

Escuela Superior Politécnica del Litoral

Facultad de Ingeniería en Electricidad y Computación

Desarrollo de una solución tecnológica para el envío de notificaciones instantáneas
mediante el procesamiento de señales de audio

Proyecto Integrador

Previo la obtención del Título de:

Ingeniero/a en Ciencias Computacionales

Presentado por:

Claudia Sofía Asanza Romero

Luis Alfredo Rodríguez Chávez

Guayaquil - Ecuador

Año: 2023


Agradecimientos

Queremos expresar nuestro profundo agradecimiento a todas las personas y entidades que contribuyeron de manera invaluable en la realización de este proyecto. Sus apoyos, consejos y aliento fueron esenciales para alcanzar los objetivos que nos propusimos. Sin su dedicación y compromiso, este proyecto no habría sido posible.

Especialmente, queremos agradecer al Mgst. Ronald Criollo, nuestro tutor, que gracias a sus oportunas intervenciones el proyecto se mantuvo en el camino correcto.

Declaración Expresa

“Los derechos de titularidad y explotación, me(nos) corresponde conforme al reglamento de propiedad intelectual de la institución; (nombre de los participantes) y doy(damos) mi(nuestro) consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”



Claudia Sofía Asanza
Romero



Luis Alfredo Rodríguez
Chávez

Evaluadores

Ronald Criollo

Profesor de Materia

Ronald Criollo

Tutor de proyecto

Resumen

En respuesta a la necesidad de dispositivos para emitir alertas inmediatas en situaciones de emergencia, este proyecto se enfoca en la creación de un dispositivo portable con capacidad de reconocimiento de voz. El objetivo principal es proporcionar una solución efectiva que permita a los usuarios pedir ayuda de manera rápida. Se plantea que el dispositivo mejore la capacidad de los usuarios para comunicarse en situaciones críticas, aumentando la seguridad y tranquilidad de los individuos. La carcasa del dispositivo se ha fabricado mediante impresión 3D, diseñada para albergar una batería de 3.7V y la placa Nicla Voice. Esta última estuvo encargada de ejecutar el modelo de reconocimiento de la palabra clave “Sacha”, una vez la placa reconoce la palabra envía una señal a la aplicación móvil. La aplicación se encargó de recibir la señal y de enviar la alerta a las autoridades. El dispositivo Sacha logra reconocer la palabra clave con un nivel de precisión del 91%, a la par se evidencia que la alerta producida después del reconocimiento tarda, en promedio, 4 segundos. Además, la distancia ideal para el reconocimiento es de no más de 30cm en espacios con poco ruido ambiental.

Palabras Clave: TinyML, Alerta, Seguridad, Arduino

Abstract

In response to the need for devices to emit immediate alerts in emergency situations, this project focuses on creating a portable device with voice recognition capability. The main objective is to provide an effective solution that allows users to request assistance quickly. It is proposed that the device enhances users' ability to communicate in critical situations, thereby increasing individuals' safety and peace of mind. The device's casing has been manufactured using 3D printing, designed to accommodate a 3.7V battery and the Nicla Voice board. The latter was responsible for executing the "Sacha" keyword recognition model; once the board recognizes the keyword, it sends a signal to the mobile application. The application receives the signal and dispatches the alert to the relevant authorities. The Sacha device achieves a keyword recognition accuracy level of 91%; it is observed that the alert generated after recognition takes an average of 4 seconds. Additionally, the ideal recognition distance is no more than 30 cm in environments with low ambient noise.

Keywords: TinyML, Alert, Security, Deep Learning, Arduino.

Índice general

Evaluadores	IV
Resumen	V
Abstract	VI
Índice general	VII
Índice de tablas.....	X
Índice de figuras	XI
Abreviaturas	XII
Simbología	XII
1. Capítulo 1	1
1.1 Introducción.....	2
1.2 Descripción del problema.....	3
1.3 Justificación del problema	4
1.4 Objetivos.....	5
1.4.1 Objetivo general	5
1.4.2 Objetivos específicos.....	5
1.5 Marco teórico.....	6
1.5.1 Estado del arte	6
1.5.2 Placas.....	8
1.5.3 Modelos de Deep Learning	9
1.5.4 Plataformas.....	12
2. Capítulo 2	15
2.1 Metodología.....	16
2.1.1 Análisis.....	16
2.1.1.1 Requerimientos.....	16
2.1.1.1.1 Requerimientos funcionales.	16

2.1.1.1.2	Requerimientos no funcionales	18
2.1.1.2	Alcance y limitaciones de la solución.....	21
2.1.1.2	Riesgos y beneficios de la solución	21
2.1.1.4	Usuarios de la solución.	22
2.1.2	Prototipado.	23
2.1.2.1	Dispositivo portable.	23
2.1.2.2	Aplicación móvil.....	23
2.1.2.3	Evaluación.....	25
2.1.3	Diseño de la solución	26
2.1.3.1	Hardware	26
2.1.3.1.1	Características dispositivo portable.....	26
2.1.3.2	Software.	28
2.1.3.2.1	Arquitectura.	28
2.1.3.2.2	Diagramas.	29
2.2	Normativas.....	31
2.2.1	Consideraciones éticas y legales.	31
3.	Capítulo 3	33
3.1	Pruebas.....	34
3.1.1	Pruebas sin ruido ambiente	34
3.1.2	Pruebas con ruido ambiente	34
3.1.3	Pruebas de tiempo de envío.....	35
3.2	Plan de implementación.....	35
3.3	Resultados.....	37
3.3.1	Resultados del modelo de inteligencia artificial.	37
3.3.2	Pruebas y validación del dispositivo portable.	39
3.3.2.1	Tiempo de envío de alerta.....	39

3.3.2.2	Reconocimiento de palabra clave y envío de alerta.....	39
3.3.2.3	Comparativa de reconocimiento del comando de voz “Sacha” en entornos ... ruidosos y sin ruido.	40 40
3.3.2.4	Comparativa de envío de alerta en escenarios con ruido ambiental.	41
3.4	Análisis de costo.....	41
3.4.1	Desglose de costos.....	41
3.4.2	Análisis de mercado	42
3.4.2.1	Modelos de negocio del proyecto.	42
3.4.1.1.1	Oferta de servicio de seguridad a través de suscripciones.	42
3.4.1.1.2	Adquisición personal del producto.....	42
3.4.1.1.3	Empresas que requieran el producto como servicio.....	42
4.	Capítulo 4.....	43
4.1	Conclusiones y recomendaciones	44
4.1.1	Conclusiones.....	44
4.1.2	Recomendaciones.....	45
4.1.2.1	Recomendaciones al desarrollar el proyecto.....	45
4.1.2.2	Recomendaciones para trabajos futuros.....	46
	Referencias.....	48

Índice de tablas

Tabla 2.1	16
Tabla 2.2.....	16
Tabla 2.3.....	17
Tabla 2.4.....	17
Tabla 2.5.....	18
Tabla 2.6.....	18
Tabla 2.7.....	18
Tabla 2.8.....	19
Tabla 2.9.....	19
Tabla 2.10.....	20
Tabla 2.11.....	20
Tabla 2.12.....	26
Tabla 2.13.....	27
Tabla 2.14.....	27
Tabla 2.15.....	27
Tabla 3.1.....	37
Tabla 3.2.....	39
Tabla 3.3.....	41

Índice de figuras

Figura 2.1	23
Figura 2.2	23
Figura 2.3	25
Figura 2.4	28
Figura 2.5	28
Figura 2.6	30
Figura 2.7	30
Figura 2.8	31
Figura 2.9	31
Figura 3.1	36
Figura 3.2	36
Figura 3.3	39
Figura 3.4	40
Figura 3.5	41

|

Abreviaturas

ESPOL Escuela Superior Politécnica del Litoral

TINYML Tiny Deep Learning

LoRaWAN Long Range Wide Area Network

GPS Global Positioning System

RAM Random Access Memory

FPU Floating-Point Unit

BLE Bluetooth Low Energy

PWM Pulse Width Modulation

USB Universal Serial Bus

UART Universal Asynchronous Receiver/Transmitter

SPI Serial Peripheral Interface

ADC Analog-to-Digital Converter

GPIO General-Purpose Input/Output

PDM Pulse Density Modulation

API Application Programming Interface

PLA Polylactic Acid

SRAM Static Random Access Memory

NDP Neural Decision Processors

IA Inteligencia Artificial

Simbología

Cm Centímetros

V Voltios

MHz Megahertzios

Mm Milímetros

MB Megabytes

KB Kilobytes

Mbps Megabits por segundo

mAh Miliamperios-hora

1. CAPÍTULO 1

1.1 Introducción

En el contexto actual del país, la seguridad y bienestar de los ciudadanos se ha convertido en una preocupación prioritaria debido a la escasez de alertas inmediatas para proteger la integridad física de los ciudadanos día a día. Aunque la tecnología ha empleado un papel fundamental en la mejora de los sistemas de alertas, como el "mecanismo de localización de personas atrapadas utilizando dispositivos inalámbricos del Servicio Nacional de Gestión de Riesgos y Emergencias en Ecuador" [1], todavía se enfrentan desafíos significativos en áreas críticas como secuestros, robos, delincuencia organizada, emergencias médicas, accidentes de tránsito, violencia urbana y violencia de género. Por esta razón, es crucial actuar con cautela y de manera oportuna para evitar que estas situaciones conlleven consecuencias más graves.

Con el objetivo de abordar estos desafíos, esta investigación se centra en el desarrollo de un dispositivo portable con capacidades de procesamiento de audio y envío de notificaciones instantáneas. Su objetivo principal es proporcionar alertas inmediatas y efectivas en situaciones de emergencia, mejorando así la capacidad de respuesta ante situaciones de riesgo y aumentando la protección de la integridad física de las personas en Ecuador. Para lograr esto, se usaron tecnologías innovadoras y se llevaron a cabo pruebas exhaustivas para verificar la eficacia del dispositivo.

La urgencia de abordar estos desafíos se ve reflejada en las alarmantes estadísticas de inseguridad, referenciadas en el medio digital Primicias [2], donde se incluye una entrevista con el general Wilson Pavón. Se identifica que entre el 1 de enero y el 23 de abril de 2023, el distrito de Guayaquil registró un número de 729 muertes violentas. La implementación de este dispositivo puede contribuir de manera significativa a mejorar la seguridad y bienestar de los ciudadanos, al proporcionar una herramienta eficaz para alertar y responder rápidamente en situaciones de emergencia.

