

FuSA Roads

Sebastián Montecinos, Renato Atencio, Orlando Contreras, Javier Reyes y Handel Venegas

Abstract

El crecimiento urbano y el aumento del tráfico vehicular han intensificado la contaminación acústica en entornos urbanos, afectando la calidad de vida y la salud de la población. Fusa Roads propone una solución innovadora para el monitoreo y análisis de la contaminación sonora y el tráfico mediante inteligencia artificial. A través del uso de modelos avanzados de visión por computadora (YOLO) y análisis de audio (PANN), el sistema procesa videos capturados por cámaras con micrófonos estratégicamente ubicadas en la ciudad. La plataforma permite la carga de videos, su procesamiento distribuido y la generación de mapas de calor del ruido urbano, facilitando la toma de decisiones en planificación y gestión del tráfico. La arquitectura de Fusa Roads se basa en tecnologías escalables como MinIO para el almacenamiento de objetos, FastAPI para la gestión de APIs y React con TypeScript para la interfaz web. Este enfoque busca proporcionar una herramienta accesible y eficiente para la recopilación de datos urbanos, contribuyendo a la reducción de la contaminación acústica y la optimización del flujo vehicular.

Index Terms

Inteligencia Artificial, YOLO, PANN, MinIO, mapa de calor, contaminación acústica

I. INTRODUCCIÓN

EL crecimiento urbano y el aumento del tráfico vehicular han generado un impacto significativo en la calidad de vida de las personas, especialmente debido a la contaminación acústica. La medición y análisis de los niveles de ruido en entornos urbanos es fundamental para desarrollar estrategias de mitigación y regulación ambiental. Sin embargo, los métodos tradicionales de monitoreo suelen ser costosos, requieren una gran cantidad de recursos y no siempre permiten obtener un registro detallado y preciso del impacto del ruido generado por el tráfico.

En respuesta a esta problemática, surge Fusa Roads que busca ofrecer una solución eficiente y escalable mediante un enfoque basado en visión por computadora y análisis de audio, utilizando modelos de aprendizaje profundo.

El sistema funcionará a partir de videos capturados por cámaras equipadas con micrófonos, ubicadas estratégicamente en distintos puntos de la ciudad. Estos videos podrán ser subidos a una plataforma web en la cual se tendrá la opción de crear 'proyectos de IA', donde se seleccionarán videos según la necesidad y estos serán procesados utilizando modelos avanzados de inteligencia artificial, como YOLO (You Only Look Once) para la detección de vehículos en las imágenes y PANN (Pre-trained Audio Neural Networks) para el análisis de los niveles de ruido en el audio. Los datos obtenidos permitirán la generación de mapas de calor del ruido urbano, proporcionando información clave para la toma de decisiones en la planificación y gestión del tráfico.

El desarrollo de Fusa Roads se basa en el uso de tecnologías modernas y escalables, incluyendo MinIO para el almacenamiento de objetos, FastAPI para la gestión de las APIs, y React con TypeScript para la creación de una interfaz web dinámica e intuitiva. Con esta infraestructura, el sistema busca proporcionar una herramienta accesible y eficiente para la recopilación y análisis de datos urbanos, contribuyendo a la reducción de la contaminación sonora y a la optimización del flujo vehicular en la ciudad.

II. PROBLEMA DETECTADO

La gestión del tráfico y la contaminación acústica en las ciudades se ha convertido en un desafío creciente, especialmente en áreas urbanas con alta densidad de población y flujo vehicular intenso. En Valdivia, como en muchas otras ciudades, la ausencia de datos precisos y en tiempo real sobre el tráfico y los niveles de ruido dificulta la implementación de políticas públicas eficaces para mejorar la movilidad y mitigar el impacto ambiental. Sin una infraestructura adecuada de monitoreo, es complejo evaluar con exactitud la distribución del tráfico y la contaminación sonora en distintos sectores de la ciudad, lo que impide la toma de decisiones informadas sobre planificación urbana y medidas de reducción del ruido ambiental.

