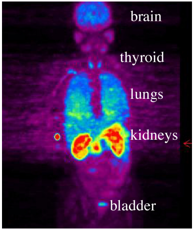
The pyramid quantized Weisfeiler-Lehman graph representation

Katerina Gkirtzou & Matthew Blaschko

January 8, 2014

Medical Imaging

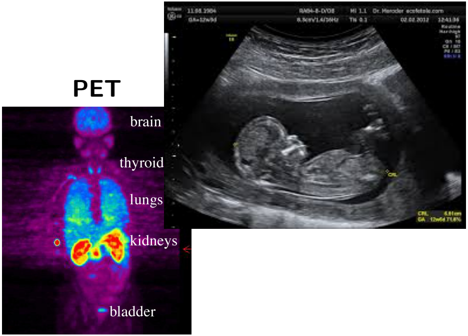
PET



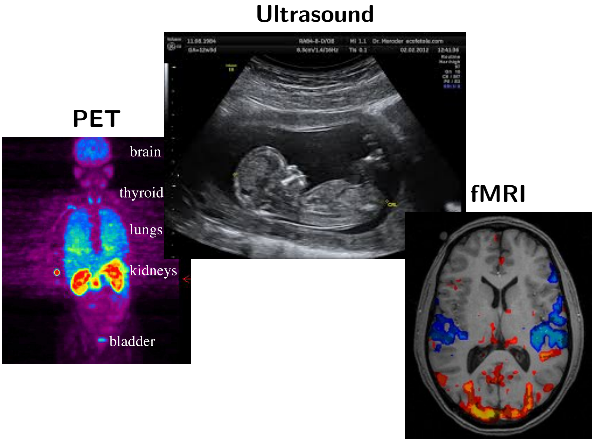
Medical Imaging

Ultrasound

PET



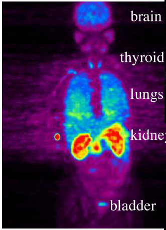
Medical Imaging



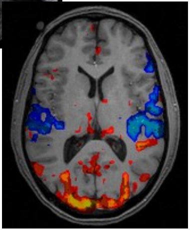
Medical Imaging

Ultrasound

PET

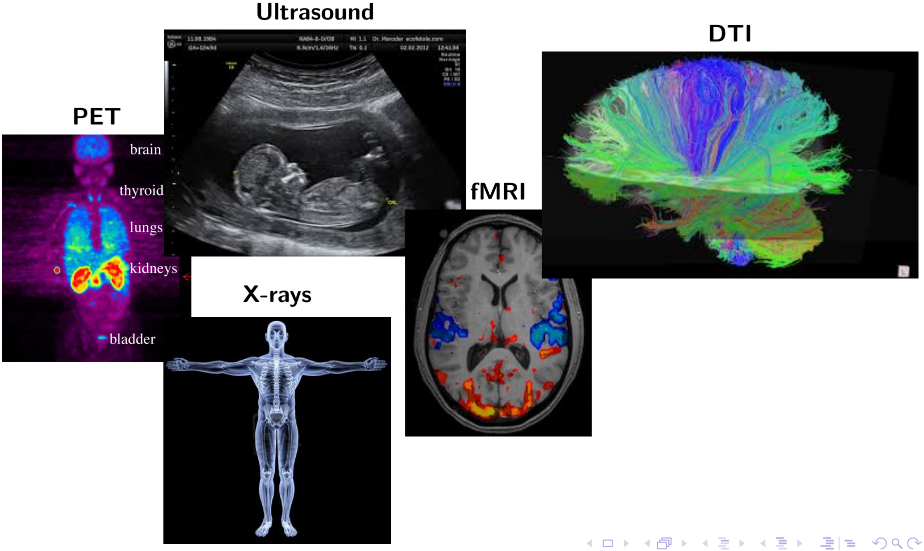


fMRI

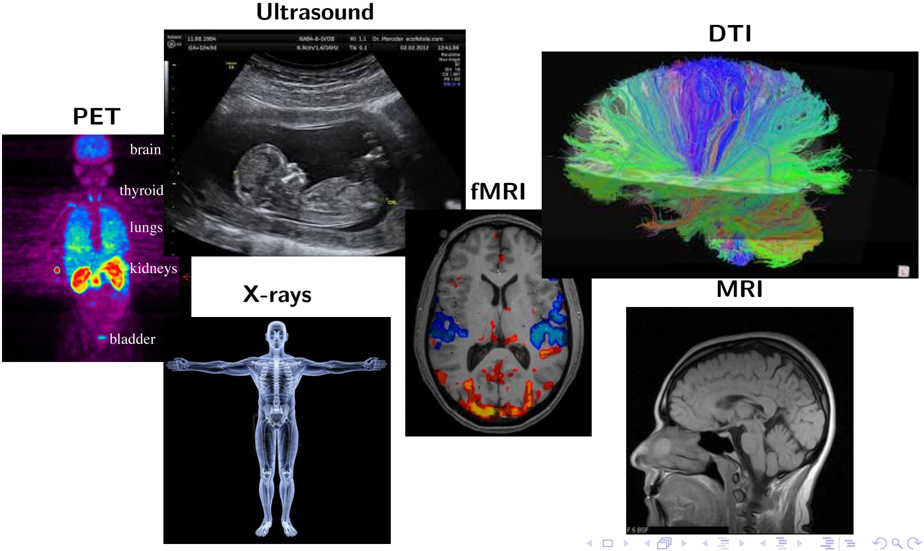


X-rays

Medical Imaging

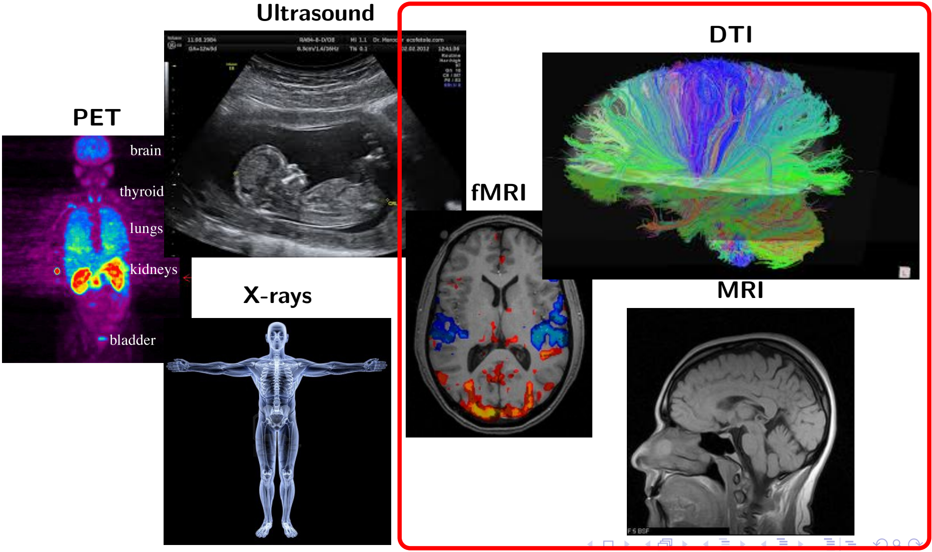


Medical Imaging



Medical Imaging

Image Classification



Medical image classification

Given a set of n paired observations $\{(\mathbb{I}_i, y_i)\}_{1 \leq i \leq n}$ where

- $\mathbb{I}_i$ is an medical image and
- $y_i \in \mathbb{R}$ is the classification label

the goal is to learn a *classification function* f.

Medical image classification

Given a set of n paired observations $\{(\mathbb{I}_i, y_i)\}_{1 \leq i \leq n}$ where

- $\mathbb{I}_i$ is an medical image and
- $y_i \in \mathbb{R}$ is the classification label

the goal is to learn a *classification function* f.

Problems

1. The representation of $\phi(\mathbb{I})$.
2. The learning process.

Medical image classification

Given a set of n paired observations $\{(\mathbb{I}_i, y_i)\}_{1 \leq i \leq n}$ where

- $\mathbb{I}_i$ is an medical image and
- $y_i \in \mathbb{R}$ is the classification label

the goal is to learn a *classification function* f.

Problems

1. The representation of $\phi(\mathbb{I})$.
    - Bag of Words approach
    - Graph representation.
2. The learning process.

Medical image classification

Given a set of n paired observations $\{(\mathbb{I}_i, y_i)\}_{1\leq i\leq n}$ where

- $\mathbb{I}_i$ is an medical image and
- $y_i \in \mathbb{R}$ is the classification label

the goal is to learn a *classification function* f.

