

# Reconstrucción de audio a partir de MRI

## Proyecto SGBI

Por Pablo González  
Santamarta



# Table of contents

01

## Definición del problema

Definición de los  
objetivos del problema

02

## Ideas generadas

Subdivisión del problema  
en tareas

03

## Objetivos de aprendizaje

Conocimientos por adquirir  
para solucionar el problema

04

## Investigación

Medios por los que se han  
adquirido los  
conocimientos

05

## Solución al problema

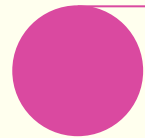
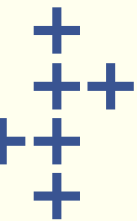
Solución final del  
problema

06

## Resultado

Resultados obtenidos



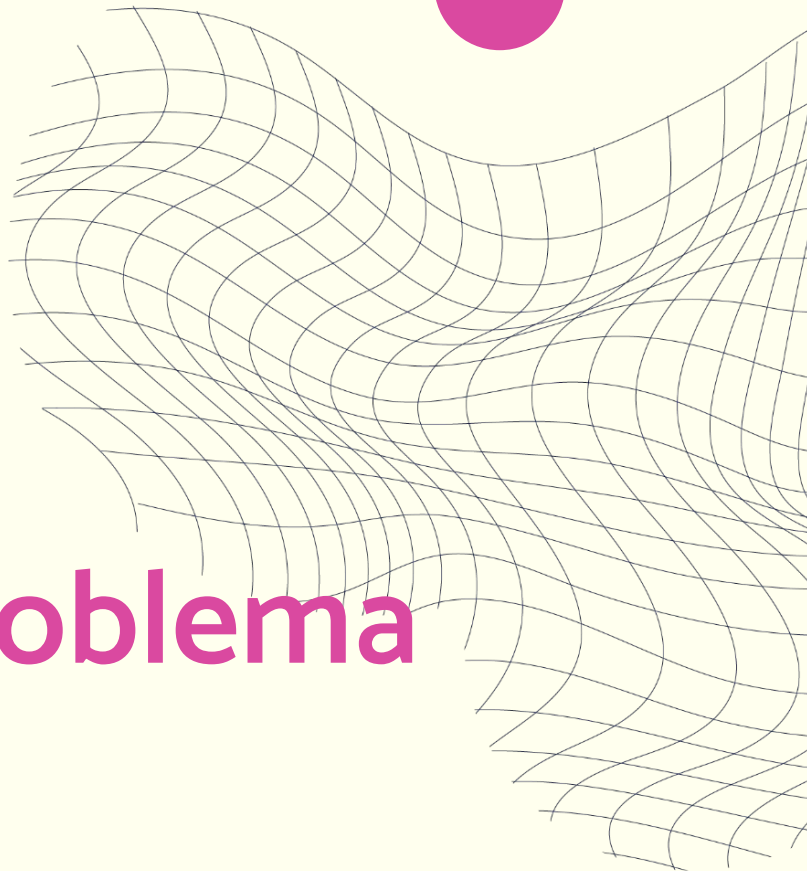


# 01

# Definición del problema

---

Definición de los objetivos del problema



# Reconstrucción de audio a partir de MRI

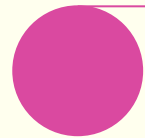
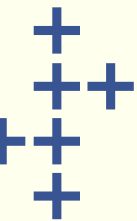
## ¿Qué es un MRI?

Un MRI es una imagen médica obtenida por la técnica de Resonancia Magnética. En éste caso hablamos específicamente de MRIs del cerebro

## ¿Qué quiere decir reconstrucción de audio?

La idea es utilizar un MRI tomado a un sujeto mientras éste escucha un audio para reconstruir ese audio.

Si el MRI se toma mientras el sujeto escucha un ladrido, el programa debe reconstruir ese ladrido a partir del MRI