La exposición prolongada al ruido ambiental no solo afecta la calidad de vida de los ciudadanos, sino que también representa un riesgo significativo para la salud. Según estudios científicos, el ruido excesivo está asociado con trastornos cardiovasculares, endocrinos, gástricos y respiratorios, además de impactos en la salud mental como alteraciones del sueño, estrés, depresión, irritabilidad y aumento de la mortalidad por enfermedades relacionadas. La Organización Mundial de la Salud (OMS) ha establecido límites recomendados para la exposición al ruido de tránsito vehicular, fijando en 53 dBA el nivel máximo para el período diurno y 45 dBA para la noche. Sin embargo, en muchas ciudades, estos valores son ampliamente superados.

A nivel internacional, la Comunidad Europea ha implementado desde 2002 la obligatoriedad de diagnosticar las condiciones acústicas urbanas mediante la elaboración y actualización de mapas de ruido, una herramienta fundamental para la planificación y gestión del ambiente sonoro. En Chile, el Ministerio del Medio Ambiente estima que más de 1.2 millones de personas en el Gran Santiago están expuestas a niveles críticos de ruido ambiental, y la Superintendencia de Medio Ambiente señala que el 40% de las denuncias recibidas corresponden a contaminación acústica.

Actualmente, no existen soluciones accesibles ni automatizadas que permitan recopilar y analizar estos datos de manera eficiente. Los métodos tradicionales de monitoreo de tráfico y ruido suelen ser costosos y requieren personal especializado, lo que limita su implementación a gran escala.

Fusa Roads emerge como una solución innovadora a esta problemática, ofreciendo una plataforma que facilita el análisis automatizado y distribuido de datos sobre tráfico y contaminación acústica. A diferencia de los sistemas de monitoreo en tiempo real, este enfoque permite que los videos, capturados por cámaras con micrófono estratégicamente ubicadas en la ciudad, sean almacenados y procesados posteriormente. Utilizando modelos avanzados de inteligencia artificial, el sistema analizará grandes volúmenes de datos con alta precisión. El modelo YOLO permitirá identificar y contabilizar los vehículos en cada grabación, mientras que PANN analizará los registros de audio para evaluar los niveles de ruido en distintos momentos del día.

Esta solución no solo facilitará la detección de zonas con alta contaminación acústica y tráfico vehicular, sino que también contribuirá a la formulación de estrategias más efectivas para la reducción del ruido urbano y la optimización del flujo vehicular en Valdivia.

III. SOLUCIÓN PROPUESTA

Nuestra solución consiste en el desarrollo de un sistema capaz de analizar el tráfico vehicular y la contaminación acústica en la ciudad mediante el uso de inteligencia artificial y almacenamiento distribuido. Este sistema integrará modelos de Machine Learning (ML) para procesar videos y audios capturados en distintos puntos estratégicos, generando información clave sobre la cantidad de vehículos y los niveles de ruido en cada zona.

Uno de los principales objetivos de nuestra propuesta es la generación de archivos en un formato compatible que pueda ser utilizado como entrada para los modelos de machine learning. Esto garantizará que el procesamiento de datos se realice de manera eficiente y estructurada.

Para gestionar eficientemente los datos, se implementarán dos bases de datos: una relacional, que permitirá organizar y consultar información estructurada sobre los procesos del sistema, y una NoSQL, que almacenará metadatos de los videos y audios procesados. Además, el sistema contará con almacenamiento distribuido a través de MinIO, garantizando una gestión escalable y accesible de los archivos multimedia.

El sistema proporcionará una interfaz interactiva donde los usuarios podrán visualizar mapas de ruido y consultar información procesada de manera intuitiva. También incluirá un módulo de generación de reportes, facilitando la toma de decisiones basada en datos precisos y actualizados.

IV. DISEÑOS

A. Modelo Físico

El modelo físico se refiere a la infraestructura tangible necesaria para implementar y ejecutar un sistema distribuido. Este modelo describe cómo se organiza y distribuye el hardware, los recursos y los dispositivos en el entorno físico, abarcando todos los componentes que permiten el funcionamiento efectivo del sistema. En términos prácticos, el modelo físico detalla la distribución de servidores, redes de comunicación, almacenamiento de datos y los dispositivos de entrada/salida que interactúan dentro del sistema.

