Learning to Classify Seismic Images with Deep Optimum-Path Forest

Luis Afonso\textsuperscript{1}
Michelle Kuroda\textsuperscript{2}, Alexandre Vidal\textsuperscript{2}
Alexandre Falcão\textsuperscript{3}
João Papa\textsuperscript{4}

\textsuperscript{1}Department of Computing, UFSCar-Brazil
\textsuperscript{2}Institute of Geosciences, UNICAMP-Brazil
\textsuperscript{3}Institute of Computing, UNICAMP-Brazil
\textsuperscript{4}Department of Computing, UNESP/Bauru-Brazil
Outline

• Introduction

• Optimum-Path Forest

• Learning Deep Representations

• Methodology

• Experimental Results

• Conclusions
Introduction

- Image classification plays important role in wide range of applications:
  - Remote sensing-driven tools.
Introduction

- Image classification plays important role in wide range of applications:
  - Remote sensing-driven tools.
  - Medical image analysis.
Introduction

• Issues:
  ○ Lack of labeled data.
Introduction

• Issues:
  ○ Lack of labeled data.
  ○ Huge amount of images.
Introduction

- **Issues:**
  - Lack of labeled data.
  - Huge amount of images.

- **Active learning-based techniques:**
Introduction

• Issues:
  ○ Lack of labeled data.
  ○ Huge amount of images.

• Active learning-based techniques:
  ○ Also requires human interaction.
Introduction

• Issues:
  ○ Lack of labeled data.
  ○ Huge amount of images.

• Active learning-based techniques:
  ○ Also requires human interaction.

• Deep learning: Unsupervised learning.
Introduction

- A few existing solutions:
  - K-means
Introduction

- A few existing solutions:
  - K-means
  - Mean-shift
Introduction

- A few existing solutions:
  - K-means
  - Mean-shift
  - Self-Organizing Map
Introduction

- A few existing solutions:
  - K-means
  - Mean-shift
  - Self-Organizing Map

- Issue:
Introduction

• A few existing solutions:
  ○ K-means
  ○ Mean-shift
  ○ Self-Organizing Map

• Issue:
  ○ Number of clusters
Introduction

• A few existing solutions:
  ○ K-means
  ○ Mean-shift
  ○ Self-Organizing Map

• Issue:
  ○ Number of clusters
  ○ Get trapped in local optima
Introduction

• Graph-based clustering algorithms - Optimum Path Forest:
  ○ Samples are represented by nodes of a graph.
Introduction

- Graph-based clustering algorithms - Optimum Path Forest:
  - Samples are represented by nodes of a graph.
  - Competition-based learning process.
Introduction

- Graph-based clustering algorithms - Optimum Path Forest:
  - Samples are represented by nodes of a graph.
  - Competition-based learning process.
  - Application in bag-of-visual words.
Introduction

• Graph-based clustering algorithms - Optimum Path Forest:
  ○ Samples are represented by nodes of a graph.
  ○ Competition-based learning process.
  ○ Application in bag-of-visual words.
  ○ Similar centroids when compared to K-means.
Introduction

- Graph-based clustering algorithms - Optimum Path Forest:
  - Issue: "what if the value of $k$ is known?"
Introduction

• Graph-based clustering algorithms - Optimum Path Forest:
  ◦ Issue: "what if the value of $k$ is known?"
  ◦ $k_{max}$ parameter.
Introduction

• Graph-based clustering algorithms - Optimum Path Forest:
  ○ Issue: “what if the value of $k$ is known?”
  ○ $k_{\text{max}}$ parameter.
  ○ High cost to find optimum value of $k$. 
Introduction

- Graph-based clustering algorithms - Optimum Path Forest:
  - Deep-driven approach to reach a number of clusters close to the desired.
Introduction

- Graph-based clustering algorithms - Optimum Path Forest:
  - Deep-driven approach to reach a number of clusters close to the desired.
  - Unsupervised learning applied in different views of the data.
Optimum-Path Forest

- Nodes are connected by edges weighted by their distance (e.g., Euclidean distance).
Optimum-Path Forest

- Nodes are weighted by a probability density function $\rho$. 
Optimum-Path Forest
Optimum-Path Forest

- Create a priority queue $Q$ ordered by the value of $V$. 

