

Примена на машинско учење во обработка на слики (машинска визија)

Препознавање на регистарски таблички од слика

1. Вовед во темата/проблемот и неговата значајност

Вештачката интелигенција со своите гранки постојано помага да напредуваат различни индустрии, а една од позначајните гранки е машинската визија. Машинската визија ја вклучува можноста машините да ги разберат визуелните информации, слично како визијата на човекот. Една од многуте примени на машинското учење, која е особено значајна, е препознавањето на регистарски таблички преку употреба на алгоритми за машинско учење.

Препознавањето на регистарски таблички, има значителна важност во различни области, од полиција до разни иницијативи за развој на „паметни“ градови. Способноста за автоматско извлекување на текстуални информации од регистарските таблички во слики обезбедува многу практични предности:

А) Подобрена безбедност;

Системите за препознавање на регистарски таблички допринесуваат за подобрување на мерките за безбедност со овозможување на брзо и точно идентификување на возилата. Ова е особено важно во надзорот и контролата на границите и превенцијата на криминални активности.

Б) Управување во сообраќајот;

Во градските средини, препознавање на регистарски таблички, е доста важно во надгледувањето и управувањето со сообраќајот. Помага во автоматизацијата на наплатата на патарини, идентификацијата на сообраќајни нарушувања и слично.

В) Паметни градови;

Со градовите кои се развиваат кон станување „паметни “ и поврзани, препознавањето на регистарски таблички игра голема улога во создавањето на интелигентни системи за транспорт. Ова допринесува во намалувањето на загадувањето, подобрувањето на квалитетот на воздухот и зголемувањето на општата урбана ефикасност.

Г) Решенија за паркинг;

Автоматското препознавање на регистарски таблички го олеснува управувањето на паркинг областите. Го оптимизира процесот на влегување и излегување од паркинзите, правејќи го удобен за корисниците и помагајќи во спроведувањето на правилата.

2. *Преглед на сродни истражувања/литература*

Во областа на препознавањето на регистарски таблички и машинското учење, бројни истражувања и литература се фокусирани на подобрување на алгоритмите за препознавање и повеќе примени на оваа технологија. Во следниот дел, се разгледуваат слични истражувања:

А) Истражувања за претпроцесирање на слика;

Техниките за претпроцесирање на слики вклучуваат конверзија во сиво, bilateral filtering за smoothing (измазнување на сликата) и Canny edge detection за оцртување на остри контури на сликата. Добро дизајниран систем за претпроцесирање е неопходен за извлекување на релевантните карактеристики и обезбедување на точност во препознавањето на регистарски таблички во различни услови.

Б) Користење на рекурентни невронски мрежи за препознавање;

Современи истражувања истражуваат употреба на рекурентни невронски мрежи за подобрување на перформансите во препознавањето на текст од регистарски таблички. Овој пристап покажува одлични резултати за процесирањето на податоците.

В) Оптимизација на OCR Алгоритми;

Има разни истражувања за како да се оптимизира процесот на Optical Character Recognition (OCR) преку промени во параметрите и конфигурациите на алгоритмите. Ова вклучува истражување на различни PSM вредности (Page Segmentation Mode) за подобрување на точноста при препознавањето на текст.

Г) Процена на точност за предвидените вредности на таблички;

Литературата се фокусира и на методи за проценка на сличноста помеѓу препознатите (предвидените) и вистинските регистарски таблички. Користењето на

методи за мерење на сличност, како што е SequenceMatcher, обезбедува квалитетна евалуација на препознавањето.

Во главно има направено многу истражувања на овие теми кои се големо значење. Константното истражување и усовршување на овие методи има за цел да ги подобри перформансите и да прошири примената на технологиите во различни области. На следните линкови можеме да разгледаме неколку статии, проекти и литература со слични истражувања:

<https://learnopencv.com/automatic-license-plate-recognition-using-deep-learning/>

<https://pyimagesearch.com/2020/09/21/opencv-automatic-license-number-plate-recognition-anpr-with-python/>

<https://kccemsr.edu.in/public/files/technovision/1/Number%20Plate%20Recognition%20Using%20Machine.pdf>

<https://iopscience.iop.org/article/10.1088/1757-899X/1084/1/012027/pdf>

3. Опис на користена методологија

За постигнување на посакуваната цел, односно за препознавање на регистарски таблички од слика и претворање на истите во текст, користени се податоци врз кои се изработени разни методологии и алгоритми.

