Machine Learning I

Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

Contents:

- This part of the course is about an application of structured learning to the task of pixel classification.

Contents:

- This part of the course is about an application of structured learning to the task of pixel classification.
- No additional definitions or algorithms are introduced in this lecture. Instead, this lecture illustrates the definitions introduced in previous lectures on structured learning.

We consider:

- A grid graph (V, A) whose nodes are called **pixels**

We consider:

- A grid graph (V, A) whose nodes are called **pixels**
- a non-empty set C called a set of **colors**

We consider:

- A grid graph (V, A) whose nodes are called **pixels**
- a non-empty set C called a set of **colors**
- a function $c\colon V \to C$ called a **digital image**.

We consider:

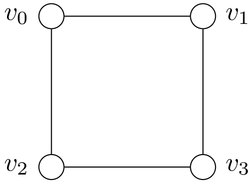
- A grid graph (V, A) whose nodes are called **pixels**
- a non-empty set C called a set of **colors**
- a function $c\colon V \to C$ called a **digital image**.

The task of **pixel classification** is concerned with making decisions at the pixels, e.g., decisions $y : P \to \{0, 1\}$ indicating whether a pixel $v \in V$ is of interest ($y_v = 1$) or not of interest ($y_v = 0$).

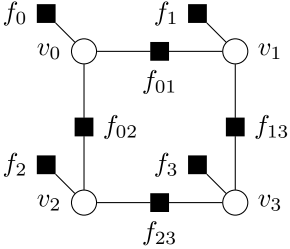
Source: `https://www.pexels.com/photo/nature-flowers-garden-plant-67857/`

For instance, we wish to map to 1 precisely those pixels of images like the one above that belong to the yellow part of any of the flowers.

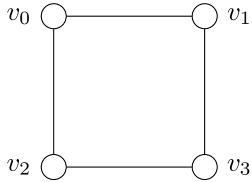
Structured Learning for Pixel Classification



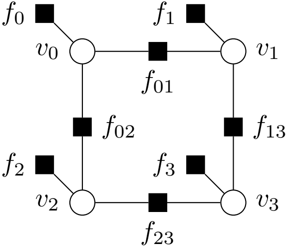
Pixel grid graph (V, A)



Factor graph (V, F, E)



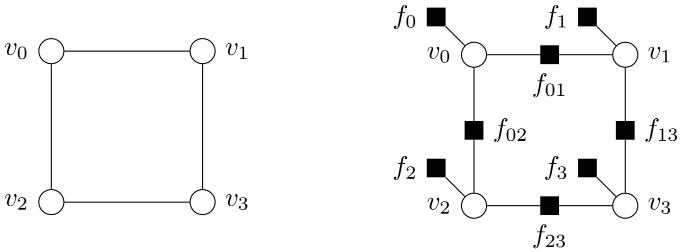
Pixel grid graph (V, A)



Factor graph (V, F, E)

For every **factor node** f_j, we consider:

- Two **attributes**: The constant $x_{f_j 0} = 1$ and the distance $x_{f_j 1}$ between the color $c(v_j)$ of the pixel v_j and pure yellow.

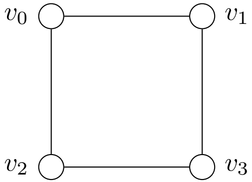


Pixel grid graph (V, A) Factor graph (V, F, E)

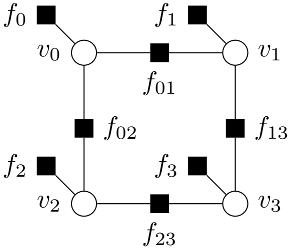
For every **factor node** f_j, we consider:

- Two **attributes**: The constant $x_{f_j 0} = 1$ and the distance $x_{f_j 1}$ between the color $c(v_j)$ of the pixel v_j and pure yellow.
- The **factor**

$$\begin{aligned} h_{f_j \theta}(x_{f_j}, y_j) &= (\theta_0 \, x_{f_j 0} + \theta_1 \, x_{f_j 1}) \, y_j \\ &= (\theta_0 + \theta_1 \, x_{f_j 1}) \, y_j \end{aligned}$$



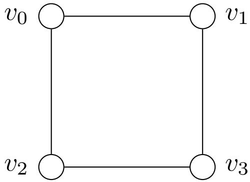
Pixel grid graph (V, A)



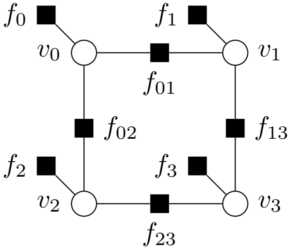
Factor graph (V, F, E)

For every **factor node** f_{jk}, we consider:

- One **attribute**: The similarity $x_{f_{jk}0} = \exp(-|c(v_j) - c(v_k)|)$ of the colors $c(v_j)$ and $c(v_k)$



Pixel grid graph (V, A)

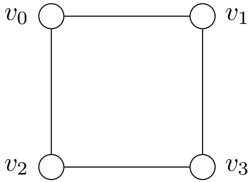


Factor graph (V, F, E)

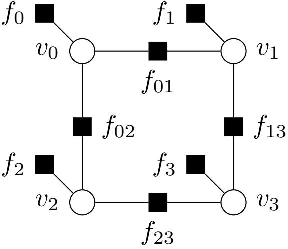
For every **factor node** f_{jk}, we consider:

- One **attribute**: The similarity $x_{f_{jk}0} = \exp(-|c(v_j) - c(v_k)|)$ of the colors $c(v_j)$ and $c(v_k)$
- The **factor**

$$h_{f_{jk}\theta}(x_{f_{jk}}, y_j, y_k) = \theta_3 \, x_{f_{jk}0} \, |y_j - y_k|$$



Pixel grid graph (V, A)

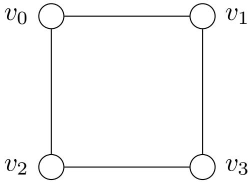


Factor graph (V, F, E)

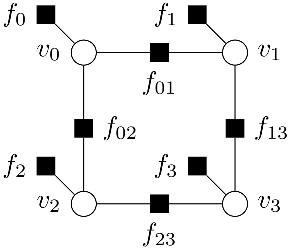
For the entire image, we obtain the **energy function**

$$H_\theta(x, y) = \sum_{v_j \in V} h_{f_j \theta}(x_{f_j}, y_j) + \sum_{\{v_j, v_k\} \in A} h_{f_{jk} \theta}(x_{f_{jk}}, y_j, y_k)$$

Structured Learning for Pixel Classification



Pixel grid graph (V, A)



Factor graph (V, F, E)

For the entire image, we obtain the **energy function**

$$
\begin{aligned}
H_\theta(x, y) &= \sum_{v_j \in V} h_{f_j \theta}(x_{f_j}, y_j) + \sum_{\{v_j, v_k\} \in A} h_{f_{jk} \theta}(x_{f_{jk}}, y_j, y_k) \\
&= \sum_{v_j \in V} (\theta_0 + \theta_1 \, x_{f_j 1}) \, y_j + \sum_{\{v_j, v_k\} \in A} \theta_3 \, x_{f_{jk} 0} \, |y_j - y_k|
\end{aligned}
$$

Summary.

- We have seen an application of structured learning with a conditional graphical model to the task of pixel classification.

Summary.

- We have seen an application of structured learning with a conditional graphical model to the task of pixel classification.
- In this application, the energy function of the conditional graphical model can express the fact that neighboring pixels are more likely to obtain the same label than distinct labels, as well as a dependency of this increased likelihood on the local contrast in the image.