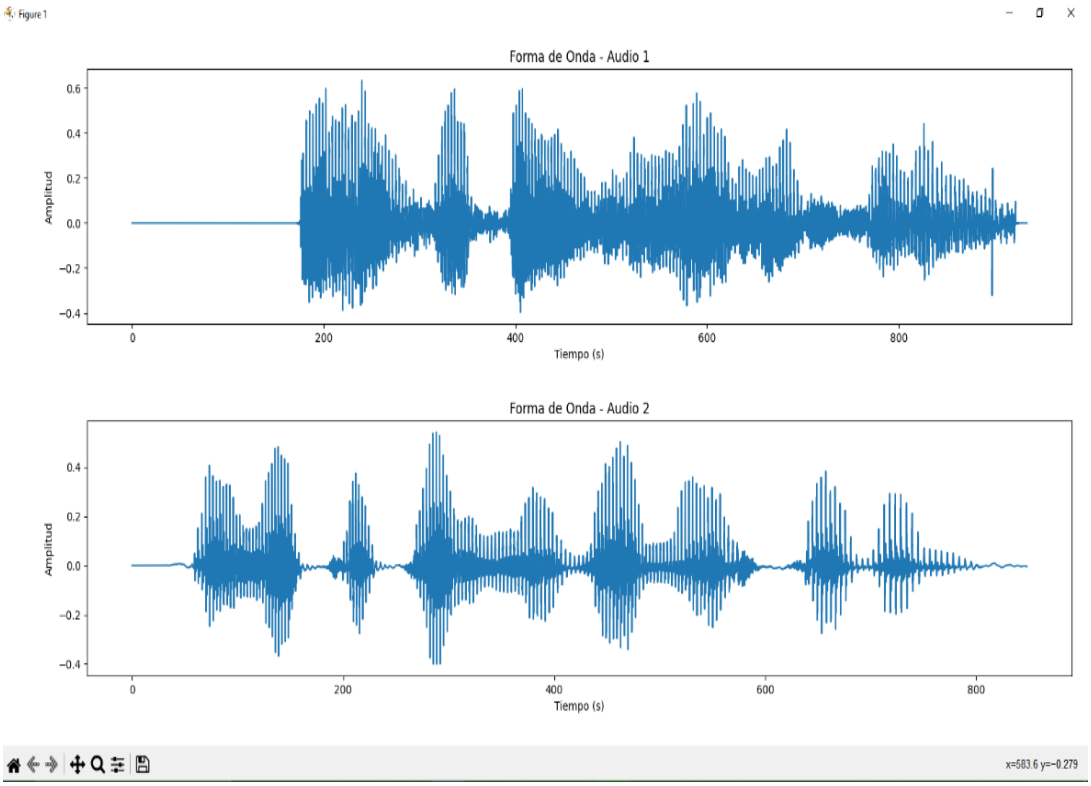



**Imagen 76: Grafico comparativo de las muestras de audio analizadas**

# ANÁLISIS ACÚSTICO – HERRAMIENTA PRAAT

1. Como primer paso vamos a revisar el espectro del audio deepfake mediante el programa praat.
    - Abrimos el audio que esta en la carpeta de archivos (open)
    - Seleccionamos la opción read from file