1) *Infraestructura*: Las cámaras serán las encargadas de capturar video y audio en diferentes ubicaciones. Inicialmente, los videos se podrán subir de manera manual a MinIO a través de la página web. MinIO estará alojado en servidores locales, proporcionando almacenamiento distribuido y escalable para gestionar los archivos de manera eficiente.

El procesamiento de los videos y audios será realizado por los proyectos de inteligencia artificial (PIA), que ejecutarán modelos de IA para analizar tanto las imágenes como el sonido. Estos modelos estarán disponibles como APIs y recibirán los archivos almacenados en MinIO para su análisis.

Los datos generados por los modelos de IA, serán almacenados inicialmente en bases de datos en formato JSON. Posteriormente, esta información será consolidada en una base de datos unificada que permitirá la consulta y el análisis eficiente de los resultados, facilitando su integración en el sistema de visualización y toma de decisiones.

B. Modelo Fundamental

El modelo fundamental se refiere a la lógica interna y las reglas que determinan cómo interactúan los componentes del sistema, cómo se manejan las tareas y procesos, cómo se gestionan los errores, y cómo se asegura la sincronización.

1) *Modelo de interacción*: En cuanto a la interacción dentro del sistema, los videos y audios pueden ser subidos en cualquier momento, lo que otorga flexibilidad a los usuarios. Dado que el sistema permite el procesamiento de archivos de diversas duraciones, los videos pueden variar significativamente en su extensión, pudiendo durar desde un par de horas hasta incluso todo un día. Esta variabilidad en la duración de los archivos implica que el tiempo de procesamiento también será impredecible y dependiente del tamaño y complejidad de los mismos. Debido a esto, se opta por un enfoque **asíncrono** en el que los archivos se suben y procesan de manera independiente, sin necesidad de esperar a que uno termine para iniciar otro, permitiendo que el sistema gestione múltiples tareas simultáneamente y de manera eficiente.

