

GIMI: A Geographical Generalizable Image-to-Image Search Engine with Location-explicit Contrastive Embedding

Hao Li¹, Jiapan Wang¹, Balthasar Teuscher¹, Peng Luo¹, Gengchen Mai², Danfeng Hong³, Martin Werner¹,
¹ Technical University of Munich, Germany; ² University of Georgia, USA, ³ Chinese Academy of Sciences, China

Location Explicit Embedding

- Motivation: *the cluster hypothesis says "closely associated documents tend to be relevant to the same requests"*
- GIMI - a geographical generalizable image-to-image neural search engine:
 - High-dimensional vector embedding from both geo-locations and image representations
 - Flexible similarity search with a customized index

Distance-Penalized Triplet

- DPT extends Triplet loss with a geographical distance term
- Idea: *"the influence of different positive and negative samples may differ when spatially clustered or co-located"*

$$L_{DPT} = [||f(t_a) - f(t_p)||_2 - ||f(t_a) - f(t_n)||_2 + P(x_i) + a]_+$$

$$P(x_a, x_p, x_n) = q(x_a, x_p) + q(x_a, x_n) - q(x_p, x_n)$$

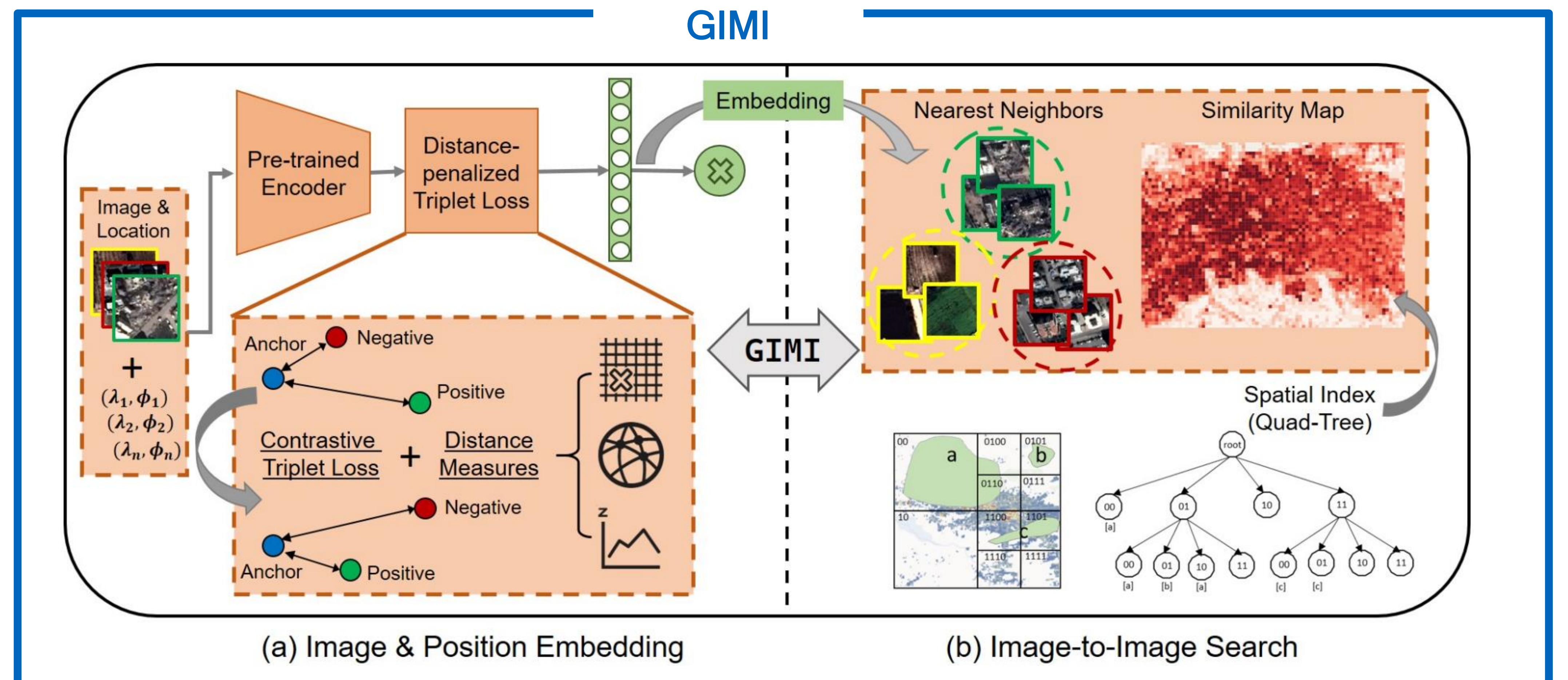
Disaster Mapping Damage Building

- Problem: Global mapping inequalities
- Trend: Disaster mapping with Earth Observation and OpenStreetMap (OSM)
- Study Area: the city of Adiyaman affected by the 2023 Kahramanmaras Earthquake in Turkey



© Satellite Image before and after Earthquake from Pleiades 1A and 1B

Overview of GIMI



Preliminary Results

Encoder	Method	Fine-tuned	Top 5% (%)	NDCG at 5%	Top 10% (%)	NDCG at 10%
ResNet	Base	✗	68.97 ± 19.87	0.908 ± 0.081	66.30 ± 17.29	0.912 ± 0.072
	Softmax	✓	92.22 ± 23.19	0.985 ± 0.055	91.68 ± 23.44	0.974 ± 0.084
	Triplet Loss	✓	94.06 ± 17.24	0.982 ± 0.060	94.11 ± 14.65	0.985 ± 0.048
	DPT Loss	✓	94.03 ± 17.61	0.983 ± 0.062	95.32 ± 13.24	0.987 ± 0.048
ViT	Base	✗	67.39 ± 13.83	0.921 ± 0.059	61.21 ± 10.99	0.907 ± 0.053
	Softmax	✓	93.46 ± 15.49	0.982 ± 0.055	91.50 ± 17.95	0.977 ± 0.069
	Triplet Loss	✓	96.09 ± 12.00	0.990 ± 0.035	93.97 ± 13.81	0.988 ± 0.040
	DPT Loss	✓	98.04 ± 5.63	0.995 ± 0.015	96.96 ± 7.07	0.993 ± 0.024

