

**УНИВЕРЗИТЕТ “СВ. КИРИЛ И МЕТОДИЈ” – СКОПЈЕ**  
**ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ НАУКИ И**  
**КОМПЈУТЕРСКО ИНЖЕНЕРСТВО**



**-СЕМИНАРСКА РАБОТА-**  
**Дигитално Процесирање на Слика**

Тема

**Детекција на регистарски таблички и препознавање на  
карактери**

Изработил:

Атанас Витанов (221128)

Ментор:

проф. д-р Ивица Димитровски

## Содржина

Апстракт .....	3
Вовед .....	4
Историски преглед.....	5
Архитектура и работа на ANRP(Automatic Number-Plate Recognition) системот.....	6
Основни компоненти на ANPR систем.....	6
Принцип на работа (чекор-по-чекор).....	6
Имплементирани алгоритми .....	9
YOLO модел за детекција на објекти.....	9
Зошто YOLO е погоден за детекција на регистарски таблички? .....	9
Како функционира YOLO (скратен преглед).....	9
Обработка на слика со OpenCV .....	10
Процес на обработка: .....	11
Tesseract OCR за препознавање текст .....	12
Карактеристики и функционалности .....	12
Предности, недостатоци и подобрување .....	14
Случаи кога системот работи добро.....	14
Непрепорачлива употреба .....	15
Сценарија каде детекцијата може да биде проблематична.....	15
Лимитации на YOLO и Tesseract .....	15
Како може да се подобрат резултатите? .....	16
Заклучок .....	17

## Апстракт

Овој проект претставува сеопфатна студија за автоматска детекција на регистарски таблички и препознавање на карактери, применувајќи современи методологии и техники од областа на машинското учење и компјутерска визија. Главната цел е развој на систем кој е способен прецизно да ги лоцира табличките во статични слики, а потоа да ги препознае и извлече карактерите со висока точност. Низ документацијата е прикажана историската позадина на оваа проблематика, почнувајќи од првите обиди за автоматско препознавање на таблички, па сè до денешните комплексни модели со длабоко учење. Дополнително, детално е опишан процесот на интеграција на YOLO моделот за детекција на објекти и Tesseract OCR за препознавање на карактери, вклучувајќи ги алгоритмите, архитектурните избори и параметрите кои влијаат врз перформансите. Се анализираат условите под кои ваков систем дава висока точност (на пример, добро осветлување и остри слики) како и сценаријата каде е помалку ефикасен (при затемнети, заматени или комплексни позадини). Документацијата, исто така, се осврнува на предизвици поврзани со брзината на извршување, мемориските барања и потенцијалните грешки при препознавање на специфични типови карактери или различни стилови на таблички. Покрај тоа, се издвојуваат препораки за подобрување на точноста и робушноста на системот, како и идни насоки за проширување на функционалностите, како интеграција со системи за интелигентен транспорт или автоматска наплата на патарини.

Клучни зборови: *детекција на регистарски таблички, препознавање на карактери, машинско учење, YOLO, Tesseract OCR, компјутерски вид, обработка на слики, длабоко учење, објектна детекција.*

## Вовед

Детекцијата на регистарски таблички и препознавањето на карактери претставуваат клучни процеси во современите системи за автоматизација и безбедност. Благодарение на технологиите за компјутерски вид и обработка на слики, овие процеси добиваат сè поголемо значење во различни сектори, како што се транспортната индустрија и јавната безбедност. Ваквите системи овозможуваат ефикасно следење на сообраќајот, автоматска наплата на патарини, електронски системи за паркирање и многу други применливи сценарија кои бараат брза и доверлива идентификација на возилата.

Во рамките на овој проект, претставуваме решение кое се заснова на напредни методи за машинско учење и компјутерска визија, овозможувајќи прецизна детекција и препознавање на карактерите од регистарските таблички во статични слики. Системот е дизајниран да биде лесно интегриран во постоечки апликации или инфраструктури, со цел да ја зголеми нивната ефикасност и сигурност. Тој е изграден врз основа на моќни модели за детекција на објекти и технологии за оптичко препознавање на карактери (OCR), кои заедно овозможуваат брза и точна обработка на табличките дури и при варијации во големина, агол на гледање или степен на осветлување.