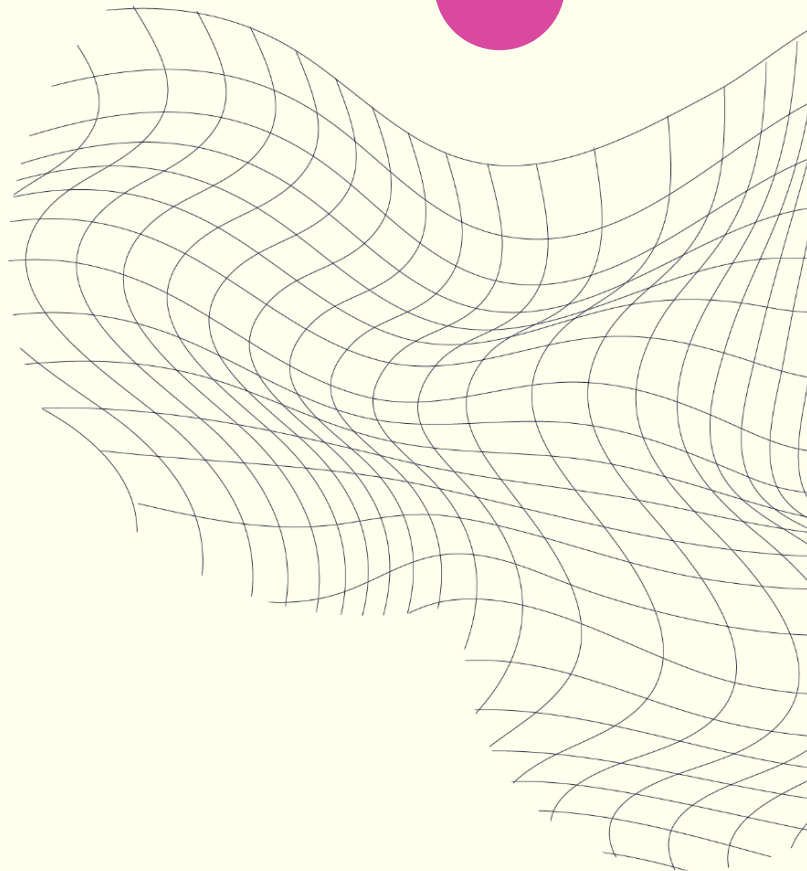


# 02

## Ideas Generadas

---

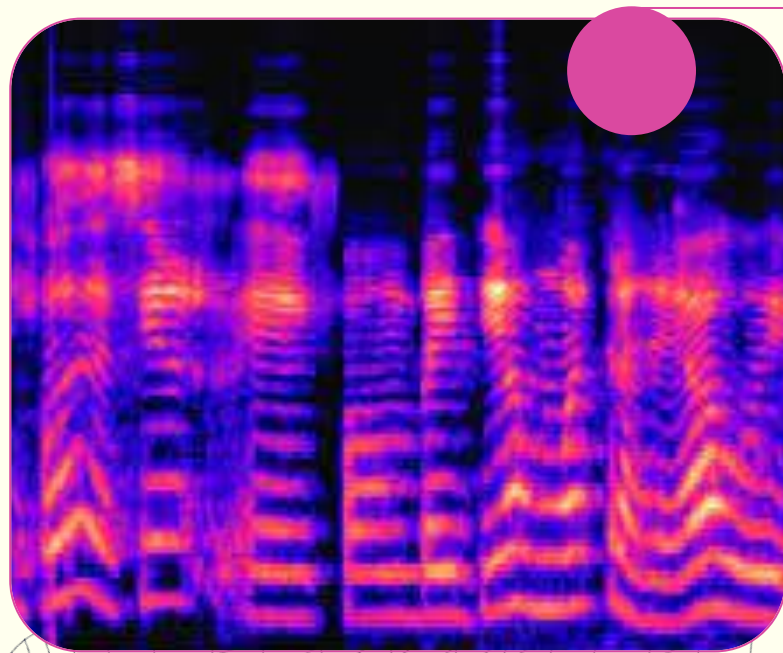
Subdivisión del problema en tareas

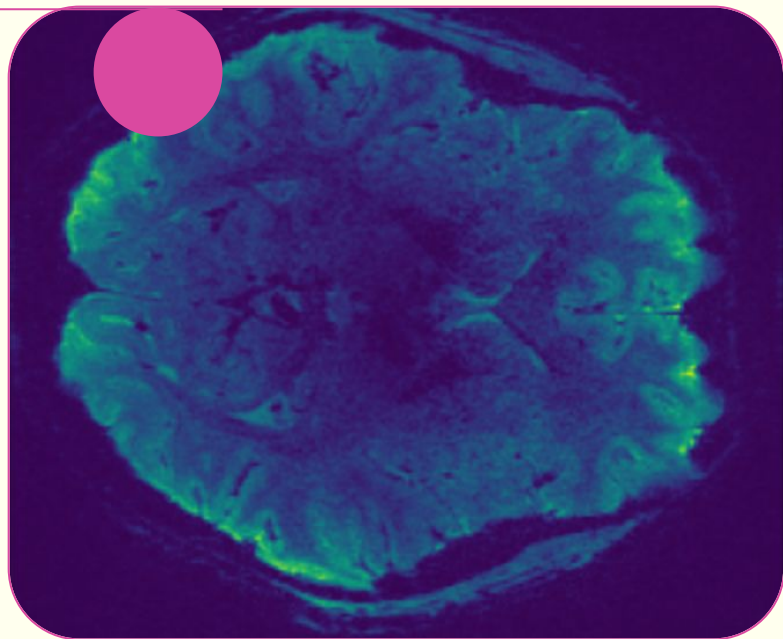




# Representación y procesamiento del audio

Se debe representar el audio de una forma adecuado que permite extraer sus características correctamente