No obstante, es fundamental tener en cuenta que abordar los desafíos de seguridad no recae únicamente en la tecnología, sino que también requiere el compromiso y la colaboración de toda la sociedad, así como el fortalecimiento de las instituciones encargadas de salvaguardar la seguridad ciudadana.

1.2 Descripción del problema

En 2021 se dieron un total de 802 secuestros en el país, además de 6 264 asaltos y aproximadamente 28 000 delitos sexuales [3]. Según el personal de la Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC), menciona que Ecuador sigue siendo punto de origen, tránsito y destino, tanto para víctimas de explotación sexual, como laboral, de origen nacional como extranjero [4].

A pesar de estas estadísticas, la inversión en tecnología específicamente en el ámbito de seguridad en redes, asistencia social y servicios de cuidado social es limitada. De acuerdo con información del Banco Mundial, solo se han implementado 6 proyectos en esta área [5]. Por otro lado, el 25 de noviembre de 2022, la policía Nacional del Ecuador recibió una capacitación internacional sobre uso de fuerza y derechos humanos [6]. Sin embargo, ninguno de estos casos ha considerado la incorporación de tecnologías actuales como herramientas para fortalecer la seguridad ciudadana, como por ejemplo mejorar el portal de recepción de alertas inmediatas.

En la última reunión interinstitucional para la elaboración de la Estrategia Nacional contra el crimen organizado, llevada a cabo del 7 al 9 de febrero 2023 en la ciudad de Cuenca, en colaboración con la Oficina de las Naciones Unidas contra la Droga y el Delito, se plantearon diversas medidas. Entre las propuestas, se destaca la necesidad de implementar marcos jurídicos, fortalecer las zonas de riesgos y mejorar la coordinación entre las diferentes funciones del Estado. Además, se planteó la generación de protocolos para el intercambio de información interinstitucional incluyendo centros educativos de nivel superior [7]. En este sentido, la

implementación de un sistema de alertas inmediatas podría suplir esta necesidad, ofreciendo mecanismos de control y evaluación para mejorar la seguridad ciudadana.

Como es notorio, es imperativo abordar la necesidad de dispositivos que generan alertas inmediatas. Esto requiere una mayor inversión en tecnología, la incorporación de enfoques innovadores y la cooperación interinstitucional para fortalecer la seguridad.

La implementación de un dispositivo portable permitirá enviar alertas inmediatas, como una medida adicional, contribuiría a mitigar los riesgos y mejorar la capacidad de respuesta ante situaciones de peligro.

1.3 Justificación del problema

En los últimos años, la ciudad de Guayaquil ha experimentado un preocupante incremento en el número de asaltos a mano armada, incluso a plena luz del día. La universidad en la que se planea implementar el dispositivo portátil que se desarrolla en este estudio se encuentra en la urbe porteña, debido a una potencial vulnerabilidad de estudiantes, profesores y personal administrativo en términos de seguridad. Aunque hasta el momento no se han reportado incidentes de este tipo en el centro de estudio, es fundamental anticiparse a esta situación y estar preparados para brindar una respuesta efectiva en caso de que ocurran eventos de esta naturaleza en el campus.

Es importante destacar que no solo en Guayaquil, sino en otras universidades del país, como la Universidad Técnica de Machala y la Universidad Central [8], se han registrado casos de robos a mano armada, lo que confirma la necesidad de implementar medidas de seguridad proactivas y eficientes en entornos académicos.

La presente investigación propone el desarrollo de un dispositivo portátil basado en TinyML [9], una tecnología de vanguardia que permite la ejecución de modelos de deep learning (aprendizaje profundo) en dispositivos de recursos limitados. Este contará con una

placa especializada capaz de reconocer una palabra clave específica y, al detectarla, enviará una alerta inmediata al teléfono de las autoridades pertinentes dentro de la universidad. De esta manera, se busca proporcionar una respuesta rápida y precisa en situaciones de riesgo, facilitando la intervención oportuna y minimizando el potencial impacto de los sucesos delictivos.

El enfoque en el uso de TinyML para potenciar la seguridad es una perspectiva innovadora, dado que esta tecnología aún no ha sido ampliamente explotada en este ámbito. Su capacidad para ejecutar modelos de deep learning en dispositivos portátiles brinda la oportunidad de mejorar significativamente la efectividad de las medidas de seguridad existentes, superando las limitaciones de otros dispositivos convencionales que no son fácilmente accesibles en situaciones de emergencia o eventos a mano armada.

Además de los casos específicos de robos a mano armada, es importante tener en cuenta que el ámbito de aplicación del dispositivo se extiende a otros eventos de seguridad, como emergencias médicas y accidentes de tránsito. Esto amplía su utilidad y relevancia, permitiendo una respuesta efectiva en diferentes escenarios de riesgo.

1.4 Objetivos

1.4.1 Objetivo general

Crear un dispositivo portable para la emisión de alertas inmediatas de eventos que atenten contra la seguridad de las personas.

1.4.2 Objetivos específicos

1. Recopilar muestras de audio para la construcción de un conjunto de datos representativo.

2. Desarrollar un modelo de reconocimiento de palabras clave utilizando técnicas de inteligencia artificial.
3. Entrenar el modelo utilizando el conjunto de datos recopilado para mejorar su precisión y capacidad de reconocimiento.
4. Diseñar un plan de pruebas rigurosas y métricas establecidas.
5. Desplegar el modelo entrenado en un sistema embebido, asegurando su funcionamiento óptimo y eficiente.
6. Realizar la validación experimental del dispositivo portátil, evaluando su rendimiento en escenarios reales.

1.5 Marco teórico

El presente marco teórico aborda la implementación de innovaciones tecnológicas en la reducción y prevención de emergencias. Se destaca el uso de dispositivos portátiles como una solución eficaz para mejorar la detección y respuesta en situaciones de riesgo. Estos dispositivos permiten la detección de palabras clave y emiten alertas instantáneas. El objetivo es explorar el estado del arte, conceptos clave y metodologías empleadas en el diseño y desarrollo de estos dispositivos, centrándose en su aplicación en entornos de seguridad.

1.5.1 Estado del arte

Entre las soluciones existentes que se han desarrollado para salvaguardar la seguridad, se destaca un dispositivo desarrollado en junio 2020 que consiste en una banda inteligente, usando ARDUINO MEGA, este dispositivo contiene sensores que mantienen en observación al usuario todo el tiempo, comprobando situaciones incómodas por medio de la frecuencia cardíaca, su temperatura y vibración. Al ocurrir un evento anómalo los dispositivos activan un módulo GPS y

también alertan a una organización de servicio público para que tome medidas inmediatas, aun así, este dispositivo, este dispositivo no cuenta con tecnología de inteligencia artificial[9].

Por otro lado, en septiembre 2020, se ha desarrollado un dispositivo portátil de bajo costo con inteligencia artificial, el cual es capaz de reconocer con precisión señales de socorro en presencia de ruido de fondo, permitiendo así su procesamiento de alertas en tiempo real. Este dispositivo ha sido implementado utilizando una base de datos en español, Raspberry Pi y Tensorflow. Los resultados fueron sorprendentes, ya que las redes neuronales convolucionales (CNN) lograron captar señales importantes de manera efectiva. Es importante mencionar que, no está enfocado a que personas lo lleven consigo[10].

Otra solución presente en la India se trata de un dispositivo inteligente portable llamado Suraksha Kawach desarrollado en enero 2022. Este dispositivo es pequeño, fácil de esconder y diseñado para que la persona lo traiga consigo, este activa mediante patrones o con un código de activación que requiere una frecuencia determinada, al activarse inmediatamente inicia la grabación y envía un SMS estándar con la ubicación de la víctima a la policía y sus familiares. Entre las mejoras futuras de este dispositivo, se recomienda el estudio de técnicas de inteligencia artificial para reducir la complejidad de la red, además de reducir los tiempos de procesamiento [11].

El presente proyecto utiliza tecnología con inteligencia artificial, que permita el reconocimiento de señales de audio, esto tiene el potencial de simplificar y mejorar la eficiencia del proceso de generación de alertas. Adicionalmente, se podrían reducir las métricas por evaluar, ya que se centraría en detectar y reconocer la palabra clave específica que activa la

alerta, evitando la necesidad de evaluar múltiples parámetros biométricos. Esto permitirá una respuesta más eficiente y precisa en situaciones de emergencia.

1.5.2 Placas

Las placas de desarrollo son dispositivos electrónicos versátiles y compactos utilizados para prototipar y desarrollar proyectos electrónicos. Son herramientas indispensables para crear soluciones tecnológicas innovadoras de las cuales se mencionarán tres.

Tabla 1.2

Comparación de características entre las placas Arduino Nano 33 BLE Sense, Nicla Voice y Seeed nRF52840 Sense

Características	Arduino Nano 33 BLE Sense	Nicla Voice	Seeed nRF52840 Sense
Procesador	ARM Cortex-M4 (64 MHz) con FPU	ARM Cortex-M4F (64MHz)	nRF52840 M4F@64MHz
Dimensiones	45 x 18 mm	22,86 x 22,86 mm	21 x 17.5 mm
Memoria Flash	1MB	512 KB	1MB
Memoria RAM	256 KB RAM	64 KB SRAM	256 KB RAM
Comunicación inalámbrica	Bluetooth 5 multiprotocolo radio (2 Mbps) con soporte de IEEE 802.15.4 (Thread, Zigbee)	ANNA-B112 Bluetooth® Module	Bluetooth 5.0 / BLE
Sensores incorporados	Acelerómetro, giroscopio, magnetómetro, sensor de luz, sensor de temperatura, sensor de proximidad, luz ambiental, color RGB	IMU 6-axis (BMI270) con acelerómetro y giroscopio 3 ejes, Magnetómetro 3 ejes (BMM150), Micrófono de alta calidad (IM69D130).	IMU, Micrófono
Pines GPIO o E/S	14 pines digitales, 4 pines PWM, 8 pines de	Castellated pads con I2C, UART, SPI,	26 pines digitales (incluyendo UART, SPI,

interrupción externa, 8 pines analógicos, pines de alimentación y 1 pin de comunicación USB.	ADC, GPIOs configurable, Entrada de batería, Conector para micrófono PDM externo	I2C), 8 pines PWM, 6 pines analógicos, pines de alimentación (3.3V, GND) y pines de comunicación USB.
--	--	---

Nota. La información se ha extraído de las datasheets correspondientes [12]–[14].

