

Počítačové vidění: od hledání vzoru po Deep Learning

Obor počítačového vidění (*computer vision*) využívá obraz a video k detekci, klasifikaci a sledování objektů nebo událostí za účelem pochopení reálné situace. Úlohy počítačového vidění lze řešit různými přístupy, od porovnávání vzoru se snímkem až po metodu *deep learning*, nabízející velkou přesnost a spolehlivost. Software Matlab, jakožto inženýrský nástroj s jazykem a interaktivním prostředím pro vědecké a technické výpočty, poskytuje množství funkcí, které lze v oboru počítačového vidění s výhodou využít.

Techniky zpracování obrazu lze rozdělit do dvou kategorií – klasické zpracování obrazu a počítačové vidění. Klasické zpracování obrazu se zaměřuje především na tradiční úlohy, jako např. odstranění šumu v obrazu a filtrace, nastavení kontrastu, práci s barevnými prostory nebo měření a transformaci obrazu. Vedle úpravy a předzpracování snímků patří k technikám klasického zpracování obrazu také detekce objektů založená na barevné informaci, morfologických operacích, prahování a tvorbě masky.

Počítačové vidění vnáší do analýzy obrazu hlubší rozměr se zaměřením na detekci, klasifikaci a sledování objektů na základě jejich charakteristických rysů. Nástroje obsažené v prostředí Matlab obsahují velmi mnoho specializovaných algoritmů použitelných k efektivnímu řešení úloh z oboru počítačového vidění.

Hledání vzorového objektu

První skupinou úloh počítačového vidění jsou úlohy spočívající v nalezení vzorového objektu na snímku nebo ve videosekvenci. K tomu se využívají algoritmy, které vyhledávají podobné prvky na snímku a v hledaném vzoru. Výsledkem může být poloha hledaného objektu ve scéně nebo transformace, která vzor do snímku scény mapuje.

K řešení úlohy nalezení vzorového objektu jsou v prostředí Matlab využívány tzv. příznaky, které popisují klíčová místa v obrazu. Vedle základních příznaků (hrany, rohy, vzory) je možné rovněž využít příznaky tzv. SURF (*Speeded-Up Robust Features*), které nejsou citlivé ke změnám měřítka, natočení i změnám jasu a kontrastu. Detekce objektu je tudíž mnohem robustnější.

Nalezené příznaky na snímku a ve vzoru jsou vzájemně porovnány. Umístění párů odpovídajících příznaků určuje polohu vzorového objektu na snímku. Hledání příznaků i jejich porovnávání probíhají zcela automaticky a tato funkce disponuje mnoha uživatelsky nastavitelnými parametry. Na obr. 1 je ukázáno hledání značky *Přechod pro chodce* ve snímáné scéně. Zelené kružnice označují příznaky SURF, žlutá úsečka ukazuje na nalezenou polohu vzoru na snímku.

Detekce objektů

Vedle již popsaného hledání vzorového objektu lze také detekovat obecnější kategorie objektů pojímajících určitou diverzitu. Vhod-

předem naučených detektorů je trénování nového detektoru pro uživatelské třídy objektů.

Sledování objektů

Sledování objektů není pouhý opakovaná detekce objektů aplikovaná na jednotlivé snímky ve videu. Při sledování se odborníci snaží nalézt souvislosti a změny scény ve snímcích jdoucích po sobě. V prostředí Matlab lze k tomuto účelu využít sledování obdélníkové oblasti snímku na základě analýzy histogramu nebo sledování skupiny bodů (obr. 3).



Obr. 1. Nalezení vzorového objektu na snímku

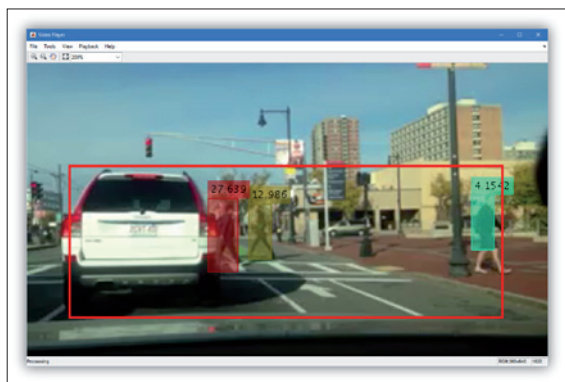
né funkční schopnosti v této oblasti poskytují detekční moduly *cascade object detector* (využívající algoritmus Viola-Jones) a *ACF object detector*. Detektory je nutné nejprve s použitím množiny vzorových dat naučit na konkrétní typ objektů a poté je lze použít na nové snímky.

V prostředí Matlab jsou k dispozici především naučené detektory pro detekci postav (obr. 2), detekci obličejů (obr. 3), nosu, očí a dalších částí lidského těla. Použití je v praxi je tudíž velmi snadné. Alternativou k využití

Vedle uvedených technik nabízí prostředí Matlab také funkce pro detekci popředí či predikci pohybu objektu pomocí Kalmanova filtru.