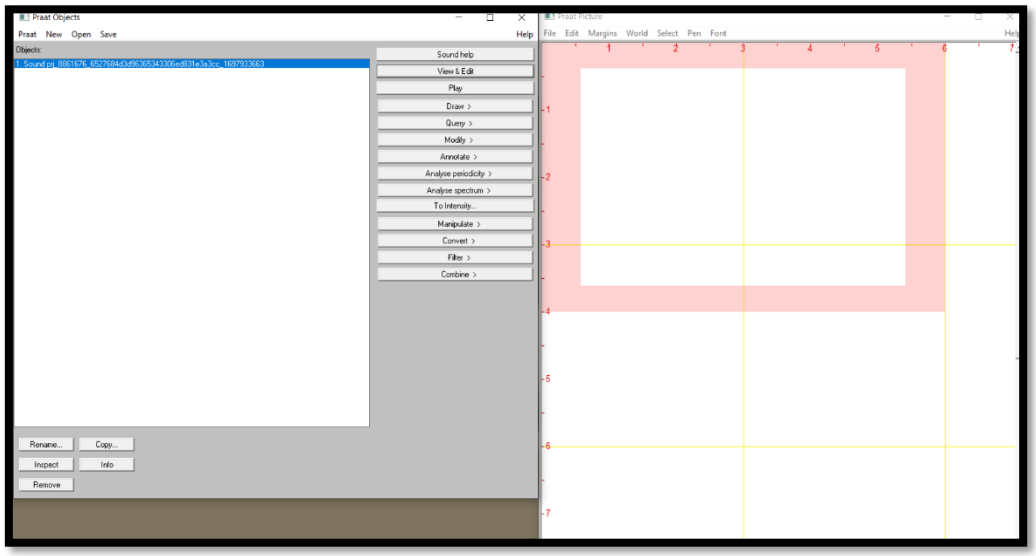



**Imagen 77: Open file audio**

2. Se observa el análisis de la forma de onda del audio

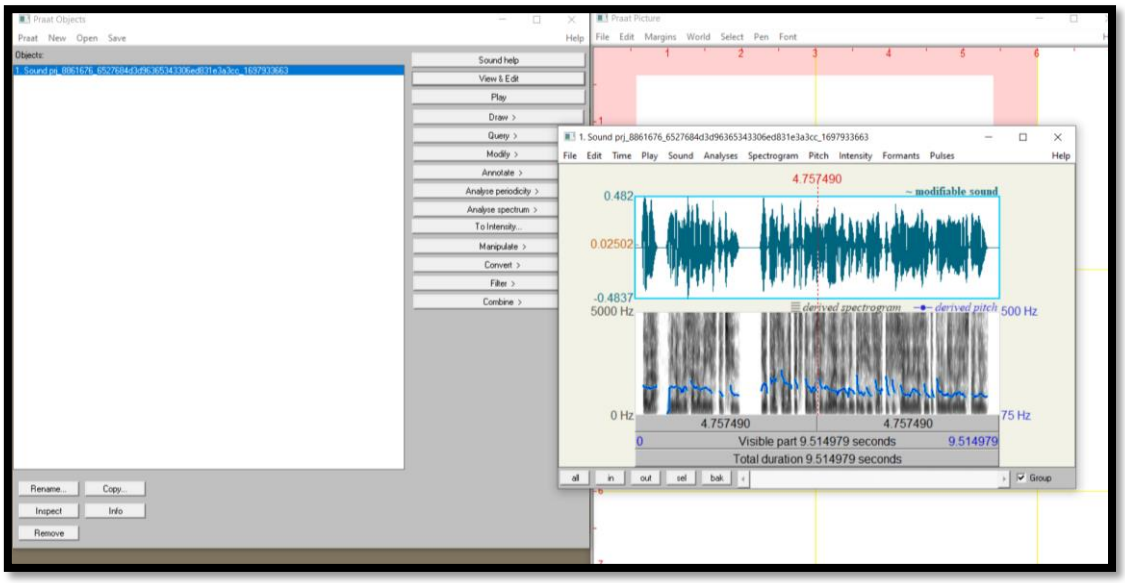



**Imagen 78: Forma de onda audio – análisis**

3. Seleccionar el apartado To List

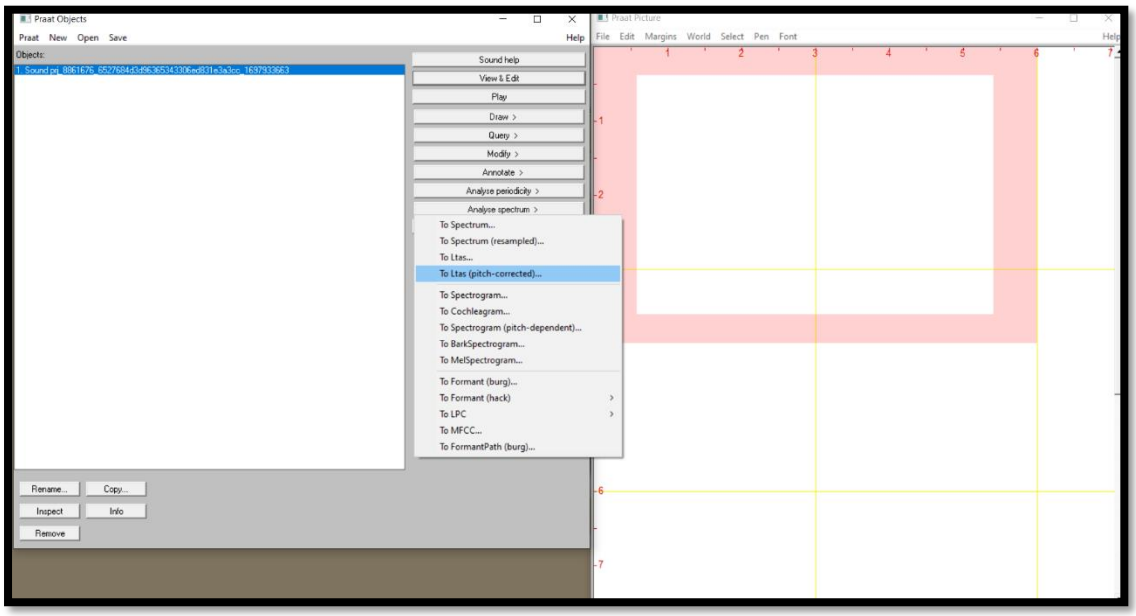