Tras el comparativo previo, se determinó que una de las metas a lograr es que el dispositivo sea pequeño y portable, enfocado en el reconocimiento de una palabra clave; debido a esto, la mejor opción es el Nicla Voice. Su procesador brinda un equilibrio entre rendimiento y eficiencia energética y cuenta con un micrófono de alta calidad, fundamental para la detección de la palabra clave y captura de audio en varios entornos; además de que ofrece pines que permiten la flexibilidad para integrar a otros dispositivos. Por otro lado, el Arduino Nano 33 BLE Sense y el Seeed Nrf52840 tienen sensores y características que no están bien enfocadas al objetivo del proyecto.

1.5.3 Modelos de Deep Learning

En este apartado se exploran diferentes modelos de deep learning. Se analizan el transfer learning, el modelo híbrido DNN-HMM (Red Neuronal Profunda-Modelo de Markov Oculto) y el enfoque CTC (Conexión Temporal de Clasificación), considerando sus ventajas en términos de utilización de modelos preentrenados, captura de dinámicas temporales y entrenamiento de extremo a extremo.

- Transfer learning (transferencia de aprendizaje).
 - Utilización de modelos preentrenados: El aprendizaje transferido permite aprovechar modelos preentrenados que han sido entrenados en conjuntos de datos a gran escala, como datos de habla multilingües o en español. Esto proporciona

una base sólida al capturar patrones acústicos y lingüísticos generales, reduciendo la necesidad de una extensa recopilación de datos.

- Adaptabilidad a nuevos idiomas: El aprendizaje transferido permite adaptar modelos preentrenados a nuevos idiomas, como el quichua, incluso con datos limitados. A través del fine-tuning (ajuste fino) del modelo preentrenado en un conjunto de datos objetivo más pequeño, se puede beneficiar de las representaciones aprendidas al mismo tiempo que se adaptan al modelo y características específicas del idioma objetivo.
- Potencial para un mejor rendimiento: El aprendizaje transferido a menudo conduce a un mejor rendimiento al aprovechar el conocimiento de tareas o idiomas relacionados. El modelo preentrenado captura patrones de habla generales, lo cual puede ser beneficioso para tareas específicas posteriores como el reconocimiento de género y palabras [15].
- Híbrido DNN-HMM
 - Captura de dinámicas temporales: Los modelos híbridos DNN-HMM son excelentes para capturar las dinámicas temporales del habla al combinar redes neuronales profundas con el marco de HMM. El HMM proporciona un mecanismo para modelar la naturaleza secuencial del habla e incorporar modelos de lenguaje, mientras que la DNN captura los patrones acústicos complejos.
 - Enfoque establecido: Los modelos híbridos DNN-HMM han sido ampliamente estudiados y utilizados en sistemas tradicionales de reconocimiento del habla. Su marco y técnicas bien establecidas los convierten en una opción confiable, especialmente cuando se dispone de suficientes datos de entrenamiento etiquetados [16].

- CTC
 - Entrenamiento de extremo a extremo: Los modelos basados en CTC ofrecen un enfoque simplificado de entrenamiento de extremo a extremo para el reconocimiento del habla. Eliminan la necesidad de alineaciones explícitas entre las secuencias de entrada y salida, lo que hace que el proceso de entrenamiento sea más sencillo.
 - Efectividad con transcripciones limitadas: CTC es particularmente útil cuando las transcripciones son limitadas o incompletas. Puede aprender a alinear el habla de entrada con las etiquetas de salida, incluso en casos donde la alineación exacta es desconocida o incierta [17].

El factor determinante en la selección de un modelo para este proyecto reside en la complejidad y dificultad asociadas a la construcción del conjunto de datos de entrenamiento. Según la información presentada, se observa que los modelos híbridos y CTC requieren grandes volúmenes de datos variados para lograr un entrenamiento adecuado, mientras que el aprendizaje por transferencia reduce dicha necesidad.

El aprendizaje por transferencia ha sido utilizado de manera efectiva en aplicaciones de reconocimiento de voz para adaptar modelos acústicos de dominios/idiomas con recursos abundantes a dominios/idiomas con recursos limitados. Este enfoque aborda el desafío de los recursos limitados y reduce el esfuerzo requerido para la preparación de datos, lo cual es especialmente crucial para idiomas con recursos limitados como el quichua. Mediante la transferencia de conocimiento desde modelos preentrenados, se pueden lograr mejoras significativas en el rendimiento, permitiendo el desarrollo de sistemas precisos de reconocimiento de voz para idiomas con recursos limitados [18].

1.5.4 Plataformas

En este apartado se exploran diferentes plataformas para el desarrollo de modelos de aprendizaje automático. Entre ellas se encuentran Edge Impulse, TensorFlow y ONNX, cada una con ventajas distintas en términos de facilidad de uso, versatilidad e interoperabilidad.

- Edge Impulse
 - Fácil de usar: Edge Impulse ofrece una plataforma amigable con una interfaz visual que simplifica el proceso de desarrollo de modelos de aprendizaje automático para dispositivos edge (de borde).
 - Flujo de trabajo integrado: Ofrece un flujo de trabajo integral que incluye la recolección de datos, entrenamiento de modelos y despliegue, lo que facilita el desarrollo de soluciones basadas en dispositivos edge.
 - Optimización para dispositivos edge: Edge Impulse se especializa en optimizar modelos para dispositivos de edge, garantizando un despliegue eficiente y que no requiere muchos recursos.
 - Amplio soporte de dispositivos: Edge Impulse se enfoca en permitir el desarrollo de TinyML, proporcionando bibliotecas y herramientas optimizadas específicamente diseñadas para ejecutar modelos de aprendizaje automático en microcontroladores [19].
- TensorFlow
 - Versatilidad: TensorFlow es una plataforma potente y versátil de aprendizaje automático que admite una amplia gama de arquitecturas y aplicaciones de modelos.

- Gran comunidad y ecosistema: TensorFlow cuenta con una amplia comunidad de desarrolladores, una extensa documentación y un rico ecosistema de bibliotecas y modelos preentrenados, lo que facilita encontrar soporte y recursos.
- Amplia adopción: Es ampliamente adoptado tanto en el ámbito académico como en la industria, y su popularidad garantiza un desarrollo y actualizaciones continuas.
- Flexibilidad: TensorFlow permite un control detallado sobre el desarrollo de modelos y ofrece diversas opciones de implementación, incluyendo dispositivos edge [20].
- ONNX (Intercambio abierto de redes neuronales).
 - Interoperabilidad de frameworks (marco de trabajo): ONNX permite la interoperabilidad entre diferentes frameworks de aprendizaje profundo, lo que te permite entrenar modelos en un framework y desplegarlos en otro, brindando flexibilidad y compatibilidad.
 - Intercambio de modelos: Ofrece un formato estandarizado para representar modelos de aprendizaje profundo, facilitando el intercambio y uso de modelos en diferentes plataformas y herramientas.
 - Integración: ONNX es compatible con frameworks populares como TensorFlow y PyTorch, lo que proporciona compatibilidad y facilita el uso de modelos en diferentes frameworks [21].

Considerando el enfoque en TinyML y el flujo de trabajo fácil de usar con herramientas especializadas y bibliotecas optimizadas para microcontroladores, Edge Impulse destaca como la opción predilecta para desarrollar el modelo pensado para el presente proyecto. Edge Impulse ofrece ventajas significativas para las aplicaciones de TinyML. Aborda las limitaciones en las

capacidades computacionales, las estrictas restricciones de memoria y la diversidad de hardware que se encuentran en los sistemas TinyML. Con técnicas eficientes de compresión y optimización, Edge Impulse permite adaptar los modelos a dispositivos con recursos limitados de memoria y energía. Su plataforma integrada simplifica el desarrollo y despliegue, admitiendo una amplia gama de hardware y fomentando la colaboración y la reproducibilidad. En el ámbito de TinyML, Edge Impulse destaca como la opción preferida para aprovechar su potencial en entornos con limitaciones de recursos [19].

2. CAPÍTULO 2

2.1 Metodología

2.1.1 Análisis

2.1.1.1 Requerimientos. En esta sección, se documentan las necesidades y expectativas del proyecto, sentando una base sólida para el desarrollo del producto final. Se describen detalladamente los requerimientos funcionales y no funcionales con el objetivo de garantizar que el dispositivo cumpla con las expectativas y necesidades de los usuarios finales.

2.1.1.1.1 *Requerimientos funcionales.*

Tabla 2.1

Dispositivo portable para notificación de alertas

Requerimiento funcional N°	RQF01
Tipo	Organizacional
Prioridad	Alta
Descripción	Permite enviar alertas en situaciones de riesgos a través de diferentes medios analizados por el grupo que realiza el desarrollo del api.

Nota. Autoría propia.

Tabla 2.2

Reconocimiento de palabra clave

Requerimiento funcional N°	RQF02
Tipo	Producto
Prioridad	Alta
Descripción	El dispositivo debe ser capaz de reconocer

una palabra clave específica pronunciada por el usuario para activar el envío de alertas.

Nota. Autoría propia.

Tabla 2.3

Integración con API externo

Requerimiento funcional N°	RQF03
Tipo	Organizacional
Prioridad	Media
Descripción	El dispositivo debe poder conectarse a un API desarrollado por estudiantes de ESPOL para enviar las alertas generadas, esto debe permitir el envío de información relevante.

Nota. Autoría propia.

Tabla 2.4

Autonomía y duración de la batería

Requerimiento funcional N°	RQF04
Tipo	Producto
Prioridad	Media
Descripción	El dispositivo debe tener una autonomía y duración de la batería adecuada para garantizar su funcionalidad durante un periodo prolongado de tiempo.

Nota. Autoría propia.

Tabla 2.5*Comunicación inalámbrica*

Requerimiento funcional N°	RQF05
Tipo	Producto
Prioridad	Alta
Descripción	El dispositivo debe ser capaz de establecer una conexión inalámbrica, como Bluetooth, para comunicarse con otros dispositivos o sistemas, incluyendo la API mencionada anteriormente.

Nota. Autoría propia.

Tabla 2.6*Alertas silenciosas*

Requerimiento funcional N°	RQF06
Tipo	Producto
Prioridad	Media
Descripción	El dispositivo debe enviar alertas de forma silenciosa, sin emitir sonidos o luces visibles, para evitar alertar a los agresores.

Nota. Autoría propia.

