Multimodal Deep Learning for Disaster Classification from Social Media Data in Brazil and Peru

Mariana Risco[1], Cláudio R. Jung[2], Sharon Quispe[3]

[1]Universidade Federal do Rio Grande do Sul
[2]Universidade Federal do Rio Grande do Sul
[3] Universidad San Ignacio de Loyola

Motivation

Natural disasters in Latin America cause recurrent human and economic losses (Obregón Pérez 2024). In 2024, Brazil faced historic floods (Wink Junior et al. 2024), and Peru remains vulnerable to El Niño-related mudslides and frost (Carbonel et al. 2018). Despite their frequency, disaster data is fragmented and underrepresented in global datasets. Social media platforms like Facebook and X offer real-time images (Arrobo Agila 2019), but the content is often unstructured and difficult to analyze.

Dataset

DisasterDataset

- 4,039 images of natural disasters from Social Media.
- Seven types of disasters: huaico, forest fire, frost hail, flood, landslide, drought, electrical storm

Data Collection and Annotation

- Used Selenium web scrapper to collect posts from Facebook and X (Selenium 2025).
- Searching for specific disaster names, we retrieved posts related to each type of disaster.
- The posts texts are in Spanish and Portuguese.

#AREQUIPA. INCENDIO FORESTAL ARRASA 4 HECTÁREAS
EN EL SANTUARIO NACIONAL DE MEJÍA Y OBLIGA
CIERRE DE LA COSTANERA #RadioLibertadArequipa
#IncendioForestalMejía #SantuarioMejía #Arequipa.
Desde el mediodía del miércoles 9 de abril,
un incendio forestal de gran magnitud viene
afectando los pastizales del Santuario
Nacional de Mejía, en la provincia de
Islay, Arequipa.

O BEBÊ É NOSSO SOL Fabrício Carpinejar
É proibido rir no Rio Grande do Sul,
devido ao pesar pela meia centena de mortes
até então e pelas dezenas de desaparecidos.
Mas eu tenho um distúrbio. E muitos são como eu.
Em situações de extremo nervosismo, começo a rir.
As lágrimas já não fazem efeito, tal remédio que
se mostra fraco pelo uso constante.
Eu choro pelo riso. Choro borbotões pelo riso...

Methodology

We explored zero-shot classification and compared the performance of pretrained and fine-tuned SigLIP models using the same prompts in Spanish, English, and Portuguese.

SigLIP Models (VLMs)

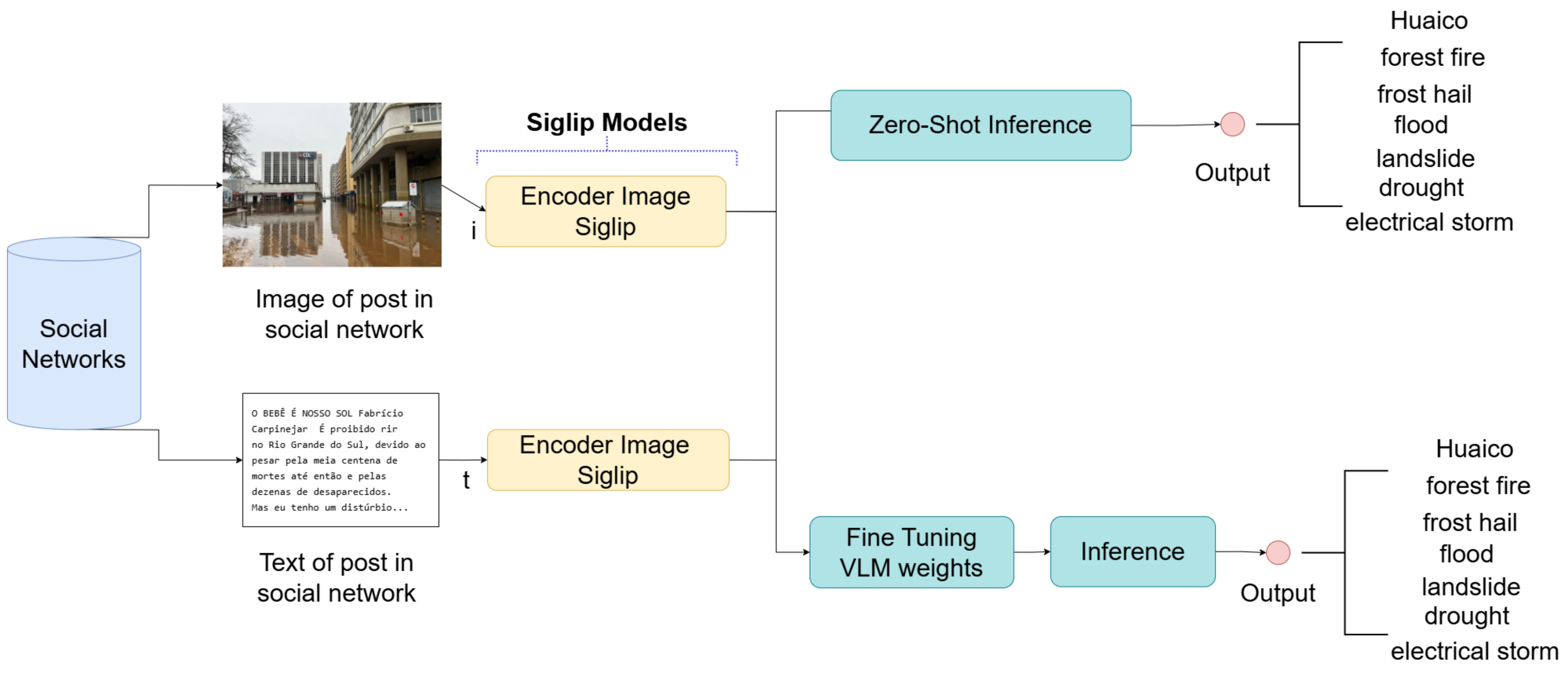
- SigLIP Base Multilingual (Zhai et al. 2023)
- SigLIP 2 Naflex enable different resolutions (Tschannen et al. 2025)