# Representación y procesamiento de los MRI

Hay que encontrar una forma adecuada de leer y representar los datos de MRI que muestre la actividad relacionada con el audio

# ++ Correspondencia entre MRI y Audio



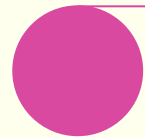
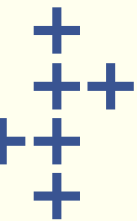
Hay que encontrar una forma adecuada de leer y representar los datos de MRI que muestre la actividad relacionada con el audio





# Reconstrucción del audio a partir de representación MRI





# 03

## Objetivos de aprendizaje

---

Conocimientos por adquirir para solucionar el problema



# Conocimientos

## Situación actual

En lo relevante al problema:

- Programación en python
- Programación en Matlab
- Visión por computador básica
- Procesamiento de imágenes
- Deep Learning básico



1

Herramientas de programación



2

Procesamiento de señales de audio



3

Diseño de RNN generativas

Procesamiento de archivos MRI

4

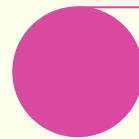
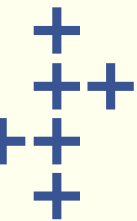


Comprensión de datasets

5



Necesarios para el problema



# 04

# Investigación

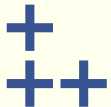
---

Medios por los que se ha adquirido el conocimiento



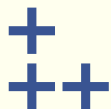
# Artículos de investigación I

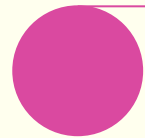
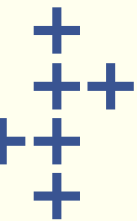
- Vaswani, A., et al. (2017). *Attention is all you need*. Advances in neural information processing systems, 30.
  - Redes neuronales transformer, capas de atención.
- Yuan, L., et al. (2021). *Tokens-to-token vit: Training vision transformers from scratch on imagenet*. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 558-567).
  - Aplicación de transformadores a problemas de visión
- Huo, Y. (ed). (2021). *Sec. Brain Imaging Methods*. Front. Neurosci., 15 - 2021.  
<https://doi.org/10.3389/fnins.2021.795488>
  - Métodos de reconstrucción de imágenes a partir de MRI
- Ren, Z., et al. (2021). *Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning*. NeuroImage, 228, 117602.
  - Aplicación de arquitectura VAE-GAN a reconstrucción de imágenes a partir de MRI. Transfer Learning



# Artículos de investigación II

- Razghandi, M., et al. (2022). *Variational autoencoder generative adversarial network for Synthetic Data Generation in smart home*. In ICC 2022-IEEE International Conference on Communications (pp. 4781-4786).
  - Arquitectura VAE-GAN y usos.
- Oord, A. V. D., Vinyals, O., & Kavukcuoglu, K. (2017). *Neural discrete representation learning*. arXiv preprint arXiv:1711.00937.
  - Extensión de la arquitectura VAE usando Cuantización Vectorial
- Iashin, V., et al. (2021). *Taming visually guided sound generation*. arXiv preprint arXiv:2110.08791.
  - Aplicación de la arquitectura VQ-GAN a la generación de audio





# 05

# Solución del problema

---

Solución final al problema



# Herramientas utilizadas



## Google Colab

Programación en python  
utilizando cuadernos y  
ejecución cloud en una T4



## Pytorch

Librería con la que he  
desarrollado la  
arquitectura y el  
entrenamiento



OpenNEURO

## OpenNeuro

Repositorio de datos  
experimentales de  
neurología



# Dataset utilizado

Michelle Moerel and Essa Yacoub (2023). *High-res gradient echo EPI and 3D GRASE data of auditory cortex*. OpenNeuro. [Dataset] doi: [doi:10.18112/openneuro.ds004814.v1.0.0](https://doi.org/10.18112/openneuro.ds004814.v1.0.0)