![Graph with nodes A, B, C, D, E, F and weighted edges between them. The weights are labeled on the edges, and the values $\rho$ and $V$ are associated with each node.]}
Optimum-Path Forest

- Remove from $Q$ the node with the highest value.
Optimum-Path Forest

- Check if node’s predecessor is null. Update $V$ and create new label $l$, if so.
Optimum-Path Forest

- Check if node’s predecessor is null. Update $V$ and create new label $l$, if so.
- Node becomes a prototype.
Optimum-Path Forest
Optimum-Path Forest
Optimum-Path Forest

A \rightarrow B

tmp = \min\{V(A), \rho(B)\}

tmp = \min\{5, 3\}
Optimum-Path Forest

\[ \text{A} \rightarrow \text{B} \]
\[ \text{tmp} = \min\{V(A), \rho(B)\} \]
\[ \text{tmp} = \min\{5, 3\} \]
\[ \text{tmp} = 3 \]
Optimum-Path Forest

\[ \text{tmp} = \min\{V(A), \rho(B)\} \]
\[ \text{tmp} = \min\{5, 3\} \]
\[ \text{tmp} = 3 \]

\[ \text{tmp} > V(B)? \]
Optimum-Path Forest

\[ A \rightarrow B \]
\[ \text{tmp} = \min\{V(A), \rho(B)\} \]
\[ \text{tmp} = \min\{5, 3\} \]
\[ \text{tmp} = 3 \]
\[ \text{tmp} > V(B)\? \]
\[ 3 > 2 \? \]
Optimum-Path Forest

A → B

tmp = min\{V(A), \rho(B)\}
tmp = min\{5, 3\}
tmp = 3

tmp > V(B)\
3 > 2 ?
Yes!
Optimum-Path Forest

- \( V(B) \leftarrow \text{tmp} \)
Optimum-Path Forest

- \( V(B) \leftarrow \text{tmp} \)
- \( l(B) \leftarrow l(A) \)

```
\[
\begin{align*}
\text{tmp} &= \min\{V(A), \rho(B)\} \\
\text{tmp} &= \min\{5, 3\} \\
\text{tmp} &= 3 \\
n\text{tmp} > V(B)? \quad 3 > 2? \\
\text{Yes!}
\end{align*}
\]
```
Optimum-Path Forest

A – C

tmp = min\{V(A), \rho(C)\}
tmp = min\{5, 3\}
tmp = 3

tmp > V(C) ?
3 > 2 ? Yes
Optimum-Path Forest
Optimum-Path Forest

\[ \text{tmp} = \min\{V(A), \rho(4)\} \]
\[ \text{tmp} = \min\{5, 4\} \]
\[ \text{tmp} = 4 \]
\[ \text{tmp} > V(E) \? \]
\[ 4 > 3 \? \text{Yes} \]
Optimum-Path Forest

Diagram:
- Node A: \( \rho = 5 \), \( V = 5 \), \( l = 1 \)
- Node B: \( \rho = 3 \), \( V = 3 \), \( l = 1 \)
- Node C: \( \rho = 3 \), \( V = 3 \), \( l = 1 \)
- Node E: \( \rho = 4 \), \( V = 4 \), \( l = 1 \)

Edges:
- A to B: 0.2
- A to C: 0.4
- A to E: 0.5
Optimum-Path Forest

F → E → B → C → D
Optimum-Path Forest

\[ \begin{align*}
\text{F} & \quad \rho = 5 \\
\text{V} & \quad V = 5 \\
\text{l} & \quad l = 2 \\
\text{D} & \quad \rho = 2 \\
\text{V} & \quad V = 1 \\
\text{E} & \quad \rho = 4 \\
\text{V} & \quad V = 4 \\
\text{l} & \quad l = 1 \\
\text{B} & \quad \rho = 3 \\
\text{V} & \quad V = 3 \\
\text{l} & \quad l = 1 \\
\end{align*} \]

\[ \text{F} - \text{E} \]
\[ \text{tmp} = \min\{V(F), \rho(E)\} \]
\[ \text{tmp} = \min\{5, 4\} \]
\[ \text{tmp} = 4 \]
\[ \text{tmp} > V(E) ? \]
\[ 4 > 4 ? \text{No} \]
Optimum-Path Forest

\[ \rho = 5 \quad V = 5 \quad l = 2 \]

\[ \rho = 4 \quad V = 4 \quad l = 1 \]

\[ \rho = 2 \quad V = 1 \]

\[ F \quad D \]

\[ 0.5 \]

\[ 0.6 \]

\[ E \quad B \]

\[ \text{F - D} \]

