

Sémantická segmentace obrazu pomocí konvolučních neuronových sítí

Diplomová práce

Bc. Filip Špila
filifspila@gmail.com

Ústav mechaniky těles, mechatroniky a biomechaniky
Vysoké učení technické v Brně

5. července 2020

- ① Nastudování problematiky segmentace obrazu pomocí konvolučních neuronových sítí
- ② Výběr perspektivní architektury sítě spolu s její implementací
- ③ Vytvoření vlastní trénovací množiny obrázků
- ④ Vytvoření segmentovaného obrazu
- ⑤ Vyhodnocení úspěšnosti segmentace

Cíle sémantické segmentace

- Přiřadit každému pixelu v obrázku právě jednu třídu objektu (auto, člověk, zvíře, ...)
- Provádět segmentaci co nejpřesněji
- Zajistit, aby algoritmus uměl generalizovat

Úloha neuronové sítě

- Aproximace obecné funkce $y = f(x; \phi)$ funkcí $y^* = f^*(x; \phi)$ řešením:

$$\phi \leftarrow \arg \min L(y, f^*(x; \phi)) \quad (1)$$

kde ϕ jsou trénovatelné parametry.

Funkce L se nazývá **hodnotící funkce**. Parametry ϕ jsou nalezeny učením sítě tzv. **učením s učitelem**.

VYSVETLIT, ZE NEURONY JSOU JEN SLOZENE FUNKCE

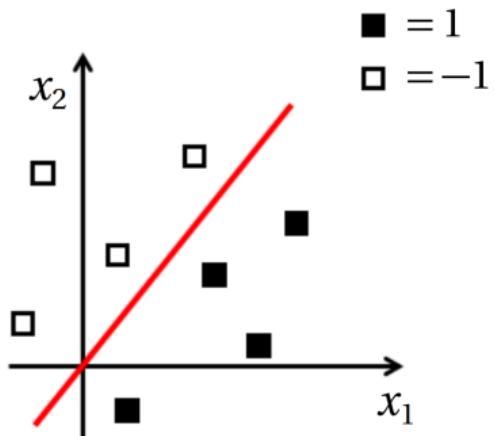
Proč pro segmentaci místo klasických metod použít neuronovou síť?

- Dovede velmi dobře approximovat silně nelineární funkce
- Potřebuje jen mírně předzpracovaná data
- Detekuje obecné vzory v datech = dovede lépe generalizovat
- Algoritmus je robustnější

