

UANL



FCFM

Authors

Marco Antonio Obregón Flores

Affiliations

Universidad Autónoma de Nuevo León
Facultad de Ciencias Físico Matemáticas

1 Introducción

Este estudio se dirige a un desafío muy concreto dentro del reconocimiento de voz: la interpretación precisa de comandos de voz para controles direccionales. La finalidad es adaptar esta tecnología para ser especialmente útil para personas con discapacidades motoras, quienes podrían beneficiarse de una interacción más efectiva con dispositivos electrónicos mediante comandos de voz.

Mi teoría es que a través de la combinación de técnicas avanzadas de procesamiento de señales y modelos de aprendizaje profundo, se puede desarrollar un sistema que interprete comandos direccionales con buena precisión. Este avance sería una contribución a la calidad de vida de las personas con discapacidades motoras.

2 Objetivos

Principal: Desarrollar un sistema de reconocimiento de voz de alta precisión para interpretar comandos direccionales, destinado a mejorar la interacción de personas con discapacidades motoras con dispositivos electrónicos.

Secundarios:

- Investigar y aplicar técnicas de preprocesamiento y extracción de características avanzadas para optimizar la precisión del sistema.
- Aplicar y validar modelos de aprendizaje profundo para interpretar comandos de voz de manera más efectiva que los enfoques tradicionales.
- Diseñar e implementar estrategias para mitigar el impacto del ruido externo en la precisión del sistema, permitiendo una mayor robustez en entornos no ideales.

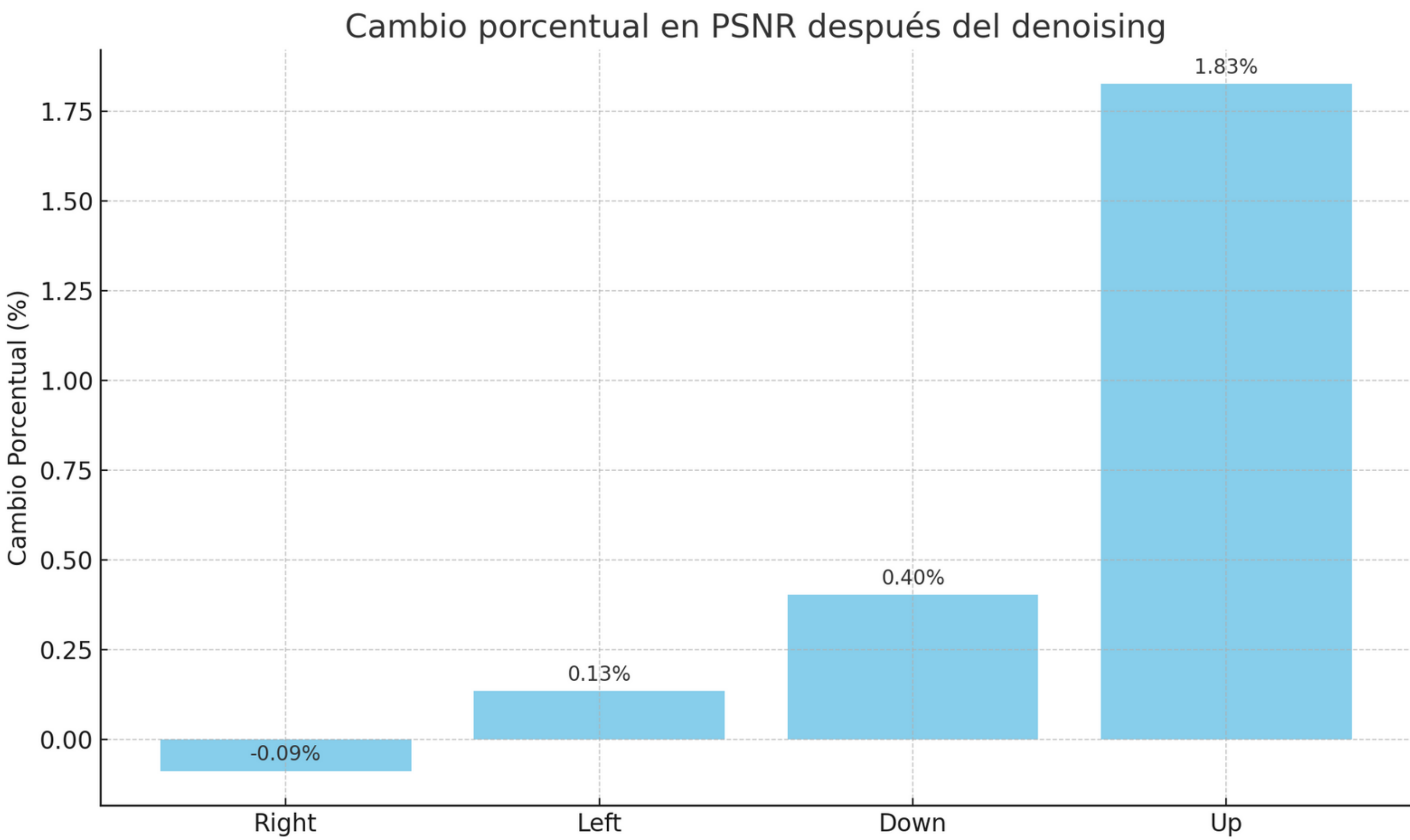
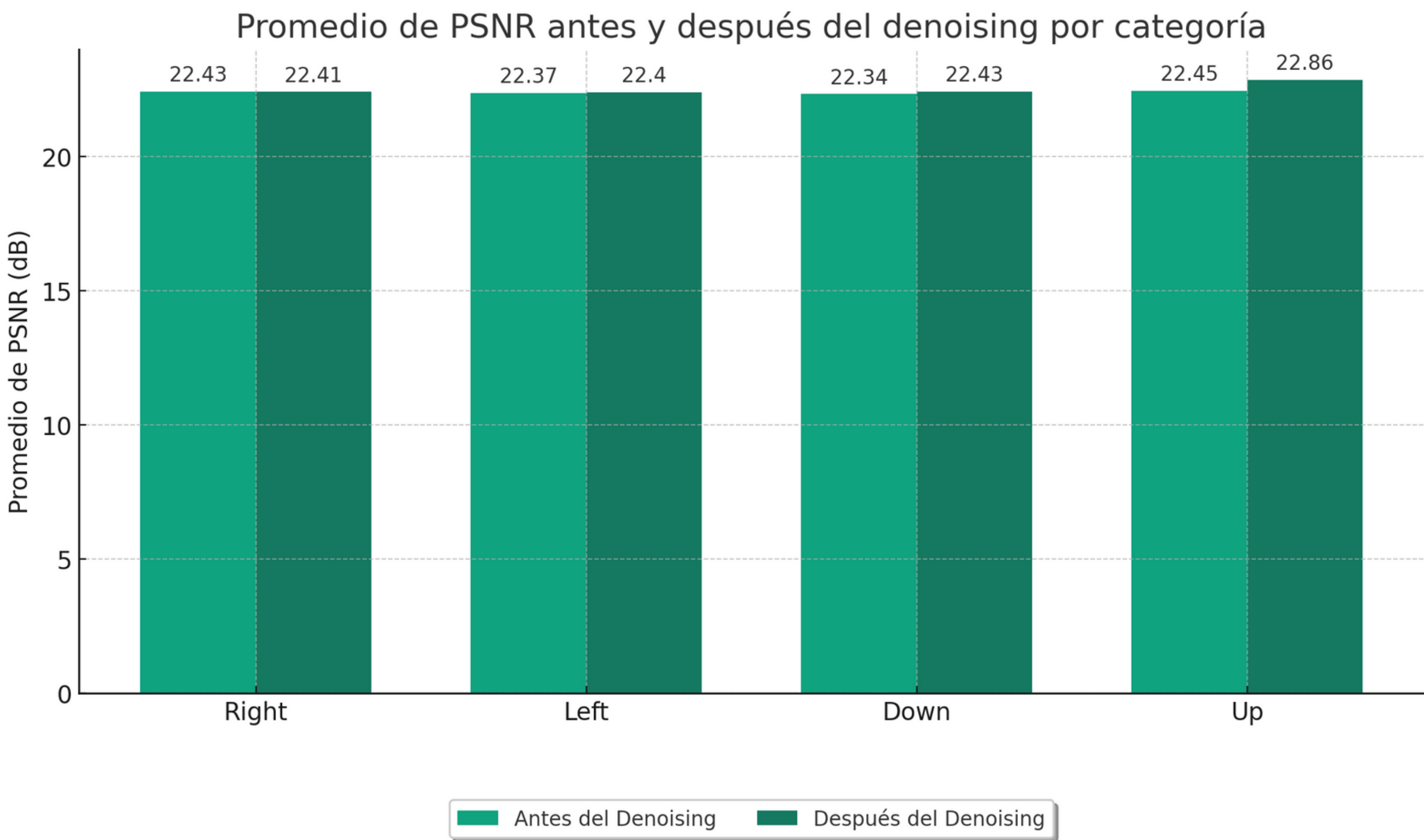
3 Metodología