а) Податоци

За овој проект, користени се слики од моторни возила на кои се гледаат регистарските таблички. Овие податоци рачно ги преземав од разни веб-сајтови. Овие слики варираат во различни услови во однос на осветлување, агол и слично. Овие влезни податоци како свое име ја носат регистарската табличка (односно доколку регистарската табличка на автомобилот на сликата е КА 008 JR тогаш сликата би била со име КА008JR.jpg). Од овие влезни податоци се обработени и како излезни податоци добиваме текст кој ја означува регистарската табличка. Како влезно множество се земено околу 100 слики.

За споредба земав и друго множество на податоци, каде има слики кои се повеќе на број (3000). Овие слики се веќе процесирани, односно веќе е пронајден делот каде се наоѓа регистарската таблица, а само врз нив извршуваме конверзија во сива боја и измазнување на сликата. Додека претходното множество содржеше само македонски регистарски табlici, ова множество содржи регистарски табlici од разни држави низ светот. Вистинските вредности од овие регистарски табlici се зачувани во фајл со име Ipr.csv во следниот формат:

,images,labels
0,1.jpg,RZ0047
1,10.jpg,V95246
2,100.jpg,6258TU
3,10000.jpg,B88082

Најпрво е индексот на сликата, потоа името на фајлот кој ја означува сликата, и реалната вредност на регистарската табличка на сликата.

b) Пристапи за обработка на податоци (алгоритми, архитектури, ...)

Претпроцесирање на сликата:

Овие влезни податоци (сликите) се конвертирани на разни начини со цел полесно добивање на делот од регистарската табличка. Конвертирани се во сиви нијанси, потоа се применува bilateral filter кој се користи за добивање на ефект за измазнување на сликата. Потоа се извршува edge detection кој ги оцртува сите ивици на сликата. Целта е подобрување на квалитетот на сликите и идентификација на контурите на регистарските таблички.

За второто множество бидејќи не се потребни сите овие конверзии на сликата, се користи само конвертирање во сиви нијанси и измазнување на сликата.

Откривање на контури:

Со примена на Canny алгоритам се откриваат сите ивици на сликата и се земаат само крајните точки од овие линии (ивици). Се сортираат овие контури во зависност од големината на пределот кој го опфаќаат и се земаат најголемите 10.

Овој дел не се применува во второто множество.

Препознавање на Текст (OCR):

Користење на Tesseract OCR за издвојување на текстот од регистарските таблички. Оваа методологија се користи два пати, првиот пат за земање на оној предел откриен со контурите на кој се препознава некој текст со должина поголема од 4, а вториот пат за откривање на овој текст од овие слики. Конфигурацијата на OCR е подесена со одредени параметри како PSM (Page Segmentation Mode) и ограничување на карактерите за подобрување на точноста.

За второто множество препознавањето за текст се користи при одвојување на 3000 слики од множеството, со цел да не се земат оние слики во кои не е препознаен текст или е препознаен текст кој има помалку од 4 карактери. Потоа препознавањето на текст се користи врз овие слики кои ги исполнуваат условите (има текст со големина поголема од 4) за да се препознае текстот на овие слики со различни PSM вредности.

Оптимизација на PSM вредности:

Со цел добивање на подобри резултати, искористено е динамичко истражување на различни вредности за PSM во Tesseract OCR за постигнување на оптимална конфигурација, што допринесува кон подобрување на препознавањето на текстот на овие таблички. Се тестираат вредностите од 6 до 13.

Пресметување на сличност:

Пресметување на сличност помеѓу вистинските и предвидените регистарски таблички се одвива преку SequenceMatcher. Ова овозможува вреднување на перформансите на системот и добивање на информации за точноста на препознавањето. Во зависност од сличноста на овие две податоци (предвиден и реален), овој алгоритам задава вредност помеѓу 0 и 1, каде 1 означува идентично. Овие сличности се пресметуваат за секоја една слика и потоа со функција се сумираат резултатите за да се открие точноста на целиот модел. За секоја една слика е прикажана точноста на предвидувањето како и сумираната просечна сличност.