2.1.1.1.2 Requerimientos no funcionales**Tabla 2.7***Tiempo máximo de retardo en segundos*

Requerimiento no funcional N°:	RQNF01
Tipo	Rendimiento

Prioridad	Alta
Descripción	El sistema debe ser capaz de reconocer la palabra en quichua en tiempo real, con un retardo máximo de 5 segundos desde la captura del audio hasta la generación de la alerta a las autoridades.
<i>Nota. Autoría propia.</i>	

Tabla 2.8*Interfaz de usuario*

Requerimiento no funcional N°:	RQNF02
Tipo	Usabilidad
Prioridad	Media
Descripción	La interfaz de usuario del dispositivo debe ser intuitiva y fácil de usar, incluso para personas con poca experiencia técnica. Se deben seguir las pautas de diseño y accesibilidad para garantizar que los usuarios puedan interactuar de manera efectiva con el dispositivo.
<i>Nota. Autoría propia.</i>	

Tabla 2.9*Facilidad de mantenibilidad*

Requerimiento no funcional N°:	RNGQF03
Tipo	Mantenibilidad
Prioridad	Baja

Descripción	El sistema debe estar diseñado de manera modular y bien documentada para facilitar el mantenimiento y futuras actualizaciones.
-------------	--

Nota. Autoría propia.

Tabla 2.10

Disponibilidad de dispositivo y aplicación móvil

Requerimiento no funcional N°:	RNGF04
Tipo	Confiabilidad
Prioridad	Alta

Descripción	El sistema debe estar disponible de manera continua y confiable para garantizar una respuesta oportuna ante posibles actos delictivos. El tiempo de inactividad del sistema debe ser mínimo y se deben implementar medidas de redundancia y recuperación ante fallos para garantizar que el dispositivo esté operativo la mayor parte del tiempo.
-------------	---

Nota. Autoría propia.

Tabla 2.11

Integración con sistemas externos

Requerimiento no funcional N°:	RNQF05
Tipo	Integración
Prioridad	Media

Descripción	El dispositivo de reconocimiento de palabras en quichua debe ser capaz de integrarse con los sistemas de alerta existentes o futuros.
-------------	---

Nota. Autoría propia.

2.1.1.2 Alcance y limitaciones de la solución. El alcance del proyecto contempla el envío exitoso de alertas. Estas alertas se generan cuando el dispositivo portátil detecta el comando de voz “Sacha”. Para llevar a cabo esta comunicación, se empleó un dispositivo celular que se conectó mediante bluetooth y requirió acceso a internet. Una vez que el dispositivo portátil se encontró vinculado con la aplicación móvil, esta última estableció una conexión con una API externa, la cual, se encargó de enviar la alerta al contacto designado, quien recibió estas notificaciones de seguridad para actuar lo más pronto posible.

Una de las limitaciones más importantes que presentó el dispositivo portátil fue la dependencia de la conexión a internet. Para su funcionamiento óptimo, requirió de una conexión estable como la red WIFI. No obstante, en entornos urbanos, que representarían los escenarios ideales para la utilización de este dispositivo, se hace necesario recurrir a conexiones de datos móviles, sin embargo, en muchas ocasiones no garantizan la rapidez para la transmisión inmediata de alertas.

Otra restricción por considerar se encontró en el micrófono de la placa Nicla Voice. Es preciso señalar que este componente no ha sido específicamente diseñado para filtrar el ruido ambiental al momento de la captación auditiva.

2.1.1.2 Riesgos y beneficios de la solución. La solución puede conllevar ciertos riesgos. Uno de ellos se relaciona con la posibilidad de falsos positivos, incluso si el modelo de inteligencia artificial es altamente preciso, siendo un riesgo inherente de que este escenario ocurra. Otro riesgo es la dependencia de conectividad. La aplicación móvil, que está vinculada al dispositivo a través

de Bluetooth, necesita tener acceso a internet para asegurar el envío efectivo de las alertas. La falta de conexión podría resultar en la incapacidad de enviar alertas en momentos críticos.

No obstante, la solución presenta muchos más beneficios que superan los riesgos potenciales. Una de las ventajas es la respuesta rápida. Esto permite la detección rápida y automatizada de asistencia en situaciones de emergencia mediante el comando de voz “Sacha”. Además, asegura que las notificaciones de seguridad lleguen rápidamente al contacto designado.

Otra ventaja es la accesibilidad y facilidad de uso. Esta cualidad la hace apta para varios usuarios de diferentes edades y niveles de destreza tecnológica. También hay que destacar el potencial para salvar vidas. El uso de estas alertas puede ser esencial para prevenir accidentes o emergencias graves, lo que podría tener un impacto sumamente positivo en la seguridad de los usuarios. No debe pasarse por alto la alta capacidad de adaptabilidad del dispositivo a otras tecnologías y su facilidad de conexión a APIs externas. Estos factores contribuyen a reducir el riesgo asociado con la dependencia exclusiva de la conectividad a internet.

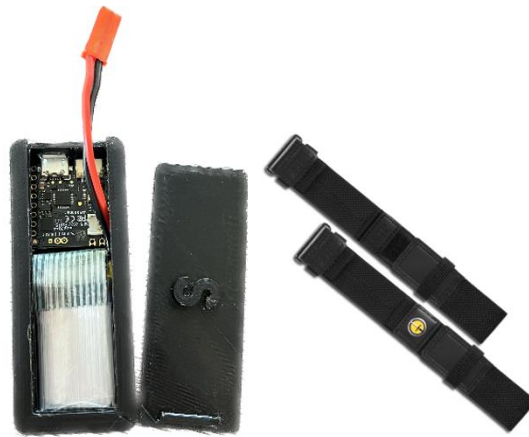
2.1.1.4 Usuarios de la solución. El proyecto contempla a varios perfiles de usuarios, entre ellos se encuentran los usuarios finales, quienes serán los portadores del dispositivo y utilizarán la detección de la palabra clave en situaciones de riesgo. Además, los desarrolladores involucrados en la creación y mejora del dispositivo constituyen un grupo esencial para garantizar su funcionalidad y rendimiento óptimos. No menos importante, las autoridades y organismos encargados de la respuesta a emergencias forman parte integral del proyecto, ya que recibirán las alertas generadas por el dispositivo y tomarán las medidas necesarias.

2.1.2 Prototipado.

2.1.2.1 Dispositivo portable.

Figura 2.1

Dispositivo portable



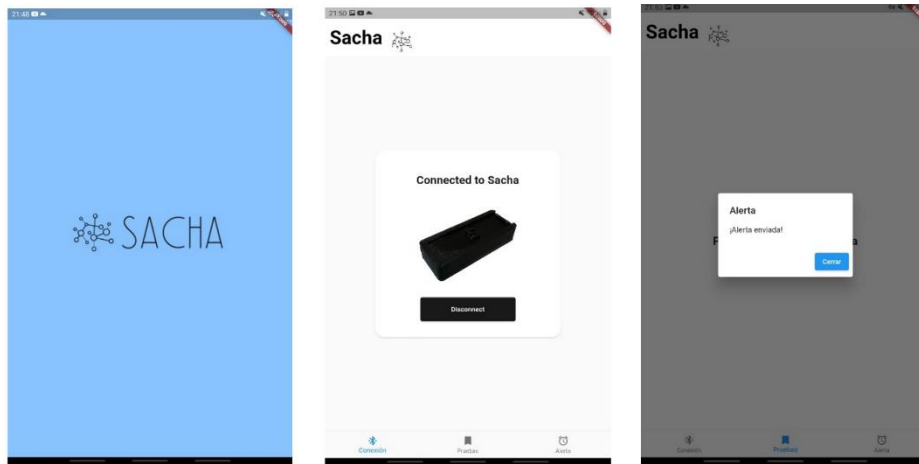
Nota. Autoría propia.

Para el prototipado del dispositivo portable capaz de reconocer la palabra clave “Sacha” se contó con una placa de desarrollo con un micrófono de alto rendimiento, una caja de PLA, una batería de litio y una correa de velcro. El propósito de este dispositivo es que se puede usar debajo de la ropa como un accesorio más. Se ha determinado que la ubicación del dispositivo podría ser en el hombro y entorno al bíceps. La idea consiste en asegurar que el dispositivo se encuentre a una distancia prudente que le permita capturar la voz del usuario.

2.1.2.2 Aplicación móvil

Figura 2.2

Aplicación móvil



Nota. Autoría propia.

Para el desarrollo de la aplicación se optó por priorizar su funcionalidad por encima del diseño, es por esto por lo que en la Figura 2.2, se aprecia un estilo sobrio en las diferentes pestañas. La aplicación es bastante sencilla, esta cuenta con dos apartados importantes: el de conexión y el de pruebas.

- Conexión

La segunda vista en la Figura 2.2 corresponde al apartado de conexión. En este se aprecia la conexión con el dispositivo, si ese no es el caso, se presenta un listado con los dispositivos compatibles cercanos. De igual manera si no se encuentran dispositivos simplemente se presenta el siguiente mensaje: “No se encuentran dispositivos”.

- Pruebas

La tercera vista en la Figura 2.2 corresponde al apartado de pruebas. En este se muestra una ventana afirmativa ante una alerta enviada exitosamente, en caso de no encontrarse una conexión activa se presenta el siguiente mensaje: “Conéctese al dispositivo primero”. Si se encuentra conectado al dispositivo, pero no se ha enviado la alerta, se muestra el siguiente mensaje: “Pruebe a dispositivo”.

2.1.2.3 Evaluación. Durante el proceso de desarrollo del artefacto, se inició con un diseño que sirvió como base para la creación del producto final. Este primer bosquejo, se sometió a pruebas con varios usuarios, con el propósito de observar cómo interactuaban con el dispositivo y detectar oportunidades de mejora que pudieran aplicarse en la versión definitiva.

Figura 2.3

Maquetación y diseño final del dispositivo



Nota. Autoría propia

En la Figura 2.3, se presentan los dos diseños, el de color blanco es el prototipo sometido a pruebas. Los usuarios que participaron en las pruebas coincidieron en que este diseño era bastante voluminoso, lo que resultaba incómodo al ponerse la ropa después de colocarse el prototipo. Además, la posición inicial del prototipo, que era horizontal, generaba un aspecto llamativo cuando se colocaba alrededor del bíceps. Esto ocasionaba una protuberancia notable y evidente, lo cual no fue bien recibido por los usuarios.

Inicialmente, se consideró un diseño que incluía dos arcos para encajar con dos aberturas, con el fin de tener una caja cerrada de manera segura. Sin embargo, esto resultó en un prototipo más grueso, lo que fue criticado por los usuarios debido a su visibilidad incluso bajo la ropa.

Otro aspecto que se mencionó con frecuencia en las evaluaciones fue el material utilizado. Varios usuarios destacaron que el prototipo generaba un ruido considerable al rozar con la ropa, debido al material empleado en la impresión. Este ruido causaba molestias entre quienes lo percibían.