Покрај основните функционалности, системот вклучува и механизми за справување со предизвици, како што се различните формати на таблички, оштетувања на површината, рефлексии или неповолни временски услови. Сепак, колку и да е технолошки напреден, системот има свои ограничувања, особено кога се работи за многу комплексни позадини, екстремно ниска резолуција или целосно покриени таблички.

Преку оваа документација, детално ќе ги разгледаме опфатените алгоритми, процесот на нивна имплементација и оптимизација, како и достапните конфигурациски параметри кои можат да влијаат врз конечните резултати. Дополнително, ќе бидат прикажани реални примери и студии на случај кои ги илустрираат предностите и ограничувањата на системот, како и препораки за идни подобрувања и проширувања на неговите функционалности.

## Историски преглед

Историјата на автоматската детекција на регистарски таблички датира од првите обиди за обработка на слики и компјутерска визија во 1970-тите години. Во овој период, истражувањата во полето на дигиталната обработка на слики не било доволно развиено, а достапната компјутерска моќност била значително ограничена. Првите системи за препознавање на регистарски таблички биле базирани на едноставни алгоритми за прагови(thresholding), бинирање и морфолошки операции, кои функционирале само во строго контролирани услови: идеално осветлување, минимално движење на возилото и еднолична позадина. Ваквите предизвици го ограничуваат успехот на ранливите системи и честопати бараат мануелно подесување на параметрите за да се постигне подобра точност.

Сепак, со текот на времето и паралелно со развојот на хардверот, алгоритмите станале сè посоефицирани. Во 1980-тите и 1990-тите години, се појавиле комерцијални системи кои се потpirале врз подобрени техники за препознавање, комбинирајќи го напредокот на дигиталната обработка на сигнали со фундаментални статистички методи, како на пример Хафова трансформација за детекција на линии и форми, или шаблони базирани на корелација. Овие системи започнале да наоѓаат примена во специфични домени, како автоматска наплата на патарини и контрола на влезно-излезни пунктови.

Во почетокот на 2000-тите, се забележал значаен напредок со воведувањето на методи за машинско учење, како поддржувачки векторски машини (SVM) и невронски мрежи со ограничена длабочина, кои овозможиле подобро препознавање на карактери и поголема робустност при варијации во осветлувањето и аголот на снимање. Сепак, ефективността на овие алгоритми била сè уште ограничена, бидејќи барала мануелна екстракција на карактеристики (features), а перформансите зависеле од квалитетот и разновидноста на тренинг-податоците.

Голем напредок се случил со развојот на длабокото учење и појавата на конволуциски невронски мрежи (CNN), кои драстично ја сменија динамиката во компјутерскиот вид. Еден од позначајните претставници на овој тренд е алгоритмот YOLO (You Only Look Once), кој овозможува реално-временска детекција на објекти со висока точност и брзина. За разлика од традиционалните методи, YOLO ги третира процесите на идентификација и класификација на објектите како единствена задача, со што значително се намалува комплексноста и времето за извршување.

Денес, автоматската детекција на регистарски таблички е веќе воспоставена технологија која има широко поле на примена: од безбедносни системи (камери за надзор на сообраќајот и објектите), до паметни паркинзи и системи за брза наплата на патарини. Воедно, напредните системи ги користат придобивките од големите сетови на податоци и моќните графички единици (GPUs) за постојано подобрување на точноста на препознавањето и робустноста при варијации во условите на снимање.



# Архитектура и работа на ANRP(Automatic Number-Plate Recognition) системот

## Основни компоненти на ANPR систем

### Камера/Сензор

Камерата обично е поставена на локација каде што е возможно јасно да се видат предните или задните регистарски таблички на возилата (на пример, на рамка над патот, до наплата рампa, на полициско возило и сл.). Во некои случаи се користат инфрацрвени (IR) или специјални камери за да се подобри видливоста во услови на слаба осветленост или ноќе.

