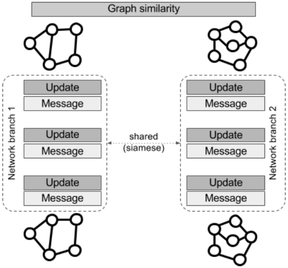
Learning Graph Distances with Message Passing Neural Networks

Pau Riba, Andreas Fischer*, Josep Lladós, Alicia Fornés

Computer Vision Center, *Université de Fribourg

ICPR, Beijing, August 23rd, 2018.



Centre de Visió per Computador

Universitat Autònoma de Barcelona

Generalitat de Catalunya

Parc de Recerca UAB

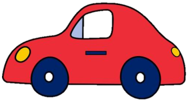
Outline

Introduction

Motivation

Graph representations

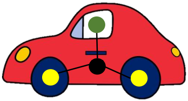
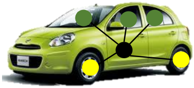
1. Increasing relevance in **visual object recognition and retrieval**, beyond classical pure appearance-based approaches.



Motivation

Graph representations

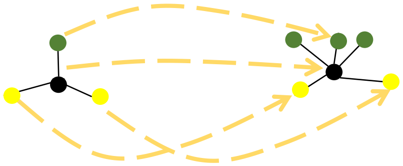
1. Increasing relevance in **visual object recognition and retrieval**, beyond classical pure appearance-based approaches.

Motivation

Graph representations

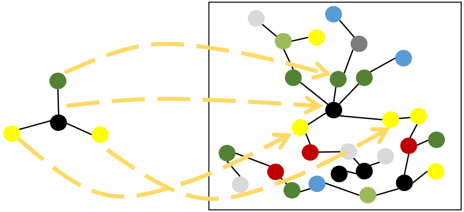
1. Increasing relevance in **visual object recognition and retrieval**, beyond classical pure appearance-based approaches.
2. Visual object detection using graphs involves an **inexact subgraph matching formulation**.



Motivation

Graph representations

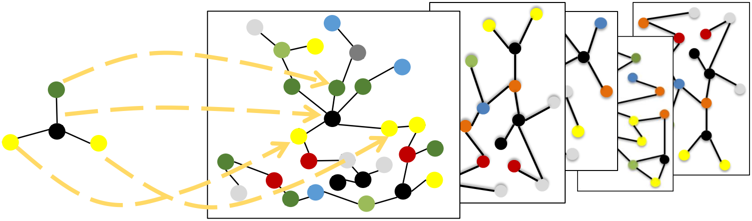
1. Increasing relevance in **visual object recognition and retrieval**, beyond classical pure appearance-based approaches.
2. Visual object detection using graphs involves an **inexact subgraph matching formulation**.



Motivation

Graph representations

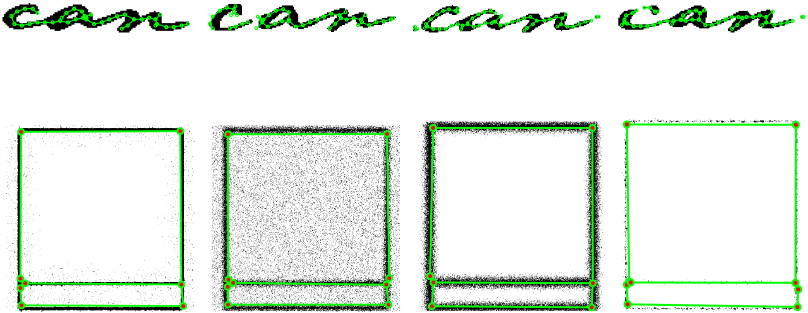
1. Increasing relevance in **visual object recognition and retrieval**, beyond classical pure appearance-based approaches.
2. Visual object detection using graphs involves an **inexact subgraph matching formulation**.
3. It is unavoidable in **large scale retrieval** (i.e. subgraph matching).



Motivation

Document Analysis

- A graph is a powerful representation, both for text and graphics



Motivation

Geometric Deep Learning

Geometric Deep Learning

Extension of Deep Learning techniques to graph/manifold structured data.