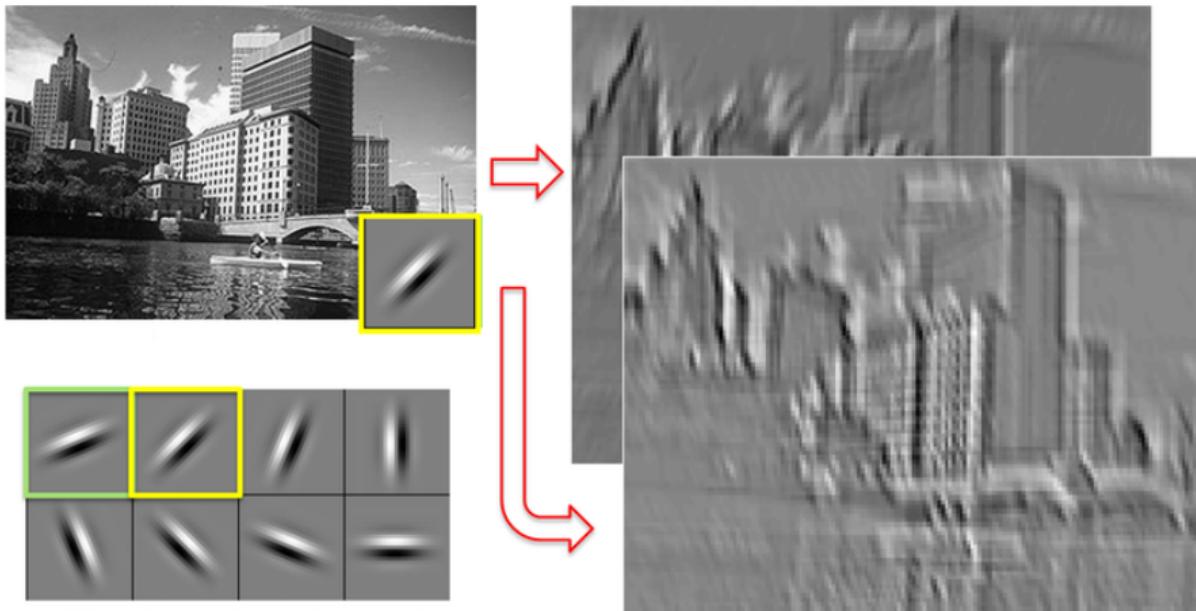
Definice klasifikační úlohy

- Aproximovaná funkce nabývá pouze určitých diskrétních hodnot

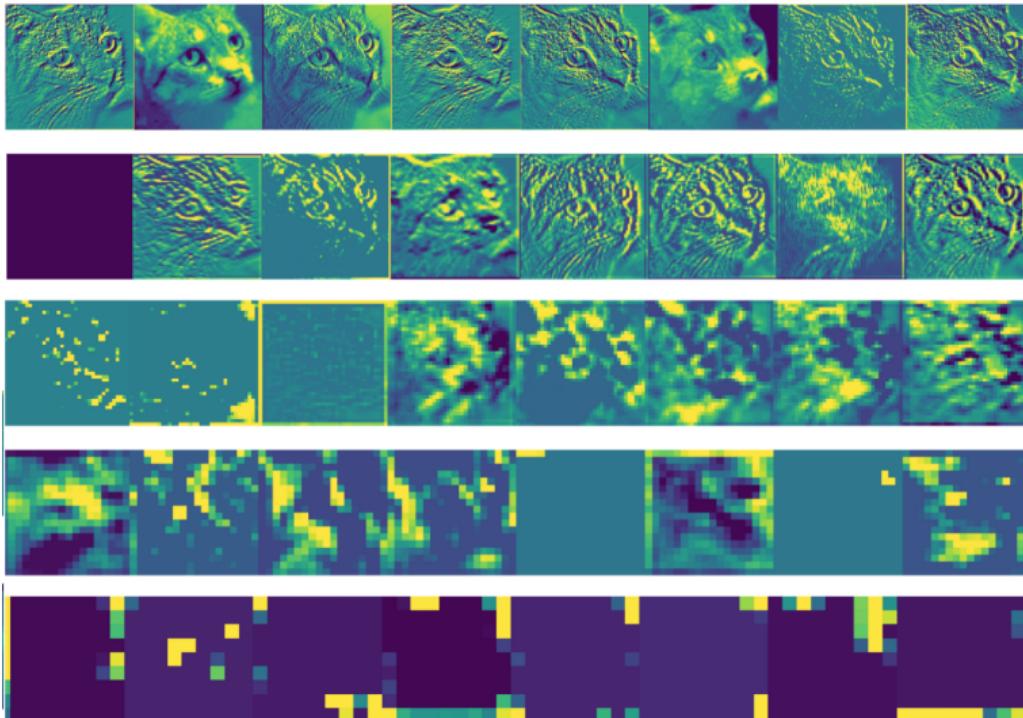
$$f(x_1, x_2; \phi) = \begin{cases} 1 \\ -1 \end{cases} \quad (2)$$