Problems

1. The representation of $\phi(\mathbb{I})$.
   - Bag of Words approach
   - Graph representation.
2. The learning process.

Medical image classification

Given a set of n paired observations $\{(\mathbb{I}_i, y_i)\}_{1 \leq i \leq n}$ where

- $\mathbb{I}_i$ is an medical image and
- $y_i \in \mathbb{R}$ is the classification label

the goal is to learn a *classification function* f.

Problems

1. The representation of $\phi(\mathbb{I})$.
    - Bag of Words approach
    - Graph representation.
2. Supervised statistical learning framework

$$\arg\min_{f \in \mathcal{F}} \lambda\Omega(f) + \overbrace{\frac{1}{n}\sum_{i=1}^{n} \mathbb{L}(f(\phi(\mathbb{I}_i)), y_i)}^{\text{Empirical Risk}}$$

where $\mathbb{L}$ is the *loss function* and $\lambda\Omega(f)$ is the regularization term.

Table of Content

Table of Contents

What is a graph and why is it interesting?

Definition

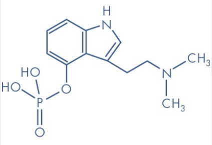
A *labeled graph* G is defined as a triplet $(V, E, \mathcal{L})$, where V is the *vertex set* and $E \subseteq V \times V$ is the *edge set* which represents a binary relation on V and $\mathcal{L} : X \mapsto \Sigma$ is a function assigning a label from an alphabet Σ to each element of the set X, which can be either V, E or $V \cup E$.

What is a graph and why is it interesting?

Definition

A *labeled graph* G is defined as a triplet $(V, E, \mathcal{L})$, where V is the *vertex set* and $E \subseteq V \times V$ is the *edge set* which represents a binary relation on V and $\mathcal{L} : X \mapsto \Sigma$ is a function assigning a label from an alphabet Σ to each element of the set X, which can be either V, E or $V \cup E$.

Areas of application



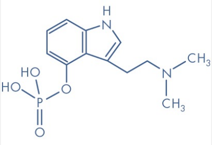
Chemoinformatics

What is a graph and why is it interesting?

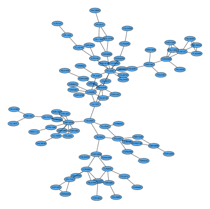
Definition

A *labeled graph* G is defined as a triplet $(V, E, \mathcal{L})$, where V is the *vertex set* and $E \subseteq V \times V$ is the *edge set* which represents a binary relation on V and $\mathcal{L} : X \mapsto \Sigma$ is a function assigning a label from an alphabet Σ to each element of the set X, which can be either V, E or $V \cup E$.

Areas of application



Chemoinformatics



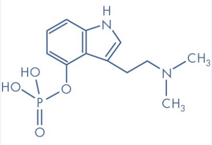
Bioinformatics

What is a graph and why is it interesting?

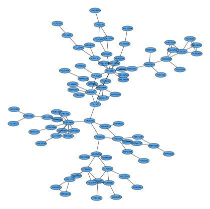
Definition

A *labeled graph* G is defined as a triplet $(V, E, \mathcal{L})$, where V is the *vertex set* and $E \subseteq V \times V$ is the *edge set* which represents a binary relation on V and $\mathcal{L} : X \mapsto \Sigma$ is a function assigning a label from an alphabet Σ to each element of the set X, which can be either V, E or $V \cup E$.

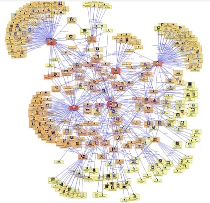
Areas of application



Chemoinformatics



Bioinformatics



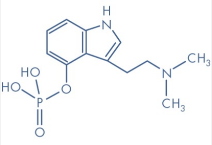
SNA

What is a graph and why is it interesting?

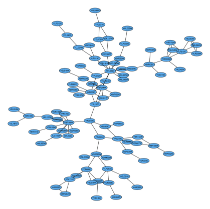
Definition

A *labeled graph* G is defined as a triplet $(V, E, \mathcal{L})$, where V is the *vertex set* and $E \subseteq V \times V$ is the *edge set* which represents a binary relation on V and $\mathcal{L} : X \mapsto \Sigma$ is a function assigning a label from an alphabet Σ to each element of the set X, which can be either V, E or $V \cup E$.

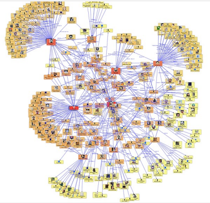
Areas of application



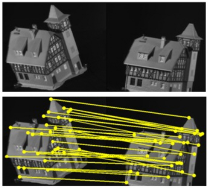
Chemoinformatics



Bioinformatics



SNA



Computer Vision

Graph Comparison Problem

Definition

Given a set $\mathcal{G}$ of graphs, the problem of graph comparison is defined as a function

$$k : \mathcal{G} \times \mathcal{G} \mapsto \mathbb{R}$$

such that $k(G, G')$ for $G, G' \in \mathcal{G}$ quantifies the similarity of G and G'.

Graph Comparison Problem

Definition

Given a set $\mathcal{G}$ of graphs, the problem of graph comparison is defined as a function

$$k : \mathcal{G} \times \mathcal{G} \mapsto \mathbb{R}$$

such that $k(G, G')$ for $G, G' \in \mathcal{G}$ quantifies the similarity of G and G'.

1st Approach

- Graph Isomorphism
- Subgraph Isomorphism
- Largest common subgraph

Graph Comparison Problem

Definition

Given a set $\mathcal{G}$ of graphs, the problem of graph comparison is defined as a function

$$k : \mathcal{G} \times \mathcal{G} \mapsto \mathbb{R}$$

such that $k(G, G')$ for $G, G' \in \mathcal{G}$ quantifies the similarity of G and G'.

1st Approach

- Graph Isomorphism - No efficient algorithm is known
- Subgraph Isomorphism - Proven to be NP-complete
- Largest common subgraph - Proven to be NP-hard

Graph Comparison Problem

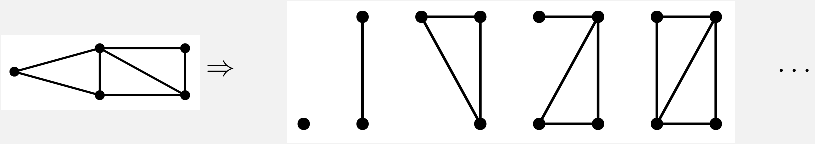
Definition

Given a set $\mathcal{G}$ of graphs, the problem of graph comparison is defined as a function

$$k : \mathcal{G} \times \mathcal{G} \mapsto \mathbb{R}$$

such that $k(G, G')$ for $G, G' \in \mathcal{G}$ quantifies the similarity of G and G'.

2nd Approach - R-convolution Kernels



Graph Comparison Problem

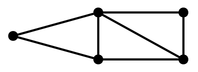
Definition

Given a set $\mathcal{G}$ of graphs, the problem of graph comparison is defined as a function

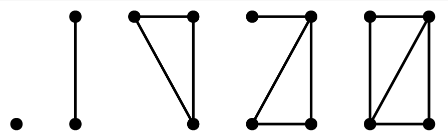
$$k : \mathcal{G} \times \mathcal{G} \mapsto \mathbb{R}$$

such that $k(G, G')$ for $G, G' \in \mathcal{G}$ quantifies the similarity of G and G'.

2nd Approach - R-convolution Kernels



$\Rightarrow$



$\cdots$

Calculating all subgraphs is at least as hard as deciding whether two graphs are isomorphic [Gärtner 03]

Graph kernels

	Algorithm	Complexity[1]	Unlabeled	Discrete	Continuous	Vector
Walks	[Gärtner 03]	$\mathcal{O}(n^2 v^6)$	✓	✓	✓	✓
	[Mahé 04]		✓	✓		
	[Vishwanathan 10]	$\mathcal{O}(n^2 v^3)$	✓	✓	✓	✓
Paths	[Borgwardt 05]	$\mathcal{O}(n^2 v^4)$	✓	✓	✓	✓
	[Ralaivola 05]		✓	✓		
Graphlets	[Horváth 04]		✓	✓		
	[Shervashidze 09]	$\mathcal{O}(vd^{k-1})$	✓			
	[Costa 10]		✓	✓		
Subtree Patterns	[Ramon 03]	$\mathcal{O}(n^2 v^2 h 4^d)$	✓	✓		
	[Bach 08]		✓	✓		
	[Mahé 09]		✓	✓		
	[Shervashidze 11]	$\mathcal{O}(nhe + n^2 hv)$	✓	✓		

---

[1] where n is the number of graphs, v is the maximal number of nodes, e is the maximal number of edges, h is the height of subtree patterns, d is the maximum degree and k is the size of graphlets.

