



Факултет за Информатички Науки и Компјутерско Инженерство

Детекција на регистарски таблички и препознавање на карактери



Изработил: Миле Пеливанов 201057

Ментор: Проф. Ивица Димитровски

Содржина

Абстракт	3
I. Вовед	3
II. Користени методи	5
1. Детекција на регистарски таблички	5
1.1. Податочно множество.....	5
1.2. Користење на претходно трениран модел за детекција на регистарски таблички.....	5
2. Препознавање на регистарски таблички	9
2.1. Сегментација на карактери	10
2.2. Препознавање на карактери со користење на Keras OCR	12
III. Резултати.....	14
IV. Заклучок.....	17
Референци	17

АБСТРАКТ

Задачата на откривањето и препознавање на регистарски таблички има голема важност во Интелигентните Транспортни Системи, со примена во безбедноста, контролата на сообраќајот и електронските системи за плаќање како што се плаќање на патарина на автопат и плаќање на паркинг место. Покрај бројните алгоритми развиени за препознавање на регистарски таблички, секои со свои предности и недостатоци, некои методи бараат да бидат интензивни при пресметување и одземаат многу време. Во овој труд, се запознаваме со директен и ефикасен пристап за справување со предизвиците за откривање на регистарски таблички и препознавање на знаци. Користејќи дводимензионална трансформација, прво ги откриваме регистарските таблички со извлекување на вертикални рабови од влезните слики. Густината на вертикални рабови помага во идентификување на потенцијални области на регистарски таблички, потоа валидирано преку конволуциска невронска мрежа. По успешното откривање на табличката, применет е едноставен метод на сегментација базиран на растојанието помеѓу знаците. Кандидатите за знаци потоа се класифицирани користејќи друга конволуциска невронска мрежа. Резултатите од детектираните регистарски таблички покажуваат висока точност, достигнувајќи 99.43% во детекција на регистарските таблички и 92.46% во препознавање на карактерите. Ефикасноста и точноста на методите се потврдуваат преку споредба со постоечки податочни множества. Едноставноста на архитектурата овозможува лесна инсталација на стандарден компјутер, покажувајќи ја разновидноста и ефикасноста на методите во справувањето со различни задачи на откривање на регистарски таблички и препознавање на знаци.

I. ВОВЕД

Доаѓањето и развојот на технологиите за компјутерска визија и дигитално процесирање на слики отвора нова можност за напредок во автоматизирани системи, а една од забележителните примени во овој домен е откривањето на регистарски таблички и последователното препознавање на карактери. Оваа документација има за цел да обезбеди детално истражување на методологиите, алгоритмите и технологиите вклучени во точно идентификување на регистарски таблички и извлекување и препознавање на карактерите кои се наоѓаат на нив. Мотивацијата на ова истражување произлегува од многу важната улога што ја имаат откривањето на регистарски таблички и препознавањето на карактери во различни апликации во реалниот свет. Најчесто, кај системите за детекција и препознавање на регистарски таблички постојат три главни фази. Првата фаза е предобработка, откако сликата е снимена, се извршува дополнителна обработка на сликата како конвертирање на

сликата од еден во друг простор, промена на резолуцијата на сликата и отстранување на шумови. Втората фаза е локализација на регистарската табличка. Делот од сликата кој ни е од интерес, односно самата регистарска табличка се открива на основа на некои карактеристики на регистарската табличка и карактеристиките на сликата. Конечната фаза е оптичко препознавање на карактерите и оваа фаза се смета за најважниот чекор бидејќи помага да се прочита бројот на табличката и да се идентификува возилото.



Сл. 1: Колекција од слики собрани за детекција и препознавање на регистарски таблички