**Imagen 79: Top List**

4. Se detalla el análisis del audio presentando valores como frecuencia máxima, longitud, amplitud, entre otros

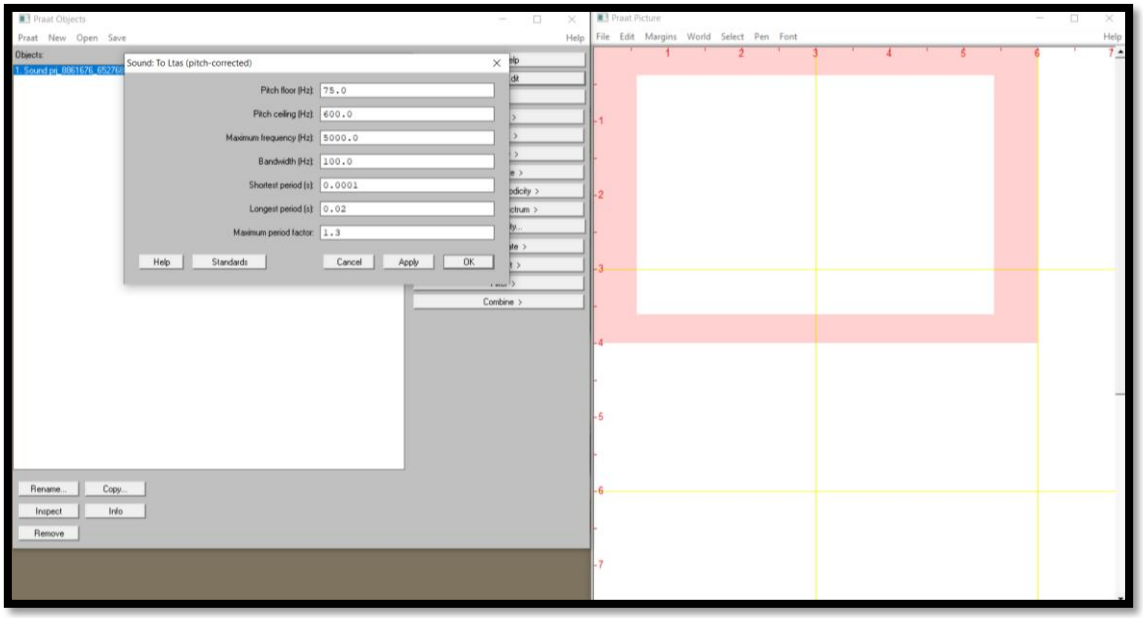



**Imagen 80: Resultados del análisis**

5. Dar en opción de Get Slope

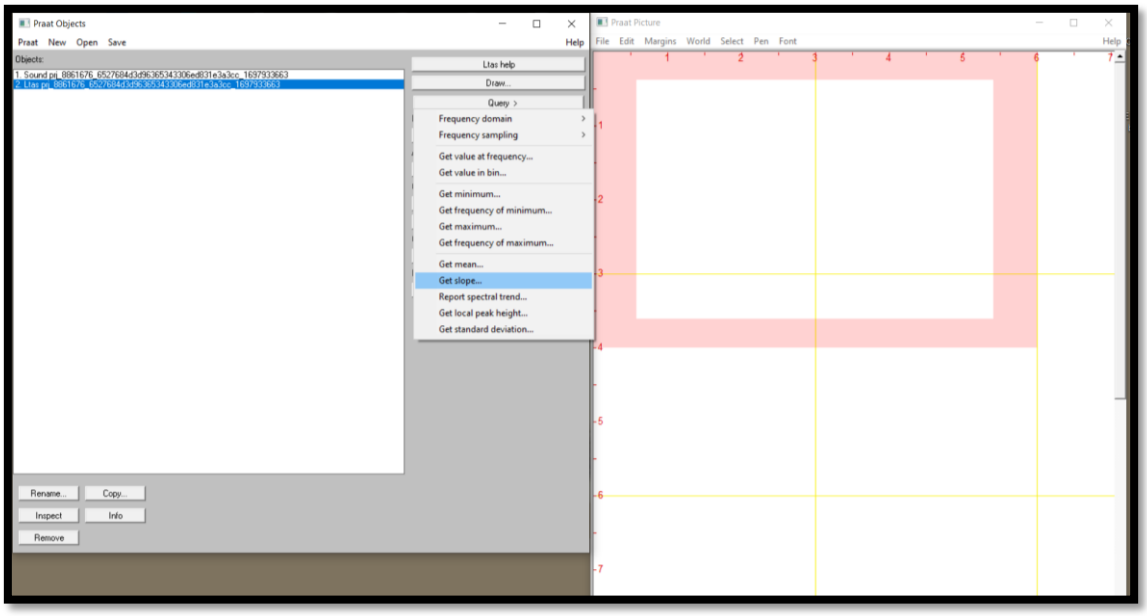



**Imagen 81: Get Slope**

6. Se presenta el valor de banda de ancha