Klasifikace objektů

Klasifikace snímaných objektů představuje odlišný typ úlohy než hledání konkrétního objektu na snímku. Při klasifikaci je celému snímku přiřazována jedna klasifikační kategorie, která charakterizuje snímaný objekt. Úlo-



Obr. 2. Detekce chodců

hy se mohou týkat typických objektů okolního světa nebo mohou být specializovány na rozeznání vybraného objektu z dané skupiny.

Klasifikátor je třeba nejprve natrénovat na vzorových datech a poté jej použít na nový snímek. V prostředí Matlab je k dispozici klasifikátor využívající přístup *bag-of-visual words*. Přístup zahrnuje extrakci příznaků SURF, *k-means clustering* pro získání vizuálního „slovníku“ a klasifikační algoritmus SVM (*Support Vector Machine*).

Deep learning a počítačové vidění

Detekci a klasifikaci objektů na snímcích lze řešit i metodami ze skupiny deep learning. V prostředí Matlab jsou v tomto oboru využívány konvoluční neuronové sítě (*Convolutional Neural Network – CNN*), které jsou v současné době velmi populární.

Konvoluční neuronová síť je uspořádána do mnoha vrstev obsahujících sady propojených uzlů. Výpočetní algoritmus CNN je založen na konvoluci mezi vstupními daty a 2D filtry, které reprezentují hledané charakteristické rysy. Na základě těchto rysů je objekt na

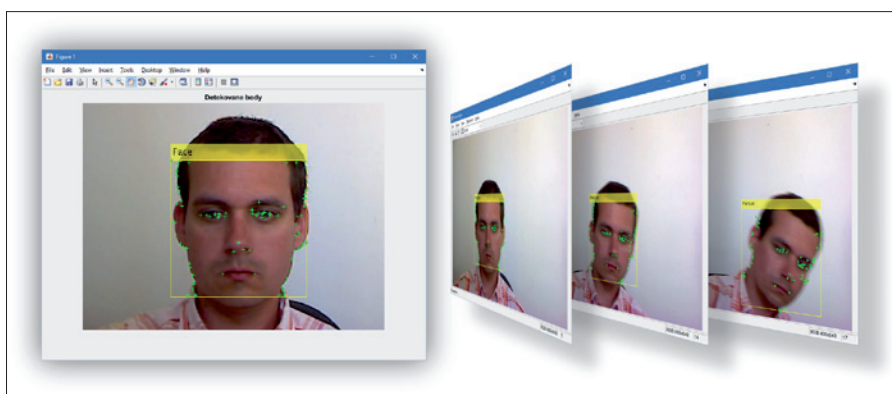
obrázku klasifikován. Úspěšnost sítí CNN při klasifikaci obrázků nezděříka přesahuje 95 %, čímž se mezi známými metodami řadí k nejspolehlivějším.

Sítě CNN lze využít nejen ke klasifikaci obrázku jako celku, ale též k detekci specifických objektů ve snímané scéně. Hledaný objekt zde obvykle představuje pouze malou část snímku, přičemž v ostatních částech snímku se nacházejí jiné objekty. Proto není možné snímek klasifikovat jako celek. V prostředí

Podrobněji je o technikách CNN a R-CNN a jejich využití pojednáno např. v [1] a [2].

Stavební kameny

Algoritmy popsané v článku nepředstavují hotové uzavřené řešení, ale stavební kameny, které může uživatel využít k budování vlastních komplexních algoritmů detekce a sledování objektů či osob. V prostředí Matlab je nabízeno mnoho příkladů, jak tyto stavební kameny skládat a řešit s jejich použitím mj. i náročné úlohy a situace, jako např. problém ztráty a opětovného nalezení sledovaného objektu, sledování mnoha objektů zároveň, stereovize atd.



Obr. 3. Detekce a sledování obličeje

Matlab se v těchto případech využívají algoritmy typu R-CNN (*Regions with CNN*), kdy je konvoluční neuronová síť použita ke klasifikaci výřezů (regionů) v daném snímku. Namísto klasifikace všech výřezů získaných pomocí „plovoucího“ okénka zpracovává detektor R-CNN pouze ty výřezy, v nichž je výskyt hledaného objektu pravděpodobný. Tím se významnou měrou redukuje požadavky na výpočetní výkon.

Literatura:

- [1] Deep Learning v prostředí Matlab. *Automa*. Děčín: Automa – ČAT, 2017, (5), 12–14. ISSN 1210-9592.
- [2] JIRKOVSKÝ, Jaroslav. Metody Deep Learning k segmentaci obrazu. *Automa*. Děčín: Automa – ČAT, 2017, (11), 37–39. ISSN 1210-9592.

Jaroslav Jirkovský,
Humusoft s. r. o.

MATLAB® & SIMULINK®

ROZBĚHNĚTE ÚSPĚŠNĚ
VÁŠ STARTUP

„Získejte za **výhodných podmínek** profesionální nástroj, který se stal celosvětovým standardem v oblasti technických výpočtů, modelování a simulací, generování kódu a certifikace.“

www.humusoft.cz/startups

Humusoft s.r.o.
Pobřežní 20, Praha 8

E-mail: info@humusoft.cz
Tel: +420 284 011 730