### Модул за детекција на табличка

Овој дел од системот ги анализира сликите (или видео-стримот) и ја лоцира точната позиција на регистарската табличка. Современи решенија често користат YOLO (You Only Look Once) или други модели за длабоко учење кои можат со висока брзина и точност да ги издвојат регионите каде што има таблички.

### Модул за препознавање на карактери (OCR)

Откако ќе се издвои табличката од сликата, потребно е да се препознаат индивидуалните букви и броеви. Најчесто се користат модели за оптичко препознавање на карактери (OCR), како Tesseract, или други невронски мрежи кои се специјализирани за препознавање на текст.

### База на податоци и централен сервер

Системот за ANPR обично е поврзан со база на податоци каде што се чуваат прочитаните таблички, заедно со соодветни информации: време, локација, слика од возилото, итн. Оваа база може да биде дел од системот за наплата на патарини, полициска база, паркинг систем, или друг централен сервер кој врши валидација и обработка на добиените податоци.

### Интерфејс/Апликација

Резултатите (препознаените таблички) најчесто се прикажуваат во кориснички интерфејс: тоа може да биде монитор во контролна соба, екран во полициско возило или софтвер за управување со паркинзи. Системот може и автоматски да реагира, на пример, да отвори рампа за патарина, да испрати сигнал за аларм ако е детектирано возило со истечени документи, и сл.

## Принцип на работа (чекор-по-чекор)

### Земање на слика (Image Acquisition)

Камерата е позиционирана така што возилото минува низ нејзиното видно поле. Во системите за патарина, камерата е фиксно поставена и секое возило што поминува се фотографира. Во полициските возила, камерата е мобилна и може да скенира возила околу себе. По потреба, користат блиц или IR-осветлување за да се обезбеди добра видливост ноќе.

### Предобработка на слика

Сликите се праќаат до компјутерски модул кој врши предобработка: намалување шум, корекција на контраст/осветленост, отстранување на позадина или засилување на контурите на табличката. Целта

е да се нагласи регион кој потенцијално би можел да биде табличка, олеснувајќи го процесот на детекција.

### **Детекција на табличка**

Користејќи алгоритми за детекција (YOLO или друг CNN модел), системот ја пронаоѓа локацијата на табличката во рамките на сликата.

Обично се добиваат координати на „bounding box“ – правоаголна област која ги опфаќа табличките. Системите за патарина често имаат дополнителни механички или софтверски бариери, со цел возилото да биде на оптимално растојание и агол за снимање.

### **Сегментација и препознавање (OCR)**

Откако ќе се знае регионот на табличката, системот го исекува тој регион и го подготвува за OCR. OCR алгоритмот ги распознава буквите и броевите врз основа на претходно истренирани модели за различни јазици или стилови на букви. Важно е коректно да се препознаат карактерите, и системите често користат валидација врз база на формати на таблички, нумерички правила и сл.

### **Проверка и складирање на податоците**

Прочитаната табличка, заедно со информациите (време, локација, можеби фотографија од возилото) се праќа до централниот сервер или локална база во полициското возило.