## Sujetos

Es dataset está compuesto de MRIs de 6 sujetos distintos



## Sonidos

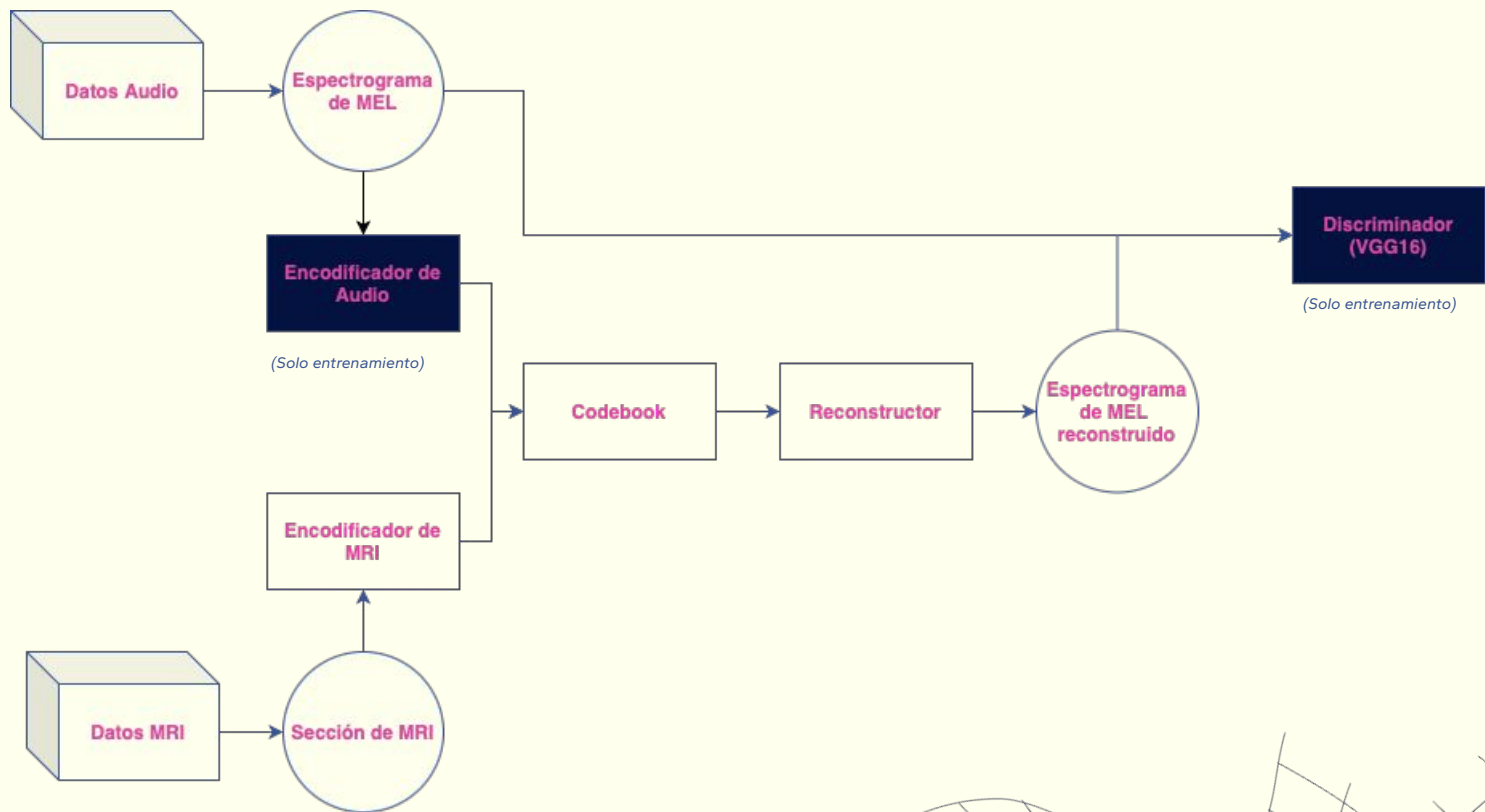
Los sujetos escuchan sonidos de entre 144 sonidos naturales de 1 segundo



## Experimento

En cada ejecución del experimento se toma un MRI del sujeto mientras escucha sonidos marcados por la tabla de eventos de la ejecución





# Entrenamiento

Fase	Encoder MRI	Encoder Audio	Codebook	Reconstructor	Discriminador
1					
2					
3					
Eval					



: No se usa

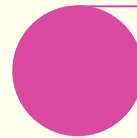
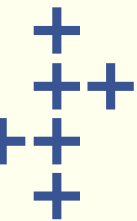


: Se entrena



: No se entrena





# 06

# Resultados

---

Resultado del proyecto



# Resultados del proyecto

## ¿Hasta dónde he llegado?

Debido a limitaciones en mi conocimiento, el hardware que utilizo y los datos de los que dispongo, no he conseguido crear un prototipo que funcione de la solución.

Lo que he conseguido es:

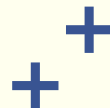
- Procesar el dataset
- Implementar una arquitectura VQ-GAN
- Entrenar la fase 1

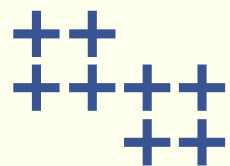
## 1 Problemas a solucionar

- El tamaño del espectrograma reconstruido depende del tamaño del input.
- El espectrograma reconstruido tiene pocos detalles

## 2 Objetivos de futuro

- Sustituir el discriminador por uno mejor
- Utilizar SPM para preprocesar los MRI.





# iGracias!

¿Alguna pregunta?

[pgonzs08@estudiantes.unileon.es](mailto:pgonzs08@estudiantes.unileon.es)

+34 650 11 34 32