Con base en los comentarios recopilados, se llegó a una versión final de la caja. En la Figura 2.3 se muestra el diseño definitivo en color negro. Esta versión redujo su tamaño al ajustar las partes internas más cercanamente y al emplear una batería más compacta. Además, se optó por un diseño con tapa deslizable, lo que disminuyó su grosor y lo volvió menos visible bajo la ropa. Se cambió el material y se redondearon los bordes de la caja para evitar cualquier ruido al entrar en contacto con las prendas de vestir. También se modificó la orientación a una posición vertical para evitar protuberancias cuando el dispositivo se colocará alrededor del bíceps.

2.1.3 *Diseño de la solución*

2.1.3.1 **Hardware**

2.1.3.1.1 *Características dispositivo portable.*

- Nicla Voice

Tabla 2.12

Características Nicla Voice

Microprocesador	Syntiant® NDP120 Neural Decision Processor™ (NDP)
Sensor	Micrófono de alto rendimiento (IM69D130)
Interfaz	Interfaz USB con funcionalidad de depuración
Memoria	512KB Flash / 64KB SRAM

	16MB SPI Flash para almacenamiento 48KB SRAM dedicada para NDP120
Conectividad	Bluetooth® de baja energía (ANNA-B112)

Nota. Autoría propia

- Batería

Tabla 2.13

Características batería

Capacidad	300mAh
Voltaje	3.7V
Tipo	Recargable de litio

Nota. Autoría propia

- Caja dispositivo portable.

Tabla 2.14

Características caja del dispositivo

Material	PLA
Dimensiones	6.55x1.95x1.40cm

Nota. Autoría propia

- Correa

Tabla 2.15

Dimensiones correa del dispositivo

Material	Velcro
Dimensiones	30x2cm

Nota. Autoría propia

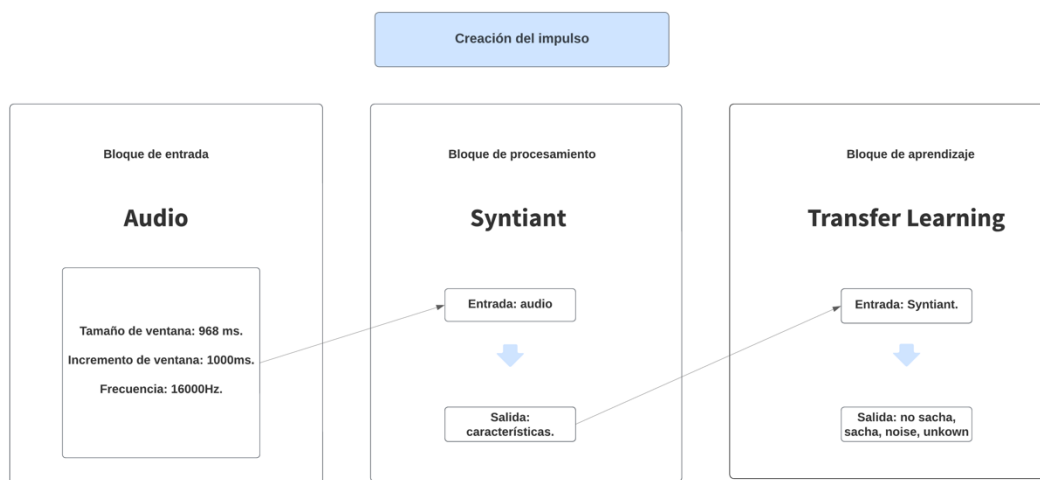
2.1.3.2 Software.

2.1.3.2.1 Arquitectura.

- Edge Impulse

Figura 2.4

Creación del impulso



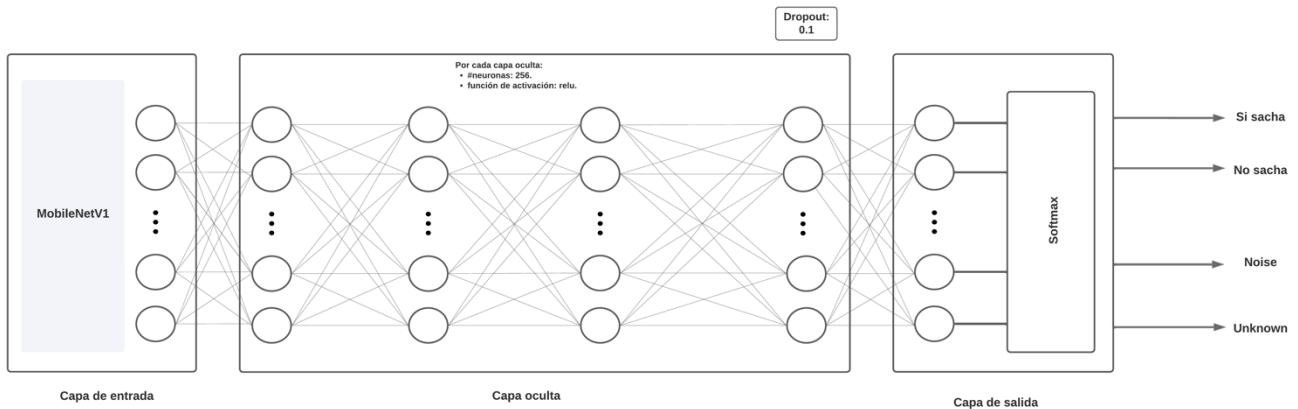
Nota. Autoría propia.

Debido a que la placa Nicla Voice incorpora un microprocesador Syntiant NDP120, fue importante elegir la opción correspondiente en el bloque de procesamiento al generar el impulso, como se indica en la Figura 2.4. Además, se configuró con precisión los parámetros en el bloque de entrada. En el bloque de procesamiento se usaron los parámetros por defecto recomendados por la plataforma Edge Impulse, se extrajeron características de acuerdo al algoritmo de Syntiant, las cuales fueron aprendidas por la red neuronal en el bloque de aprendizaje.

- Modelo de inteligencia artificial

Figura 2.5

Arquitectura modelo inteligencia artificial



Nota. Autoría propia

En el bloque de aprendizaje, se optó por la estrategia de transferencia de conocimiento. La arquitectura se basó en MobileNetV1 como se ilustra en la Figura 2.5 en la capa de entrada, previamente entrenada en el conjunto de datos de ImageNet. Esta red preentrenada fue ajustada para adaptarse a las particularidades de los datos específicos del proyecto.

En el modo experto de Edge Impulse se ajustó a cinco capas densas, cada una compuesta por 256 neuronas. En la quinta capa, se implementó una técnica de Dropout con una tasa del 10% para la desactivación aleatoria de neuronas, agregando robustez al modelo. La culminación del diseño se materializó con una capa densa adicional que empleó una función de activación Softmax, como se detalla visualmente en la Figura 2.5 en la capa de salida.

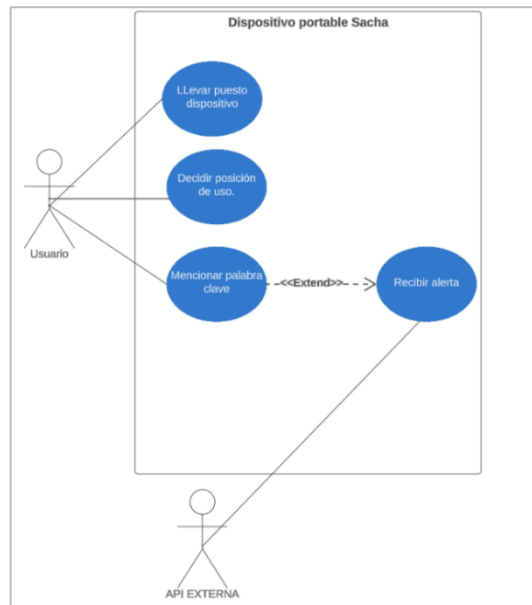
La configuración global del proceso incluyó la definición de 35 épocas de entrenamiento, una tasa de aprendizaje establecida en 0.0001 y un tamaño de lote (Batch Size) de 28.

2.1.3.2.2 Diagramas.

- Casos de uso.

Figura 2.6

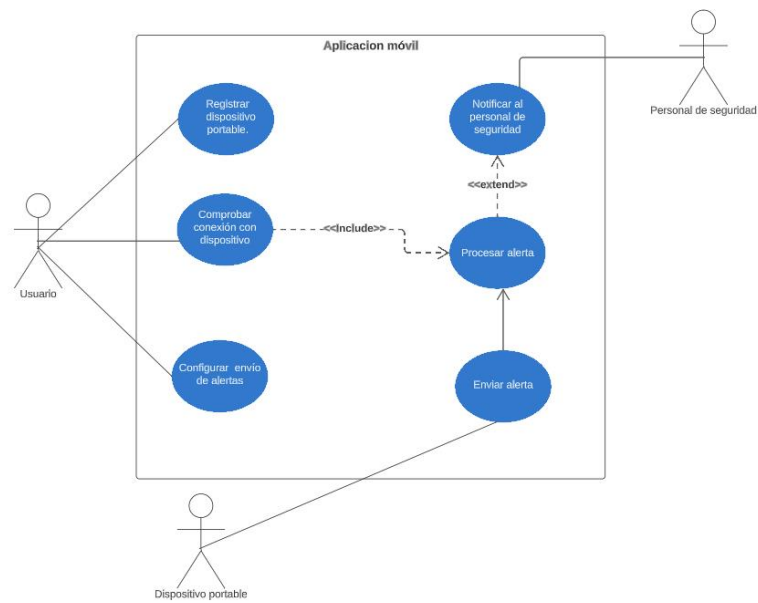
Caso de uso: Dispositivo portable.