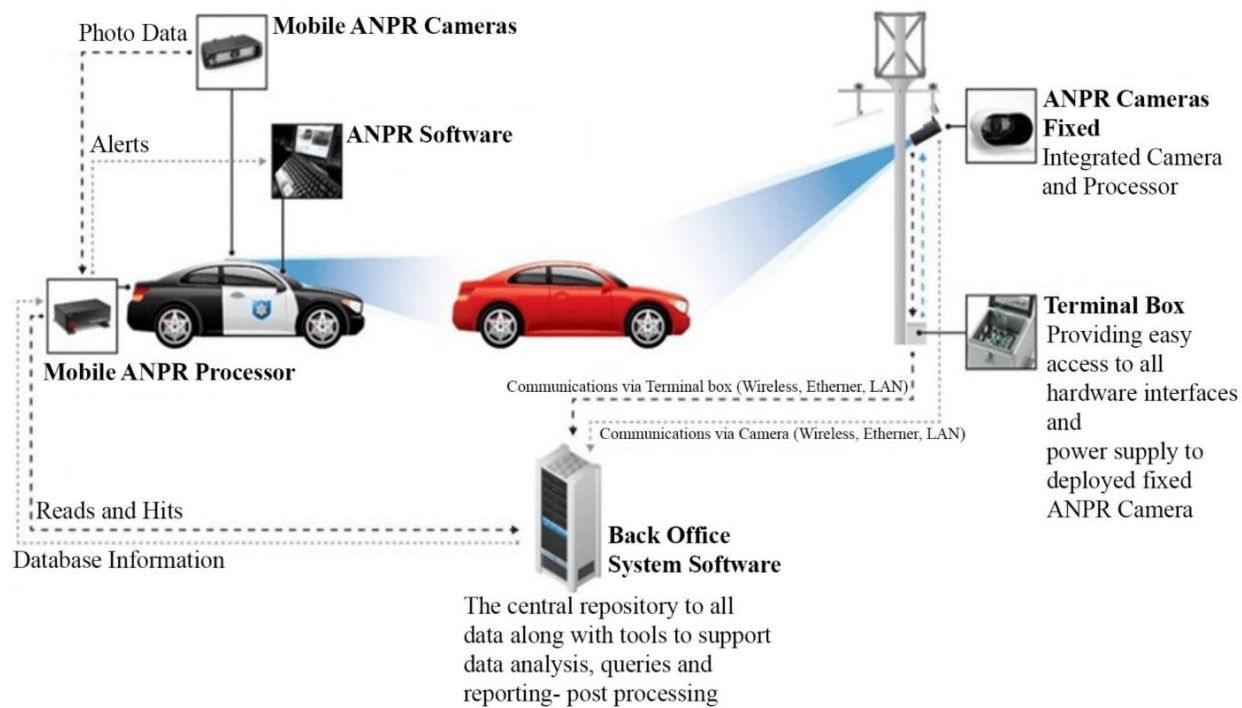
Системот прави проверка во одредена база:

- Дали регистарското возило е осигурано?
- Дали има неплатени казни или патарини?
- Дали возилото е пријавено за кражба?
- Ако има проблем, системот може да генерира аларм, да испрати нотификација до оператор или да преземе автоматска акција (на пример, да не ја подигне рампата на патарина).

### **Реакција/Известување**

Во случај на патарина, системот може да направи автоматско процесирање на плаќањето (доколку возилото е регистрирано во електронски систем).

Во полициски контекст, системот може да испрати предупредување ако детектира несоодветна состојба на табличката или посебен статус на возилото.



Слика: Архитектура на ANRP (Automatic Number-Plate Recognition) систем



## Имплементирани алгоритми

Во овој дел се опишуваат главните алгоритми и методи кои се користат за да се изгради системот за детекција на регистарски таблички и препознавање на карактери. Комбинацијата на YOLO моделот за детекција, OpenCV за обработка на слики и Tesseract OCR за препознавање на текст ја претставува основата на проектот, обезбедувајќи висока точност и робустност.

### YOLO модел за детекција на објекти

YOLO (You Only Look Once) е еден од најпопуларните и најефикасни алгоритми за детекција на објекти во реално време. Главната предност на YOLO е тоа што третира целата слика како една целина и прави само едно поминување (forward pass) низ мрежата за да ги детектира сите објекти, наместо да разгледува различни региони или прозорци од сликата.

### Зошто YOLO е погоден за детекција на регистарски таблички?

- ✓ **Реално-временска изведба:** Во ситуации како патарини, полициски пунктови или системи за надзор на сообраќај, брзината е клучна. YOLO нуди детекција во реално време, дури и на релативно послаб хардвер.
- ✓ **Висока прецизност:** Современите варијанти на YOLO постигнуваат големи стапки на точност (mAP – mean Average Precision), што е од суштинско значење за доверлив систем кој скенира и евидентира сообраќај.
- ✓ **Едноставна интеграција со OCR:** Откако ќе се детектира табличката, лесно е да се изврши исекување (crop) и да се проследи таа подслика кон алатка за OCR (како Tesseract). Ова значи дека се заштедуваат ресурси, бидејќи не се анализира целата слика карактер по карактер, туку само малиот исечок.