4. Резултати

Резултатите кои ги дава овој модел комбиниран со разни алгоритми можеме да го забележиме на следните слики.

Пример за влез на слика од првото множество:



Слика 1. Пример за влезен податок од првото множество на податоци.

По обработка на оваа слика го добиваме следниот резултат:



Слика 2. Пример за влезен податок по обработка на сликата.

Врз втората слика се извршува предвидување и овој алгоритам го дава следниот резултат:

Actual: SK1705BD Predicted: SK1705BD Similarity: 1.00

Но сепак резултатите на овој код не се идеални поради тоа што овој модел успева да пронајде регистарски таблички само во околу 40% од првото влезното множество на слики, а од пронајдените регистарски таблички, овој модел предвидува сличност со околу 73%. Главниот проблем е во препознавањето на пределот каде се наоѓа регистарската табличка во влезните информации.

Пример за излез од првото множество:

Actual: SK9889AM	Predicted: SK9889AN	Similarity: 0.88
Actual: VE2376AD	Predicted: BVE2376AD	Similarity: 0.94
Actual: PP0001DD	Predicted: PP00010D	Similarity: 0.88
Actual: TE5420AF	Predicted: WEE5220009	Similarity: 0.44
Actual: SK2772BJ	Predicted: SK2722B8J	Similarity: 0.82
Actual: RA7112AB	Predicted: PRA112A8	Similarity: 0.75
Actual: BT0315AG	Predicted: B0	Similarity: 0.40
Actual: KS1962AB	Predicted: BS81962AB	Similarity: 0.82
Actual: SK8460AG	Predicted: SK88460AG	Similarity: 0.94
Actual: SK5342BG	Predicted: SK25342B6	Similarity: 0.82
Actual: SK5294BE	Predicted: SK55294BEI	Similarity: 0.89
Actual: BT7855AC	Predicted: BT27855AC	Similarity: 0.94
Actual: SK619VD	Predicted: SK619VD	Similarity: 1.00
Actual: BT7649AB	Predicted: T7EVSAB	Similarity: 0.53
Actual: KU5604AD	Predicted: KUB5602AN	Similarity: 0.71
Actual: SK2988AK	Predicted: MK2900A	Similarity: 0.53
Actual: SK3824AB	Predicted: ROTLAB	Similarity: 0.29
Actual: SK6123AR	Predicted: BSK6123AP	Similarity: 0.82
Actual: GV0001AO	Predicted: V000TTO	Similarity: 0.67
Actual: SK2043AL	Predicted: ISK82023AL	Similarity: 0.78
Actual: SK1705BD	Predicted: SK1705BD	Similarity: 1.00
Actual: SK006NM	Predicted: SK2006N	Similarity: 0.86
Actual: SK0049BA	Predicted: SK004958	Similarity: 0.75
Actual: SK5481AF	Predicted: SK95481AF	Similarity: 0.94
Actual: SR5552AE	Predicted: SR5552AE	Similarity: 1.00
Actual: ST9169AD	Predicted: BT9	Similarity: 0.36
Actual: SK955SU	Predicted: SK2955SU	Similarity: 0.93
Actual: SK262NV	Predicted: S267NV	Similarity: 0.77
Actual: PS8000AB	Predicted: IEPS8000A8J	Similarity: 0.74
Actual: PP8103AC	Predicted: T	Similarity: 0.00
Actual: SK1245BE	Predicted: YE1	Similarity: 0.18
Actual: OH611BS	Predicted: OH611BS	Similarity: 1.00
Actual: BT0284AB	Predicted: T028A8	Similarity: 0.71
Actual: SK0135BK	Predicted: KSU1S9BK	Similarity: 0.50
Actual: PP0801AG	Predicted: PP20801AG	Similarity: 0.94
Average Similarity: 72.98%		

Со цел да го решиме овој проблем со пронаоѓањето на делот со регистарските таблички во сликите и да добиваме подобро предвидување, работиме со второто множество каде веќе се препознаени деловите со регистарски таблички во сликите и потребно е само да го препознаеме текстот кој е напишан во истите.



