

ViSketch-GPT: Collaborative Multi-Scale Feature Extraction for Hand-Drawn Sketch Retrieval

Giulio Federico^{1,2,*}, Fabio Carrara², Claudio Gennaro² and Marco Di Benedetto²

¹University of Pisa, Pisa (PI), Italy

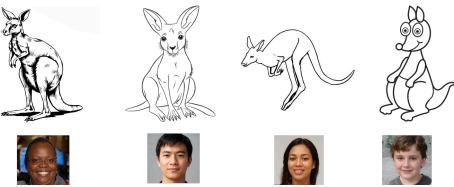
²Institute of Information Science and Technologies (ISTI-CNR), Pisa (PI), Italy
 {name.surname}@isti.cnr.it

The problem we want to solve

Natural/Realistic Images of a Kangaroo



Hand-Drawn Sketches of a Kangaroo



1

Natural images of a category (e.g., kangaroo) look alike—regardless of the device. A kangaroo is always a kangaroo.

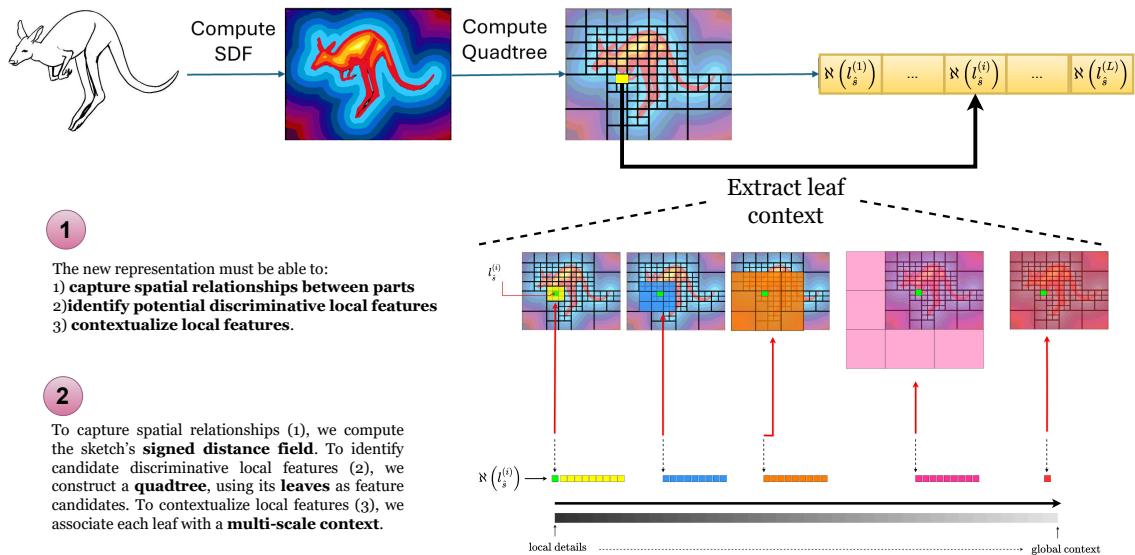
2

Each sketch reflects its author's background, perception, and skill, making the same object appear in highly diverse and subjective ways.

3

In recognition or retrieval tasks, embeddings from models trained on realistic images are highly informative. For sketches, however, **no representation yet captures their essence or handles the wide variability in style and conceptual interpretation**.

A novel representation for sketches



Sketch Retrieval

1

We trained our new architecture to classify sketches, then tested whether the **final extracted embedding** could be used for retrieval.