Во полето на компјутерска визија и обработка на слики, процесот често започнува со предобработка на слики. Овој почетен степен гарантира дека сликите се соодветно формирани и оптимизирани. Во контекстот на нашето истражување, ние започнуваме со конверзија на просторот на боја на сликата и промена на големината на сликата. Овие мерки придонесуваат за точна и ефикасна обработка во понатамошните фази. Откако сликите ќе поминат низ фаза на предобработка, ние преминуваме на фазата за откривање на регистарски таблички. Тука ќе користиме пре-трениран модел, способен да разликува и локализира регистарски таблички во рамки на слики. Овој модел користи современи техники, вклучувајќи конволуциски невронски мрежи, за да ги испита визуелните содржини и да ги локализира контурите на дадената регистарска табличка. По успешното откривање на регистарската табличка, нашиот фокус се пренесува кон два пристапи за препознавање на регистарската табличка. Во првиот пристап, го користиме Keras OCR за изведување на препознавање на целата регистарска табличка, без да се вклучува сегментација на карактерите. Во вториот пристап, првично се прави сегментација на карактерите, процес кој вклучува изолирање на поединечни карактери од регистарската табличка. Овој процес на сегментација овозможува подетален преглед на секој карактер. Следно, повторно се користи Keras OCR, но овој пат се применува врз сегментираните карактери

Целокупно, оваа документација претставува практичен водич за сите оние кои се заинтересирани со детекција на регистарски таблички и системи за препознавање на карактери, нудејќи објаснувања за користените методологии, алгоритми и технологии.

II. КОРИСТЕНИ МЕТОДИ

1. Детекција на регистарски таблички

1.1. Податочно множество

За процесот на детекција на регистарските таблички користиме претходно истрениран модел. Податочното множество кое е искористено за тренирање на овој модел разновидна колекција на фотографии од регистарски таблички во различни сценарија. Се состои од слики со еден или повеќе автомобили, вклучувајќи различни перспективи, како што се различни агли од регистарски таблички, различни осветлувања и поставки на околината. Исто така вклучени се различни фонтови и големини на регистарските таблички осигурувајќи модел кој е робустно истрениран за препознавање и адаптирање на најразличните фотографии од регистарски таблички

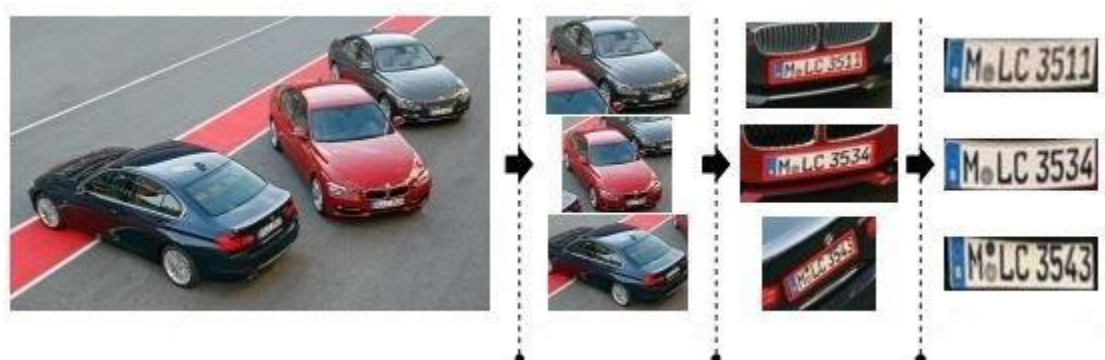
Ова големо податочно множество има клучна улога во подобрувањето во адаптивноста на моделот, дозволувајќи му ефективно да се справува со различен спектар на сценарија што се среќаваат во ситуации од реалниот свет. Преку изложување на моделот со фотографии од регистарски таблички сликани од различни агли и со различни големини, процесот на обука осигурува дека моделот може да ги детектира регистарските таблички без оглед на нивната позиција или големина. Оваа разновидност во податочното множество ги одсликува реалните предизвици со кои се соочуваат системите за препознавање на регистарски таблички, подготвувајќи го моделот правилно да ја извршува неговата задача без разлика на условите од животната средина.

Разновидноста на податочното множество не само придонесува кон адаптивноста на моделот, туку и го овозможува моделот ефикасно да генерализира, правејќи го способен да се справува со нови ситуации кои не биле сретнати за време на процесот на тренирање. Овој пристап се совпаѓа со целта за креирање на робустен и сигурен систем за препознавање на регистарски таблички, способен за справување со непредвидени сценарија од реалниот свет.

