**Introduction**

- **Parametric methods**
  - Achieve good performance
  - on datasets with a **moderate size** of labels
- **Nonparametric methods**
  - More efficient and **adaptive** to real-world data

**Proposed scenario**

- **Given one FCN-based parametric model**
  - Off-line trained with a **fixed set** of labels
- **Given an annotated real-world dataset**
  - with labels unknown to parametric model

## Content-adaptive parameter $\alpha(I)$

- Adaptive to different query image

\[
\alpha(I) = \frac{|C_f(I) - C_r(I)|}{|C_f(I) \cup C_r(I)|} 
\]

(4)

### Observation

- **Consistent $\psi_f$** is reliable
- **Inconsistent $\psi_f$** is unreliable

### Number of pixels estimated as new labels

\[
\alpha(I) = \frac{\#\text{pixels}(C_f(I) \cap C_r(I))}{\text{in}(m \times n)} 
\]

(5)

## Proposed scenario

### Challenge

- FCN model is unaware of unknown labels $c \in (C_r - C_f)$

### Observation

- False positive new labels

### Number of pixels estimated as new labels

\[
\alpha(I) = \frac{\#\text{pixels}(C_f(I) \cap C_r(I))}{\text{in}(m \times n)} 
\]

### Label-aware parameter $\beta(I)$

- **Inverse proportion** to # of pixels estimated as the label $c$

\[
\beta(I) = \left\{ \begin{array}{ll}
2 - \frac{\#\text{pixels}(c)}{\max_{c \in C_f} \#\text{pixels}(c)}, & \text{if } c \in C_f, \\
0, & \text{otherwise}
\end{array} \right.
\]

### Experimental Results

<table>
<thead>
<tr>
<th>Method</th>
<th>$#\text{labels}$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>Pre-pixel acc. (%)</th>
<th>Pre-class acc. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCN-He-offline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1MSpec dataset</td>
<td>$</td>
<td>C_f</td>
<td>= 232$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposed method</td>
<td>$</td>
<td>C_f</td>
<td>= 33$</td>
<td>0.5</td>
<td>55.2</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>C_f</td>
<td>= 32$</td>
<td>Eq (4)</td>
<td>65.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eq (5)</td>
<td>66.9</td>
<td>16.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eq (6)</td>
<td>65.9</td>
<td>16.4</td>
<td></td>
</tr>
</tbody>
</table>

**SIFT Flow dataset**

<table>
<thead>
<tr>
<th>Method</th>
<th>$#\text{labels}$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>Pre-pixel acc. (%)</th>
<th>Pre-class acc. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCN-He-offline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1MSpec dataset</td>
<td>$</td>
<td>C_f</td>
<td>= 232$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposed method</td>
<td>$</td>
<td>C_f</td>
<td>= 33$</td>
<td>0.5</td>
<td>55.2</td>
</tr>
<tr>
<td></td>
<td>$</td>
<td>C_f</td>
<td>= 32$</td>
<td>Eq (4)</td>
<td>65.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eq (5)</td>
<td>66.9</td>
<td>16.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eq (6)</td>
<td>65.9</td>
<td>16.4</td>
<td></td>
</tr>
</tbody>
</table>

**Conclusion**

- By jointly exploiting supervised and transferrable knowledge, the proposed method achieves good performance on both datasets with a moderate size of labels.

---

**Markov Random Field (MRF) Framework**

- **Supervised potential**
  - Off-line trained FCN

\[
\psi_{fs}(c, p) = \sum_{e \in \delta(c, p)} e^W_{fs}(e) 
\]

- **Label transfer potential**
  - Window-based label transfer
  - Fine-tuned Faster R-CNN

- **Pairwise potential term**

\[
\theta(c, p, q) = -\log(P(c | p) | c(q)) + P(c | q) | c(p)))/2 \cdot \delta(c | p) \neq c(q) 
\]

---

**Take advantage of both methods for increasing real-world data**