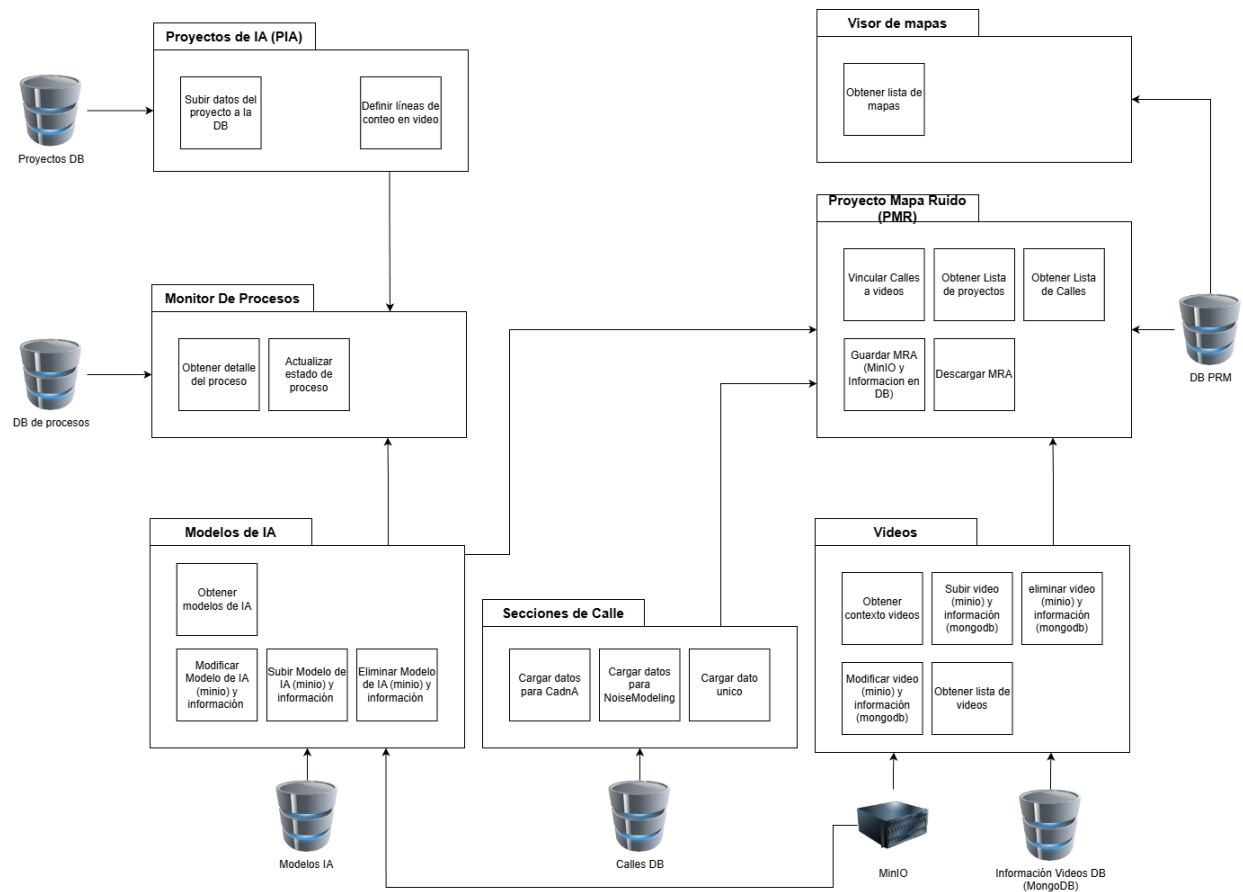


Fig. 1. Arquitectura de micro-servicios V1

2) *Modelo de fallas*: El manejo de fallas es un aspecto esencial dentro de cualquier sistema distribuido. En este caso, se deben prever posibles fallas tanto en la subida de archivos como en su procesamiento. Si ocurre un error durante la subida de un video o audio, el sistema debe ser capaz de identificar el problema y permitir que el usuario reintente la carga sin perder los datos previamente subidos. Para los casos en los que el procesamiento falle, el sistema debe incluir mecanismos de reintento o redirección automática a otro nodo disponible que pueda continuar con el análisis. Además, un sistema de notificación será clave para alertar a los usuarios sobre cualquier tipo de fallo en el proceso, brindando transparencia y control sobre el estado de los archivos subidos y procesados.

3) *Modelo de seguridad*: Inicialmente, se contemplan mecanismos como middleware de validación, que verificará la estructura y coherencia de los mensajes en las APIs antes de su procesamiento, asegurando que cumplan con los formatos esperados. Además, se utilizará autenticación basada en JWT (JSON Web Tokens) para gestionar el acceso seguro a los servicios del sistema.

Para el almacenamiento de archivos en MinIO, se definirán perfiles de acceso con permisos específicos, asegurando que solo usuarios autorizados puedan interactuar con los datos almacenados. De igual manera, en la interfaz web se implementará un sistema de autenticación y gestión de perfiles, permitiendo asignar diferentes niveles de acceso según el rol del usuario.

Dado que la seguridad es un aspecto dinámico y crítico, durante el desarrollo del proyecto se evaluará continuamente la necesidad de reforzar o ampliar las medidas de protección. En caso de identificar riesgos adicionales o nuevas vulnerabilidades, se implementarán mecanismos de seguridad complementarios para fortalecer la protección del sistema y sus datos.

C. Diagramas

1) *Arquitectura por micro-servicios*: Ver Fig 1

V. PREGUNTAS PARA EL CLIENTE

Para diseñar una solución eficiente y alineada con las necesidades del cliente, es fundamental definir con precisión los requisitos del sistema. Para ello, se plantean una serie de preguntas clave que permiten comprender mejor los objetivos del proyecto, las condiciones operativas y las expectativas en términos de procesamiento y análisis de datos. Estas preguntas abarcan aspectos técnicos y funcionales, proporcionando una base sólida para el desarrollo de la plataforma Fusa Roads.