1.2. Користење на претходно трениран модел за детекција на регистарски таблички

Со оглед на распространетата употреба на возила во широко користените податочни множества за детекција и препознавање на објекти, нашиот пристап избегнува целосно тренирање на детектор. Наместо тоа користиме модел за детекција на возила, земајќи ги во предвид критериумите кои ги имаме дефинирано во нашиот просец на селекција. Од една страна посакувана е висока recall стапка, поради тоа што секоја детекција на промашок од страна на моделот, односно препознавање дека нема регистарска табличка на сликата, кога всушност постои регистарска табличка директно влијае врз целокупното откривање на промашоци при детекција на регистарски таблички. Од друга страна, исто така посакуваме

и висока прецизност, така што секое погрешно детектирано возило мора да биде валидирано од WPOD-NET.



Сл. 2: Илустрација од предложениот процес

Врз основа на овие потреби, користиме YOLOv2 мрежа поради нејзиното брзо извршување (околку 70 FPS) и добар компромис помеѓу прецизност и recall стапка (76.8% mAP за PASCAL-VOC dataset). Не правиме никаква промена кај YOLOv2 мрежата, правиме само спојување на мрежата со WPOD-NET, поврзувајќи ги излезите поврзани со возила и игнорирање на сите останати класи.

Кај позитивните детекции потоа правиме промени во големината пред тие да бидат предадени на WPOD-NET. Поголемите влезни слики дозволуваат детекција на помали објекти, но со тоа се зголемува потрошувачката на ресурси. Доколку сликите од возилата се од напред или позади, тогаш односот помеѓу големината на регистарската табличка и рамката околку детектираното возило е голем и со тоа лесно може да се детектира регистарската табличка. Меѓутоа односот тежи да биде многу помал кога сликите од возилата се странично или косо, бидејќи рамката околку возилото е поголема и поиздолжена. Поради тоа, косите погледи треба да се променат во поголема димензија, за разлика од сликите од напред или од назад, со цел да се задржи регионот на регистарската табличка да биде препознатлив.

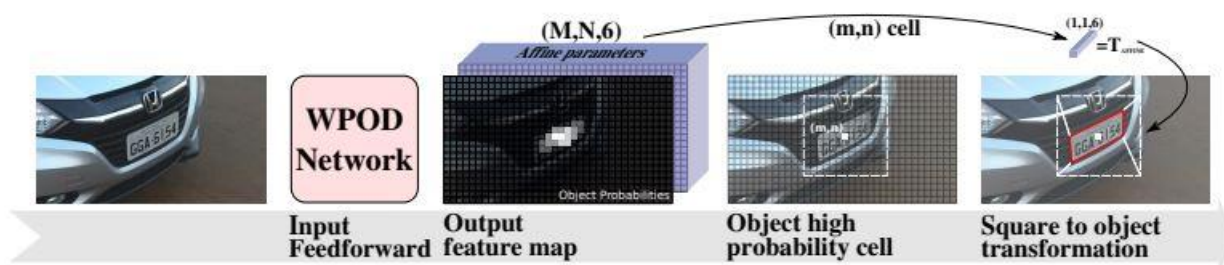
Иако 3D методите за проценска на положбата можат да се користат за да се утврди скалата за промена на големината, пристапот кој се користи е едноставен и брз, заснован на односот на позицијата на рамката околку детектираното возило. Ратата на промена на големината е дефинирана како

$$f_{sc} = \frac{1}{\min\{W_v, H_v\}} \min \left\{ D_{min} \frac{\max(W_v, H_v)}{\min(W_v, H_v)}, D_{max} \right\},$$

каде што W_v и H_v се ширина и висина на рамката околу возилото. Може да се забележи дека D_{min} и D_{max} се гранични вредности за опсегот на најмалата димензија на ратата на промена на големината на рамката. Според направени истражување и обид за задржување на добар соодност помеѓу прецизноста и време на извршување, вредностите се дефинирани како $D_{min} = 288$ и $D_{max} = 608$. Асд ас ас ас

Регистарските таблички се правоаголни и рамни објекти, кои се прикачени на возилата заради идентификација. За да се искористи нивната форма, ние користиме CNN (Convolutional Neural Network) наречена Warped Planar Object Detection Network (WPOD-NET). Оваа мрежа учи да ги детектира регистарските таблички во различни дисторзии и користи коефициенти на трансформација со која косата слика од регистарската табличка ја доведува во правоаголна форма која наликува на фронтален поглед од возилото.

