

**Automatsko prepoznavanje teksta sa  
tablica automobila koristeći  
Jedinstveni Vizuelni Model za  
Prepoznavanje Teksta na Sceni**

**Student:** Andrija Urošević

**Mentor:** dr Nemanja Ilić

Računarski fakultet,  
Univerzitet Union

Maj 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 Prepoznavanje teksta</b>	<b>3</b>
2.1 Uvod u prepoznavanje teksta . . . . .	3
2.2 Istorijski pregled prepoznavanja teksta . . . . .	4
2.3 Korišćenje Jednog Vizuelnog Modela za Prepoznavanje Teksta na Sceni . . . . .	4
<b>3 Primene</b>	<b>6</b>
<b>4 Prikupljanje i rad sa podacima</b>	<b>7</b>
4.1 Upoznavanje sa podacima . . . . .	7
4.2 Razvrstavanje i čišćenje podataka . . . . .	7
4.3 Pravljenje sintetičkog data seta . . . . .	7
4.3.1 Generisanje pozadina tablica . . . . .	7
4.3.2 Generisanje teskta na tablicama . . . . .	7
<b>5 Implementacija</b>	<b>8</b>
5.1 Treniranje modela prepoznavanja teksta . . . . .	8
5.2 Komponente sistema . . . . .	8
5.2.1 Detektor tablica . . . . .	8
5.2.2 Detektor teksta . . . . .	8
5.2.3 Prepoznavanje teksta . . . . .	8
5.3 Tehnologije . . . . .	8
5.3.1 PaddlePaddle . . . . .	8
5.3.2 Docker . . . . .	8
5.3.3 FastAPI . . . . .	8
<b>6 Rezultati</b>	<b>9</b>
<b>7 Buduća poboljšanja</b>	<b>10</b>
<b>8 Zaključak</b>	<b>11</b>
<b>Literatura</b>	<b>12</b>

## Sažetak

Automatsko prepoznavanje teksta sa registarskih tablica automobila od izuzetne je važnosti za savremene sisteme nadzora saobraćaja, praćenje vozila i obezbeđenje sigurnosti na putevima. Identifikacija registarskih tablica ima različite primene, uključujući praćenje ukradenih vozila, naplatu putarine, sigurnosne provere i nadzor saobraćaja. U proteklim decenijama, napredak u oblasti obrade slike i mašinskog učenja omogućio je razvoj efikasnih sistema za automatsko prepoznavanje teksta sa tablica vozila. Primena dubokih neuronskih mreža i algoritama dubokog učenja omogućila je visoku tačnost prepoznavanja teksta, čak i u složenim scenama i različitim uslovima snimanja. Ovaj rad istražuje metode i tehnike za automatsko prepoznavanje teksta sa tablica automobila, sa ciljem razvoja sistema koji može precizno identifikovati registarske tablice u realnom vremenu. Eksperimentalni rezultati prikazuju performanse sistema u stvarnim uslovima i ukazuju na mogućnosti za primenu u različitim oblastima, uključujući nadzor saobraćaja, bezbednosne provere i identifikaciju vozila.

# 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.

Automatsko prepoznavanje teksta sa registarskih tablica automobila obuhvata nekoliko ključnih koraka koji se odvijaju u procesu od prikupljanja podataka do konačne integracije sistema u softver za prepoznavanje tablica automobila.