A continuación, se presentan las preguntas junto con las respuestas esperadas:

- 1) ¿Cuál es el objetivo principal del análisis de tráfico y ruido?
R: Obtener datos detallados sobre el flujo vehicular y los niveles de ruido en la ciudad para mejorar la planificación urbana y reducir la contaminación acústica.
- 2) ¿Cuál es el formato de video y audio con el que se trabajará?
R: Los videos podrán estar en formato ... y los audios en formato ...
- 3) ¿Qué nivel de precisión se espera en el conteo de vehículos y análisis de ruido?
R: Se requiere al menos un 90% de precisión en la detección de vehículos y una clasificación del ruido con un margen de error del 5 dB.
- 4) ¿Cuál es la infraestructura disponible para el procesamiento de los videos?
R: (Actualmente los modelos de IA estan como API)
- 5) ¿Qué capacidad de almacenamiento se necesita para los videos y mapas de ruido?
R: Un video de calidad SD (entre 240p y 480p) de 24 horas pesa alrededor de 8.5 gigas (Según la IA de google). Considerando que podrán subir múltiples videos asumimos que el mínimo de almacenamiento, solo considerando los videos, seria de 100 gigas.
- 6) ¿Dónde se alojará el sistema: en servidores locales, en la nube o una combinación de ambos?
R: Servidores locales inicialmente, pero el paso a una nube debería de ser posible.
- 7) ¿Se requiere redundancia o respaldo de los datos en caso de fallos?
R: Lo ideal es que exista redundancia para evitar una perdida de datos permanente, pero dependerá de la infraestructura disponible.
- 8) ¿Cómo se notificará a los usuarios sobre el estado de procesamiento de sus videos?
R: A través de notificaciones en la web y correo electrónico cuando el análisis esté listo.
- 9) ¿Se permitirá el acceso público a los datos o solo estará disponible para ciertos usuarios?
R: Inicialmente se tiene pensado tener perfiles de usuario y que cada usuario pueda hacer uso solamente de los videos subidos por el mismo.
- 10) ¿Habrá diferentes niveles de acceso y roles de usuario dentro del sistema?
R: Sí, se definirán roles como administrador, analista y usuario general, cada uno con permisos específicos.

VI. HERRAMIENTAS UTILIZADAS

Fusa Roads se basa en una arquitectura diseñada con tecnologías de software libre que permiten el procesamiento eficiente de videos y audios, garantizando escalabilidad y flexibilidad en su implementación. Para la detección de vehículos en los videos se utilizará YOLO (You Only Look Once), un modelo avanzado de deep learning especializado en la identificación rápida y precisa de objetos en imágenes y secuencias de video. Este modelo es ampliamente utilizado en aplicaciones de visión por computadora debido a su alto rendimiento y capacidad de procesamiento en tiempo real. Para el análisis del ruido ambiental, se empleará PANN (Pre-trained Audio Neural Networks), una red neuronal preentrenada optimizada para la clasificación y detección de patrones en señales de audio. Su uso permitirá identificar niveles de contaminación acústica con alta precisión, facilitando la generación de mapas de calor del ruido en la ciudad.

En cuanto a la infraestructura del sistema, se implementará Docker, una plataforma de contenedorización que facilitará el despliegue y la gestión de los distintos componentes del sistema en entornos distribuidos. Su uso permitirá garantizar la portabilidad y escalabilidad de los servicios, evitando conflictos entre dependencias y asegurando un entorno homogéneo de ejecución. En complemento, se considera el uso de Kubernetes como herramienta de orquestación de contenedores, lo que permitirá distribuir la carga de trabajo de manera eficiente en múltiples nodos, optimizando así los recursos disponibles para el procesamiento de datos.

Para el almacenamiento de videos y audios, se utilizará MinIO, una solución de almacenamiento distribuido compatible con el protocolo S3. Esta herramienta es ideal para manejar grandes volúmenes de datos multimedia, como videos y audios capturados en distintas ubicaciones de la ciudad. MinIO permite almacenar y recuperar archivos de manera eficiente, garantizando tanto la persistencia como la accesibilidad de los datos procesados a través de un sistema escalable. Gracias a su compatibilidad con S3, se puede integrar fácilmente con diversas aplicaciones y servicios, asegurando que los datos estén disponibles para su análisis de forma rápida y segura.