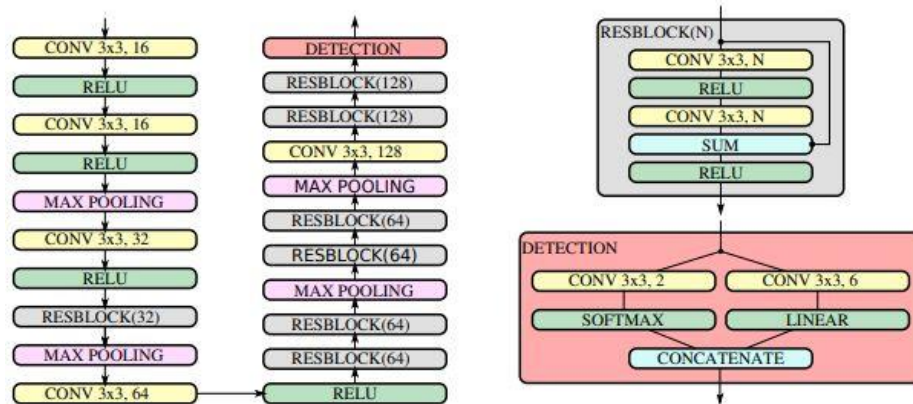
WPOD-NET е развиена користејќи ги YOLO, SSD и Spatial Transformer Network (STN). YOLO и SSD прават брзо откривање и препознавање на повеќе објекти одеднаш, но не земаат во предвид просторни трансформации, генерирајќи само рамки околу секое нивно откривање. Од друга страна STN може да биде искористена за откривање на неправоеаголни објекти, но не може да се справи со повеќе трансформации во исто време, правејќи само една просторна трансформација на целиот влез.



Сл. 3: Целосно конволуциско откривање и препознавање на рамни објекти

Процесот на детекција користејќи WPOD-NET е претставен во Слика 3. Првично на мрежата се дава излезот од YOLOv2 модулот за детекција на возила заедно со направените промени во големината на сликата. Препраќањето на информации во рамките на WPOD Network резултира со 8-канална мапа на карактеристики која содржи информации за тоа дали објект (во нашиот случај регистарска табличка) е присутен, заедно со параметри за трансформација на сликата за доведување на сликата во фронтален поглед. Идејата е да се дади фокус на одредени ќелии од мапата на карактеристики. Доколку веројатноста за присуство на објектот во одредена ќелија е поголема од дефинираната граница, тогаш мрежата ги користи доделените параметри за да направи матрица на трансформација. Оваа матрица потоа се применува врз ќелијата, ефективно трансформирајќи ја ќелијата во регион соодветен на регистарска табличка, ја намалува дисторзијата и ја доведува до хоризонтално и вертикално порамнет објект.

Користената архитектура содржи 21 конволуциски слоеви, каде 14 од нив се во главни блокови, каде големината на сите конволуциски филтри е фиксна и има вредност од 3×3 . Во рамките на целата мрежа се користат ReLU активациски функции, освен во блокот за детекција. Постојат 4 pooling слоеви со големина од 2×2 кои ја намалуваат влезната димензионалност за факториел од 16, но задржувајќи ги сите важни информации. На крај во блокот за детекција постојат два паралелни конволуциски слоеви, едниот е за одлучување на веројатноста, активиран преку softmax функција, и втор за дефинирање на параметри за трансформација, кај кој нема активација.



Сл. 4: Детална WPOD-NET архитектура

При дефинирање на функцијата на чинење, во предвид се зема геометриската прецизност и проценетата веројатност за присуство на регистарската табличка. Обележаните агли на регистарската табличка и рамката околу објектот ги дефинираат координатите и референтните точки кои се користат за проценка на точноста на трансформацијата. Излезот на мрежата, организиран во мапа на карактеристики, поминува низ сложен процес на трансформација и нормализација. Овој сложен процес гарантира дека моделот не само ќе научи точно да го предвидува присуството на регистарската табличка, туку и просторните трансформации потребни за правилно порамнување во фронтален поглед. Земајќи ги во предвид геометриската прецизност и проценетата веројатност за присуство на регистарската табличка, функцијата на чинење го води моделот кон сигурно и робустно откривање на регистарски таблички, важно за практични примери од реалниот свет.

