

# Automatsko prepoznavanje teksta sa tablica automobila koristeći Jedan Vizuelni Model za Prepoznavanje Teksta na Sceni

**Student:** Andrija Urošević

**Mentor:** dr Nemanja Ilić

Računarski fakultet,  
Univerzitet Union

April 2024

# Predgovor

Prostor za predgovor.

# Sadržaj

<b>Predgovor</b>	<b>i</b>
<b>Sažetak</b>	<b>1</b>
<b>1 Uvod</b>	<b>2</b>
<b>2 Istorijski pregled prepoznavanja teksta</b>	<b>3</b>
2.1 Tesseract OCR . . . . .	3
2.2 CNN . . . . .	3
2.3 SVTR . . . . .	3
<b>3 Primene</b>	<b>4</b>
<b>4 Prikupljanje podataka i rad sa podacima</b>	<b>5</b>
4.1 Upoznavanje sa podacima . . . . .	5
4.2 Razvrstavanje i čišćenje podataka . . . . .	5
4.3 Pravljenje sintetičkog data seta . . . . .	5
<b>5 Implementacija</b>	<b>6</b>
5.1 Komponente sistema . . . . .	6
5.1.1 Detektor tablica . . . . .	6
5.1.2 Detektor teksta . . . . .	6
5.1.3 Prepoznavanje teksta . . . . .	6
5.2 Tehnologije . . . . .	6
5.2.1 PaddlePaddle . . . . .	6
5.2.2 FastAPI . . . . .	6
5.2.3 Docker . . . . .	6
5.3 Rezultati . . . . .	6
<b>6 Buduća poboljšanja</b>	<b>7</b>
<b>7 Zaključak</b>	<b>8</b>
<b>Literatura</b>	<b>9</b>

## Sažetak

Popularni modeli za prepoznavanje teksta sa scene su najčešće građeni od iz dva bloka: vizuelnog modela za ekstrakciju karakteristika i modela za transkripciju teksta iz ekstraktovanih karakteristika. Iako je takva hibridna arhitektura precizna ona ima dve očigledne mane, a to su kompleksnost i smanjena efikasnost. Autori rada predlažu korišćenje Jednog Vizuelnog modela za prepoznavanje teksta na sceni sa prisputom tokenizacije slike, gde se u potpunosti izostavlja sekvencijalno modelovanje. Takav pristup prepoznavanju teksta prvo rastavi tekst sa slike na deliće gde svaki delić predstavlja jednu karakter-komponentu.

Predložena metoda, nazvana SVTR, prvo dekomponuje sliku teksta u male segmente nazvane komponente karaktera. Zatim se hijerarhijske faze ponavljaju putem mešanja, spajanja i/ili kombinovanja na nivou komponenti. Globalni i lokalni blokovi mešanja osmišljeni su kako bi percipirali među-karacterske i unutar-karacterske obrasce, što dovodi do opažanja komponentata karaktera na različitim novima granularnosti. Na taj način, karakteri se prepoznaju jednostavnom linearnom predikcijom.

Prepoznavanje teksta na sceni ima za cilj da tekst sa slika konvertuje u digitalni niz karaktera, što prenosi semantiku visokog nivoa ključnu za razumevanje scene. Zadatak prepoznavanja teksta sa scene je izazovan zbog varijacija u: deformacijama teksta, fontovima, preklapanjima različitih tekstova, prekompleksnim pozadinama, itd. U

Automatsko prepoznavanje teksta sa tablica automobila je proces identifikacije i izdvajanja teksta koji se nalazi na registracionim tablicama vozila. Ovaj zadatak je važan u mnogim aplikacijama, uključujući nadzor saobraćaja, bezbednost, praćenje vozila i upravljanje saobraćajem. Korišćenjem SVTR (Jedan Vizuelni Model za Prepoznavanje Teksta na Sceni) pristupa, u poređenju sa tradicionalnim pristupom koji koristi dva odvojena modela - jedan za izdvajanje karakteristika slike i drugi za dekodiranje teksta - SVTR koristi jedan vizuelni model. Ovaj model koristi tokenizaciju slike po delovima kako bi podelio tekst sa slike na male delove (karakteristične komponente). Nakon toga, karakteristične komponente prolaze kroz seriju hijerarhijskih faza gde se vrše operacije mešanja, spajanja i kombinovanja. Konačno, koristi se jednostavna linearna predikcija za prepoznavanje karaktera. Ovaj pristup omogućava postizanje konkurentne tačnosti na zadacima prepoznavanja teksta sa tablica automobila, dok istovremeno pruža veću efikasnost u poređenju sa tradicionalnim metodama.

