

INFRAESTRUCTURA IoT PARA EL DESPLIEGUE DE SERVICIOS DE EMERGENCIA

OBJETIVOS:

- Crear un sistema físico capaz de detectar señales de audio en el medio ambiente de una ciudad.
- Aplicar un algoritmo de triangulación para ubicar las señales de interés en un mapa geográfico.
- Diseñar e implementar una red neuronal de clasificación que procese las señales de audio obtenidas y determine si existe una emergencia en la grabación.

RESULTADOS PRELIMINARES:

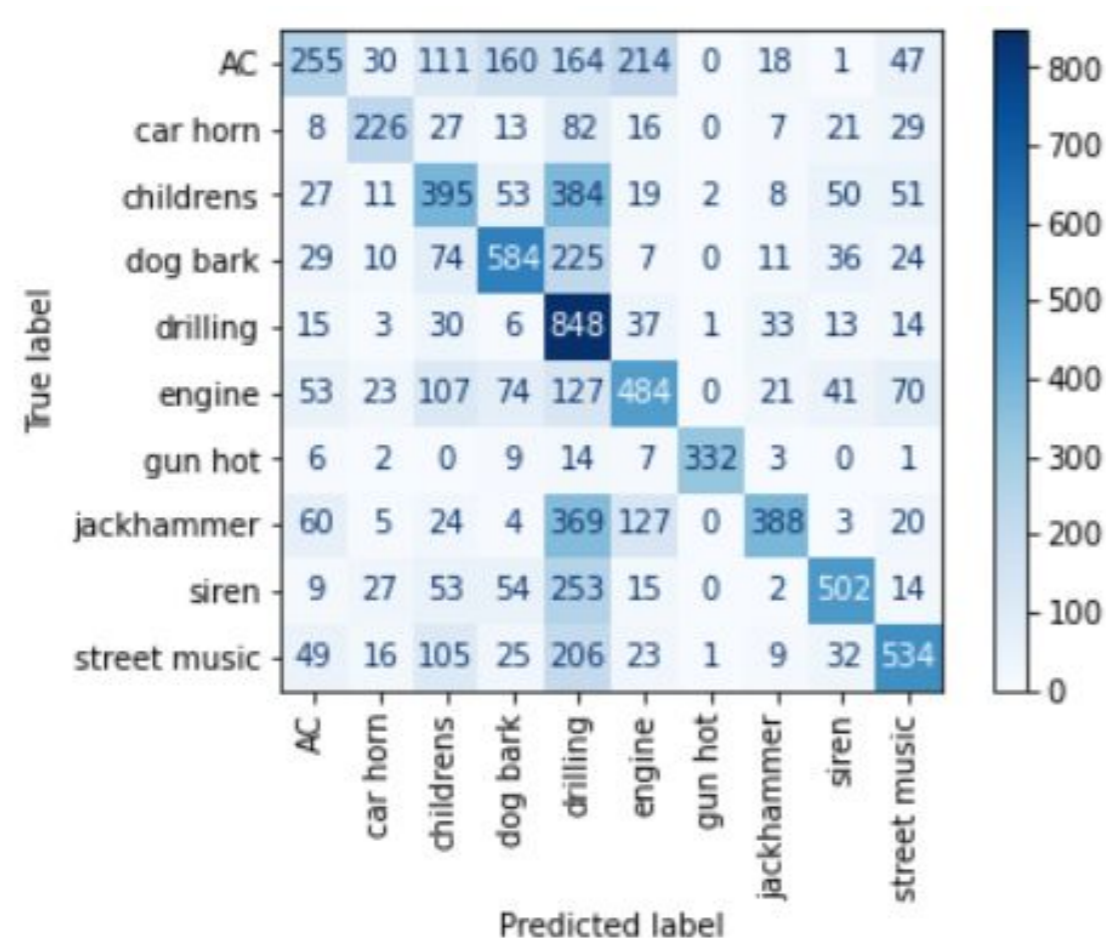


Figura : Matriz de confusión de la red neuronal

	precision	recall	f1-score	support
AC	0.50	0.26	0.34	1000.00
car horn	0.64	0.53	0.58	429.00
childrens	0.43	0.40	0.41	1000.00
dog bark	0.59	0.58	0.59	1000.00
drilling	0.32	0.85	0.46	1000.00
engine	0.51	0.48	0.50	1000.00
gun shot	0.99	0.89	0.94	374.00
jackhammer	0.78	0.39	0.52	1000.00
siren	0.72	0.54	0.62	929.00
street music	0.66	0.53	0.59	1000.00
accuracy	0.52	0.52	0.52	0.52
macro avg	0.61	0.54	0.55	8732.00
weighted avg	0.58	0.52	0.52	8732.00

Figura : Tabla de resumen de cada uno de las clases en la base de datos

DISEÑO E IMPLEMENTACIÓN:

- Un dispositivo de despliegue que consta de 4 micrófonos DAIKO de alta sensibilidad. El arreglo de los micrófonos es en formación tipo “T” para asistir a la triangulación.
- Implementación del modelo ResNet34 de red neuronal que sea capaz de clasificar los sonidos del medio ambiente y detecte si existe una señal de “grito de pánico” o “disparo de bala”, empleando las bases de datos “UrbanSounds8k” y “Mivia Audio Events”
- Una interfaz gráfica donde se muestre los datos obtenidos, es decir, una grabación del audio captado, la conclusión de la clasificación y la ubicación aproximada de la señal de emergencia.

CONCLUSIONES:

- Los resultados obtenidos pueden mostrarnos dos datos interesantes, el primero es que nuestro modelo de entrenamiento tiene una precisión del 52% en todas las clases de nuestra base de datos esto puede representar un punto a mejorar, lo que nos lleva a nuestro segundo punto de interés: la clase “gun_shot” o disparo de bala tiene una precisión del 99%, lo cual refleja que nuestro modelo tiene un buen desempeño con la clase de nuestro interés.

REFERENCIAS:

- K. Palanisamy, D. Singhanian, y A. Yao, “Rethinking cnn models for audio classification.” [Online]. Available: <https://ui.adsabs.harvard.edu/abs/2020arXiv200711154P/abstract>