За тренирање на WPOD-NET користено е податочно множество од 196 слики, од кои 105 слики се од Cars Dataset, 40 слики од SSIG Dataset и 51 слики од AOLP Dataset. За секоја слика означени се аглите на регистарската табличка на сликата. Избраните слики од Cars Dataset содржат најмногу Европски регистарски таблички, но има многу и Американски регистарски таблички. Слика од SSIG и AOLP содржат Бразилски и Тајвански регистарски таблички.



Сл. 5: Пример од означени регистарски таблички во тренирачкото податочно множество

Со оглед на намалениот број на означени слики во тренирачкото податочно множество, употребата на data augmentation е од клучно значење. Со цел зголемување на податочното множество се користат неколку трансформации.

- Rectification: целата слика се корегира врз основа на означената регистарска табличка
- Aspect-ratio: соодносот на регистарската табличка е случајно поставен во интервал [2,4] со цел да се приспособат големини од различни региони.
- Centering: центарот на регистарската табличка станува центар на сликата.
- Scaling: регистарската табличка е скалирана, така што ширината има вредност помеѓу 40px и 208px.
- Rotation: извршена е 3D ротација со случајно избрани агли, со цел да се земе во предвид различниот опсег на позиции на камери.
- Mirroring: 50% шанса.
- Colorspace: мали модификации во HSV простор на боја
- Annotation: локацијата на аглите на регистарската табличка се прилагодени со додавање на истите просторни трансформации користени за исправување на влезната слика.

Со избраното множество од горенаведените трансформации, од еден примерок може да се добие голема разновидност на тест слики со многу различни визуелни карактеристики. На пример, Слика 6 покажува различни примероци добиени од иста слика.



Сл. 6: Различни трансформации за ист примерок.

Мрежата е тренирана со 100.000 итерации на мини-бачови со големина од 32 користејќи ADAM оптимизатор. Ратата на учење е поставена на 0.001 заедно со параметри $\beta_1 = 0.9$ и $\beta_2 = 0.999$. Мини-бачовите кои се користат се генерирани преку случајно избирање и зголемување (augmentation) на примероците од тренирачкото множество, што резултира со нови влезни тензори со големина (32 x 208 x 208 x 3) при секоја итерација.

2. Препознавање на регистарски таблички

По локализацијата на регистарската табличка, истата е отсечена од нејзината оригинална слика и е класифицирана. Поминува низ процес на препознавање за класифицирање и препознавање на нејзините букви и броеви. Суштината на класифицирањето е да се најдат кандидатите за знаци во откриената регистарска табличка, а потоа класифицираните кандидати да се совпаднаат со некој знак. Во предложениот метод, се покриваат два пристапи при препознавање на знаци. Првиот пристап го прескокнува процесот на сегментирање на

карактери и директно преминува на користење на модел за препознавање на карактери. Вториот пристап прво ги користи празните линии помеѓу знаците со цел да се сегментираат знаците на регистарската табличка и да се генерираат сегментираните кандидати. Потоа овие сегментирани кандидати за знаци се класифицираат со користење на претходно трениран модел за препознавање на карактери.

2.1. Сегментација на карактери

Сегментација на карактерите од регистарската табличка е клучен чекор кај системите за оптичко препознавање на карактери (OCR), чија цел е прецизно да извлечат поединечни знаци од детектираната регистарска табличка. Во овој процес, целта е ефикасно идентификување и изолирање на карактери за да се олесни задачата на нивно препознавање. Еден ефикасен пристап вклучува вертикално скенирање на откриената регистарска табличка, колона по колона, од горниот кон долниот дел од табличката. Техниката се потпира на обзERVацијата дека знаците на регистарската табличка се добро одвоени и секој знак е на одредено постојано растојание од другите знаци. Понатаму, со поставување на праг за број на црни пиксели во секоја колона, алгоритмот ги препознава потенцијалните сегменти на знаци, намалувајќи ја комплексноста при пресметувањето. Овој процес гарантира дека колоните кои го надминуваат дефинираниот праг за веројатност да содржат знаци се предадени на понатамошна анализа, оптимизирајќи ја брзината и точноста во процесот на сегментација.