Table of Contents

Overview of the WLpyramid

Given a set $\mathcal{G} = \{G_i = (V_i, E_i, \mathcal{L}_i)\}_{1 \leq i \leq n}$ where $\mathcal{L}_i : V_i \mapsto \mathbb{R}^d$

1. A pyramid quantization of the label space.
2. Transformation of the initial graphs.
3. Produce subtree features with Weisfeiler-Lehman algorithm.
4. Learning the combination of the subtree features.

Overview of the WLpyramid

Given a set $\mathcal{G} = \{G_i = (V_i, E_i, \mathcal{L}_i)\}_{1 \leq i \leq n}$ where $\mathcal{L}_i : V_i \mapsto \mathbb{R}^d$

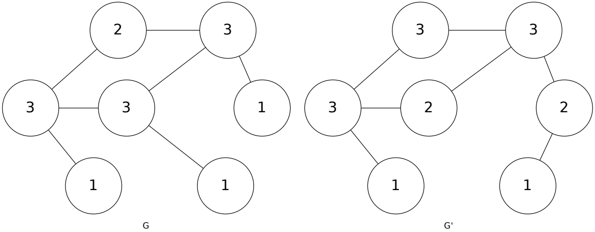
1. A pyramid quantization of the label space.
2. Transformation of the initial graphs.
3. Produce subtree features with Weisfeiler-Lehman algorithm.
4. Learning the combination of the subtree features.

Why Weisfeiler-Lehman?

1. Computational time $\mathcal{O}(nhe)$
   - n the number of graphs
   - e the maximal number of edges and and
   - h the height subtree features.
2. Competitive accuracy in several classification benchmark data sets [Shervashidze 11].

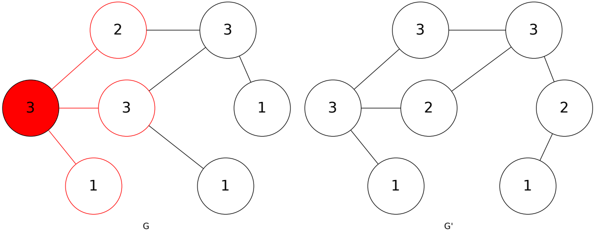
The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Given labeled graphs G and G'



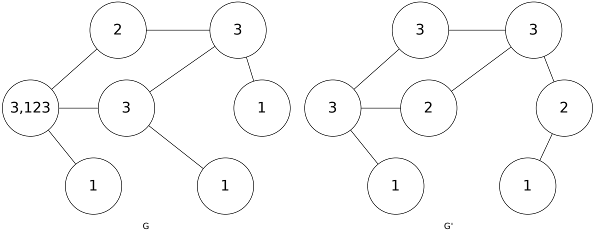
The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Multi-label determination of G and G'



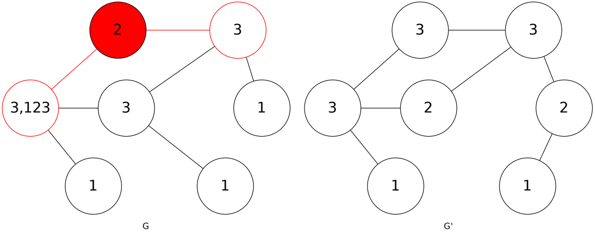
The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Multi-label determination of G and G'



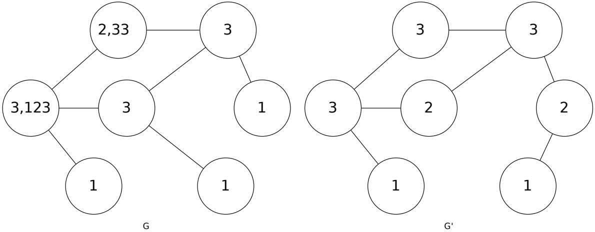
The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Multi-label determination of G and G'



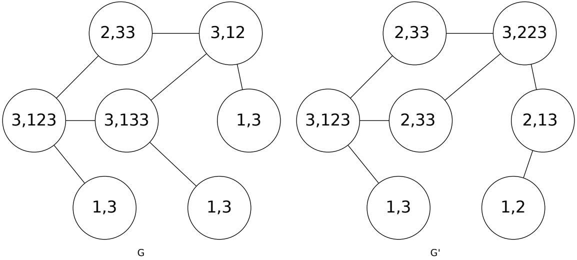
The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Multi-label determination of G and G'



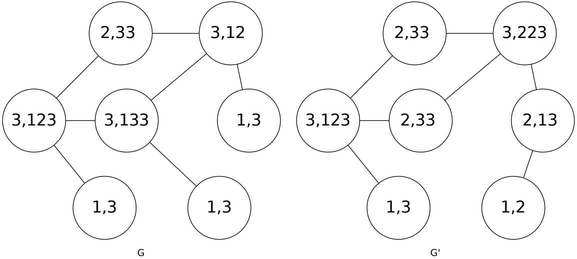
The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Multi-label determination of G and G'



The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Multi-label determination of G and G'



Label compression via hashing

1,2 ⟶ 4

1,3 ⟶ 5

2,13 ⟶ 6

2,33 ⟶ 7

3,12 ⟶ 8

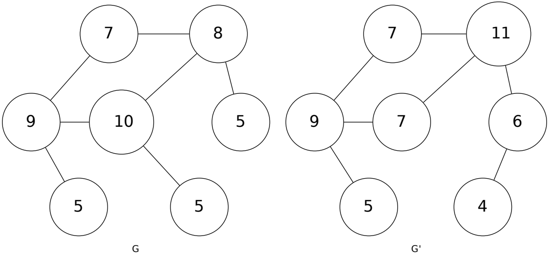
3,123 ⟶ 9

3,133 ⟶ 10

3,223 ⟶ 11

The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Relabeling graphs G and G'



Label compression via hashing

1,2 ⟶ 4

1,3 ⟶ 5

2,13 ⟶ 6

2,33 ⟶ 7

3,12 ⟶ 8

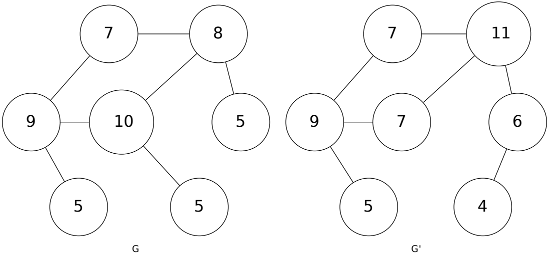
3,123 ⟶ 9

3,133 ⟶ 10

3,223 ⟶ 11

The Weisfeiler-Lehman test of isomorphism [Weisfeiler 68]

Relabeling graphs G and G'

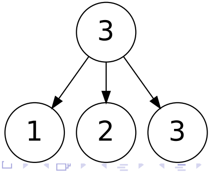


Label compression via hashing

1,2 → 4
1,3 → 5
2,13 → 6
2,33 → 7

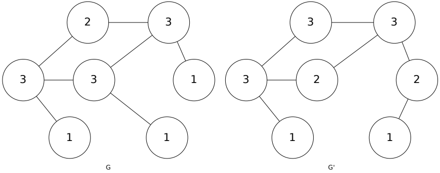
3,12 → 8
3,123 → 9
3,133 → 10
3,223 → 11

Subtree Pattern of Compressed label 9



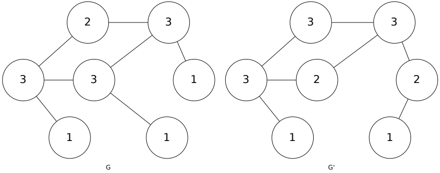
Weisfeiler-Lehman subtree features

Subtree patterns of depth 0.

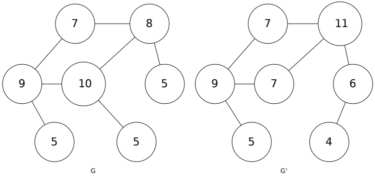


Weisfeiler-Lehman subtree features

Subtree patterns of depth 0.

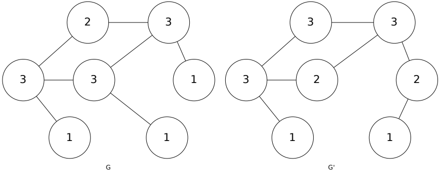


Subtree patterns of depth 1.