Слика 3. Пример за влезен податок при работа со второто множество.

При извршувањето на предвидување конкретно за овој пример го добиваме следниот резултат:

```
Actual: 7609XP Predicted: 7609XP Similarity: 1.00
```

Доколку земеме друг пример:



Слика 4. Друг пример за влезен податок при работа со второто множество.

При предвидувањето ја добиваме следната вредност:

```
Actual: 8838YK Predicted: 8838K Similarity: 0.91
```

Со ова множество на поголем број на податоци (околу 3000), добиваме полоши перформанси со околу 57% точност, но сепак работиме со поголем број на податоци.

```
...
Actual: CX1623 Predicted: CX1623 Similarity: 1.00
Actual: H50330 Predicted: H50330 Similarity: 1.00
Actual: GA2297 Predicted: GA2297 Similarity: 1.00
Actual: QR0138 Predicted: OROT39 Similarity: 0.33
Average Similarity: 57.28%
```

Можеме да заклучиме дека единствениот проблем не е во препознавањето на делот каде што се наоѓаат регистарските таблици, туку дека и препознавањето на текстот не е идеално. Но сепак и ова решение не е толку лошо.

5. Заклучок

Многу индустрии напредуваат како резултат на развивањето на машинската визија и вештачката интелигенција. Машинската визија продолжува да се развива со напредокот во вештачката интелигенција и длабокото учење, овозможувајќи посоефицирана и точна анализа на визуелните податоци. Примените на машинската визија се разновидни и придонесуваат во подобрување на ефикасноста во работата, безбедноста и донесувањето одлуки во различни индустрии (паркинг, управување во сообраќајот, безбедност на граници, паметни градови и слично).

Во доменот на препознавање на регистарски таблички и машинско учење, бројни студии се фокусираат на подобрување на алгоритмите за препознавање и како и на што поширока примена на оваа технологија. Области на истражување опфаќаат:

- Студии за претпроцесирање на слики,
- Користење на рекурентни невронски мрежи за препознавање,
- Оптимизација на OCR Алгоритми,
- Процена на точност за предвидените вредности на таблички

и многу други области.

Општо земено, на овие теми се спроведени опсежни истражувања, кои значително придонесуваат за напредок. Постојаното истражување и усовршување на овие методи има за цел да ги подобри перформансите и да ја прошири примената на овие технологии во различни полиња.

Со оглед на тоа што се работи со визуелни податоци, како влез во овие алгоритми и методологии се земаат слики, а со понатамошен развој на овој код, може да се земат и видеа. Врз овие слики се извршуваат разни алгоритми и конверзии со цел добивање на посакуваните резултати, а тоа се:

- Претпроцесирање на сликата;
- Откривање на контури;
- Препознавање на Текст (OCR);
- Оптимизација на PSM вредности;
- Пресметување на сличност.

Како краен резултат добиваме податоци за предвидувањето на регистарските таблички заедно со процентот на просечна сличност на истите.

Перформансите на овој модел може да се подобрат најмногу со подобрување на алгоритмите за претпроцесирање на сликата, а и со подобар начин за препознавање на текстот. Резултатите кои ги дава овој модел не се идеални, но прикажуваат пример како би можел еден модел да функционира, да се усоврши и потоа користи во реални проблеми.

Врска до проектот поставен на Github:

<https://github.com/Aleksandar-Stojanov/License-Plates-Project>

Врска до кодот на Google Colaboratory:

Со првото множество:

[https://colab.research.google.com/drive/1n5RjrNscitS9BY3giC1LV5bTjbJK21r ?usp=sharing](https://colab.research.google.com/drive/1n5RjrNscitS9BY3giC1LV5bTjbJK21r?usp=sharing)

Со второто множество:

<https://colab.research.google.com/drive/1Yo8QoqB9rEYPCHHCgPnBKdDzwjmis1J?usp=sharing>