### Како функционира YOLO (скратен преглед)

#### Делење на сликата на мрежа (grid):

Сликата се дели на  $N \times N$  клетки (пример  $13 \times 13$ ,  $19 \times 19$  или друго, зависно од верзијата на YOLO и влезната резолуција).

#### Предвидување на bounding boxes:

Секоја клетка „предвидува“ неколку ограничувачки рамки (bounding boxes), заедно со вредности за доверба (confidence) и класификација (која категорија објект е во прашање).

#### Non-Maximum Suppression (NMS):

Кога има повеќе рамки кои се преклопуваат за истиот објект, алгоритмот NMS го задржува само најрелевантниот (оној со највисока доверба), а останатите ги елиминира.

#### Добивање финални координати:

Резултатот е список со рамки и класи за кои YOLO е уверен дека претставуваат реални објекти во сликата (во нашиот случај, регистарски таблички). Во рамките на овој проект, YOLO се користи за локализирање на регистарските таблички во сликите или видеото, изолирање на делот од сликата каде што е пронајдена табличка, за понатамошна обработка со OCR. Со тоа, наместо да го

обработуваме целиот кадар или да применуваме OCR на целата слика, ние му даваме на Tesseract само исечок (crop) што ја содржи табличката, што е многу поефикасно.

## Обработка на слика со OpenCV

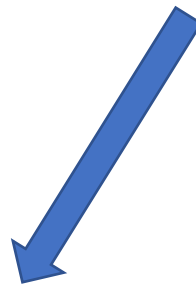
OpenCV (Open Source Computer Vision Library) е библиотека која содржи голем број функции за обработка на слики и видео. Во рамките на проектот, таа се користи за да се обезбеди претходна (preprocessing) и последователна (postprocessing) обработка на исечоците со таблички, со цел да се подобри точноста при OCR.

Главни чекори за обработка на слика (пример работен тек):

- **Конверзија во сива скала (Grayscale):**
  - Сликата од делот каде е табличката најпрво се претвора во црно-бела (8-битна слика).
  - Со ова се намалува сложеноста и шумот што често се јавува во боените канали.
- **Отстранување на шум (Noise Removal) – MedianBlur:**
  - MedianBlur или некој друг вид на филтрација (GaussianBlur, BilateralFilter) се користи за намалување на шумот без драстично да се уништат контурите на карактерите.
  - MedianBlur е корисен особено за отстранување на „сол-и-бибер“ шум (salt-and-pepper noise).
- **Morphological операции (Dilation, Erosion, Open, Close):**
  - Morphology Open се користи за отстранување на ситни „бели“ точки или објекти во позадина.
  - Morphology Close се користи за пополнување на мали „црни“ отвори (gaps) во рамките на карактерите.
  - Ваквите операции помагаат контурите на буквите и броевите да бидат појасно дефинирани и полесни за OCR.
- **Sharpening (Изострување):**
  - Со техника на „unsharp mask“ или директно со комбинирање на филтри, се постигнува заострување на рабовите.
  - Изострувањето помага OCR алгоритмот да ги препознае фините детали во карактерите.
- **Бинаризација или прагови(Thresholding):**
  - Во некои случаи, може да се примени Thresholding за дополнително да се истакнат карактерите (white) наспроти позадината (black).
  - Ова може да биде фиксно (global threshold) или адаптивно (adaptive threshold), зависно од условите на осветлување.

Со овие чекори, излезната слика е максимално „прочистена“ и подготвена за Tesseract. Колку е подобро подготвена сликата, толку OCR процесот е покоректен, со помалку грешки.

Процес на обработка:



## Tesseract OCR за препознавање текст

Tesseract OCR (Optical Character Recognition) е еден од најпознатите open-source алатки за препознавање текст од слики. Поддржува повеќе од 100 јазици, вклучувајќи македонски, и може да се обучува за подобро препознавање на специфични фонтови и формати.