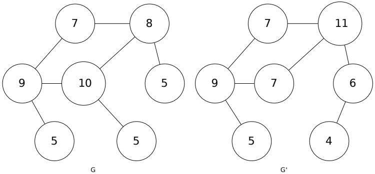


Weisfeiler-Lehman subtree features

Subtree patterns of depth 0.



Subtree patterns of depth 1.

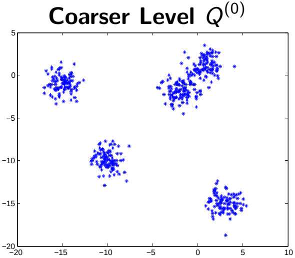


$$\begin{array}{rcl}
\text{Labels } \{\Sigma_0, \Sigma_1\} & = & \{\overbrace{1,\ \ 2,\ \ 3,}^{\text{Original node labels } \Sigma_0}\ \ \overbrace{4,\ \ 5,\ \ 6,\ \ 7,\ \ 8,\ \ 9,\ \ 10,\ \ 11}^{\text{Compressed node labels } \Sigma_1}\} \\
\phi_{(1)}(G) & = & (3,\ \ 1,\ \ 3,\ \ 0,\ \ 3,\ \ 0,\ \ 1,\ \ 1,\ \ 1,\ \ 1,\ \ 0) \\
\phi_{(1)}(G') & = & (2,\ \ 2,\ \ 3,\ \ 1,\ \ 1,\ \ 1,\ \ 2,\ \ 0,\ \ 1,\ \ 0,\ \ 1)
\end{array}$$

$\phi_{(h)}(G)$ are histograms of occurences of the subtree patterns up to depth h in graph G.

The pyramid quantization strategy

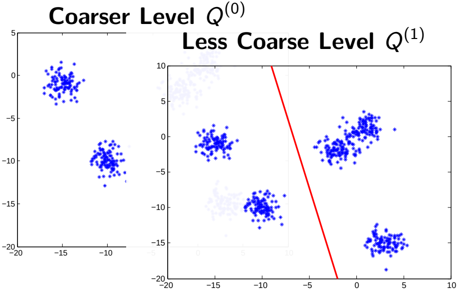
Given a set $\mathcal{G} = \{G_i = (V_i, E_i, \mathcal{L}_i)\}_{1 \leq i \leq n}$ where $\mathcal{L}_i : V_i \mapsto \mathbb{R}^d$



where $L = \lceil log_2 |V| \rceil$ and $|V| = \sum_i^n |V_i|$ [Grauman 07].

The pyramid quantization strategy

Given a set $\mathcal{G} = \{G_i = (V_i, E_i, \mathcal{L}_i)\}_{1 \leq i \leq n}$ where $\mathcal{L}_i : V_i \mapsto \mathbb{R}^d$



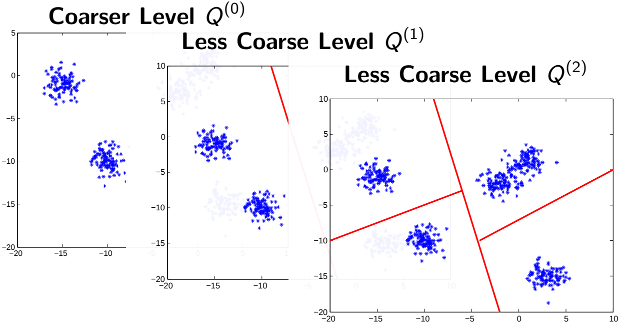
Coarser Level $Q^{(0)}$

Less Coarse Level $Q^{(1)}$

where $L = \lceil log_2 |V| \rceil$ and $|V| = \sum_i^n |V_i|$ [Grauman 07].

The pyramid quantization strategy

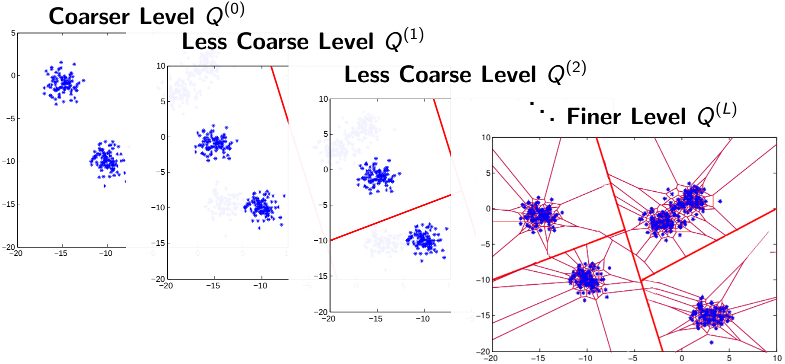
Given a set $\mathcal{G} = \{G_i = (V_i, E_i, \mathcal{L}_i)\}_{1 \leq i \leq n}$ where $\mathcal{L}_i : V_i \mapsto \mathbb{R}^d$



where $L = \lceil log_2 |V| \rceil$ and $|V| = \sum_i^n |V_i|$ [Grauman 07].

The pyramid quantization strategy

Given a set $\mathcal{G} = \{G_i = (V_i, E_i, \mathcal{L}_i)\}_{1\leq i\leq n}$ where $\mathcal{L}_i : V_i \mapsto \mathbb{R}^d$



Coarser Level $Q^{(0)}$

Less Coarse Level $Q^{(1)}$

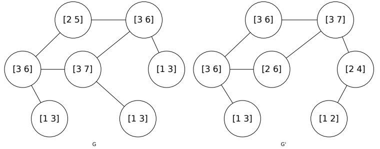
Less Coarse Level $Q^{(2)}$

$\ddots$ **Finer Level** $Q^{(L)}$

where $L = \lceil log_2 |V| \rceil$ and $|V| = \sum_i^n |V_i|$ [Grauman 07].

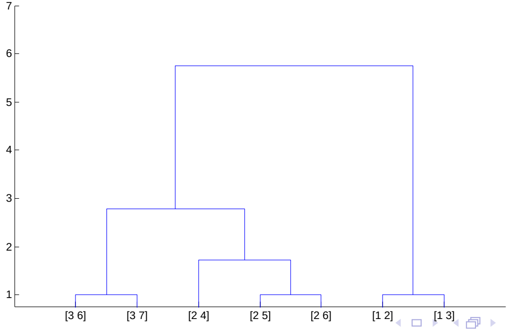
Data guided pyramid quantization scheme

Given labeled graphs G and G'



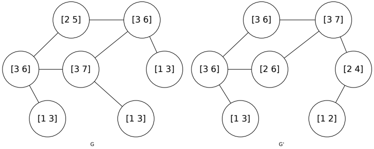
Notes

- Ward's minimum variance method over the image of V under $\mathcal{L}$.



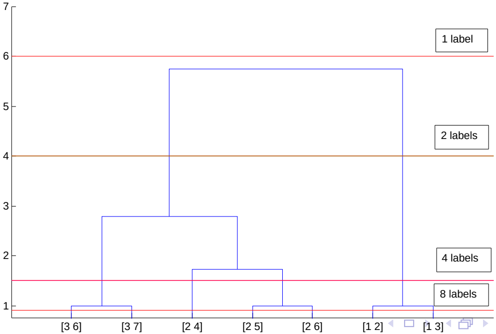
Data guided pyramid quantization scheme

Given labeled graphs G and G'



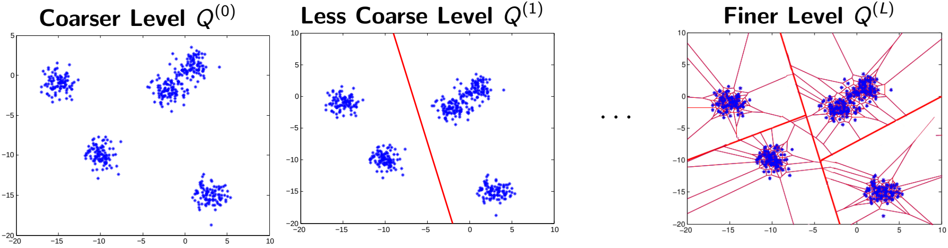
Notes

- Ward's minimum variance method over the image of V under $\mathcal{L}$.
- Selecting $L = \lceil log_2 D \rceil$, where $D \leq |V|$ the number of unique values in the image of V under $\mathcal{L}$
- Each level l has 2^l discrete labels.