# 1 Uvod

Automatsko prepoznavanje teksta je ključna tehnologija u oblasti kompjuterske vizije koja ima široku primenu u različitim aplikacijama, uključujući prepoznavanje registarskih tablica vozila, prepoznavanje rukopisa, prepoznavanje dokumenata i mnoge druge. Glavni cilj automatskog prepoznavanja teksta je pretvaranje vizuelno prikazanog teksta u format koji računari mogu razumeti i obrađivati, omogućavajući im da interpretiraju tekstualne informacije slično kao što to radi čovek.

Tokom 1990-ih, široko usvajanje ličnih računara i interneta dovelo je do značajnog povećanja korišćenja tehnologije prepoznavanja teksta. Sistemi prepoznavanja teksta su se koristili za digitalizaciju knjiga, časopisa i drugih štampanih materijala, čineći lakšim pretragu i pristup informacijama. Ova tehnologija je takođe korišćena za automatizaciju procesa unosa podataka u industrijama poput finansija, zdravstva i vlade.

Početkom 2000-ih, tehnologija prepoznavanja teksta napredovala je sa uvođenjem novih algoritama i poboljšanjem hardvera. Sistemi za prepoznavanje teksta postali su precizniji i sposobni da prepoznaju širi spektar karaktera i jezika. To je otvorilo put za široko usvajanje tehnologije prepoznavanja teksta u različitim industrijama i aplikacijama, poput upravljanja dokumentima i obrade faktura. U ovom vremenskom periodu, Google je lansirao Google Books, koristeći prepoznavanje teksta za digitalizaciju desetina miliona knjiga i čineći njihov tekst pretraživim.

Tradicionalni pristupi ovom zadatku obično uključuju korišćenje više modela i složenih arhitektura, što može dovesti do povećane složenosti, troškova obrade i vremena izvršavanja. U poslednjih nekoliko godina, napredak dubokog učenja doveo je do značajnog poboljšanja performansi sistema za prepoznavanje teksta.

## 2 Istorijski pregled prepoznavanja teksta

<https://people.ischool.berkeley.edu/~buckland/statistical.html>

<https://www.historyofinformation.com/detail.php?id=684>

### 2.1 Tesseract OCR

### 2.2 CNN

RNN

### 2.3 SVTR

Tradicionalni modeli za prepoznavanje teksta obično uključuju dve odvojene komponente: vizuelni model za izdvajanje karakteristika sa slike i sekvencijalni model za dekodiranje izdvojenih karakteristika u tekst. SVTR eliminiše potrebu za sekvencijalnim modelom u potpunosti, čineći ga jednostavnijim i efikasnijim.

Uklanjanjem komponente sekvencijalnog modeliranja, SVTR postiže konkurentnu preciznost na zadacima prepoznavanja teksta, pružajući veću efikasnost u poređenju sa tradicionalnim metodama.

### **3    Primene**

Slanje upozorenja za parking

## 4 Prikupljanje podataka i rad sa podacima

### 4.1 Upoznavanje sa podacima

### 4.2 Razvrstavanje i čišćenje podataka

### 4.3 Pravljenje sintetičkog data seta



## **5 Implementacija**

### **5.1 Komponente sistema**

#### **5.1.1 Detektor tablica**

#### **5.1.2 Detektor teksta**

#### **5.1.3 Prepoznavanje teksta**

### **5.2 Tehnologije**

#### **5.2.1 PaddlePaddle**

#### **5.2.2 FastAPI**

#### **5.2.3 Docker**

### **5.3 Rezultati**

## 6 Buduća poboljšanja

## 7 Zaključak

## Literatura