En cuanto a la gestión de datos estructurados y no estructurados, Fusa Roads empleará una combinación de dos bases de datos para optimizar el almacenamiento y la recuperación de la información. Se utilizará PostgreSQL, una base de datos relacional, para gestionar las relaciones entre los distintos componentes del sistema, como los registros de procesamiento, las métricas obtenidas y los datos de referencia que requieren integridad y consistencia. Paralelamente, se implementará MongoDB, una base de datos NoSQL orientada a documentos, que permitirá almacenar metadatos de los videos, resultados intermedios del análisis y otros datos de estructura flexible. Esta combinación garantizará un acceso rápido y eficiente a la información, permitiendo un manejo óptimo tanto de los datos estructurados como de aquellos que requieren mayor flexibilidad.

En la capa de desarrollo web y API, el sistema utilizará FastAPI, un framework de alto rendimiento basado en Python. FastAPI se destaca por su capacidad para construir APIs rápidas, escalables y fáciles de mantener, lo que facilita la optimización

del procesamiento de solicitudes y respuestas. Gracias a su enfoque en la eficiencia, el sistema podrá manejar un gran volumen de tráfico de datos sin comprometer el rendimiento, permitiendo que los usuarios accedan a la información de forma casi inmediata.

Finalmente, para la creación de la interfaz web, se empleará React.js, pero en esta implementación se utilizará TypeScript en lugar de JavaScript. TypeScript es un lenguaje que extiende JavaScript, incorporando un sistema de tipos estáticos que permite detectar errores de tipo durante el desarrollo, antes de que el código se ejecute. Esto mejora la calidad del código y reduce los posibles fallos en tiempo de ejecución. Además, TypeScript facilita la escalabilidad del proyecto, ya que ofrece una mayor claridad y seguridad, especialmente cuando el código crece en complejidad. También permite aprovechar herramientas avanzadas de autocompletado y refactorización, lo que mejora la productividad del equipo de desarrollo. Al integrarse con React.js, TypeScript aporta una robustez adicional, permitiendo la creación de interfaces dinámicas e interactivas, basadas en componentes reutilizables, que podrán visualizar de manera clara y eficiente la información procesada sobre el tráfico y la contaminación sonora en la ciudad. Con esta combinación, el proyecto se beneficiará de un flujo de trabajo más organizado y seguro, optimizando la experiencia de usuario y el rendimiento general del sistema.

VII. CONCLUSIÓN

El crecimiento urbano y el aumento del tráfico vehicular han generado una problemática significativa en términos de contaminación acústica, afectando la calidad de vida y la salud de la población. Sin embargo, las soluciones tradicionales para medir y analizar el impacto del ruido urbano suelen ser costosas y poco escalables. En este contexto, Fusa Roads se presenta como una alternativa innovadora, basada en el uso de inteligencia artificial y almacenamiento distribuido para procesar grandes volúmenes de datos de manera eficiente.

A través de la integración de modelos avanzados como YOLO para la detección de vehículos y PANN para el análisis de audio, Fusa Roads permitirá generar mapas de calor del ruido en la ciudad, proporcionando información clave para la toma de decisiones en la planificación urbana y la gestión del tráfico. Además, su arquitectura escalable, basada en tecnologías modernas como MinIO, FastAPI y React con TypeScript, garantiza un sistema eficiente y accesible.

Este enfoque no solo facilitará la detección de zonas con alta contaminación acústica y tráfico vehicular, sino que también contribuirá a la formulación de estrategias efectivas para la mitigación del ruido urbano y la optimización de la movilidad en Valdivia. En futuras iteraciones, el sistema podrá incorporar nuevas funcionalidades, como la integración con sensores IoT en tiempo real.

REFERENCES

- [1] Facultad de Ciencias de la Ingeniería UACH, "Ingeniería UACH se adjudicó proyecto que busca entregar soluciones tecnológicas para la gestión ambiental del ruido vehicular," Universidad Austral de Chile, 27 de marzo de 2025. [Online].
- [2] Acústica UACH, "FUSA Roads – Proyecto de gestión ambiental del ruido vehicular," Acústica UACH, 2025. [Online].