Transform the initial graphs as a sequence of graphs

The pyramid quantization of label space



Coarser Level $Q^{(0)}$

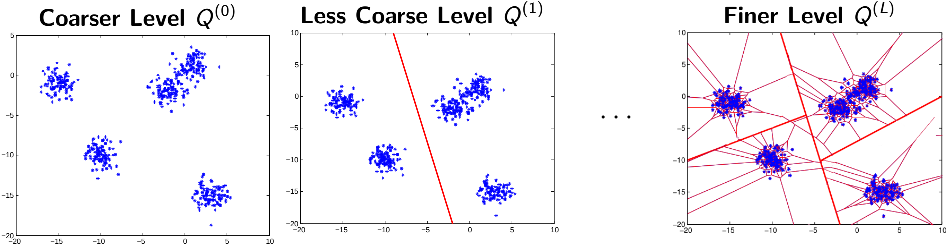
Less Coarse Level $Q^{(1)}$

. . .

Finer Level $Q^{(L)}$

Transform the initial graphs as a sequence of graphs

The pyramid quantization of label space



Coarser Level $Q^{(0)}$ **Less Coarse Level $Q^{(1)}$** . . . **Finer Level $Q^{(L)}$**

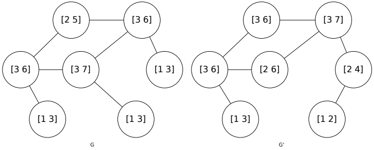
Sequence of discretely labeled graphs

$$G = (V, E, \mathcal{L}) \overset{Q^{(l)}\circ\mathcal{L}}{\underset{\forall l}{\approx}} \underset{\text{Increasing granularity}}{\underrightarrow{\left(G^{(0)}, \dots, G^{(L)}\right)}} = \underset{\text{Increasing granularity}}{\underrightarrow{\left((V, E, \mathcal{L}^{(0)}), \dots, (V, E, \mathcal{L}^{(L)})\right)}}$$

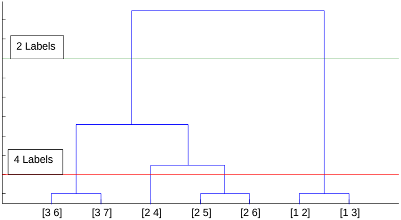
where $\mathcal{L}^{(l)} : V \rightarrow \Sigma_0^{(l)}$, $|\Sigma_0^{(l)}| = 2^l$ and $l \in \{0, \dots, L\}$.

A sequence of discretely labeled graphs

Given labeled graphs G and G'

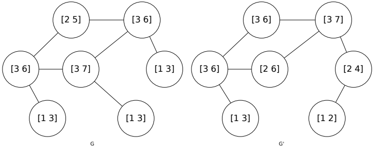


Data guided pyramid quantization.

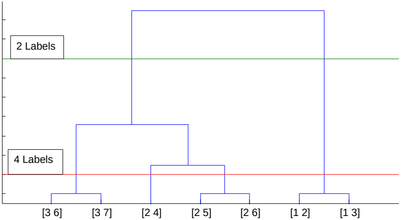


A sequence of discretely labeled graphs

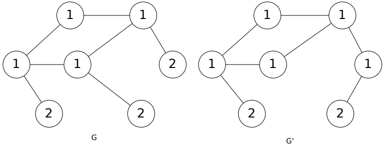
Given labeled graphs G and G'



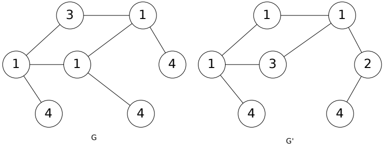
Data guided pyramid quantization.



Quantization level 1 with 2^1 number of labels.



Quantization level 2 with 2^2 number of labels



Creating and combining subtree features

Run Weisfeiler-Lehman on each quantization level

$$G = \left(G^{(0)}, \ldots, G^{(L)}\right) \xrightarrow[Lehman]{Weisfeler} \left(\phi^{(0)}_{(h)}(G^{(0)}), \ldots, \phi^{(L)}_{(h)}(G^{(L)})\right) = \widehat{\phi_{(h)}(G})$$

where $\phi^{(l)}_{(h)}(G^{(l)})$ are histograms of occurences of the subtree patterns up to depth h at the quantization level l in graph G

Creating and combining subtree features

Run Weisfeiler-Lehman on each quantization level

$$G = \left(G^{(0)}, \ldots, G^{(L)}\right) \xrightarrow[Lehman]{Weisfeler} \left(\phi^{(0)}_{(h)}(G^{(0)}), \ldots, \phi^{(L)}_{(h)}(G^{(L)})\right) = \widehat{\phi_{(h)}(G)}$$

where $\phi^{(l)}_{(h)}(G^{(l)})$ are histograms of occurences of the subtree patterns up to depth h at the quantization level l in graph G

Learning to combine the quantization levels

1. Learn the selection of the subtree features $\widehat{\phi_{(h)}(G)}$.
2. Combine the subtree features $\phi^{(l)}_{(h)}(G^{(l)})$ per level l into a kernel and then learn the combination of kernels.

Learn the subtree patterns selection

- Labeled training data $\{(\widehat{\phi_{(h)}(G_i)}, y_i)\}_{1 \leq i \leq n} \in (\mathbb{N} \times \mathbb{R})^n$ where
  - $\widehat{\phi_{(h)}(G_i)}$ is the concatination of histograms of the occurences of subtree patterns up to depth h of graph G_i across all quantization levels,
  - y_i is the ground truth label and

Learn the subtree patterns selection

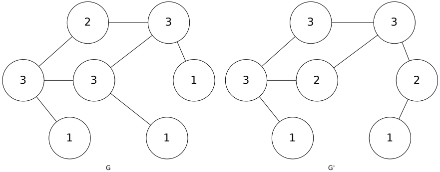
- Labeled training data $\{(\widehat{\phi_{(h)}(G_i)}, y_i)\}_{1 \leq i \leq n} \in (\mathbb{N} \times \mathbb{R})^n$ where
  - $\widehat{\phi_{(h)}(G_i)}$ is the concatination of histograms of the occurences of subtree patterns up to depth h of graph G_i across all quantization levels,
  - y_i is the ground truth label and
- Elastic Net [Zou 05]

$$\arg\min_{w \in \mathbb{R}^d} \underbrace{\lambda_1 \|w\|_1}_{\ell_1 \text{ norm}} + \underbrace{\lambda_2 \|w\|_2^2}_{\ell_2 \text{ norm}} + \underbrace{\frac{1}{n} \sum_{i=1}^{n} \left(\langle w, \widehat{\phi_{(h)}(G_i)} \rangle - y_i \right)^2}_{\text{Squared loss}}$$

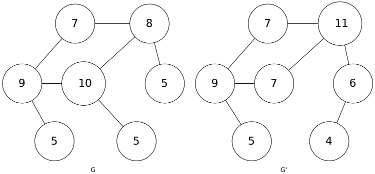
λ_1, λ_2 are scalar parameters controling the degree of regularization.

The intersection Weisfeiler-Lehman kernel

Subtree patterns of depth 0.



Subtree patterns of depth 1.



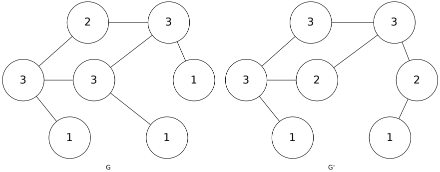
$$\text{Labels } \{\Sigma_0, \Sigma_1\} = \{\overbrace{1, \quad 2, \quad 3,}^{\text{Subtree patterns } h = 0} \quad \overbrace{4, \quad 5, \quad 6, \quad 7, \quad 8, \quad 9, \quad 10, \quad 11}^{\text{Subtree patterns } h = 1}\}$$

$$\phi^{(l)}_{(1)}(G^{(l)}) = (3, \quad 1, \quad 3, \quad 0, \quad 3, \quad 0, \quad 1, \quad 1, \quad 1, \quad 1, \quad 0)$$

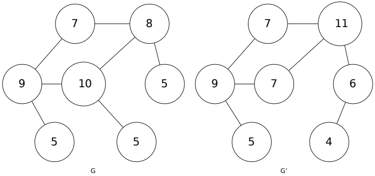
$$\phi^{(l)}_{(1)}(G'^{(l)}) = (2, \quad 2, \quad 3, \quad 1, \quad 1, \quad 1, \quad 2, \quad 0, \quad 1, \quad 0, \quad 1)$$

The intersection Weisfeiler-Lehman kernel

Subtree patterns of depth 0.



