

Resumen de tesis

Esta tesis explora el creciente campo de la Inteligencia Acústica Computacional, centrándose en la aplicación de técnicas avanzadas de Inteligencia Artificial (IA) a señales de audio no semánticas y no musicales. La investigación aborda desafíos críticos en entornos reales mediante el desarrollo de sistemas robustos para la detección, localización y análisis de sonidos en contextos industriales, automotrices y de ciencias del deporte. Los métodos tradicionales de procesamiento de audio a menudo presentan limitaciones frente a altos niveles de ruido ambiental, condiciones dinámicas y escenarios acústicos complejos. Este trabajo propone y valida soluciones innovadoras que integran modelos de aprendizaje profundo, como los Audio Spectrogram Transformers (AST) y las Redes Neuronales Convolucionales (CNN), con hardware especializado y pipelines de procesamiento de señal para superar dichas limitaciones.

Las contribuciones principales se presentan a través de tres dominios de aplicación distintos, pero interconectados. Primero, se desarrolla un Sistema Avanzado de Asistencia al Conductor (ADAS) para la detección y localización en tiempo real de vehículos de emergencia. Este sistema combina una novedosa carcasa aerodinámica para micrófonos, que mitiga el ruido del viento, con un pipeline de IA en múltiples etapas, proporcionando a los conductores información situacional crítica. Segundo, se diseña un sistema inteligente de protección auditiva que preserva la conciencia del usuario en entornos industriales ruidosos. Este sistema aborda el dilema de la “sobreprotección” mediante la detección selectiva, el aislamiento y la reintroducción espacial de sonidos críticos, como alarmas y voces, utilizando una combinación de modelos de separación de fuentes y Funciones de Transferencia Relacionadas con la Cabeza (HRTF). Tercero, como estudio fundamental, se presenta un sistema para medir la altura de salto de atletas a partir de señales de audio, demostrando el potencial de la IA en ciencias del deporte para un monitoreo de rendimiento preciso y accesible.

A lo largo de estas aplicaciones, la tesis desarrolla y evalúa de manera sistemática metodologías para la extracción de características (espectrogramas Mel, GCC-PHAT), detección de eventos sonoros (CNNs, AST), localización de fuentes sonoras (ResNet-18) y mejora de señales (U-Net, DEMUCS). La investigación culmina con la creación de sistemas completos de extremo a extremo, validados con conjuntos de datos extensos que incluyen datos del mundo real, semi-sintéticos y sintéticos. Los resultados demuestran alta precisión y robustez, abriendo el camino para el despliegue práctico de sistemas acústicos potenciados por IA que mejoran la seguridad, el rendimiento y la interacción humano-ambiente.