Nota. Autoría propia

Figura 2.7

Casos de uso: Aplicación móvil

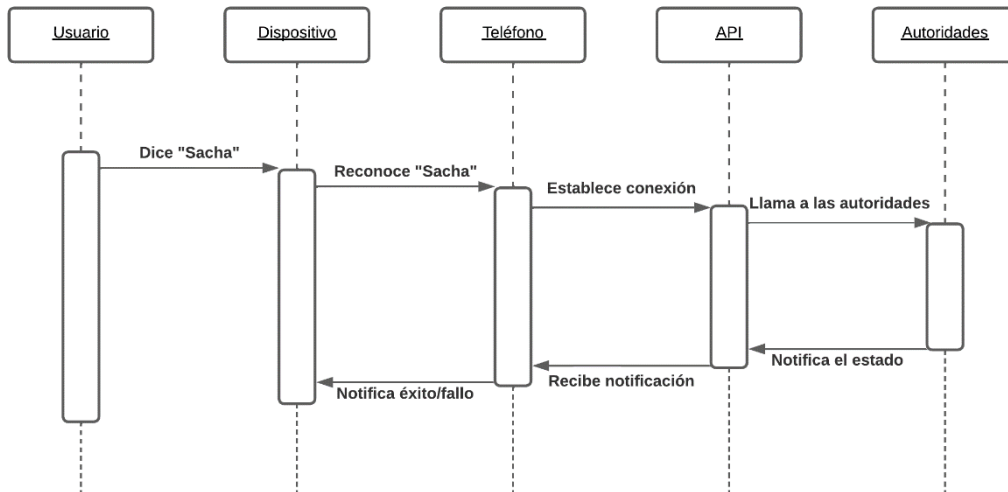


Nota. Autoría propia

- Diagrama de secuencia

Figura 2.8

Diagrama de secuencia

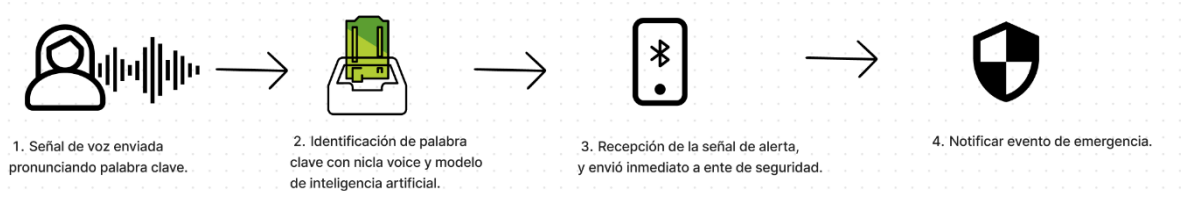


Nota. Autoría propia

- Diseño conceptual

Figura 2.9

Diseño conceptual



Nota. Autoría propia

2.2 Normativas

2.2.1 Consideraciones éticas y legales.

En este proyecto, se han abordado consideraciones éticas y legales para asegurar un desarrollo responsable y ético del dispositivo de reconocimiento de voz. Se ha priorizado la privacidad y el consentimiento de las personas cuyas voces serán procesadas. Se ha garantizado

la transparencia en el uso de datos y se han aplicado medidas de seguridad para proteger la privacidad. El diseño del dispositivo se orienta exclusivamente hacia usos legales y éticos, evitando infracciones de privacidad y daños. Se ha trabajado para prevenir sesgos y discriminación en el modelo de IA, promoviendo la imparcialidad. La transparencia en el funcionamiento del dispositivo y del modelo de IA ha sido esencial, permitiendo a los usuarios entender y cuestionar las decisiones. En términos legales, se ha cumplido con las leyes de protección de datos y se ha respetado la propiedad intelectual.

El cumplimiento normativo ha sido asegurado a través de la revisión de regulaciones relacionadas con el reconocimiento de voz y dispositivos portátiles. Se ha enfocado en la responsabilidad y seguridad en el diseño, implementando medidas para proteger a los usuarios y la energía.

3. CAPÍTULO 3

3.1 Pruebas

En este apartado se explica la evaluación exhaustiva que es llevo a cabo para el dispositivo desarrollado con el afán de validar su correcto funcionamiento y medir su nivel de eficacia. A continuación, se detalla el proceso seguido en estas pruebas donde el objetivo fue verificar la capacidad del artefacto para reconocer de manera confiable la palabra clave en los diferentes entornos donde se pueden dar situaciones de riesgo. Para esto, las pruebas se diseñaron en base a escenarios realistas que simulan dichas situaciones.

3.1.1 Pruebas sin ruido ambiente

El primer conjunto de pruebas se centró en evaluar el dispositivo en condiciones de ausencia de ruido ambiente. Este conjunto constó de grupos conformados de 10 pruebas cada uno. Se aplicó una metodología de combinación para explorar diversas combinaciones de parámetros, que incluyeron:

- Ubicación del Nicla Voice: Dentro de la caja y fuera de la caja.
- Distancia del dispositivo al usuario: 15 cm y 30 cm.

Esto resultó en un total de 40 pruebas, diseñadas para medir la respuesta del dispositivo en diferentes situaciones controladas.

3.1.2 Pruebas con ruido ambiente

En el segundo conjunto de pruebas, se evaluó el dispositivo en condiciones de ruido ambiente simulado. Al igual que en el conjunto anterior, se realizaron 40 pruebas, agrupadas en grupos de 10 pruebas cada uno. Se mantuvo la misma metodología de combinación de parámetros, pero se introdujeron escenarios de ruido ambiente, que incluyeron:

- Entorno urbano.
- Cafetería.
- Parque.

- Vehículo.

Estos escenarios representan entornos realistas donde el dispositivo podría ser utilizado. En total, se realizaron 160 pruebas para evaluar el rendimiento del dispositivo en condiciones de ruido ambiente.

3.1.3 Pruebas de tiempo de envío

Finalmente, se llevaron a cabo pruebas de tiempo de envío para medir la eficiencia de la aplicación móvil en el proceso de envío de alertas al backend. Estas pruebas se dividieron en dos grupos, cada uno compuesto por 30 pruebas:

- Grupo de Pruebas de Wifi: En este grupo, se midió el tiempo requerido para enviar alertas utilizando una conexión WiFi.
- Grupo de Pruebas de Datos Móviles: Aquí, se evaluó el tiempo de envío de alertas utilizando una conexión de datos móviles.

Estas pruebas se realizaron para comprender la variabilidad en el tiempo de envío en diferentes condiciones de conectividad, lo que resultó en un total de 60 pruebas.

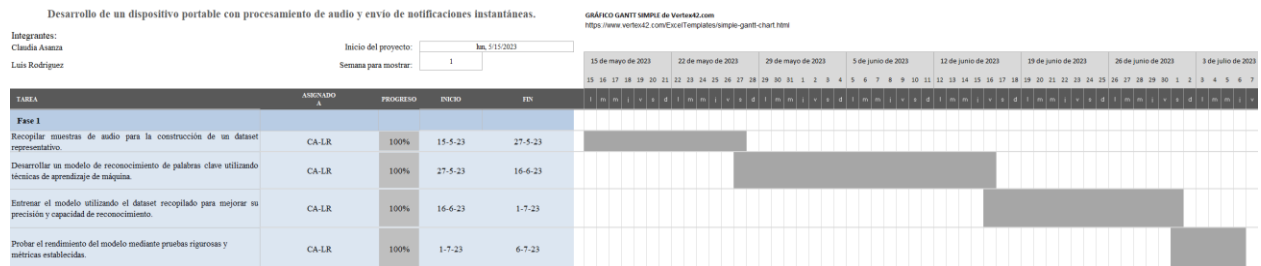
En resumen, las pruebas proporcionan una evaluación exhaustiva del rendimiento del dispositivo en diversas situaciones, incluyendo tanto condiciones controladas como escenarios realistas.

3.2 Plan de implementación

Este plan representa la culminación de la planificación del proceso de desarrollo del proyecto. Dividido en dos fases fundamentales, se aborda tanto la construcción y entrenamiento inicial del modelo como la fase de análisis de resultados y mejora continua.

Figura 3.1

Diagrama de Gantt fase 1

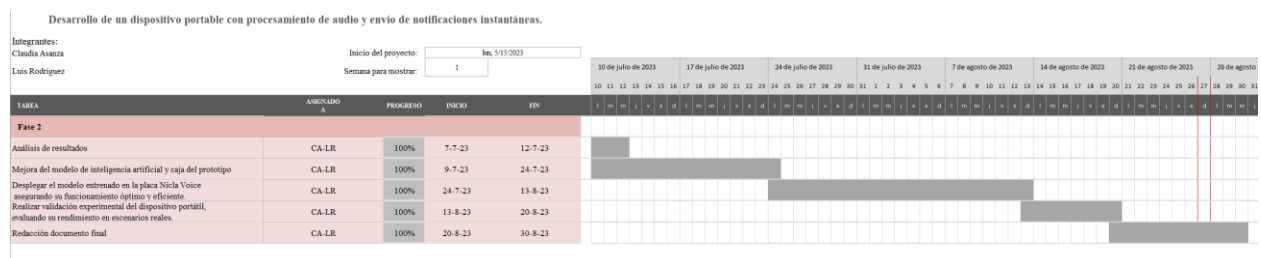


Nota. Autoría propia

En la Figura 3.1 presenta el plan relativo a la primera fase. En esta etapa inicial, se establecieron las bases para la creación del modelo de inteligencia artificial. Lo que involucró la recopilación y preparación exhaustiva de datos como la tarea de definición y toma de decisiones acerca de la arquitectura del modelo. Además, se llevó a cabo el entrenamiento del modelo y su posterior evaluación de rendimiento.

Figura 3.2

Diagrama de Gantt fase 2



Nota. Autoría propia

Una vez concluidas las pruebas de rendimiento del modelo, se procedió con la fase 2 del proyecto. En esta etapa, se destaca el análisis de los resultados previos y a la retroalimentación obtenida a partir de la primera versión de la solución. Como resultado de esta evaluación, se propuso un plan de mejora cuya ejecución se estimó en aproximadamente dos semanas, tal como se evidencia en la Figura 3.2.

Una vez implementado el plan de mejora, que incluyó la ampliación de la colección de datos, la realización de ajustes en los parámetros y arquitectura, además de la impresión 3D de la caja que portará el dispositivo portable, se dio paso al despliegue del modelo en la placa Nicla Voice. Esta fase fue llevada a cabo con el propósito de someter el modelo a pruebas finales en escenarios reales.

3.3 Resultados

En esta sección se exponen los resultados obtenidos en el contexto del proyecto, junto con un análisis de costos asociados. Además, se resalta la relevancia intrínseca de la solución y la contextualización dentro del panorama económico y empresarial.

3.3.1 Resultados del modelo de inteligencia artificial.

Una vez entrenado el modelo con los parámetros mostrados en el Capítulo 2 se procedió a evaluar su precisión, para esto se hizo uso de una matriz de confusión.

Tabla 3.1

Matriz de confusión

Matriz de Confusión				
	NO_SACHA	NOISE	SI_SACHA	UNKNOWN
NO_SACHA	87.0%	0.4%	9.4%	3.2%
NOISE	0.6%	87.0%	0%	12.3%
SI_SACHA	0.6%	0.5%	90.7%	1.0%
UNKNOWN	6.3%	30.30%	2.8%	60.6%
F1 SCORE	0.89	0.81	0.90	0.64

Nota. Autoría propia

En la Tabla 3.1, se presentan los valores obtenidos durante la evaluación. Las filas representan las clases reales, mientras que las columnas indican las clases predichas por el modelo. Para el análisis, se centró en la diagonal principal de esta matriz. Los valores en la diagonal principal señalan los Verdaderos Positivos obtenidos, se observaron valores relativamente altos, con excepción de la clase UNKNOWN, que registró un 60.6%. Aunque este último valor podría parecer inicialmente significativo debido a que se trata de un modelo de detección de palabras clave, se observó que la clase con más Falsos Positivos en relación directa con UNKNOWN es NOISE. Dado que estas dos clases comparten características similares, la coincidencia en sus resultados era esperada y se consideró de poca relevancia debido al contexto del proyecto.