Subtree patterns of depth 1.



$$\begin{array}{rcl}
\text{Labels } \{\Sigma_0, \Sigma_1\} & = & \{\overbrace{1,\quad 2,\quad 3,}^{\text{Subtree patterns } h=0}\quad \overbrace{4,\quad 5,\quad 6,\quad 7,\quad 8,\quad 9,\quad 10,\quad 11}^{\text{Subtree patterns } h=1}\} \\
\phi^{(l)}_{(1)}(G^{(l)}) & = & (3,\quad 1,\quad 3,\quad 0,\quad 3,\quad 0,\quad 1,\quad 1,\quad 1,\quad 1,\quad 0) \\
\phi^{(l)}_{(1)}(G'^{(l)}) & = & (2,\quad 2,\quad 3,\quad 1,\quad 1,\quad 1,\quad 2,\quad 0,\quad 1,\quad 0,\quad 1) \\
\min\left(\phi^{(l)}_{(1)}(G^{(l)}), \phi^{(l)}_{(1)}(G'^{(l)})\right) & = & (2,\quad 1,\quad 3,\quad 0,\quad 1,\quad 0,\quad 1,\quad 0,\quad 1,\quad 0,\quad 0)
\end{array}$$

The intersection Weisfeile-Lehman kernel is defined :

$$k^{(h)}_{i-WLsubtree}(G^{(l)}, G'^{(l)}) = \sum_{j}^{|\Sigma_0 \cup \Sigma_1|} \min\left(\phi^{(l)}_{(1)}(G^{(l)}), \phi^{(l)}_{(1)}(G'^{(l)})\right)_j$$

Multiple Kernel Learning

Problem

For each pair of graphs $G^{(l)}, G'^{(l)}$ for all the pyramid levels:

$$\left(K_{(h)}^{(0)}(G^{(0)}, G'^{(0)}), \ldots, K_{(h)}^{(L)}(G^{(l)}, G'^{(L)})\right)$$

we would like to learn a linear combination of them:

$$K_{(h)}(G, G') = \sum_{l=0}^{L} d_l K_{(h)}^{(l)}(G^{(l)}, G'^{(l)}), \text{ with } d_l \geq 0, \sum_{l=0}^{L} d_l = 1.$$

Multiple Kernel Learning

Problem

For each pair of graphs $G^{(l)}, G'^{(l)}$ for all the pyramid levels:

$$\left(K_{(h)}^{(0)}(G^{(0)}, G'^{(0)}), \ldots, K_{(h)}^{(L)}(G^{(l)}, G'^{(L)})\right)$$

we would like to learn a linear combination of them:

$$K_{(h)}(G, G') = \sum_{l=0}^{L} d_l K_{(h)}^{(l)}(G^{(l)}, G'^{(l)}), \text{ with } d_l \geq 0, \sum_{l=0}^{L} d_l = 1.$$

Solutions

- Multiple kernel learning
- Average weighted kernel

Table of Contents

fMRI Analysis

Key information

1. Inherent spatial structure brains
2. Voxel activation is a continuous value

fMRI Analysis

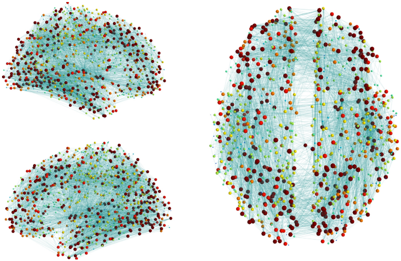
Key information

1. Inherent spatial structure brains
2. Voxel activation is a continuous value

⇓

Solution!

Represent fMRI as graphs with continuous labels.



Dataset

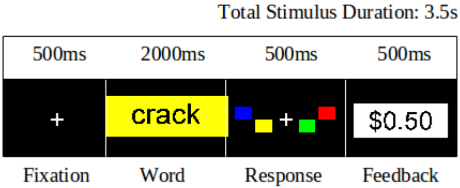
Cocaine Addiction Dataset

- 16 cocaine addicted vs 17 healthy subjects
- Drugstroop experiment with two varying conditions
  - the cue shown and
  - the monetary reward.

Input One contrast map per subject that is transformed into a graph.

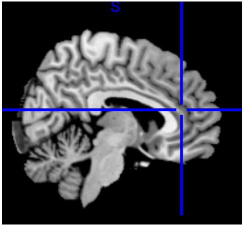
Objective The classification between cocaine abuser and control group.