Почетната фаза на процесот на сегментација вклучува претходна обработка на сликата на исечената регистарска табличка. Оваа преопработка има за цел да ја подобри видливоста на знаците и да го намали шумот, обезбедувајќи посигурен резултат на сегментација. Техники како промена на големината, трансформација во сива скала на бои, заматување на сликата придонесуваат за создавање на почиста слика. Последователно, се користи метод на адаптивен праг за бинаризација на сликата, трансформирајќи ја во бинарна репрезентација со јасни разлики помеѓу карактерите и позадината. По овој процес се добива резултат прикажан на Слика 7.

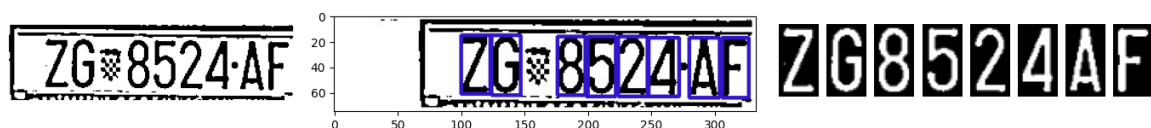


Сл. 7: Обработка на детектираната регистарска табличка за добивање на истакнати знаци.

За да се идентификуваат потенцијалните колони за сегментација на карактери, алгоритмот ја скенира бинарната слика, колона по колона, пресметувајќи го бројот на црни

пиксели во секоја колона. Поставен е предефиниран праг за да се одреди кога колоната содржи значителен број на последователни црни пиксели, при што укажува на потенцијален дел од знакот. Овој праг помага да се разликуваат вистинските сегментирани знаци од шумот. Откако ќе се идентификува колоната која што се квалификува за можен карактер, алгоритмот се преместува на следната колона, намалувајќи ги ресурсите за пресметување преку избегнување на непотребна обработка.

Колоните идентификувани како потенцијални сепаратори се означени со рамка, служејќи како визуелна претстава на процесот на сегментација. Овие потенцијални сепаратори колективно ја формираат основата за одредување на просечното растојание на раздвоеност помеѓу карактерите. Се земаат во предвид последователните колони кои се сметаат како колони за раздвојување и се пресметува нивната просечна вредност со цел да се добие веродостојна мерка за типично растојание меѓу карактерите.



Сл. 8: Поставување на сепаратори за можни знаци и сегментација на карактери.

Резултирачката сегментирана регистарска табличка прикажана на Слика 8 ја покажува ефикасноста на овој пристап во прецизно изолирање на поединечните карактери. Алгоритмот не само што ги идентификува карактерите туку ги разгледува и просторните соодноси помеѓу нив. Ова просторно разбирање е од клучно значење за зачувување на редоследот и структурата на карактерите, што на крајот ја подобрува способноста на OCR системот правилно да ги препознава информациите од регистарските таблички.

Методот на сегментација објаснат погоре ја покажува рамнотежата помеѓу пресметувачката ефикасност и прецизноста. Со користење на селективно процесирање само на колоните со значителен број на црни пиксели, односно оние кои го надминуваат предефинираниот праг, алгоритмот ги минимизира непотребните пресметки и се фокусира само на области со поголема веројатност да содржат релевантни информации. Употребата на потенцијални сепарирачки колони и пресметување на просечното нивно растојани придонесува за постоење на робустен и приспособлив процес на сегментација што може да прифати различни варијации на регистарски таблички.

Како заклучок, опишаниот пристап за сегментација на карактери од регистарска табличка го користи вертикалното скенирање и броење на пиксели за ефективно да ги изолира карактерите во сликата на регистарската табличка. Комбинацијата на техниките за претпроцесирање на сликите, поставување на адаптивен праг и просторна анализа на колоните со црни пиксели резултира со сигурен и ефикасен процес на сегментација. Овој метод игра клучна улога во подобрувањето на перформансите на OCR системите,

обезбедувајќи точно и брзо извлекување на информации од регистарски таблички во различни сценарија од реалниот свет.