Motivation

Geometric Deep Learning

Geometric Deep Learning

Extension of Deep Learning techniques to graph/manifold structured data.

Image:

- Regular grid
- Operations well defined
- Same size $\rightarrow$ batch processing
- 8-neighbourhood

Graph:

- 4-Tuple $G = (V, E, L_V, L_E)$
- Operations not efficient
- Different size $\rightarrow$ batch processing
- Different neighbourhood

Introduction

Hypothesis

Local structural node information can be learned by Geometric Deep Learning and exploited by Graph Distance algorithms.

Thus, we avoid a graph embedding that may be difficult to learn.

Related Concepts

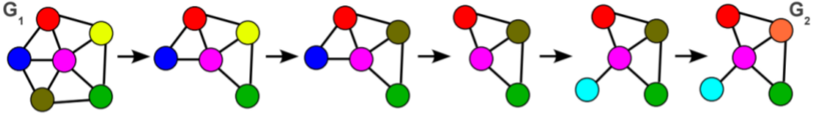
Graph Edit Distance

Definition

Given a set of *Graph Edit Operations*, the *Graph Edit Distance* (GED) between two graphs g_1 and g_2 is defined as

$$\mathrm{GED}(g_1, g_2) = \min_{(e_1,\ldots,e_k)\in\mathcal{P}(g_1,g_2)} \sum_{i=1}^{k} c(e_i)$$

where $\mathcal{P}(g_1, g_2)$ denotes the set of edit paths transforming g_1 into $g2$ and $c(e)$ is the cost of each edit operation.



Graph Edit Distance

Approximated Techniques

Computation

Exact GED is not feasible in real applications due to its complexity. Several approximations have been proposed.

* Fischer et al., "Approximation of graph edit distance based on Hausdorff matching".

† Riesen et al., "Approximate graph edit distance computation by means of bipartite graph matching".

Graph Edit Distance

Approximated Techniques

Computation

Exact GED is not feasible in real applications due to its complexity. Several approximations have been proposed.

Some approximated algorithms have been proposed.

- *Hausdorff Edit Distance* (HED)* $\mathcal{O}(n_1 \cdot n_2)$
- *Bipartite Graph Matching* (BP)† $\mathcal{O}((n_1 + n_2)^3)$

---

* Fischer et al., "Approximation of graph edit distance based on Hausdorff matching".

† Riesen et al., "Approximate graph edit distance computation by means of bipartite graph matching".

Graph Edit Distance

Approximated Techniques

Computation

Exact GED is not feasible in real applications due to its complexity. Several approximations have been proposed.

Some approximated algorithms have been proposed.

- *Hausdorff Edit Distance* (HED)* $\mathcal{O}(n_1 \cdot n_2)$
- *Bipartite Graph Matching* (BP)† $\mathcal{O}((n_1 + n_2)^3)$

The usual *Graph Edit Operations* in the GED computation are:

- Insertion and Deletion (nodes and edges)
- Substitution (nodes and edges)

* Fischer et al., "Approximation of graph edit distance based on Hausdorff matching".

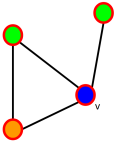
† Riesen et al., "Approximate graph edit distance computation by means of bipartite graph matching".

Geometric Deep Learning

Neural Message Passing*

Message Passing Neural Network (MPNN) is composed of 3 functions:

- Message
- Update
- Readout



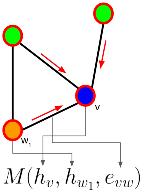
* Gilmer et al., "Neural message passing for quantum chemistry".

Geometric Deep Learning

Neural Message Passing*

Message

$$m_v^{t+1} = \sum_{w \in \mathcal{N}(v)} M_t(h_v^t, h_w^t, e_{vw})$$



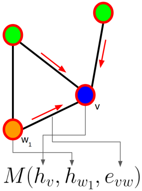
* Gilmer et al., "Neural message passing for quantum chemistry".

Geometric Deep Learning

Neural Message Passing*

Message

$$m_v^{t+1} = \sum_{w \in \mathcal{N}(v)} M_t(h_v^t, h_w^t, e_{vw})$$



Example:

$$M_t(h_v^t, h_w^t, e_{vw}) = A(e_{vw}) h_w^t$$

where $A(\cdot)$ is a NN mapping to a $d \times d$ matrix.

