Global Semantic Description of objects based on Prototype Theory

Omar Vidal Pino Erickson R. Nascimento Mario F. M. Campos
Universidade Federal de Minas Gerais (UFMG), Brazil

IEEE 2019 **W**inter Conference on **A**pplications of **C**omputer **V**ision **Doctor Consortium**

Motivation

- ✓ CNNs provide outstanding performance in image semantic processing tasks.
- ✓ CNN description models do not codify their representation based on the Cognitive Science foundations to represent the meaning.
- ✓ We bring into view the Prototype Theory as a theoretical framework to represent the semantic meaning of the visual information contained in an image.

Background

Elements of Prototype Theory

- The *prototype* stands for the *central semantic meaning* of the category.

- *Category internal semantic structure:* A category member is positioned closer to the category prototype based on its typicity degree.

- *Object categorization* is obtained based on the similarity of a new exemplar with the learned prototypes.
- Main characteristics (*prototypicality effects*):

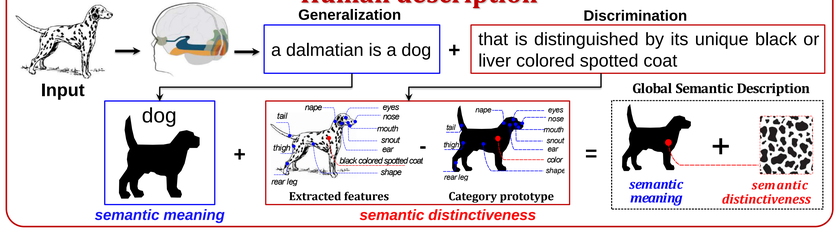
	extensional	***intensional***
non-equality *(salience effect, core/periphery)*	Difference of typicality and membership salience	Clustering into family resemblances
non-discreteness *(demarcation problems, flexibly)*	Fuzziness at the edges, membership uncertainty	Absence of necessary-and-sufficient definitions

Problem Statement

How to describe and stand for objects, semantically? Can a description model be developed in which objects are described using the same semantic features learned to classify them? How to include the semantic of a category prototype into the global semantic description of objects?

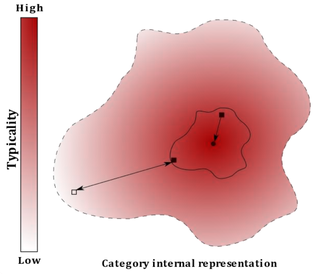
Intuition

Human description



Our Prototype Model

CNN Prototype Model



Legend

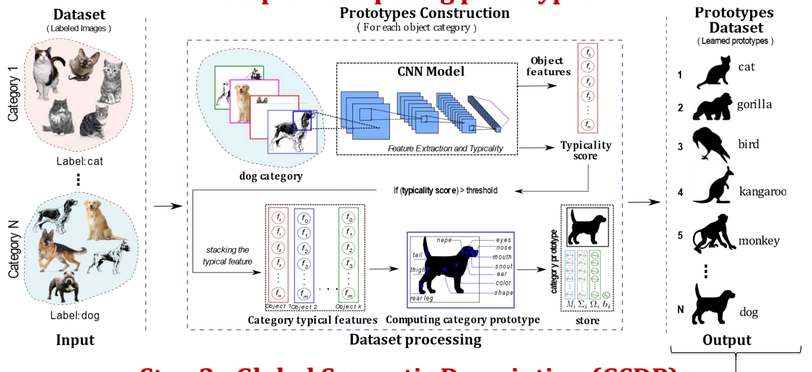
- Typical members
- Non-typical members
- Semantic Prototype
- Abstract Prototype
- Category edges (*undefined*)
- Semantic Prototype edges (*defined*)
- Prototypical distance
- Distance between objects