\[ \text{tmp} = \min\{V(F), \rho(D)\} \]

\[ \text{tmp} = \min\{5, 2\} \]

\[ \text{tmp} = 2 \]

\[ \text{tmp} > V(D) ? \]

\[ 2 > 1 ? \text{Yes} \]
Optimum-Path Forest

\[\begin{align*}
\rho &= 5 \\
V &= 5 \\
I &= 2
\end{align*}\]

\[\begin{align*}
\rho &= 5 \\
V &= 5 \\
I &= 2
\end{align*}\]

\[\begin{align*}
\rho &= 2 \\
V &= 2 \\
I &= 2
\end{align*}\]

\[\begin{align*}
\rho &= 3 \\
V &= 3 \\
I &= 1
\end{align*}\]

\[\begin{align*}
\rho &= 4 \\
V &= 4 \\
I &= 1
\end{align*}\]

- F \rightarrow B
  - tmp = min\{V(F), \rho(B)\}
  - tmp = min\{5, 3\}
  - tmp = 3
  - tmp > V(B) ?
  - 3 > 3 ? Yes
Optimum-Path Forest

E → B → C → D
Optimum-Path Forest

E - F

tmp = min\{V(E), \rho(F)\}
tmp = min\{4, 5\}
tmp = 4

tmp > V(F) ?
4 > 5 ? No
Optimum-Path Forest

E

\( \rho = 4 \)
\( V = 4 \)
\( l = 1 \)

A

\( \rho = 5 \)
\( V = 5 \)
\( l = 1 \)

C

\( \rho = 3 \)
\( V = 3 \)
\( l = 1 \)

F

\( \rho = 5 \)
\( V = 5 \)
\( l = 2 \)

0.1
0.3
0.5

E \rightarrow C

tmp = min\{V(E), \rho(C)\}
tmp = min\{4, 3\}
tmp = 4

tmp > V(C) ?
3 > 3 ? No
Optimum-Path Forest

\[ \begin{align*}
\rho &= 4 \\
V &= 4 \\
I &= 1
\end{align*} \]

\[ \begin{align*}
\rho &= 5 \\
V &= 5 \\
I &= 2
\end{align*} \]

\[ \begin{align*}
E &= A \\
tmp &= \min\{V(E), \rho(A)\} \\
tmp &= \min\{4, 5\} \\
tmp &= 4
\end{align*} \]

\[ \begin{align*}
tmp &> V(A) ? \\
4 &> 5 ? No
\end{align*} \]
Optimum-Path Forest

B → C → D

\[
\begin{align*}
\rho &= 5 \\
V &= 5 \\
l &= 1 \\
\rho &= 3 \\
V &= 3 \\
l &= 1 \\
\rho &= 3 \\
V &= 3 \\
l &= 1 \\
\rho &= 4 \\
V &= 4 \\
l &= 1 \\
\rho &= 5 \\
V &= 5 \\
l &= 2 \\
\rho &= 2 \\
V &= 2 \\
l &= 2
\end{align*}
\]
Optimum-Path Forest

$\rho = 3$
$V = 3$
$l = 1$

$\rho = 3$
$V = 3$
$l = 1$

$0.3$

$\rho = 2$
$V = 2$
$l = 2$

$0.6$

$\rho = 5$
$V = 5$
$l = 1$

$\rho = 5$
$V = 5$
$l = 2$

B - A

tmp = min\{V(B), \rho(A)\}

tmp = min\{3, 5\}

tmp = 3

tmp > V(A) ?

3 > 5 ? No
Optimum-Path Forest

\[ B - D \]

\[
\text{tmp} = \min\{V(B), \rho(D)\}
\]

\[
\text{tmp} = \min\{3, 2\}
\]

\[
\text{tmp} = 2
\]

\[
\text{tmp} > V(D) ?
\]

\[
2 > 2 ? \text{No}
\]
Optimum-Path Forest

B

\( \rho = 3 \)
\( V = 3 \)
\( l = 1 \)

D

\( \rho = 2 \)
\( V = 2 \)
\( l = 2 \)

A

\( \rho = 5 \)
\( V = 5 \)
\( l = 1 \)