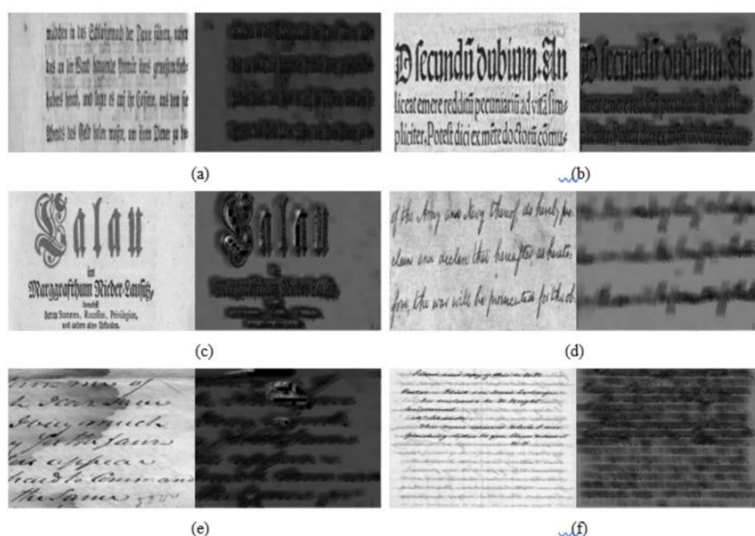
Prikupljanje raznovrsnog skupa slika registarskih tablica vozila ključno je za uspešno treniranje modela. Ove slike treba da obuhvataju različite tipove tablica, različite uslove osvetljenja i pozadine kako bi model bio što robustniji. Nakon prikupljanja, slike treba pažljivo razvrstati na one koje su pogodne za treniranje modela i one koje nisu. Ovo uključuje filtriranje slika sa veoma lošim kvalitetom, zamućenim ili nejasnim tablicama. Kako bi se obogatilo skup podataka i poboljšala generalizacija modela, potrebno je primeniti tehnike augmentacije podataka. Ovo uključuje manipulaciju sa slikama kao što su rotacija, promena osvetljenja, izobličenja i dodavanje šuma. Pored toga, sintetički podaci se mogu generisati korišćenjem programa za generisanje tablica sa tekstom. Svaka slika mora biti precizno označena sa tačnim tekstualnim sadržajem registarske tablice kako bi se koristila za obuku modela. Ovaj proces može biti ručan ili se može koristiti alat za automatsko lejbelovanje. Nakon pripreme podataka, sledi faza treniranja mreže za prepoznavanje teksta. U ovoj fazi, model se obučava nad označenim podacima kako bi naučio da prepozna tekst sa slika tablica. Kada je model obučen, integriše se u softver za prepoznavanje tablica automobila. Ovaj softver obično obuhvata module za detekciju tablica, detekciju teksta na tablicama, formatiranje izlaza i druge funkcionalnosti.

Kako bi omogućili portabilnost i lakšu distribuciju sistema za prepoznavanje tablica, koristi se Docker kontejner. Docker omogućava pakovanje softverskih aplikacija i njihovo pokretanje u izolovanim okruženjima. Još jedna od bitnih komponenata je Python web framework - FastAPI koji omogućava brzo kreiranje API-ja za komunikaciju sa softverskim komponentama. Integracija Docker-a i FastAPI modula omogućava da se servis za prepoznavanje teksta koristi nezavisno od platforme na kojoj se izvršava, čineći ga pristupačnim i jednostavnim za upotrebu u različitim okruženjima.

## 2 Prepoznavanje teksta

### 2.1 Uvod u prepoznavanje teksta

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. Dodatno, otežavajući faktor može biti i to što se tekst može pojaviti pod različitim uglovima.



Slika 1: Tekst na sceni sa različitim fontovima, pozadinama, osvetljenjem i slično

U proteklim godinama uloženi su brojni naponi kako bi se poboljšala tačnost prepoznavanja teksta. Moderni pristupi za prepoznavanje teksta, pored tačnosti, takođe uzimaju u obzir i faktore poput brzine izvršavanja modela zbog praktičnih zahteva.

Metodološki, prepoznavanje teksta na sceni može se posmatrati kao prelazak iz modaliteta slike u niz karaktera. Obično, prepoznavanje teksta se sastoji od dva osnovna dela, vizuelnog modela za ekstrakciju karakteristika i sekvencijskog modela za transkripciju teksta.