Definitions

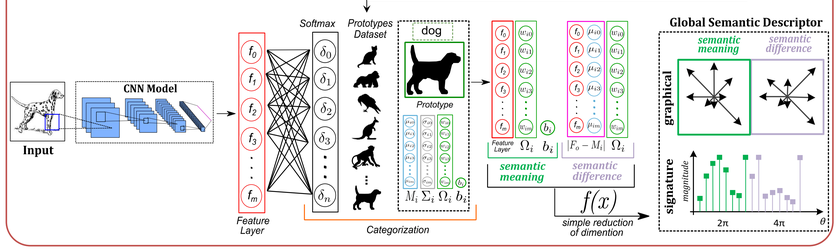
1. Semantic Prototype: $P_i = (M_i, \underset{mean}{\Sigma_i}, \underset{std}{\Omega_i}, \underset{relevance}{})$
2. Abstract Prototype: $M_i \in P_i$
3. Object semantic value: $z_o = \sum_{j=1}^{m} \omega_{ij} f_j + b_i$
4. Category semantic value: $\hat{z}_i = \sum_{j=1}^{m} \omega_{ij} \mu_{ij} + b_i$
5. Prototypical distance: $\delta(o, P_i) = \sum_{j=1}^{m} |\omega_{ij}| \, |f_j - \mu_{ij}|$
6. Distance between objects: $\delta(o_1, o_2) = \sum_{j=1}^{m} |\omega_{ij}| \, |f_j^1 - f_j^2|$

Methodology

Step 1 - Computing prototypes



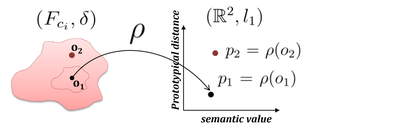
Step 2 - Global Semantic Description (GSDP)



Results

Category internal structure

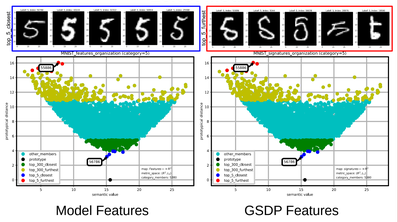
Mapping the category



$\rho : F_{c_i} \to \mathbb{R}^2 \mid \rho(o \in O_{c_i}) = \rho(F_o) = p(\hat{z}_o, \delta(o, P_i))$

$\rho : F_{c_i} \to \mathbb{R}^2$ *is continuous*

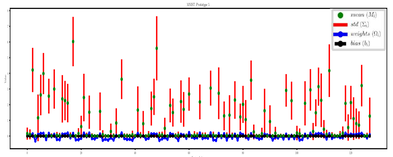
LeNet Model



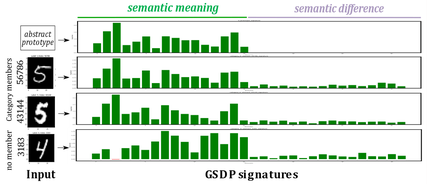
Signature taxonomies

Semantic prototype

Model: **LeNet** Dataset: **MNIST** Category: **5**



Signatures distributions

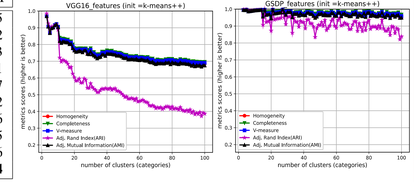


Performance Evaluation

Cluster Metrics Scores

Descriptor	Size	Time (ms)	Metrics Scores				
			H	C	V	ARI	AMI
GIST	960	1210	0.05	0.05	0.05	0.01	0.05
LBP	512	1380	0.02	0.03	0.03	0.01	0.02
HOG	1960	25	0.04	0.04	0.04	0.01	0.03
Color64	64	125	0.12	0.12	0.12	0.04	0.11
Color_Hist	512	38	0.08	0.08	0.08	0.03	0.07
Hu_H_CH	532	145	0.04	0.04	0.04	0.01	0.02
VGG16	4096	62	0.77	0.78	0.77	0.60	0.76
VGG16_PCA_25	25	80	0.76	0.77	0.76	0.60	0.75
VGG16_PCA_256	256	80	0.76	0.77	0.77	0.59	0.76
GSDP (Our)	256	120	**0.94**	**0.97**	**0.95**	**0.87**	**0.94**

Cluster Metrics History



- Our descriptor encoding significantly outperforms other image global encodings in terms of cluster metrics.

Conclusion

Our *prototype-based description model* proposes a starting point to introduce the theoretical foundation related to the *representation of semantic meaning* and the *learning of visual concepts* of the Prototype Theory in the CNN-Descriptors family.

Project Page