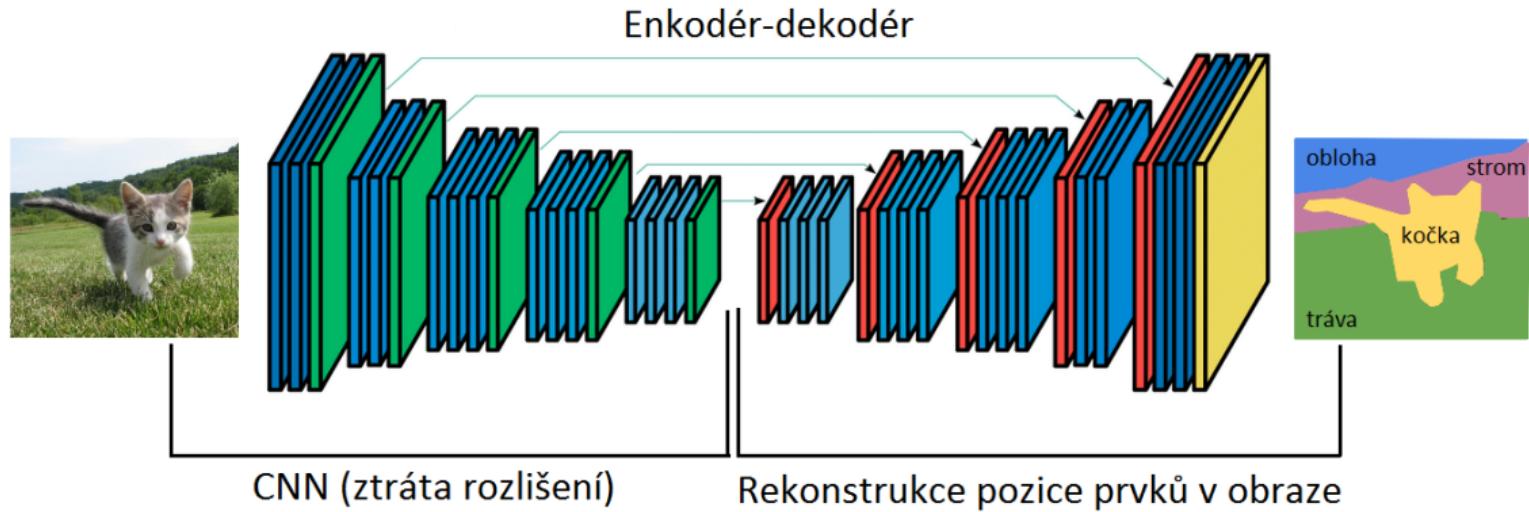
Obrázek: Úloha klasifikace



Obrázek: Mapy prvků v obrazu



Obrázek: Mapy prvků v obrazu



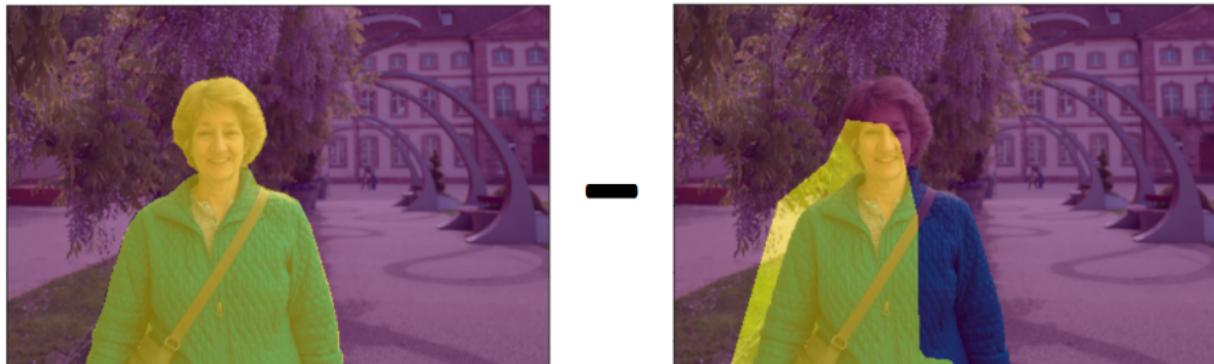
Obrázek: Architektura enkodér-dekodér

Hlavní trénovatelné parametry CNN

- Hodnoty konvolučních jader (jednotlivých filtrů)