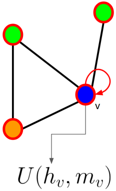
* Gilmer et al., "Neural message passing for quantum chemistry".

Geometric Deep Learning

Neural Message Passing*

Update

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1})$$



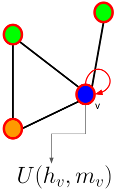
* Gilmer et al., "Neural message passing for quantum chemistry".

Geometric Deep Learning

Neural Message Passing*

Update

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1})$$



Example:

$$U_t(h_v^t, m_v^{t+1}) = GRU(h_v^t, m_v^{t+1})$$

where $GRU(\cdot,\cdot)$ is a *Gated Recurrent Unit*.

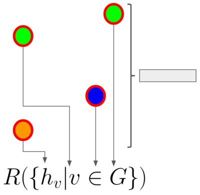
* Gilmer et al., "Neural message passing for quantum chemistry".

Geometric Deep Learning

Neural Message Passing*

Readout

$$\hat{y} = R(\{h_v^T | v \in G\})$$



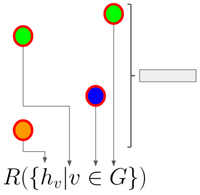
* Gilmer et al., "Neural message passing for quantum chemistry".

Geometric Deep Learning

Neural Message Passing*

Readout

$$\hat{y} = R(\{h_v^T | v \in G\})$$



Example:

$$R(\{h_v^T | v \in G\}) = \sum_{v \in V} \sigma \left(i(h_v^{(T)}, h_v^0) \right) \odot \left(j(h_v^{(T)}) \right)$$

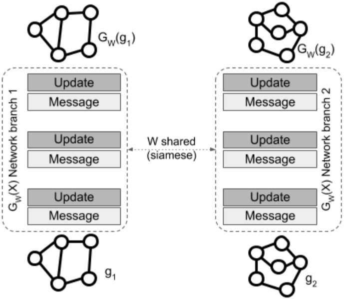
where i and j are NN and $\odot$ denotes element-wise multiplication.

* Gilmer et al., "Neural message passing for quantum chemistry".

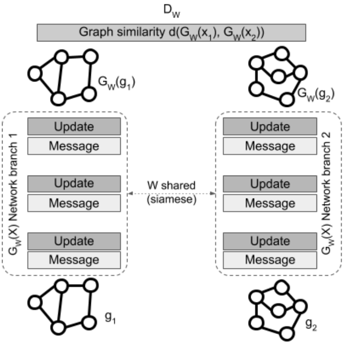
Architecture

Siamese Architecture

Siamese Architecture



Siamese Architecture



Graph Similarity

- Hausdorff Distance-based Similarity

$$\mathrm{H}(A, B) = \max\left(\max_{a \in A} \inf_{b \in B} d(a, b), \max_{b \in B} \inf_{a \in A} d(a, b)\right)$$

- More robust distance

$$\hat{\mathrm{H}}(A, B) = \sum_{a \in A} \inf_{b \in B} d(a, b) + \sum_{b \in B} \inf_{a \in A} d(a, b)$$

- Proposed distance

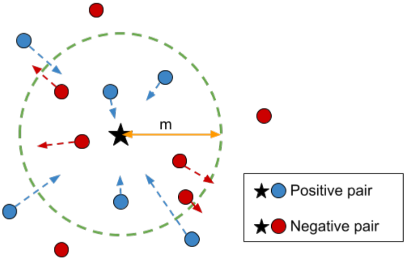
$$d(g_1, g_2) = \frac{\hat{\mathrm{H}}(V_1, V_2)}{|V_1| + |V_2|}$$

Contrastive Loss

Given $D_W = d(G_W(g_1), G_W(g_2))$ where g_1 and g_2 are graphs and W a set of specific weights W, the **Loss Function** is

$$l(D_W) = \frac{1}{2}\begin{cases} D_W^2, & \text{if } Y = 1 \text{ (positive pair)} \\ \{\max(0, m - D_W)\}^2, & \text{if } Y = 0 \text{ (negative pair)} \end{cases}$$

where $m = 1$ is the adaptive margin.

