

PŘEPIS OZDOBNÝCH NÁPISŮ

AUTOŘI: KRYŠTOF ANDRÝSEK A ONDŘEJ DACÍK

>>> DEFINICE ÚLOHY

V rámci této práce byl vytvořen model schopný řešit úlohu Scene Text Recognition (STR), jedná se rozpoznávání textu z oblastí detekovaných v obrázcích reálného světa (např. značky, etikety, nápisy atd.). Je to úloha náročnější než běžné OCR, neboť text může být zkreslený, nakloněný, různých stylů, barev a často nečitelný kvůli šumu nebo překážkám. Vstupem modelu je obrázek obsahující výřez s nápisem, který má být přečten. Jeho výstupem je přepis textu.

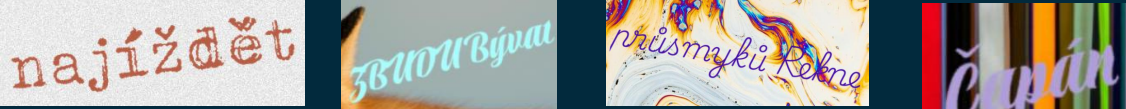
>>> DATOVÉ SADY

Kromě datových sad, na kterých se trénoval původní PARSeq, jsme využili tyto další sady:

- Sada historických nápisů - výstřižky z novin a dobové literatury



- Syntetická datová sada - sada pro učení českého jazyka a mezer

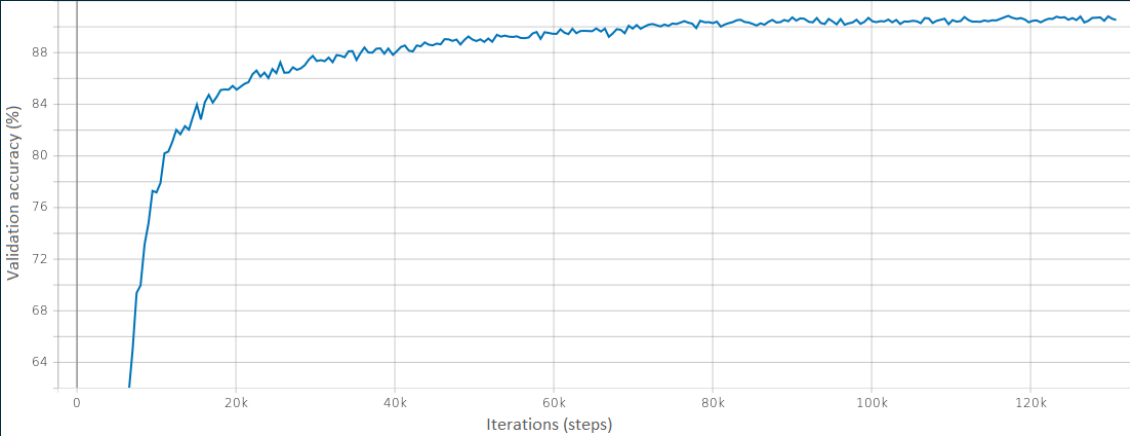


- Wordart - sada poskytnutá v rámci soutěže ICDAR24



>>> TRÉNOVÁNÍ OD NULY

Původní PARSeq neumí mezery ani český jazyk. Dotrénování modelu nepřineslo uspokojivé výsledky, a proto jsme se rozhodli model natrénovat od znova s využitím syntetické české datové sady. Vpravo lze vidět průběh trénování. Níže je tabulka srovnávající náš model s původním modelem PARSeq a jinými architekturami.



Obrázek 1: Na tomto obrázků lze vidět průběh trénování modelu od nuly. Osa x zobrazuje počet iterací, osa y validační přesnost.

>>> FINETUNING

Dalším experimentem, který jsme provedli, bylo dotrénování českého modelu na datové sadě nápisů z historických novin a časopisů. Výsledné přesnosti a srovnání s původním modelem PARSeq lze vidět v tabulce vpravo.

Model	Přesnost (%)	1-NED metrika
PARSeq	26.32	75.64
Český model	22.73	72.87
Dotrénovaný český model	58.08	87.51

Obrázek 2: Tabulka ukazující přesnost a 1-NED metriku modelů na datové sadě historických novin.

Obrázek	Skutečný text	Náš model	Parseq	ABINet	TRBA	ViTSTR-S	CRNN
	Fajntričko	Fajntričko	Fajntricko	Fayntricho	Fojntricko	Fajntricho	Fapabahe
	NašeHory	našeHory	naseHory	naseHory	nascHory	naseHory	nasellery
	Hračky Mia	Hračky Mia	HrackyMia	HrackyMia	HrackyMia	HrackyMia	HrackyMia
	LYŽE KOLA MO-TOOBLEČENÍ	LYŽE KOLA MO-TOOBLEČENÍ	LYZEKOLAINO-TOOBLECENT	LIYZEROLAMO-OOOOLEEEENN	LV2ENOLA-MOTOGBEEN	LLVIEACCAII-NTOOBLLEEEII	ATECANO-TOMACEN
	AUTO SERVIS	AUTO SERVIS	AUTOSEIVIS	AUTOSEEnS	HUTOSENIS	AUTOSENVS	MUTOSEnS
	Verbandstoffe	Verbandsteffe	Verbandsteffe	Verbandsteffe	Verbandstelle	Verbandsteffe	Verbandsleffe
	HOUSTON	HOUSTON	HOUSTON	HOUSTON	HOUSTON	MOUSTON	HOLSICH
	Kappa	Kaopa	Kappa	Kappa	Kaspa	Kappa	Kaada
	3rd Ave	3rd Ave	3rdAve	3=-Ave	3rdAve	3rdAve	Coke

Obrázek 3: V této tabulce můžete vidět přesnost našeho modelu v porovnání s modelem PARSeq a ostatními architekturami. Správné odpovědi jsou ve sloupci Skutečný text.

>>> VÝSLEDKY

Vytvořili jsme syntetickou datovou sadu vhodnou pro trénování modelů řešících úlohu přepisu těžko čitelných ozdobných a uměleckých textů. Dále se nám podařilo natrénovat model, který je na rozdíl od modelu PARSeq schopen rozpoznávat mezery a taktéž dokáže přepisovat texty v českém jazyce. Nakonec jsme provedli dotrénování našeho nového modelu na datové sadě historických textů.