F

\( \rho = 5 \)
\( V = 5 \)
\( l = 2 \)

\[ \text{B} \rightarrow \text{F} \]
\[ \text{tmp} = \min\{V(B), \rho(F)\} \]
\[ \text{tmp} = \min\{3, 5\} \]
\[ \text{tmp} = 3 \]

\[ \text{tmp} > V(F) ? \]

\[ 3 > 5 ? \text{No} \]
Optimum-Path Forest

C - E

tmp = min\{V(C), \rho(E)\}
tmp = min\{3, 4\}
tmp = 3

tmp > V(E) ?
3 > 4 ? No
Optimum-Path Forest

\[ \rho = 3 \]
\[ V = 3 \]
\[ l = 1 \]

\[ C \]
\[ 0.7 \]

\[ B \]
\[ 0.3 \]

\[ 0.4 \]

\[ A \]
\[ \rho = 5 \]
\[ V = 5 \]
\[ l = 1 \]

\[ E \]
\[ \rho = 4 \]
\[ V = 4 \]
\[ l = 1 \]

\[ C - A \]
\[ \text{tmp} = \min\{V(C), \rho(A)\} \]
\[ \text{tmp} = \min\{3, 5\} \]
\[ \text{tmp} = 3 \]

\[ \text{tmp} > V(A) \, ? \]
\[ 3 > 5 \, ? \, \text{No} \]
Optimum-Path Forest

\[
\begin{align*}
\text{C} & \quad \rho = 3 \quad V = 3 \quad l = 1 \\
\text{B} & \quad \rho = 3 \quad V = 3 \quad l = 1 \\
\text{A} & \quad \rho = 5 \quad V = 5 \quad l = 1 \\
\text{E} & \quad \rho = 4 \quad V = 4 \quad l = 1
\end{align*}
\]

\[
\text{C - B}
\]

\[
\text{tmp} = \min\{V(C), \rho(B)\}
\]

\[
\text{tmp} = \min\{3, 3\}
\]

\[
\text{tmp} = 3
\]

\[
\text{tmp} > V(B) \ ?
\]

\[
3 > 3 \ ? \ No
\]
Optimum-Path Forest

Diagram of a network with vertices and edges labeled with weights.
Optimum-Path Forest

D - B

\[
\text{tmp} = \min\{V(D), \rho(B)\}
\]
\[
\text{tmp} = \min\{2, 3\}
\]
\[
\text{tmp} = 2
\]
\[
\text{tmp} > V(B) \, ?
\]
\[
2 > 3 \, ? \, \text{No}
\]
**Optimum-Path Forest**

Diagram showing nodes D, B, F, and E connected with edges and values:
- Node D: \( \rho = 2, V = 2, l = 2 \)
- Node B: \( \rho = 3, V = 3, l = 1 \)
- Node F: \( \rho = 5, V = 5, l = 2 \)
- Node E: \( \rho = 4, V = 4, l = 1 \)

Edges with values:
- D to F: 0.3
- D to B: 0.7
- B to E: 0.4

**Algorithm**

1. Calculate \( \text{tmp} = \min\{V(D), \rho(F)\} \)
2. Calculate \( \text{tmp} = \min\{2, 5\} \)
3. Set \( \text{tmp} = 2 \)
4. Check: \( \text{tmp} > V(F) \) ?
5. 3 > 5 ? No
Optimum-Path Forest

D - E

\[ \text{tmp} = \min\{V(D), \rho(E)\} \]
\[ \text{tmp} = \min\{2, 4\} \]
\[ \text{tmp} = 2 \]

\[ \text{tmp} > V(E) \,? \]
\[ 2 > 4 \,? \text{No} \]
Optimum-Path Forest
Learning Deep Representations

- OPF does not require the number of clusters beforehand.
Learning Deep Representations

• OPF does not require the number of clusters beforehand.

• Playing with $k_{max}$ can be prohibitive for large datasets.
Learning Deep Representations

- OPF does not require the number of clusters beforehand.
- Playing with $k_{max}$ can be prohibitive for large datasets.
- Apply multiple clustering layers.
Learning Deep Representations

- OPF does not require the number of clusters beforehand.

- Playing with $k_{max}$ can be prohibitive for large datasets.

- Apply multiple clustering layers.

- Prototypes of the first layer are the samples of the second layer, and so on.
Learning Deep Representations

• OPF does not require the number of clusters beforehand.

• Playing with $k_{max}$ can be prohibitive for large datasets.

• Apply multiple clustering layers.

• Prototypes of the first layer are the samples of the second layer, and so on.