El valor más destacado para el proyecto fue el 90.7%, lo que indica que el modelo logró reconocer correctamente los datos etiquetados como SI_SACHA. Además de los Verdaderos Positivos, se evaluaron los valores del F1 Score, siendo esta métrica preferida sobre la precisión debido a las particularidades del conjunto de datos. Aunque los datos de NO_SACHA y SI_SACHA mantienen un equilibrio, al igual que los datos de NOISE y UNKNOWN, se observa una ruptura de balance entre estos dos grupos. El F1 Score refleja un rendimiento sólido en las cuatro categorías de la matriz de confusión, lo que lo hace útil en contextos de conjuntos de datos desbalanceados [22].

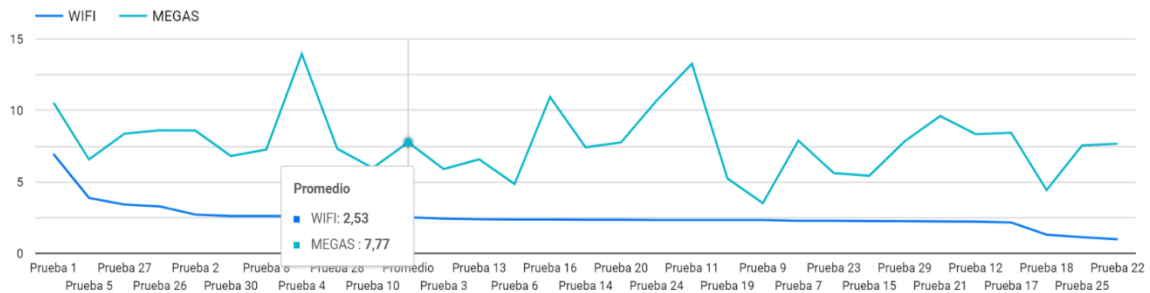
En este proyecto, la clase de mayor relevancia resultó ser SI_SACHA, ya que el modelo debe reconocerla para activar la alerta. Con un puntaje de 0.90 en el F1 Score, se determinó que el modelo estaba apto para su implementación.

3.3.2 Pruebas y validación del dispositivo portable.

3.3.2.1 Tiempo de envío de alerta.

Figura 3.3

Método de conexión vs tiempo de envío de alerta en segundos



Nota. Autoría propia

La Figura 3.3, ilustra una disparidad significativa en los tiempos de envío de alertas según el tipo de conexión. Con un promedio de tan solo 2.53 segundos. La conexión WIFI se destacó por su agilidad y eficiencia, mientras que la conexión basada en MEGAS contó con un promedio más elevado de 7.77s.

3.3.2.2 Reconocimiento de palabra clave y envío de alerta.

En la Tabla 3.2, se presenta un desglose detallado de características y capacidades del dispositivo portátil. Cabe resaltar que, si bien todos estos datos se proporcionan con precisión, es importante considerar que puede haber variaciones debido a factores externos que influyen en el rendimiento, por lo tanto, son datos aproximados a la realidad.

Tabla 3.2

Aproximación de especificaciones técnicas

Distancia conexión bluetooth	15 metros.
Duración de batería	14 días.
Tiempo de carga de batería	20 minutos cada 14 días.

Distancia máxima para reconocimiento de voz. Medio metro.

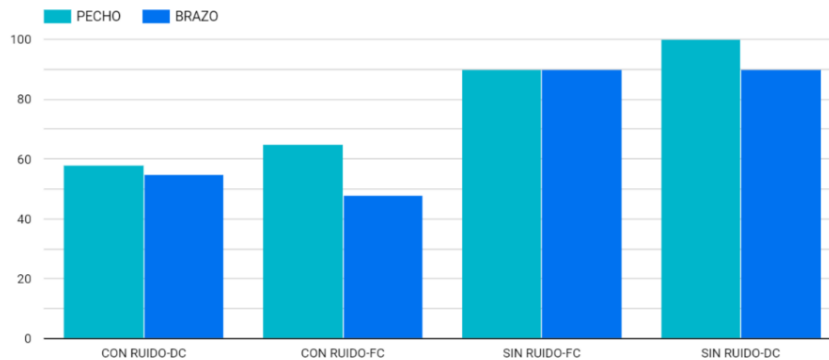
Nota. Autoría propia.

3.3.2.3 Comparativa de reconocimiento del comando de voz “Sacha” en entornos

ruidosos y sin ruido.

Figura 3.4

Reconocimiento de comando de voz “Sacha” en ambientes con ruido vs ambientes sin ruido



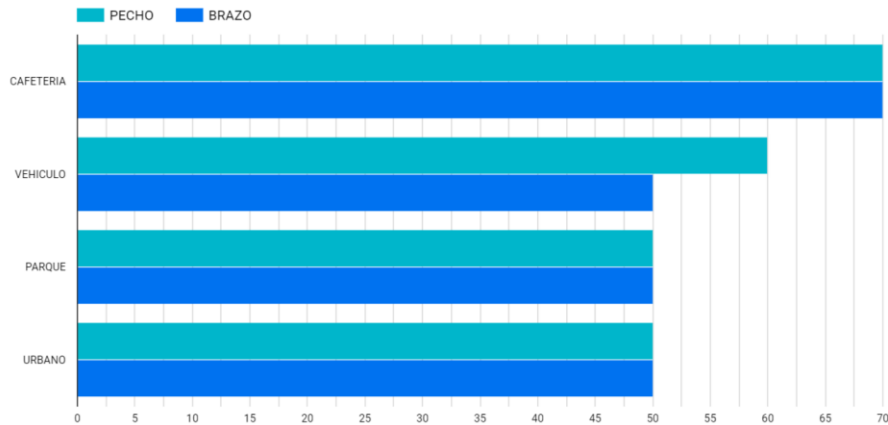
Nota. Autoría propia.

En la Figura 3.4, se puede evidenciar que el modelo presenta una mayor precisión en la identificación del comando de voz “Sacha” cuando se pronuncia a una distancia más corta, como en el caso del área del pecho, en comparación con distancias más largas, como el área del brazo. Además, los resultados de la Figura 3.4 indican que no se aprecian fluctuaciones significativas en relación con si el dispositivo se encontraba fuera o dentro de la caja portátil. También se destaca que, en entornos sin interferencias ambientales, se logró un mayor éxito de envío de alertas, incluso alcanzando un 100% de envíos exitosos.

3.3.2.4 Comparativa de envío de alerta en escenarios con ruido ambiental.

Figura 3.5

Escenarios de pruebas con ruido ambiental



Nota. Autoría propia.

La Figura 3.5 revela que, en entornos como cafeterías o vehículos, se logró un éxito en el envío de alertas igual o superior al 50%. Por otro lado, en entornos con niveles más altos de ruido ambiental, como el caso de parques o zonas urbanas, los casos mantuvieron un promedio de 50% de éxito de envío de alertas.

3.4 Análisis de costo

3.4.1 Desglose de costos

Tabla 3.3

Precio del dispositivo por unidad

Precio placa	\$82
Batería	\$6
Impresión 3D	\$18
Velcro	\$1

Total	\$107
--------------	-------

Nota. Autoría propia.

3.4.2 Análisis de mercado

Considerando los costos mencionados en la sección anterior, se ha decidido enfocar el público objetivo del proyecto en personas que utilizan vehículos propios y en empresas. Estos dos segmentos son particularmente adecuados debido a su alineación con los modelos de negocio que se detallan a continuación:

3.4.1.1 Modelos de negocio del proyecto.

3.4.1.1.1 Oferta de servicio de seguridad a través de suscripciones. Una alternativa viable para hacer que el producto sea más accesible para los grupos objetivo, dado sus considerables costos, sería presentarlo en forma de suscripción. Esto implicaría que la aplicación y el dispositivo desarrollado se incluyeran en la suscripción, garantizando su funcionamiento mientras la suscripción esté activa. Al igual que otros servicios, se aplicarían restricciones, como la devolución obligatoria del dispositivo una vez que se cancele el registro.

3.4.1.1.2 Adquisición personal del producto. Como se mencionó previamente, uno de los segmentos objetivo está compuesto por individuos con vehículos propios. Estos individuos estarán más inclinados a adquirir el producto, ya que enfrentan riesgos diarios debido a la situación actual del país. Además, se espera que tengan un mayor poder adquisitivo para cubrir el costo del producto.

3.4.1.1.3 Empresas que requieran el producto como servicio. Existe una oportunidad adicional en el ámbito empresarial. Muchas compañías cuentan con personal encargado del transporte de productos, empleados u otros fines de movilización. Dado el riesgo de la movilidad en la ciudad, ofrecer un servicio personalizado a las empresas podría ser una idea muy rentable. Esto implicaría proporcionar un número determinado de dispositivos y el acceso a la aplicación para los trabajadores de la empresa, brindando una solución integral de seguridad.

4. CAPÍTULO 4

4.1 Conclusiones y recomendaciones

4.1.1 Conclusiones

- Se recopilaron alrededor de 49 minutos de grabaciones de audio, separados en 2 363 segmentos de audio como datos de entrenamiento y 644 para pruebas.
- Mediante la aplicación de la técnica de transferencia de conocimientos, se desarrolló un modelo de inteligencia artificial, ajustando cuidadosamente los parámetros y el número de capas necesarios para lograr un proceso de aprendizaje preciso.
- El modelo se entrenó con el conjunto de datos recopilado y con las características mencionadas anteriormente, obteniendo una precisión de detección de la palabra clave de un 91%.
- Con el propósito de evaluar su rendimiento en diferentes contextos, se diseñó un plan de pruebas en ambientes ruidosos y sin ruido. Estos últimos demostraron que el dispositivo respondía de manera efectiva, permitiendo el envío de alertas con mayor rapidez y precisión.
- Se desplegó el build generado por la plataforma de Edge Impulse, en la placa Nicla Voice con éxito, donde se aseguró su funcionamiento óptimo y eficiente.
- La validación experimental se llevó a cabo en escenarios de la vida real, abarcando locaciones como cafeterías, parques, vehículos y entornos urbanos.

