#ProyectosCiber

UK AI

Spatio-temporal Latent Representations for the Analysis of Acoustic Scenes *in-the-wild*

The Second UK AI Conference, Birmingham (UK), 22 Nov 2024

Claudia Montero Ramírez

clmonter@pa.uc3m.es

Claudia Montero-Ramírez[1], Esther Rituerto-González[2,3], Carmen Peláez-Moreno[1,4]

[1] Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain
[2] Department of Psychiatry and Psychotherapy, University Hospital, Ludwig-Maximilian-University of Munich, Germany
[3] German Center for Mental Health (DZPG), partner site Munich-Augsburg, Germany
[4] University Institute on Gender Studies, Spain

Background

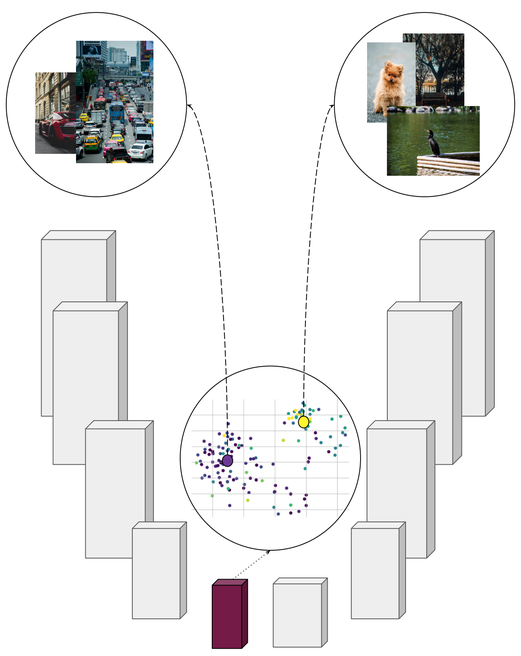
The situations in which **gender-based violence** occurs are a risk to the health of victims. In these situations, a **rapid intervention of the emergency services** can be crucial.

Bindi is an autonomous system powered by artificial intelligence and the Internet of Things (IoT) to automatically report when a woman is in a risky situation related to GBV.

This identification will be performed by automatically detecting **fear-related emotions** in the victim.

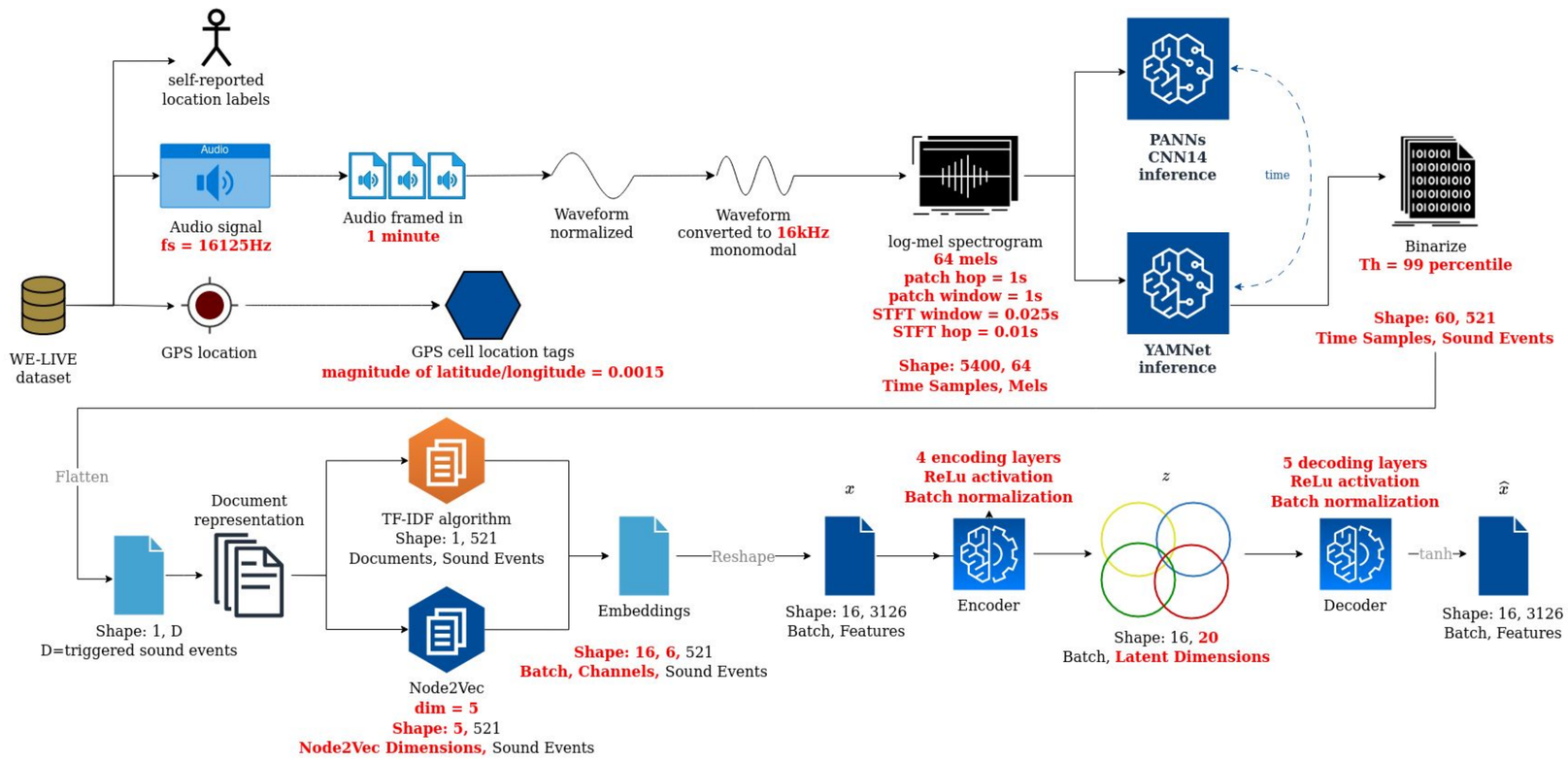
Outline of Bindi.