• Prototypes are located at regions of high density.
Learning Deep Representations
Methodology and Experimental Results

• Unlabeled dataset:
  ○ 3D seismic data

• Labeled dataset:
  ○ CIFAR10
  ○ CIFAR100
  ○ MNIST
Methodology and Experimental Results

- Seismic Images:
Methodology and Experimental Results

- Seismic Images:

<table>
<thead>
<tr>
<th>Image</th>
<th>Layer 1</th>
<th>Layer 2</th>
<th>Layer 3</th>
<th>Layer 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>924</td>
<td>4,102</td>
<td>41</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>928</td>
<td>4,135</td>
<td>41</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>932</td>
<td>4,074</td>
<td>38</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>936</td>
<td>4,144</td>
<td>41</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>940</td>
<td>4,193</td>
<td>44</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>
Methodology and Experimental Results
Methodology and Experimental Results

- Original
- K-means
- OPF
- SOM
Methodology and Experimental Results

- General-purpose Images:
  - CIFAR-10: 60,000 images of size $32 \times 32$ distributed in 10 classes.
Methodology and Experimental Results

- General-purpose Images:
  - CIFAR-100: 60,000 images of size $32 \times 32$ distributed in 100 classes (finer) grouped in 20 superclasses (coarser).
Methodology and Experimental Results

- General-purpose Images:
  - MNIST: 70,000 images of handwritten digits distributed in 10 classes.
Methodology and Experimental Results

- General-purpose Images - Metrics:
  - Homogeneity (H): each cluster contains only members of a single class. $H \in [0, 1]$, where $H = 1$ denotes the best result.
  - Completeness (C): all members of a given class are assigned to the same cluster. $C \in [0, 1]$, where $C = 1$ denotes the best result.
  - V-measure (V): this metric is the harmonic mean between homogeneity and completeness, given by:
    \[
    V = 2 \frac{(H \cdot C)}{(H + C)}. \tag{1}
    \]
Methodology and Experimental Results

- General-purpose Images:

<table>
<thead>
<tr>
<th>Dataset</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIFAR 10</td>
<td>137</td>
<td>121</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>CIFAR 100</td>
<td>216</td>
<td>163</td>
<td>24</td>
<td>15</td>
</tr>
<tr>
<td>MNIST</td>
<td>221</td>
<td>145</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>
Methodology and Experimental Results

• General-purpose Images - CIFAR10:

<table>
<thead>
<tr>
<th>Metric</th>
<th>OPF</th>
<th>k-means</th>
<th>Mean-Shift</th>
<th>SOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>0.000</td>
<td>0.054</td>
<td>0.001</td>
<td>0.049</td>
</tr>
<tr>
<td>C</td>
<td>0.153</td>
<td>0.060</td>
<td>0.039</td>
<td>0.056</td>
</tr>
<tr>
<td>V</td>
<td>0.000</td>
<td>0.057</td>
<td>0.001</td>
<td>0.052</td>
</tr>
</tbody>
</table>
Methodology and Experimental Results

- General-purpose Images - CIFAR100:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Technique</th>
<th>OPF</th>
<th>(k)-means</th>
<th>Mean-Shift</th>
<th>SOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>OPF</td>
<td>0.010</td>
<td>0.033</td>
<td>0.001</td>
<td>0.030</td>
</tr>
<tr>
<td>C</td>
<td>OPF</td>
<td>0.069</td>
<td>0.038</td>
<td>0.077</td>
<td>0.034</td>
</tr>
<tr>
<td>V</td>
<td>OPF</td>
<td>0.017</td>
<td>0.035</td>
<td>0.003</td>
<td>0.032</td>
</tr>
</tbody>
</table>
Methodology and Experimental Results

- General-purpose Images - MNIST:

<table>
<thead>
<tr>
<th>Metric</th>
<th>OPF</th>
<th>(k)-means</th>
<th>Mean-Shift</th>
<th>SOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H)</td>
<td>0.000</td>
<td>0.007</td>
<td>0.000</td>
<td>0.073</td>
</tr>
<tr>
<td>(C)</td>
<td>1.000</td>
<td>0.024</td>
<td>0.005</td>
<td>0.376</td>
</tr>
<tr>
<td>(V)</td>
<td>0.000</td>
<td>0.011</td>
<td>0.001</td>
<td>0.122</td>
</tr>
</tbody>
</table>
Conclusions

- Deep-driven approach using OPF.
- OPF provided gain in resolution in seismic data.
- OPF was able to find number of clusters close to real number in 2 out of 3 datasets.
- Flexible tool for unsupervised learning.
- Techniques are complementary.
Thank you!
Q&A