- Preprocesamiento:** Limpieza y normalización de datos de audio.
- Extracción de Características:** Uso de MFCCs y análisis de formas de onda.
- Modelado:** Implementación de modelos de aprendizaje profundo.
- Validación:** Tests empíricos y comparación con soluciones existentes.
- Mitigación de Ruido:** Estrategias para minimizar errores por ruido externo.

4 Resultados

La implementación de un modelo de autoencoder convolucional en este estudio ha demostrado ser efectiva en la reducción de ruido sin sacrificar significativamente la calidad del sonido. El análisis utilizando el Peak Signal-to-Noise Ratio (PSNR) reveló que la claridad de los comandos de voz se mantuvo o incluso mejoró ligeramente después del proceso de denoising. Las gráficas incluidas muestran los valores promedio de PSNR antes y después del denoising en las categorías 'Right', 'Left', 'Down' y 'Up'. A excepción de la categoría 'Right', que experimentó un decremento marginal de -0.09% en calidad, las demás categorías registraron un aumento leve en la claridad del sonido.

Este incremento sugiere una mejora en la calidad de los comandos de voz tras la aplicación del denoising. Este patrón de resultados demuestra la capacidad del modelo no solo para minimizar el ruido de fondo, sino también para preservar o mejorar la integridad de los comandos de voz.



6 Conclusión

To wrap up your poster, present two to three key findings. You can also add a brief explanation or narrative to these that can encourage conversation or dialogue with the audience. These findings can be actionable items that can lead to implementation, policy creation, or further study.

Related literature

References can take up a lot of space, so cite only the key references used in the study.