Finalmente, es importante resaltar el escenario óptimo para el correcto funcionamiento del dispositivo, el cual se caracteriza por contar con una conexión Wi-Fi. Esto se debe a su estabilidad, mayor velocidad y amplitud de banda. Por otro lado, la conectividad a través de MEGAS presenta desafíos, como la latencia y la congestión, lo que se traduce en tiempos de respuesta más prolongados y variables. Además, es fundamental colocar el dispositivo a una distancia no mayor de 30 cm para garantizar una mayor precisión en la detección de la palabra clave "Sacha", de acuerdo con las especificaciones técnicas mencionadas. Asimismo, se recomienda que el entorno

en el que se utilice el dispositivo sea lo más tranquilo posible, en contraste con lugares ruidosos como una cafetería o factores ambientales externos, para lograr un rendimiento óptimo.

4.1.2 Recomendaciones

4.1.2.1 Recomendaciones al desarrollar el proyecto. De la experiencia recogida a lo largo del desarrollo del presente proyecto se presentan las siguientes recomendaciones:

- **Obtención de datos:** Es recomendable obtener las grabaciones en ambientes variados, donde se simulen los escenarios de posibles usos. Es clave que las grabaciones sean expuestas al ruido con el afán de mejorar el rendimiento del modelo ante tales situaciones.
- **Cambio de plataforma:** Edge Impulse es una plataforma que ofrece el entrenamiento de modelos de manera sencilla, pero de igual manera tiene sus limitaciones, en especial por la memoria, por ello, el modelo a entrenar es poco personalizable. Se recomienda el cambio a una plataforma más versátil.
- **Modificación de la caja:** El modelado 3D de la caja fue diseñado para ser de tamaño reducido por intereses de uso, de igual manera, este es totalmente cerrado por los mismos motivos, esto provoca que no haya aberturas para el paso del sonido al micrófono. Como recomendación, se podría modificar el modelado con la intención de proveer estas aberturas para el paso de la voz al momento de realizar la alerta.
- **Cambio de la batería:** El Nicla Voice es una placa de tamaño notablemente reducido, lo que implica ciertos desafíos al soldar una batería externa. Una instalación incorrecta podría ocasionar que el estaño se extienda hacia otros pines, resultando en un cortocircuito. Por esta razón, se aconseja optar por una batería que cuente con un conector compatible con el puerto disponible en la placa para la conexión de baterías.

4.1.2.2 Recomendaciones para trabajos futuros. Mirando hacia adelante, se pueden considerar diferentes oportunidades para expandir el presente proyecto. En esta sección se exploran posibles ideas de desarrollo que puedan tomar este trabajo como base.

Considerando el objetivo del dispositivo desarrollado, una recomendación sería escoger un modelo de placa distinto. El Nicla Voice es un sistema embebido bastante completo, cuenta con varios sensores además de su sofisticado micrófono, es por ello por lo que el costo es alto. Para este proyecto, únicamente fue necesario emplear el micrófono, por lo que los demás sensores se quedaron sin utilizar. El cambio a una placa distinta puede representar la reducción en el precio de desarrollo.

De la mano del punto anterior, se podría considerar el uso de un micrófono externo. El micrófono incorporado en el Nicla Voice funciona correctamente; sin embargo, en entornos con mucho ruido, su rendimiento se ve comprometido. Si se opta por el cambio de placa, sería aconsejable implementar un micrófono con una capacidad de captación superior, de esta forma se garantizaría el correcto funcionamiento del modelo en situaciones donde predomina el ruido ambiental.

Si se piensa usar el mismo modelo de placa, existen formas de sacarle el mayor provecho con nuevas alternativas de proyecto. Debido a que esta placa cuenta con una variedad de sensores, entre los que se encuentra un acelerómetro, se podría optar por cambiar el modo en realizar la alerta. Suponiendo el uso del acelerómetro, el modelo podría ser reutilizado con cambios en los datos para realizar un nuevo proceso de entrenamiento. En esta nueva configuración, la alerta ya no estaría basada en el uso de voz, sino que el dispositivo podría distinguir entre situaciones cotidianas y accidentes, generando alertas específicamente para este último caso.

Entre las diversas consideraciones para tener en cuenta, surgen nuevas perspectivas sobre las posibles aplicaciones del dispositivo. Una de estas sería que el usuario no necesite llevar consigo el dispositivo, sino que esté instalado en el vehículo que utiliza. De esta manera, se lograría una accesibilidad constante al artefacto encargado de emitir alertas, ya sea por medio de la voz o su alternativa utilizando el acelerómetro. Esta implementación permitiría que el dispositivo encargado de gestionar las alertas esté siempre a mano y disponible.

Dicha aplicación cobra mayor relevancia en situaciones donde el acceso a redes de internet es limitado. Aquí es donde entra en juego la posibilidad de aprovechar el Internet de las Cosas (IoT). En este escenario, el dispositivo funcionaría como un periférico y, en situaciones de emergencia, enviaría alertas a través de una red de largo alcance y bajo consumo, como la red LoRaWAN, utilizando un gateway como intermediario. Esta estrategia amplía significativamente las capacidades del dispositivo y su utilidad en contextos específicos.

Referencias

- [1] R. C. Franco Reina, “Desarrollo de un mecanismo de localización e identificación de personas atrapadas en zonas catastróficas utilizando dispositivos embebidos inalámbricos de bajo consumo,” 2022, Accessed: Jun. 08, 2023. [Online]. Available: <http://dspace.ups.edu.ec/handle/123456789/22879>
- [2] Primicias, “Guayaquil registra aumento del 66% de muertes violentas en 2023,” 2022. <https://www.primicias.ec/noticias/sucesos/guayaquil-aumento-muertes-violentas/> (accessed Jun. 08, 2023).
- [3] CTS, “dp-crime-violent-offences | dataUNODC,” 2021. <https://dataunodc.un.org/dp-crime-violent-offences> (accessed Jun. 08, 2023).
- [4] United Nations Office on Drugs and Crime (UNODC), “Informe Mundial sobre Trata de Personas 2022 – Principales hallazgos,” 2022. Accessed: Jun. 08, 2023. [Online]. Available: https://www.unodc.org/documents/peruandecuador/Adjuntos/BriefGLOTIP2022_Ecuador.pdf
- [5] Banco Mundial (BIRF - AIF), “Proyecto,” *Banco Mundial*, 2022. https://projects.bancomundial.org/es/projects-operations/projects-summary?lang=es&countrycode_exact=EC&os=0 (accessed Jun. 08, 2023).
- [6] United Nations Office on Drugs and Crime (UNODC), “Policia Nacional del Ecuador recibio capacitacion internacional sobre uso de la fuerza y derechos humanos,” 2022. <https://www.unodc.org/peruandecuador/es/noticias/2021/policia-nacional-del-ecuador-recibio-capacitacion-internacional-sobre-uso-de-la-fuerza-y-derechos-humanos.html> (accessed Jun. 08, 2023).

- [7] United Nations Office on Drugs and Crime (UNODC), “Ecuador elabora de manera participativa e interinstitucional su Estrategia Nacional contra el Crimen Organizado,” 2023. <https://www.unodc.org/peruandecuador/es/noticias/2021/ecuador-elabora-de-manera-participativa-e-interinstitucional-su-estrategia-nacional-contra-el-crimen-organizado.html> (accessed Jun. 08, 2023).
- [8] Ana Maria Carvajal, “Estudiantes de la Universidad Central se movilizaron para exigir seguridad,” May 30, 2023. <https://www.expreso.ec/quito/estudiantes-universidad-central-movilizaron-exigir-seguridad-162125.html> (accessed Jun. 08, 2023).
- [9] M. Saranya and A. Professor, “WOMEN SAFETY BAND BY USING IOT WITH ARDUINO MEGA MICRO-CONTROLLER,” *International Journal of Creative Research Thoughts*, vol. 8, no. 6, pp. 2320–2882, 2020, Accessed: Jun. 08, 2023. [Online]. Available: www.ijcrt.org
- [10] J. F. Gaviria *et al.*, “Deep Learning-Based Portable Device for Audio Distress Signal Recognition in Urban Areas”, doi: 10.3390/app10217448.
- [11] P. Gupta, Y. Goyal, Di. Gulati, P. Rastogi, and S. Saboo, “Women’s Safety Device Based on Internet of Things,” *2021 IEEE International Conference on Mobile Networks and Wireless Communications, ICMNWC 2021*, 2021, doi: 10.1109/ICMNWC52512.2021.9688529.
- [12] “Nicla_Voice_DataSheet.pdf - Google Drive.” https://drive.google.com/file/d/16r_tl6nqjIFV4_9TBszWZFvvfdJ2ENly/view (accessed Jun. 08, 2023).
- [13] “Arduino_Nano_33_BLE_Sense_DataSheet.pdf - Google Drive.” <https://drive.google.com/file/d/1PWQMP8CiXTIQE6pBPh3CSr7dyt4Za1YI/view> (accessed Jun. 08, 2023).

- [14] “Seed_nrf52_DataSheet.pdf - Google Drive.”
<https://drive.google.com/file/d/1qVQIe4GQSwLrWYg7yFQYwyoMFbbUmozs/view>
(accessed Jun. 08, 2023).
- [15] Y. Woldemariam, “Transfer Learning for Less-Resourced Semitic Languages Speech Recognition: The Case of Amharic,” *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)*, pp. 61–69, 2020.
- [16] A. Graves, “Towards End-To-End Speech Recognition with Recurrent Neural Networks,” *Proceedings of the 31st International Conference on Machine Learning*, vol. 32, pp. 1764–1772, Jun. 2014.
- [17] G. Hinton *et al.*, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” *IEEE Signal Process Mag*, vol. 29, no. 6, pp. 82–97, 2012.
- [18] S. Pan and Q. Yang, “A survey on Transfer Learning,” *IEEE Trans Knowl Data Eng*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [19] S. Hymel *et al.*, “Edge Impulse: An MLOps Platform for Tiny Machine Learning,” Nov. 2022.
- [20] M. Abadi *et al.*, “TensorFlow: A System for Large-Scale Machine Learning,” *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283, Nov. 2016.
- [21] M. Nielsen, D. Amodei, and A. Graves, “ONNX: An Open Format for Interoperability in Machine Learning,” *IEEE International Conference on Computer Vision (ICCV)*, pp. 325–334, 2017.

- [22] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, no. 1, Jan. 2020, doi: 10.1186/s12864-019-6413-7.

Anexos

A. Manual de configuración dispositivo Sacha.