2.2. Препознавање на карактери со користење на Keras OCR

По успешната имплементација на процесот на сегментација на карактери од регистарска табличка, следната фаза вклучува препознавање на карактери преку примена на моќен OCR модел изграден со Keras. Сегментираните знаци кои се извлечени од сликата на регистарската табличка, сега се доделуваат за класификација и интерпретација според Keras OCR моделот. Користејќи ги способностите на длабокото учење, овој модел е обучен да ги препознава уникатните карактеристики на секој карактер, овозможувајќи точна идентификација на информацијата добиена од регистарската табличка.

Кај овој модел, дизајнирана е длабока повеќе-слојна невронска мрежа користејќи Keras за класификација на EMNIST податочното множество. EMNIST податочното множество се состои од цифри од 0 до 9, како и букви од A до Z. Содржи тренирачко множество од 112.800 примероци и тестирачко множество од 18.800 примероци. Знаците се бинарни 28x28 пиксели.

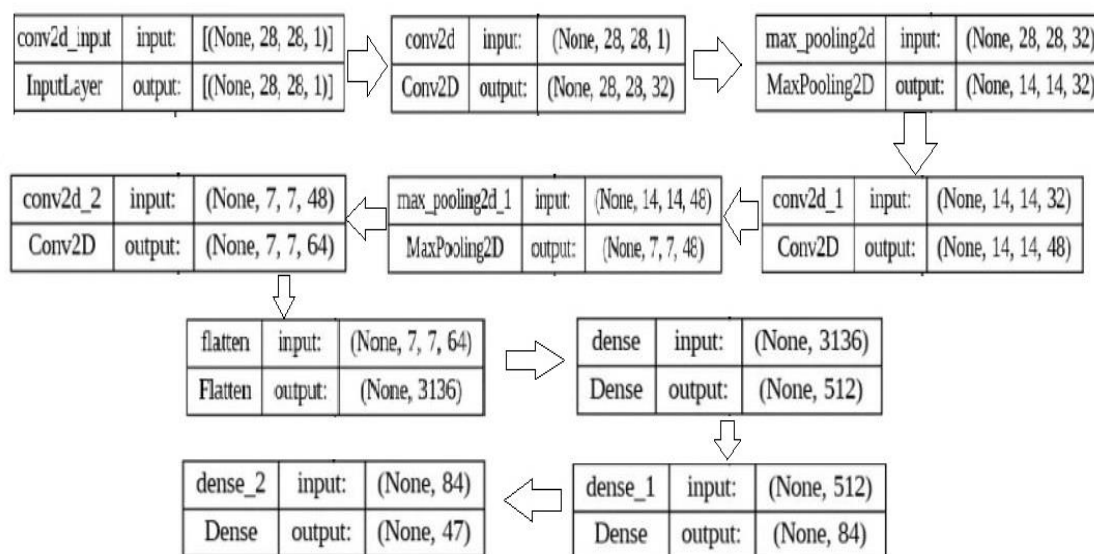


Сл. 9: EMNIST податочно множество од букви и броеви.

Моделот кој го користиме содржи почетен конволуциски блок, кој служи како основен слој, има 32 филтри со големина 5x5. Конволуциската операција, помогната од tanh активациската функција, го олеснува робусното извлекување на карактеристиките од сликата. Во овој слој се обезбедува конзистентност кај влезните и излезните димензии. Следејќи го конволуцискиот блок, слојот MaxPooling2D ги намалува димензиите на сликата додека ги зачувува најважните информации. Во овој слој иницијалната големина од 5x5 се намалува на 2x2. Мрежата потоа вклучува втор конволуциски блок, воведувајќи 48 филтри со големина 5x5 и користи tanh активациска функција. Во овој слој се прави намалување на големината на излезот во секој нареден слој. Flatten слојот претставува клучна компонента, трансформирајќи го излезот од вториот конволуциски блок во рамна структура компатибилна со следните поврзани слоеви. Потоа следуваат три поврзани слоеви, кои придонесуваат за длабочината и сложеноста на мрежата. Првиот слој интегрира 256 неврони со tanh активациска функција, поттикнувајќи сложено мапирање на карактеристиките. Следниот слој содржи 84 неврони со tanh активациска функција дополнително ги усовршува научените репрезентации. Крајниот слој со 47 неврони и

softmax активациска функција, има клучна улога во класификација на излезот на мрежата во 47 различни класи.