Tesseract поминува низ неколку чекори за препознавање на текстот:

- Претпроцесирање на сликата:
  - Конверзија во грејскала или бинаризација (OTSU thresholding).
  - Филтрирање на шум (Gaussian Blur, Median Blur).
  - Корекција на ротација и искривување.
- Сегментација на сликата
  - Разделување на сликата на блокови, линии, зборови и букви.
- Препознавање на карактери
  - Користи LSTM (Long Short-Term Memory) модели за машинско учење.
- Обработка на резултатите
  - Пост-обработка за корекција на грешки (нормализација, dictionary lookup).



## Карактеристики и функционалности

Кодот е поделен во повеќе компоненти кои соработуваат за да се изврши задачата на детекција и препознавање на регистарски таблички:

- Модул за обработка на слики:
  - Овој модул се грижи за предобработка на сликите, што вклучува конверзија во сива скала, примена на филтри за отстранување на шум, бинаризација и морфолошки операции.
- Модул за детекција на таблички:
  - Ги користи YOLO (You Only Look Once) алгоритмите за откривање на таблички на сликите.
- Модул за препознавање на карактери:
  - По детекцијата, исечените таблички се праќаат во OCR (Optical Character Recognition) систем, каде што Tesseract OCR се користи за извлекување на текстот.
- Модул за менаџмент на резултати:
  - Овозможува организирање и складирање на резултатите од OCR во структурирана форма.

## Клучни функции и нивната улога во проектот

Некои од најважните функции во проектот се:

**process\_images(image, image\_dir, output\_dir)** - Главна функција за обработка на слики и извлекување на таблички. Оваа функција ја користи YOLO мрежата за да детектира регистарски таблички на дадени слики, по што ги складира исечените делови за понатамошна анализа.

**getOutputsNames(net)** - Функција за извлекување на имињата на излезните слоеви од YOLO моделот, што е неопходно за обработка на детектираните објекти.

**postprocess(frame, outs, save\_path, counter)** - Анализира резултатите од YOLO моделот и ги исекува табличките за понатамошна обработка. Применува Non-Max Suppression (NMS) за елиминирање на слични предвидувања и осигурување на најпрецизниот исечок.

**read\_plate\_from\_images(input\_folder, output\_folder)** - Ги обработува исечените таблички и извлекува текст со OCR. Оваа функција применува различни техники за обработка на сликите како што се зумирање, медијански филтер, прагови и острење на сликата за подобрување на резултатите од Tesseract OCR.

## Предности, недостатоци и подобрување

Системот за детекција и препознавање на регистарски таблички има одредени предности и недостатоци кои зависат од условите на снимање, квалитетот на влезните податоци и применетите алгоритми.

### Случаи кога системот работи добро

Во идеални услови, системот постигнува висока точност во детекцијата и препознавањето на табличките. Некои од факторите кои придонесуваат за добри резултати се:

- **Добро осветлување**
  - Кога сликите се направени при природно или вештачко осветлување со доволен контраст, OCR алгоритмите можат полесно да ги идентификуваат карактерите на табличката.
- **Висока резолуција на сликата**
  - Колку е повисока резолуцијата на сликата, толку повеќе детали се зачувуваат, што овозможува подобро препознавање на карактерите. Ниски резолуции можат да резултираат со заматени или нечитливи букви и броеви.
- **Јасни и чисти таблички**
  - Ако табличката не е оштетена, извалкана или прекриена со кал или снег, препознавањето на карактерите е многу поточно.
- **Фронтален агол на камерата**
  - Доколку сликата е направена од директен (перпендикуларен) агол, без значителни искривувања, OCR процесот е многу поефикасен.
- **Оптимално поставување на табличката во кадарот**
  - Ако табличката зафаќа значителен дел од сликата, тоа го олеснува нејзиното исекување и обработка.

## Непрепорачлива употреба

Иако системот за детекција и препознавање на регистарски таблички базиран на YOLO и Tesseract OCR е моќен и ефикасен во многу случаи, постојат сценарија во кои неговата употреба може да биде проблематична. Овие ограничувања произлегуваат од слабости во алгоритмите, технички ограничувања на опремата и варијации во условите на снимање.