Trénování = optimalizace parametrů CNN

- Iterativní porovnávání požadovaných a approximovaných výstupů segmentace (y, y^*)
- Minimalizace hodnotící funkce $L = f(y, y^*)$



Obrázek: Hodnocení úspěšnosti segmentace

Gradientní optimalizace hodnotící funkce - kroky

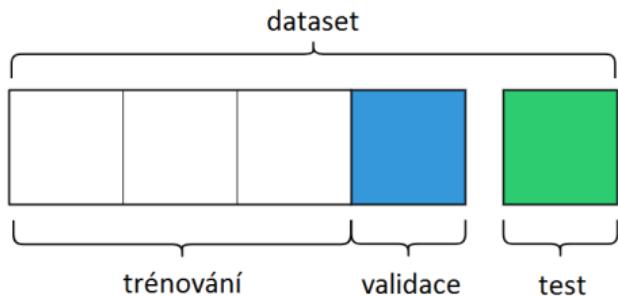
- ① Derivace hodnotící L funkce podle daného parametru ϕ : $\frac{\partial L}{\partial \phi}$
- ② Změna parametru **proti** směru derivace (hledání minima)

$$\delta\phi = -\eta \frac{\partial L}{\partial \phi} \tag{3}$$

kde η je rychlosť učení. Další detaily záleží na konkétním algoritmu.

Data pro trénování

- Cílem je zabránit přeучení sítě
 - (trénovací + validační) + testovací data
- Čím více, tím lépe (co největší definiční obor approx. funkce)



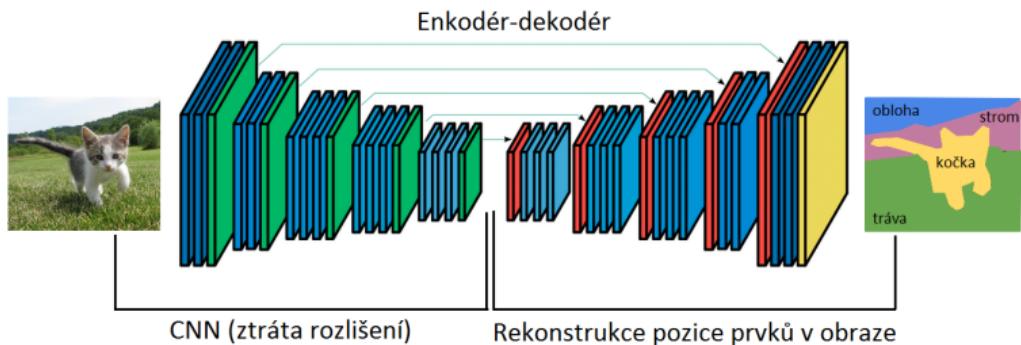
Obrázek: Hodnocení úspěšnosti segmentace

Předtrénovaná síť

- Cílem enkodéru je pouze detekovat obecné prvky v obraze
- Možnost použít již natrénovaný enkodér pro **inicializaci** parametrů
- V praxi se toto schéma používá skoro vždy!

SegNet = segmentační CNN typu enkodér-dekodér

- Vytvořeno na univerzitě v Cambridge
- Originální implementace dostupná online
- Existuje i zjednodušená verze (méně vrstev sítě)



Obrázek: SegNet

- Caffe je knihovna pro neuronové sítě
- Abstrakce vrstev CNN - snadné vytvoření architektury
- Nutné sestavit ze zdrojového kódu, nejlépe pod Linuxem

Caffe

Obrázek: Caffe logo

```
layer {
    bottom: "conv1"
    top: "conv1D"
    name: "conv1D"
    type: "Conv"
    param {
        lr_mult: 1
        decay_mult: 1
    }
    param {
        lr_mult: 2
        decay_mult: 0
    }
}
```

Tabulka: Your caption

Function name	Duration	Complexity	Length	Score
Algo 1	0.0159	0.50	125	78
Algo 2	0.0453	0.65	854	88
Algo 3	0.8642	0.77	84	95
Algo 4	0.0020	0.24	638	76

Question 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.

Answer 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.