Drugstroop Experiment



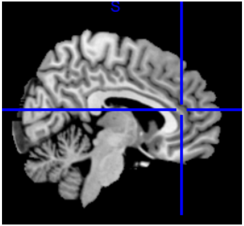
Graph Transformation

Contrast map



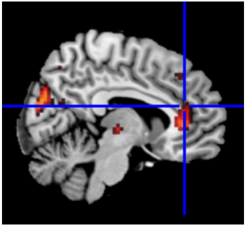
Graph Transformation

Contrast map



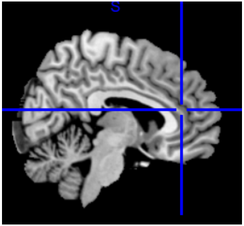
$$\xrightarrow{\text{Elastic Net}}$$

Selected voxels



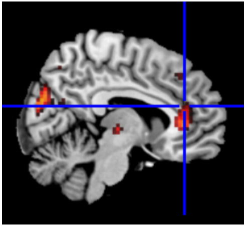
Graph Transformation

Contrast map

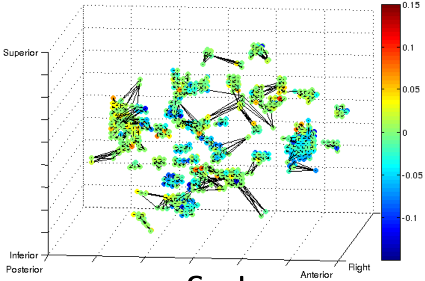


$\xrightarrow{\text{Elastic Net}}$

Selected voxels

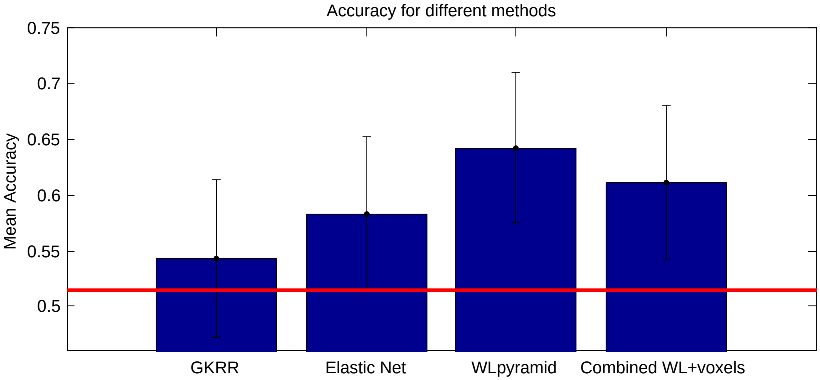


knn

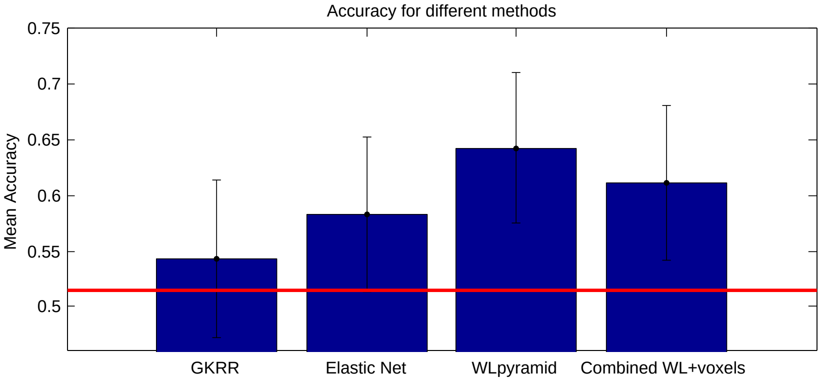


Graph

Performance



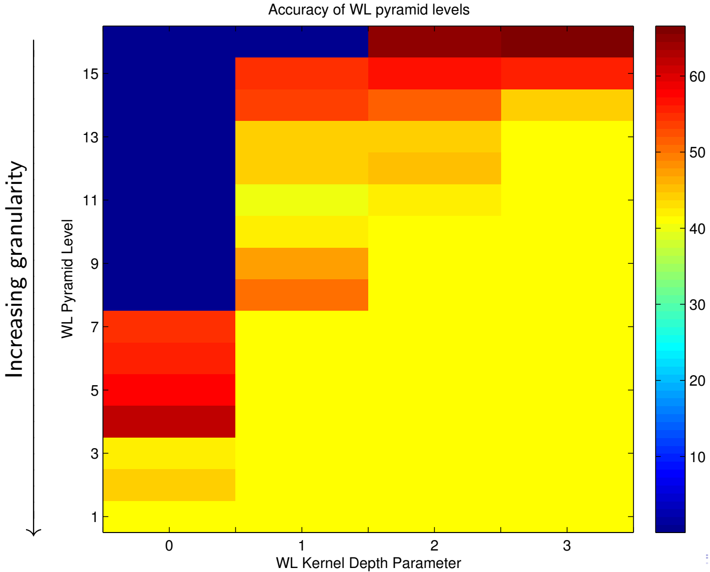
Performance



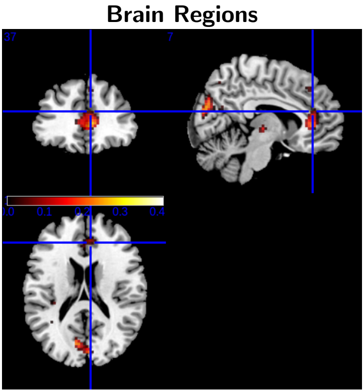
WLpyramid vs Elastic Net on raw voxels

Wilcoxon signed rank with $p = 0.02$ show statistical significance.

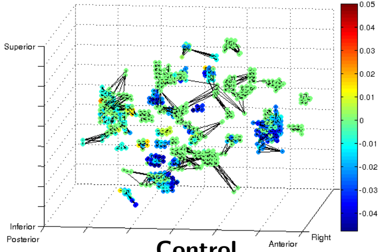
Performance per pyramid quantization level



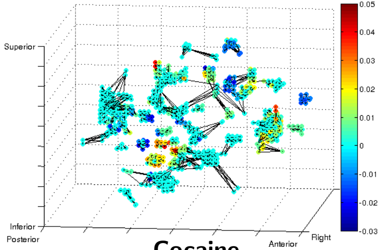
Visualization of learned function



Brain Regions



Control

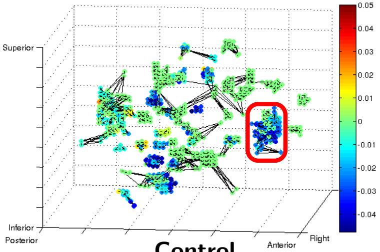


Cocaine

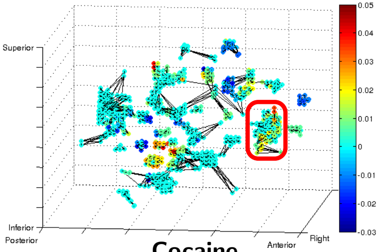
Visualization of learned function

Rostral Anterior Cingulate Cortex

- In cocaine addicted subjects deactivates during the drug Stroop experiment as compared to baseline.
- Its activity is normalized by oral methylphenidate where the dopamine transporters increase the extracellular dopamine, an increase which is associated with lower task-related impulsivity.



Control

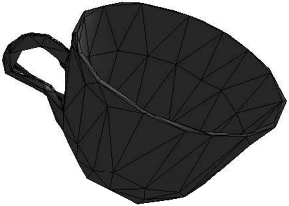


Cocaine

3D shape classification

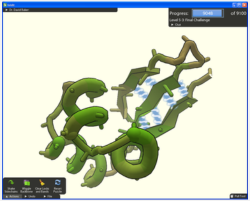
3D shape problems

- Storage
- Classification
- Retrieval

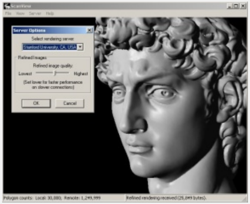


Areas of applications

3D Game



Chemoinformatics

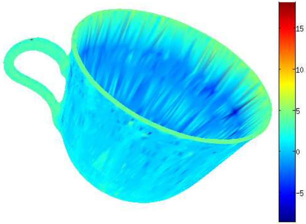


Cultural heritage

3D shape classification

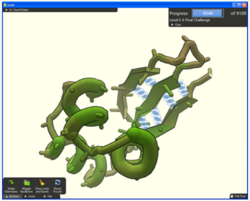
3D shape problems

- Storage
- Classification
- Retrieval



Areas of applications

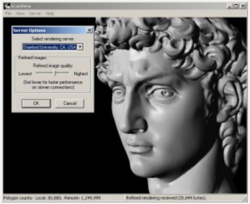
3D Game



Progress: 9048 of 9100

Level 5-3: Final Challenge

Chemoinformatics



Select rendering server: Stanford University, CA, USA

(Set lower for faster performance on slower connections)

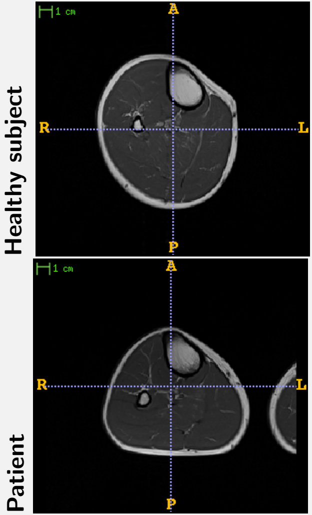
Polygon counts: Local: 30,000; Remote: 1,249,999

Refined rendering received (25,849 bytes).

Cultural heritage

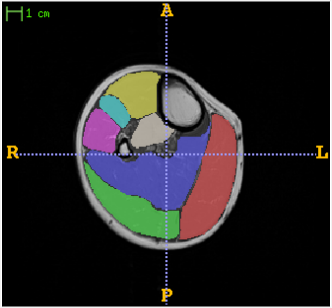
3D shape datasets

Muscle Dataset



3D shape datasets

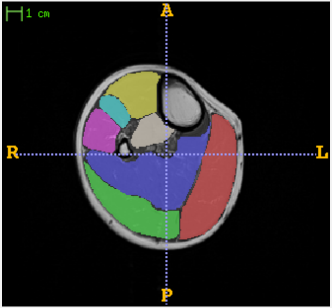
Muscle Dataset



- 27 patients vs 14 healthy subjects
- MRI images of the calf muscles
- Segmented into 7 muscles

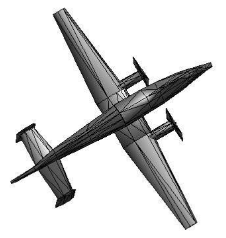
3D shape datasets

Muscle Dataset



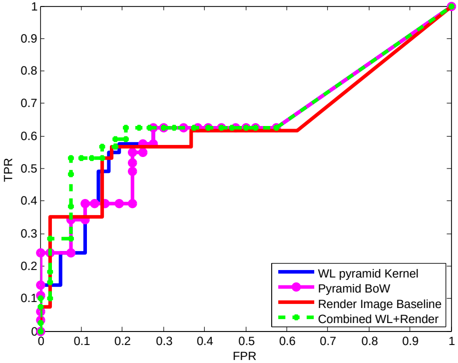
- 27 patients vs 14 healthy subjects
- MRI images of the calf muscles
- Segmented into 7 muscles

SHREC 2013 Dataset



- 20 classes of generic objects, such as bed, biplane, mug, etc.
- Each class contains 18 models.
- In total 360 3D objects.