## Сценарија каде детекцијата може да биде проблематична

### Лимитации на YOLO и Tesseract

YOLO е еден од најбрзите и најпрецизните модели за детекција на објекти, но има одредени ограничувања кога станува збор за детекција на регистарски таблички:

- **Чувствителност на осветлување**
  - Иако YOLO добро функционира во добро осветлени услови, во ноќни или слабо осветлени сцени, може да пропушти детекција или да направи повеќе грешки.
- **Големина на табличките**
  - Ако табличката е премногу мала на сликата (на пример, ако возилото е далеку), YOLO може да не ја детектира исправно.
- **Засенчени и преклопени објекти**
  - Ако делови од табличката се покриени со други објекти (на пример, друга кола или нечистотија), YOLO може да ја промаши или неточно да ја идентификува табличката.
- **Нестандардни регистарски таблички**
  - Моделот е обучен врз основа на одредени стандарди на таблички, па ако се користат специфични или модифицирани формати, YOLO може да не ги препознае како валидни.

Tesseract е еден од најчесто користените OCR модели за препознавање на текст, но има неколку слабости во контекст на регистарски таблички:

- **Чувствителност на фонтови и дизајн на табличките**
  - Tesseract најдобро работи со печатен текст со чисти фонтови. Ако табличката има декоративни елементи или нестандартни букви, препознавањето може да биде неточно.
- **Проблеми со лош квалитет на сликата**
  - OCR алгоритмите бараат јасни и остри слики за правилно препознавање. Ако карактерите се заматени или имаат низок контраст, Tesseract ќе врати погрешни резултати.
- **Недостаток на адаптација на искосени текстови**
  - Ако табличката е под агол, текстот може да биде искривен, а Tesseract нема вградени механизми за исправување на такви искривувања.
- **Преклопени или залепени букви**
  - Ако карактерите на табличката се премногу блиску еден до друг или имаат физички оштетувања, OCR може да ги препознае како погрешни симболи или да ги спои во едно.





Слика: Нестандардни таблички за кои моделот треба посебно да биде трениран



Слика: Оштетена табличка која моделот не би можел да ја прочита соодветно

### Како може да се подобрат резултатите?

За да се минимизираат недостатоците, можно е да се применат неколку техники:

- **Подобрување на осветлувањето**
  - Употреба на инфрацрвени (IR) камери за нокна детекција или поставување на дополнителни извори на светлина.
- **Примена на алгоритми за корекција на агол**
  - Користење на техники за перспектива и геометриска корекција за подобрување на читањето на искосени таблички.
- **Применување на модели обучени на оштетени таблички**
  - Додавање на примероци со оштетени, извалкани или делумно закриени таблички за да се зголеми робусноста на моделот.
- **Користење на камери со повисока резолуција**
  - Подобрување на квалитетот на влезните податоци ќе придонесе за повисока точност во препознавањето.



## Заклучок

Проектот за автоматска детекција и препознавање на регистарски таблички со употреба на YOLO и Tesseract OCR претставува значаен чекор кон автоматизација и подобрување на мониторингот на возилата. Користејќи современи алгоритми за машинско учење и обработка на слики, системот овозможува брзо и прецизно идентификување на табличките, што е од голема важност за различни индустрии и безбедносни апликации.

Главните предности на овој систем вклучуваат високата брзина на обработка, можноста за работа во реално време и адаптивилноста на различни сценарија. Преку детекција на таблички со YOLO и препознавање на симболи со Tesseract OCR, системот може да обезбеди сигурни резултати во контролирани услови. Сепак, постојат и одредени ограничувања, како што се проблемите со ниска осветленост, заматени слики или нестандартни регистарски таблички, што може да влијае на точноста на резултатите.

И покрај неговата функционалност, системот може да се унапреди со:

- Подобрување на алгоритмите за препознавање преку длабоки невронски мрежи.
- Интеграција со бази на податоци за реално време следење на сомнителни возила.
- Користење на дополнителни методи за обработка на слика за зголемување на точноста во сложени сценарија.
- Оптимизација за работа со мобилни уреди и IoT решенија.

Во иднина, оваа технологија може да најде сè поголема примена во паметните градови и напредните системи за управување со сообраќајот, зголемувајќи ја безбедноста и ефикасноста на урбаните средини.