Fine Tuning

- All layers were unfrozen, and their weights were updated during training.

Zero-Shot Prompting

- We adapted the prompts, such us *"A picture of a {classname}"*, to provide more details for each disaster.



Evaluation

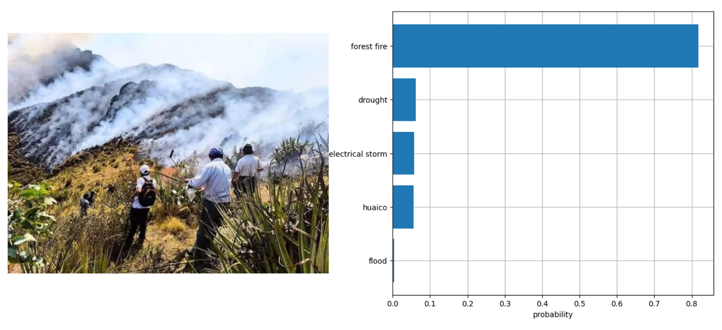
- Metrics: Accuracy, Precision, Recall, F1 score.

Results

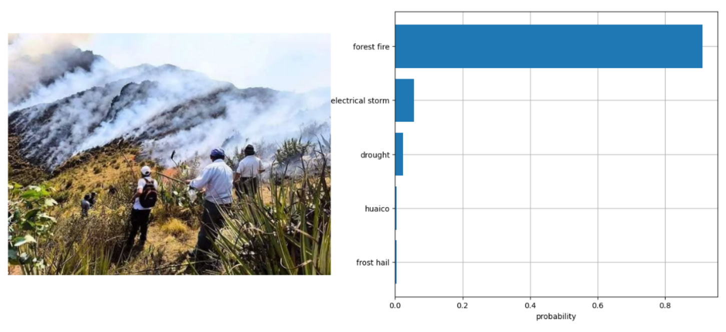
Performance Comparison of SigLIP Multilingual and SigLIP2 Naflex Before and After Fine-Tuning

Model	Language	Accuracy	Precision	Recall	F1-Score	Stage
SigLIP Multilingual	English	0.6865	0.6969	0.6865	0.6705	Before
	Spanish	0.6989	0.7215	0.6989	0.7017	Before
	Portuguese	0.6803	0.7185	0.6803	0.6864	Before
	English	**0.7410**	**0.7520**	**0.7410**	**0.7414**	After
	Spanish	**0.7150**	**0.7317**	**0.7150**	**0.7134**	After
	Portuguese	**0.7447**	**0.7481**	**0.7447**	**0.7380**	After
SigLIP2 Naflex	English	0.6642	0.7361	0.6642	0.6803	Before
	Spanish	0.5960	0.7050	0.5960	0.6101	Before
	Portuguese	0.6022	0.7349	0.6022	0.6182	Before
	English	**0.7460**	**0.7508**	**0.7460**	**0.7376**	After
	Spanish	**0.7212**	**0.7495**	**0.7212**	**0.7055**	After
	Portuguese	**0.7200**	**0.7331**	**0.7200**	**0.7179**	After

Predictions for the same *forest fire* image using English prompts. Left: SigLIP Multilingual predicts *forest fire* (82%) Right: SigLIP2 NAFLEX predicts *forest fire* (91%)



SigLIP Multilingual



SigLIP2 NAFLEX

Conclusion

- We proposed an initial approach for disaster detection in Latin America using vision-language models. Our fine-tuned multilingual SigLIP outperformed SigLIP2 Naflex. Future work includes testing other models such as CLIP (Radford et al. 2021) and LLaVA (Liu et al. 2024), as well as exploring advanced fine-tuning techniques.

Connect with me!