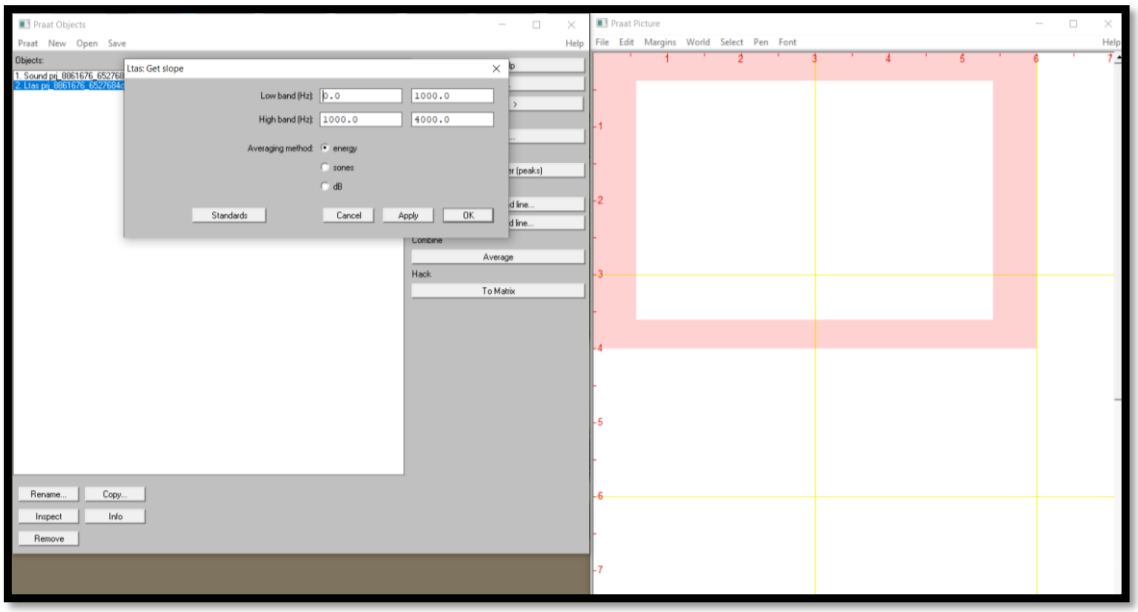



**Imagen 82: Banda de ancha - valores**

7. Se observa un valor db que significa el nivel de presión sonora del audio, en pocas palabras los decibles, como es valor negativo, el audio es deepfake. Caso contrario, si era positivo, el audio era autentico.

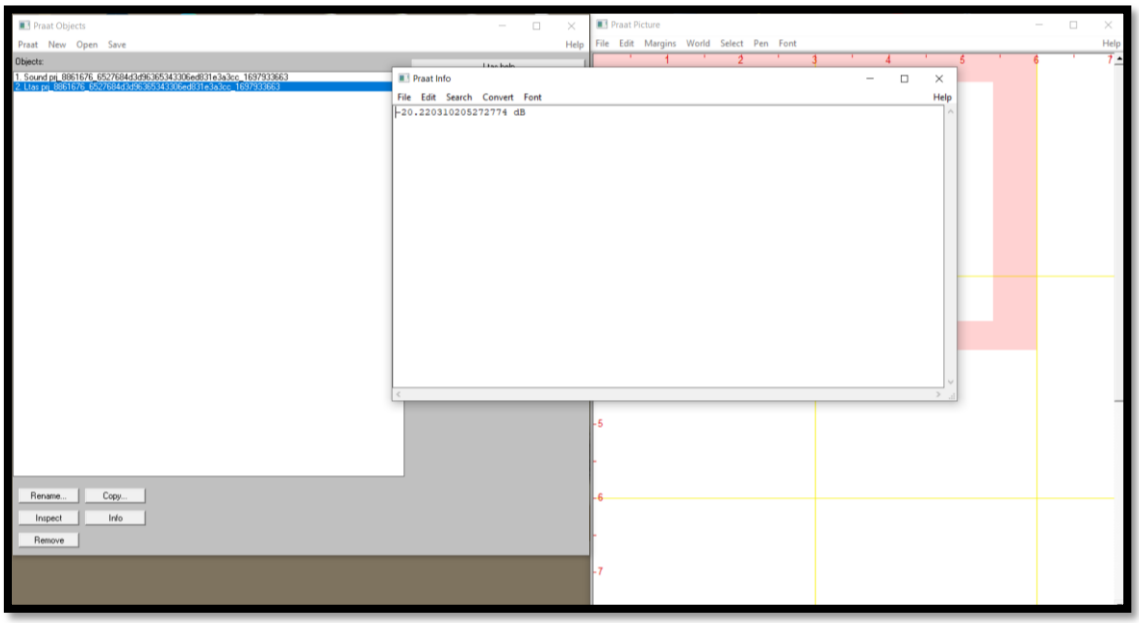



**Imagen 83: Valor de Db – Nivel de presión sonora**

# ANEXO 5
# DOCUMENTACIÓN

# INTRODUCCIÓN

En el ámbito forense, la autenticidad de los audios es crucial para determinar la veracidad de la evidencia presentada. Con el avance de la inteligencia artificial (IA) y la acústica forense, se han desarrollado técnicas para identificar audios sintéticos, los cuales pueden ser manipulados para engañar a los oyentes y distorsionar la verdad. En esta documentación, exploraremos el uso de estas técnicas para garantizar la integridad de la evidencia auditiva en investigaciones forenses.

La capacidad de discernir entre audios auténticos y sintéticos es fundamental en el contexto forense, donde la precisión y la fiabilidad de la evidencia son de suma importancia para garantizar la justicia. Con el advenimiento de tecnologías cada vez más sofisticadas, los perpetradores tienen a su disposición herramientas que les permiten manipular y fabricar audios con una calidad cada vez más convincente. Por tanto, es imperativo que los investigadores y profesionales forenses estén equipados con métodos robustos y avanzados para detectar y validar la autenticidad de los audios presentados como evidencia.

# OBJETIVO

El objetivo principal de este documento es proporcionar una guía sobre la aplicación de técnicas de IA y acústica forense para identificar audios sintéticos. Se busca ofrecer un marco de referencia para los profesionales forenses y otros interesados en la verificación de la autenticidad de los audios en contextos legales y de investigación.

# ALCANCE

- Aplicación de la acústica forense en la autenticación de audios.
- Limitaciones y desafíos asociados con la detección de audios sintéticos.
- Recomendaciones para mejorar la verificación de la autenticidad de los audios en investigaciones forenses.

# DESCRIPCION DE CONTROLES DE AUTENTICIDAD DE AUDIOS

## 1. Análisis Espectral:

- Utilización de herramientas de software especializadas para analizar el espectro de frecuencia del audio.
- Identificación de anomalías en la distribución de frecuencias que puedan indicar manipulación o síntesis del audio.
- Comparación del espectro de frecuencia del audio sospechoso con el de audios auténticos para detectar discrepancias.

## 2. Comparación de Firmas Acústicas:

- Extracción de la firma acústica del audio sospechoso y comparación con la de audios de referencia.
- Identificación de diferencias significativas en la forma de onda, el espectro de frecuencia y otros parámetros acústicos.

## 3. Análisis de Metadatos:

- Extracción y análisis de metadatos incrustados en el archivo de audio.
- Evaluación de la consistencia de los metadatos con la información contextual y la cadena de custodia del audio.

## 4. Análisis Forense de Forma de Onda:

- Examen detallado de la forma de onda del audio utilizando herramientas de software especializadas.
- Identificación de artefactos de manipulación, como cortes abruptos, inserciones o superposiciones.

## 5. Verificación de Coeficientes Cepstrales:

- Análisis de los coeficientes cepstrales del audio para caracterizar su firma única.
- Comparación de los coeficientes cepstrales del audio sospechoso con los de audios auténticos para detectar inconsistencias.

## 6. Evaluación de Coeficientes MFCC:

- Extracción de los coeficientes Mel-frequency cepstral (MFCC) del audio para representar su contenido acústico.
- Comparación de los coeficientes MFCC del audio sospechoso con los de audios auténticos para determinar su similitud.

## 7. Validación Cruzada:

- Aplicación de múltiples técnicas de autenticidad en paralelo para corroborar los resultados obtenidos.
- Integración de los hallazgos de diferentes controles para obtener una evaluación más completa y confiable del audio.

## 8. Revisión por Expertos:

- Revisión y validación de los resultados por parte de expertos en acústica forense y análisis de audio.
- Evaluación de la integridad y la coherencia de los hallazgos a la luz de las circunstancias del caso y las mejores prácticas forenses.

# RECOMENDACIONES

- Permanecer al día con las últimas técnicas y herramientas en el campo de la inteligencia artificial (IA) y la acústica forense. La evolución rápida de la

tecnología requiere una constante adaptación para mantener la efectividad en la identificación de audios sintéticos.

- Realizar una evaluación exhaustiva de la autenticidad de los audios empleando una variedad de métodos de verificación. La combinación de diferentes enfoques fortalece la confiabilidad de los resultados y minimiza la posibilidad de errores.
- Fomentar la colaboración con expertos en IA y acústica forense. La interacción entre distintas disciplinas promueve la innovación y facilita el desarrollo de técnicas más precisas y efectivas para la identificación de audios sintéticos.
- Establecer protocolos claros para el manejo y la autenticación de evidencia auditiva en investigaciones forenses. Estos protocolos deben abordar aspectos como la cadena de custodia, la preservación de la integridad de los datos y la documentación adecuada de los procedimientos utilizados.
- Emplear la validación cruzada utilizando diferentes conjuntos de datos y técnicas de verificación para corroborar los resultados. Esto ayuda a mitigar posibles sesgos y aumenta la robustez de las conclusiones obtenidas.
- Mantener registros detallados de los procedimientos utilizados y los resultados obtenidos durante el análisis de audios. Esta documentación proporciona transparencia y trazabilidad, facilitando la revisión y validación por parte de terceros.
- Involucrarse en comunidades profesionales y grupos de investigación relacionados con la IA y la acústica forense. El intercambio de conocimientos y experiencias con colegas en el campo puede impulsar la innovación y proporcionar oportunidades de colaboración.
- Realizar pruebas de estrés y simular escenarios realistas para evaluar la efectividad de las técnicas de identificación de audios sintéticos en condiciones adversas. Esto ayuda a identificar posibles limitaciones y áreas de mejora en los métodos utilizados.
- Promover la educación y la sensibilización sobre la manipulación de audios sintéticos y su impacto en investigaciones forenses. Esto incluye la capacitación de profesionales forenses y la divulgación de información sobre las últimas tendencias y amenazas en este campo.