## 2.2 Istorijski pregled prepoznavanja teksta

Prvi primeri i prva faza tehnologije optičkog prepoznavanja karaktera (OCR) pojavili su se sredinom 20. veka, pretežno tokom 1950-ih i 1960-ih godina. Ovo doba obeležilo je razvoj ranih sistema OCR-a, koji su koristili osnovne tehnike prepoznavanja obrazaca kako bi prepoznali mašinski odštampane karaktere. Ovi sistemi su često bili ograničeni na prepoznavanje određenih fontova i imali su relativno nisku stopu tačnosti u poređenju sa modernom OCR tehnologijom. Glavna primena im je bila čitanje standardizovanih obrazaca i dokumenata sa jasno štampanim tekstom i poznatim fontom.

Druga faza tehnologije optičkog prepoznavanja karaktera dogodila se krajem 20. veka i početkom 21. veka, počevši oko 1970-ih i nastavljajući se u 2000-ima. Ovo doba je obeleženo značajnim napretkom u tehnologiji OCR-a, uključujući razvoj sofisticiranih algoritama i tehnika za prepoznavanje karaktera. Ovi napredci doveli su do veće tačnosti i mogućnosti prepoznavanja šireg spektra fontova, jezika i rasporeda dokumenata. Dodatno, integracija pristupa mašinskog učenja i neuronskih mreža doprinela je daljem poboljšanju performansi OCR-a. U ovoj fazi došlo je do primene OCR sistema u širem spektru aplikacija, od skeniranja i konverzije dokumenata u digitalno arhiviranje, do automatizovanog unosa podataka i ekstrakcije teksta u različitim industrijama.

Treća faza tehnologije optičkog prepoznavanja karaktera je trenutno aktuelna i predstavlja trenutno stanje napretka u sistemima OCR-a. Ovo doba karakteriše integracija najnovijih tehnologija poput dubokog učenja, konvolucionih neuronskih mreža (CNN) i rekurentnih neuronskih mreža (RNN) u algoritme OCR-a. Ove napredne tehnike značajno su poboljšale tačnost i pouzdanost sistema OCR-a, omogućavajući prepoznavanje složenih dokumenata sa različitim fontovima, rasporedima i jezicima. Osim toga, pojava OCR usluga zasnovanih na cloud-u i integracija OCR funkcionalnosti u mobilne uređaje učinili su OCR dostupnijim i svestranijim nego ikad ranije. Treća faza takođe obuhvata napretke u analizi i razumevanju dokumenata, omogućavajući OCR sistemima da izvlače ne samo tekst već i strukturalne i semantičke informacije iz dokumenata, što dovodi do poboljšanih sposobnosti obrade dokumenata i pretraživanja informacija.

## 2.3 Korišćenje Jednog Vizuelnog Modela za Prepoznavanje Teksta na Sceni

Tradicionalni modeli za prepoznavanje teksta obično uključuju dve odvojene komponente: vizuelni model za izdvajanje karakteristika sa slike i sekven-

cijalni model za dekodiranje izdvojenih karakteristika u tekst. Jedinstveni vizuelni model za prepoznavanje teksta na sceni(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. [Du+22] itd...



### 3 Primene

## 4 Prikupljanje i rad sa podacima

### 4.1 Upoznavanje sa podacima

### 4.2 Razvrstavanje i čišćenje podataka

### 4.3 Pravljenje sintetičkog data seta

#### 4.3.1 Generisanje pozadina tablica

#### 4.3.2 Generisanje teskta na tablicama

## **5 Implementacija**

### **5.1 Treniranje modela prepoznavanja teksta**

### **5.2 Komponente sistema**

#### **5.2.1 Detektor tablica**

#### **5.2.2 Detektor teksta**

#### **5.2.3 Prepoznavanje teksta**

### **5.3 Tehnologije**

#### **5.3.1 PaddlePaddle**

#### **5.3.2 Docker**

#### **5.3.3 FastAPI**

## 6 Rezultati

## 7 Buduća poboljšanja

## 8 Zaključak

## Literatura

- [Du+22] Yongkun Du i dr. „Svtr: Scene text recognition with a single visual model”. U: *arXiv preprint arXiv:2205.00159* (2022.).