

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



Experimenty s OCR na báze hlbokého učenia

DIPLOMOVÁ PRÁCA

2020

Peter Kulcsár Szabó

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Experimenty s OCR na báze hlbokého učenia

DIPLOMOVÁ PRÁCA

Študijný program: Aplikovaná informatika
Študijný odbor: 11378 Aplikovaná informatika
Školiace pracovisko: Katedra aplikovanej informatiky
Vedúci práce: RNDr. Andrej Lúčny, PhD.

Bratislava 2020

Peter Kulcsár Szabó



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Peter Kulcsár Szabó
Študijný program: aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Experimenty s OCR na báze hlbokého učenia
Experiments with OCR using Deep Learning

Anotácia: Práca má výskumný charakter a je z oblasti hlbokého učenia a vychádza z jednoduchšej architektúry hlbkej neurónovej siete použitej pre OCR individuálneho písma. Skúša analógiu triku, ktorý pomáha zvýšiť úspešnosť klasického OCR na OCR pomocou hlbokého učenia.

Cieľ: Cieľom práce je vyskúšať tréning OCR na báze hlbkej neurónovej siete a vyhodnotiť či rozšírenie klasifikácie znakov o príznaky ako sú počet veľkých dier v stredovej vertikálnej línii, prítomnosť vonkajších a vnútorných oblúkov a rohov, zúženie v strede zľava a sprava, záseky v obvodovej línii zľava v dolnej polovici a sprava v dolnej polovici, spojitost' vnútra s ľavou či pravou stranou a podobne zlepši dosiahnutú úspešnosť. Pritom príznaky možno požadovať na výstupnej vrstve zo siete alebo priamo v jej latentom priestore. Cieľom práce nie je dosiahnuť čo najvyššiu úspešnosť, ale zistiť či danou úpravou architektúry úspešnosť klesá alebo stúpa.

Literatúra: Chollet, F.: Deep learning v jazyku Python, Grada, 2019
Learning OpenCV 3, Computer Vision in C++ with the OpenCV Library By Gary Bradski, Adrian Kaehler, O'Reilly Media, 2016
Smith, R.: An Overview of the Tesseract OCR Engine. ICDAR '07 Proceedings of the Ninth International Conference on Document Analysis and Recognition, Volume 02, pp. 629-633
Lúčny, A: Čítanie textu na pneumatike, KUZ 2018, M.U. Brno
learnopencv.com

Poznámka: Platforma: OpenCV, Keras

Kľúčové slová: OCR, hlboké učenie, počítačové videnie, OpenCV, Keras

Vedúci: RNDr. Andrej Lúčny, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, Dr.
Dátum zadania: 08.10.2019

Abstrakt v štátnom jazyku

KULCSÁR SZABÓ, Peter: Experimenty s OCR na báze hlbokého učenia [Diplo-
mová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky
a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: RNDr. An-
drej Lúčny, PhD., Bratislava, 2021, XX s

Lorem ipsum

Kľúčové slová: xx

Abstract

KULCSÁR SZABÓ, Peter: Experiments with OCR using Deep Learning [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: RNDr. Andrej Lúčny, PhD., Bratislava, 2021, XX p.

Lorem ipsum

Key words: xx

Obsah

Úvod	7
1 Popis problému	8

Úvod

Lorem ipsum

1 Popis problému

Lorem ipsum