Performance on the muscle dataset



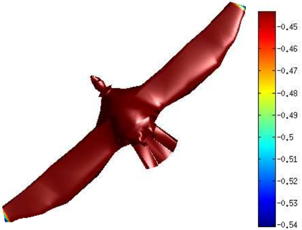
	WLpyramid Our Work	**pyramid BoW**	**Rendering**	**Combined**
Accuracy	78.00%	73.00%	75.50%	82.93%
AUC	0.6410	0.6361	0.6300	0.6648

Performance on SHREC 2013

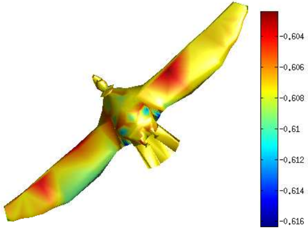
Class	WLpyramid Our Work	pyramid BoW	Rendering	Combined
Bird	0.85	0.83	0.85	**0.86**
Bicycle	0.84	0.87	0.90	0.90
Biped	0.89	0.88	0.99	0.99
Biplane	0.60	0.63	0.68	**0.69**
Bird	0.73	0.73	0.80	0.80
Bottle	0.76	0.76	0.79	**0.80**
Car	0.78	0.79	0.80	0.80
CellPhone	0.74	0.80	0.88	**0.89**
Chair	0.69	0.68	0.70	**0.72**
Cup	0.85	0.84	0.88	0.88
Desklamp	0.80	0.80	0.88	**0.89**
Fish	1.00	1.00	1.00	1.00
Floorlamp	0.80	0.77	0.89	0.89
Insect	0.64	0.60	0.62	**0.66**
Monoplane	0.84	0.82	0.88	**0.90**
Mug	0.82	0.82	0.85	**0.87**
Phone	0.83	0.74	0.72	**0.83**
Quadruped	0.89	0.86	0.97	**0.98**
Sofa	0.76	0.75	0.74	**0.75**
Wheelchair	0.81	0.79	0.88	**0.90**
Average	0.80	0.79	0.84	0.85

SHREC 2013 - Visualization of the learned weights

Subtree patterns of depth 0



Subtree patterns up to depth 2



Subtree patterns up to depth 1

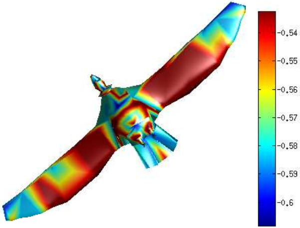


Table of Contents

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.

k-support regularized SVM

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.

***k*-support regularized SVM**

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.

k-support regularized SVM

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains

k-support regularized SVM

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains
    - fMRI analysis and

k-support regularized SVM

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains
  - fMRI analysis and
  - 3D shape classification.

k-support regularized SVM

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains
    - fMRI analysis and
    - 3D shape classification.
- Visualizations of the learned functions provide interpretability.

k-support regularized SVM

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains
  - fMRI analysis and
  - 3D shape classification.
- Visualizations of the learned functions provide interpretability.

k-support regularized SVM

- A novel regularized SVM algorithm.

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains
  - fMRI analysis and
  - 3D shape classification.
- Visualizations of the learned functions provide interpretability.

k-support regularized SVM

- A novel regularized SVM algorithm.
- Correlated sparse solution under the SVM framework.

Contributions - Methodological

The *pyramid quantized Weisfeiler-Lehman graph representation*

- A novel algorithm for comparing graphs with vector labels.
- Based on subtree patterns.
- Linear computation time in the number of graphs, in the number of edges in the graphs and in the depth of subtree patterns.
- Evaluation on two domains
  - fMRI analysis and
  - 3D shape classification.
- Visualizations of the learned functions provide interpretability.

k-support regularized SVM

- A novel regularized SVM algorithm.
- Correlated sparse solution under the SVM framework.
- Evaluation on a neuromuscular disease task.

Contributions

Methodological

Code from both algorithms is available online under GNU-GPL at
`http://cvc.centrale-ponts.fr/personnel/gkirtzou/code`

Contributions

Methodological

Code from both algorithms is available online under GNU-GPL at `http://cvc.centrale-ponts.fr/personnel/gkirtzou/code`

Clinical and Applications

- Investigate the applicability of sparsity regularizers in fMRI analysis.

Contributions

Methodological

Code from both algorithms is available online under GNU-GPL at `http://cvc.centrale-ponts.fr/personnel/gkirtzou/code`

Clinical and Applications

- Investigate the applicability of sparsity regularizers in fMRI analysis.
- In the fMRI analysis, we saw that the interconnections between voxels can contain additional information about brain structure.

Contributions

Methodological

Code from both algorithms is available online under GNU-GPL at `http://cvc.centrale-ponts.fr/personnel/gkirtzou/code`

Clinical and Applications

- Investigate the applicability of sparsity regularizers in fMRI analysis.
- In the fMRI analysis, we saw that the interconnections between voxels can contain additional information about brain structure.
- In the neuromuscular dystrophy classification task, we saw that features extracted from DTI images provide significant information.

Contributions

Methodological

Code from both algorithms is available online under GNU-GPL at `http://cvc.centrale-ponts.fr/personnel/gkirtzou/code`

Clinical and Applications

- Investigate the applicability of sparsity regularizers in fMRI analysis.
- In the fMRI analysis, we saw that the interconnections between voxels can contain additional information about brain structure.
- In the neuromuscular dystrophy classification task, we saw that features extracted from DTI images provide significant information.
- Interpretation of 3D shape meshes as annotated graphs.

Future Work

Medical image analysis

- Evaluation of k-support norm regularization on fMRI analysis problem in larger scale.
- Evaluation of k-support regularized SVM on neuromuscular disease discrimination in larger scale.
- Exploration of different constructions of the graphs from fMRI.

Future Work

Medical image analysis

- Evaluation of k-support norm regularization on fMRI analysis problem in larger scale.
- Evaluation of k-support regularized SVM on neuromuscular disease discrimination in larger scale.
- Exploration of different constructions of the graphs from fMRI.

Graph kernels

- Comparison on partially matching subtree patterns.
- Comparison on partially labeled graphs.

Table of Contents - Appendix

Francis R. Bach.
Graph kernels between point clouds.
In Proceedings of the 25th international conference on Machine learning, International Conference on Machine Learning '08, pages 25–32, 2008.

Karsten M. Borgwardt & Hans-Peter Kriegel.
Shortest-Path Kernels on Graphs.
In Proceedings of the Fifth IEEE International Conference on Data Mining, ICDM '05, pages 74–81, Washington, DC, USA, 2005. IEEE Computer Society.

Fabrizio Costa & Kurt De Grave.
Fast neighborhood subgraph pairwise distance kernel.
In Proceedings of the 26th International Conference on Machine Learning, pages 255–262, 2010.

Thomas Gärtner, Peter Flach & Stefan Wrobel.
On Graph Kernels: Hardness Results and Efficient Alternatives.
In Bernhard Schölkopf & Manfred K. Warmuth, editeurs, Learning Theory and Kernel Machines, volume 2777 of *Lecture Notes in Computer Science*, pages 129–143. Springer Berlin Heidelberg, 2003.

Kristen Grauman & Trevor Darrell.
The Pyramid Match Kernel: Efficient Learning with Sets of Features.
Journal of Machine Learning Research, vol. 8, pages 725–760, May 2007.

Tamás Horváth, Thomas Gärtner & Stefan Wrobel.
Cyclic pattern kernels for predictive graph mining.
In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '04, pages 158–167, 2004.

- P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret & J.-P. Vert.
  Extensions of marginalized graph kernels.
  In Proceedings of the Twenty-First International Conference on Machine Learning (ICML 2004), pages 552–559, 2004.

- Pierre Mahé & Jean-Philippe Vert.
  Graph kernels based on tree patterns for molecules.
  Machine Learning, vol. 75, no. 1, pages 3–35, 2009.

- Liva Ralaivola, Sanjay Joshua Swamidass, Hiroto Saigo & Pierre Baldi.
  Graph kernels for chemical informatics.
  Neural Networks, vol. 18, no. 8, pages 1093–1110, 2005.

- Jan Ramon & Thomas Gaertner.
  Expressivity versus efficiency of graph kernels.
  In Proceedings of the First International Workshop on Mining Graphs, Trees and Sequences, pages 65–74, 2003.

N. Shervashidze, S. V. N. Vishwanathan, T. Petri, K. Mehlhorn & K. Borgwardt.
Efficient graphlet kernels for large graph comparison.
In Proceedings of the International Workshop on Artificial Intelligence and Statistics. Society for Artificial Intelligence and Statistics, 2009.

Nino Shervashidze, Pascal Schweitzer, Erik Jan van Leeuwen, Kurt Mehlhorn & Karsten M. Borgwardt.
Weisfeiler-Lehman Graph Kernels.
Journal of Machine Learning Research, vol. 12, pages 2539–2561, November 2011.

S. Vichy N. Vishwanathan, Nicol N. Schraudolph, Risi Imre Kondor & Karsten M. Borgwardt.
Graph Kernels.
Journal of Machine Learning Research, vol. 11, pages 1201–1242, 2010.

Boris Weisfeiler & A.A. Lehman.
A reduction of a graph to a canonical form and an algebra arising during this reduction.
Nauchno-Technicheskaya Informatsia, vol. 2, no. 9, pages 12–16, 1968.

Hui Zou & Trevor Hastie.
Regularization and variable selection via the Elastic Net.
Journal of the Royal Statistical Society, Series B, vol. 67, pages 301–320, 2005.