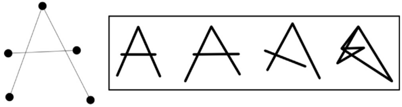


Experimental Validation

Datasets

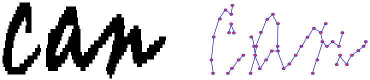
Letters

- Classification of Synthetic Graphs
- 15 classes
- 750 graphs per class
- 3 different distortion levels



George Washington

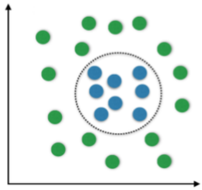
- Retrieval of Handwritten Words
- Several graph constructions
- 105 keywords
- 4894 instances
- **HistoGraph** subset for classification



Experimental Setup

Classification

- *k*-Nearest Neighbor Classifier
- Accuracy + Standard Deviation (5 runs)
- Tested with well-known Aproximated Graph Edit Distance algorithms



Letters

Classification

Table: Accuracy ± Std for 5 runs.

	LOW	MED	HIGH
BP*	99.73	94.27	89.87
HED†	**97.87**	**86.93**	**79.2**
Embedding‡	99.80	94.90	92.90
MPNN	95.04 ±0.7224	83.20 ±1.2189	72.27 ±2.0060
Siamese MPNN	**98.08 ± 0.1068**	**89.0136 ± 0.1808**	**74.77 ± 6.4505**
Test BP	98.19 ±0.1361	88.37 ±0.41	79.65 ±6.4345
Test HED	98.00 ±0.1461	89.79 ±0.3110	77.07 ±5.6106

* Riesen et al., "Approximate graph edit distance computation by means of bipartite graph matching".

† Fischer et al., "Approximation of graph edit distance based on Hausdorff matching".

‡ Gibert et al., "Graph embedding in vector spaces by node attribute statistics".

HistoGraph

Classification

Table: Classification accuracy for the HistoGraph dataset.

			Siamese MPNN	
Subset	**BP***	**PSGE†**	3-NN	5-NN
Keypoint	77.62	80.42	**85.31 ± 1.6552**	82.80 ± 0.5600
Projection	81.82	80.42	**73.15 ± 2.6014**	69.65 ± 1.5064

* Stauffer et al., "A Novel Graph Database for Handwritten Word Images".

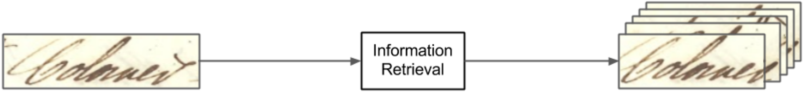
† Dutta et al., "Pyramidal Stochastic Graphlet Embedding for Document Pattern Classification".

Experimental Setup

Retrieval

- Mean Average Precision + Standard Deviation (5 runs)

$$\mathrm{mAP} = \frac{\sum_{q=1}^{Q} \mathrm{AP}(q)}{Q},$$



George Washington

Retrieval

Table: mAP from different approaches on GW dataset.

Method		mAP
PHOC*		64.00
BOF HMM†		80.00
DTW	DTW'01	42.26
	DTW'08	63.39
	DTW'09	64.80
	DTW'16	68.64
Mean Ensemble BP‡		69.16
Siamese MPNN		75.85±3.64

* Ghosh et al., "Query by string word spotting based on character bi-gram indexing".

† Rothacker et al., "Segmentation-free query-by-string word spotting with bag-of-features HMMs".

‡ Stauffer et al., "Ensembles for Graph-based Keyword Spotting in Historical Handwritten Documents".

Conclusion and Future Work

Final thoughts

Conclusions

- Enriched graph representation, incorporating the local context
- Fast similarity measure based on the Hausdorff Distance
- It emphasises the structure
- Improvements in real applications

Future Work

- To explore uses of graph structures to model relations among several images (each image encoded as a node)

Thank you for your attention!

Pau Riba
Computer Vision Center
priba@cvc.uab.cat
www.cvc.uab.cat/people/priba