Goal



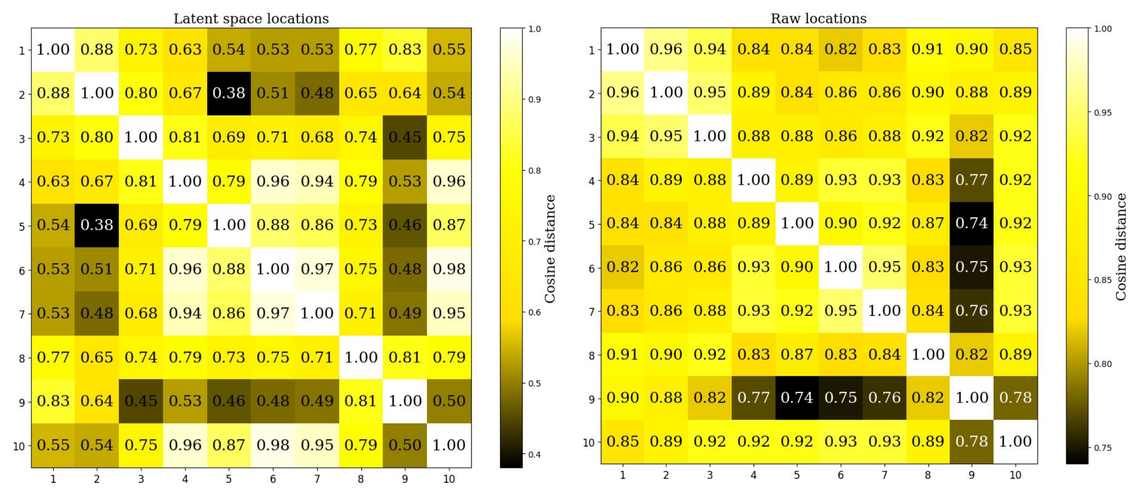
Find acoustic self-supervised latent representations

Methodology



Results

Cosine distance between spatio temporal representations



Latent space locations *Raw locations*

	GPS cell location tags	self-reported situational label
●	1	Home
●	2	Work
●	3	Near home
●	4	Metro
●	5	Metro
●	6	Metro
●	7	Metro
●	8	-
●	9	-
●	10	Metro

Sound Event Retrieval

1 Home
1. Hair dryer
2. Bang
3. Whack, thwack
4. Chopping (food)
5. Child speech, kid speaking

2 Work
1. Chirp, tweet
2. Truck
3. Fire alarm
4. Breaking
5. Slam

5 Metro
1. Drum machine
2. Slam
3. Gears
4. Raindrop
5. Squish

Conclusions

Results presented seem to have achieved a **regrouping** or **clustering** of data on the **VAE's latent space**, based on the **space-time that the acoustic scene** is related to.

It is important to continue studying the relationship between the acoustic scene and the emotions related to fear, in order to be able to still contributing to the fight against gender violence.

References

[1] Executive Summary of the 2019 Macro-survey on Violence Against Women 2021. Delegación de Gobierno contra la Violencia de Género, Ministerio de Igualdad, Gobierno de España.

[2] Luis-Mingueza, Clara, Esther Rituerto-González, and Carmen Peláez-Moreno. "Bridging the Semantic Gap with Affective Acoustic Scene Analysis: an Information Retrieval-based Approach." *Proc. IberSPEECH 2022* (2022): 91-95.

[3] J. A. Miranda Calero *et al*., "Bindi: Affective Internet of Things to Combat Gender-Based Violence," in *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21174-21193, 1 Nov.1, 2022, doi: 10.1109/JIOT.2022.3177256.

Acknowledgements

The authors would like to acknowledge project Vital-IoT funded by **INCIBE** (Ministry of Digital Transformation and Public Function) and the European Union NextGenerationEU in the framework of the Recovery and Resilience Facility (RRF).

Financiado por la Unión Europea NextGenerationEU

Collaborating entity: