

# Detekce specifických objektů v digitálním snímku pro potřeby určení druhu obsahu scény

Tereza Štanglová<sup>1</sup>

## 1 Úvod

Cílem této práce je navržení a implementování konvoluční neuronové sítě s použitím existující knihovny, která bude spolehlivě identifikovat pornografické scény. Po pečlivém zvážení byl z několika možných alternativ vybrán framework *CNTK*<sup>1</sup>. V rámci práce byly vytvořeny i webové stránky, které slouží k uživatelskému testování modelu. Stránky jsou umístěny na adrese <http://147.228.64.42/pdetect/>.

## 2 Konvoluční neuronové sítě

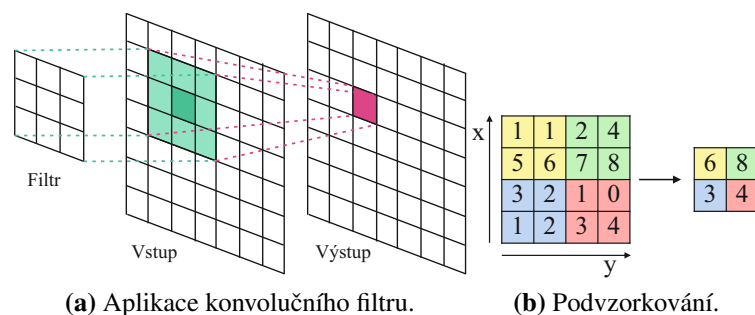
Konvoluční neuronové sítě jsou speciálním druhem vícevrstevných dopředných neuronových sítí. Byly navrženy pro rozpoznávání snímků přímo z pixelů s minimálním předzpracováním Karpathy (2016). Skládají se z vrstev, z nichž každá má svou specifickou funkci. Popis vrstev:

- **Vstupní** – Vstupem je matice obrazových bodů ve formátu *výška snímku* × *šířka snímku* × *počet barevných kanálů*.
- **Konvoluční** – Slouží k extrahování příznaků ze snímků. Obsahuje sadu konvolučních filtrů. Provádí se operace konvoluce mezi vstupem a filtry. Výstupem jsou příznakové mapy. Konvoluce je zobrazena na obrázku 1 v části (a).
- **Aktivační** – Slouží k aplikaci nelineární aktivační funkce na výstup konvoluční vrstvy.
- **Podvzorkovací** – Slouží k redukci dimenze pro snížení počtu parametrů. Příklad aplikace je ilustrován na obrázku 1 v části (b).
- **Plně propojená** – Každý neuron této vrstvy je propojený se všemi neurony vrstvy předchozí.
- **Výstupní** – Je plně propojená s předchozí. Obsahuje tolik neuronů, kolik je klasifikačních tříd.

Pro natrénování neuronové sítě byly vytvořeny dvě množiny dat. Množina cílových snímků obsahujících pornografii a množina necílových snímků, které pornografii neobsahují. Pro porovnání bylo implementováno několik architektur. Zdokumentovány byly architektury pro nedoučenou a přeúčenou síť. Dále byl zkoumán vliv různého nastavení počátečních hodnot vah a prahů sítě a vliv různé velikosti trénovací množiny na průběh učení sítě.

<sup>1</sup> studentka navazujícího studijního programu Inženýrská informatika, obor Inteligentní počítačové systémy, e-mail: [stanglov@students.zcu.cz](mailto:stanglov@students.zcu.cz)

<sup>1</sup><https://github.com/Microsoft/CNTK/>



**Obrázek 1:** Operace v konvolučních sítích.

### 3 Výsledky

V tabulce 1 je uvedeno pro představu srovnání dvou architektur s pěti a třemi konvolučními vrstvami pro různě velké trénovací sady dat. V tabulce jsou uvedeny počty cílových a necílových vzorků. Počet viděných vzorků reprezentuje velikost datové množiny po jejím umělém zvětšení.

**Tabulka 1:** Srovnání použitých architektur pro různě velké sady trénovacích dat.

Architektura	Počet celkem	Počet cílových	Počet necílových	Počet viděných	Chyba v testovací fázi [%]
Architektura 1 (5 konv. vrstev)	4052	2036	2016	121560	8.202
Architektura 2 (3 konv. vrstvy)					8.954
Architektura 1 (5 konv. vrstev)	8403	4199	4204	252090	3.085
Architektura 2 (3 konv. vrstvy)					3.010

V rámci uživatelského testování od 15.3.2017 do 20.4.2017 bylo celkem evidováno 473 nahraných vzorků. Výsledky ilustruje tabulka 2. Správně byly klasifikovány vzorky, u nichž se predikovaná třída shoduje s třídou označenou uživatelem.

**Tabulka 2:** Výsledky uživatelského testování.

		Odpověď uživatele	
		1	0
Predikce	1	20	66
	0	15	372

### 4 Závěr

V průběhu práce se podařilo experimentálně vytvořit konvoluční neuronovou síť, která uspokojivě klasifikovala digitální snímky. Model sítě byl nasazen k reálnému testování uživateli na webových stránkách. Problém při realizaci byl například v nepochopení účelu stránek uživateli. Několika uživatelům musel být účel práce individuálně vysvětlen. I přesto se v rámci uživatelského testování podařilo nasbírat dostatečné množství vzorků.

### Literatura

Karpathy, A. (2016) *Convolutional Neural Networks for Visual Recognition*. Dostupné na <http://cs231n.github.io/> [Citováno 3.4.2017]