Ова се совпаѓа потребата од повеќе-класна класификација кај проблемот, каде моделот има задача да ја додели највероватната лабела од сите 47 можни класи. Softmax активациската функција го олеснува ова трансформирање на нумерички излези во веројатноста распределба, каде што секој елемент ја претставува веројатноста за одредена класа. Збирот на сите елементи од веројатноста е ограничен да биде 1, со што се обезбедува кохерентна дистрибуција на веројатности низ целото множество на класи. Оваа структурирана невронска мрежа покажува комбинација на конволуциски, pooling и целосно поврзани слоеви, приспособени за ефикасно извлекување на карактеристики и класифицирање на влезни податоци во повеќе класи.



Сл. 10: Keras OCR архитектура

При тренирањето на моделот избран е ADAM оптимизацискиот алгоритам, познат по својата ефикасност во забрзувањето на конвергенцијата при тренирањето и надминување на ограничувањата кои се присутни кај другите оптимизатори. Приспособливоста заедно со неговата пресметувачка ефикасност го прави ADAM како широко префериран оптимизатор во различни сценарија за длабоко учење.

Користената невронска мрежа користи категоријска cross-entropy како функција на чинење, идеална за повеќе-класна класификација. За време на тренирањето на моделот, целта е да се осигура дека моделот генерира предвидени веројатности кои блиску се совпаѓаат со вистинската распределба на класите. Категориската cross-entropy ја истакнува разликата помеѓу предвидените веројатности и целните лабели, управувајќи го моделот да ја минимизира дивергенцијата.

Оваа функција на чинење е од клучно значење за прилагодување на мрежата, што на крајот ја подобрува нејзината способност да прави точни предвидувања во повеќе-класни сценарија.

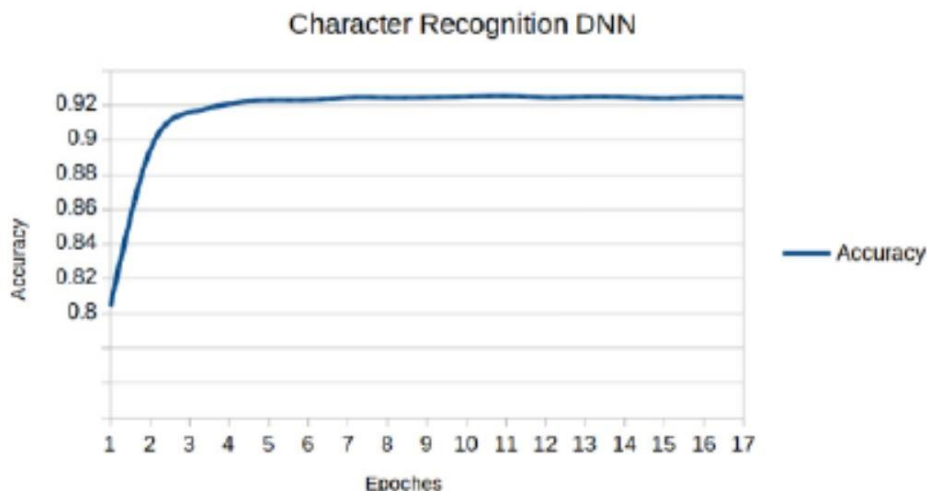
$$-\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C \mathbf{1}_{y_i \in C_c} \log p_{model}[y_i \in C_c]$$

Сл. 11: Функција на чинење кај Keras OCR – категоријска cross-entropy

Процесот на користење на истренираниот Keras OCR модел за препознавање на знаци вклучува pipeline кој ефикасно ги обработува сегментираниите знаци. Користејќи ја библиотеката `keras_ocr`, се инстанцира pipeline за да се олесни задачата за препознавање на карактерите. Pipeline-от е конфигуриран да се справува со целиот работен тек, од читање на слики до препознавање на текст во нив. За да се состават информациите за регистарската табличка, препознаениот знак од секоја ограничувачка рамка се спојува, постепено градејќи ја целосната низа на знаци. Оваа низа ја доловува суштината на препознаените знаци во сегментираните региони, формирајќи го конечниот излез на регистарската табличка.

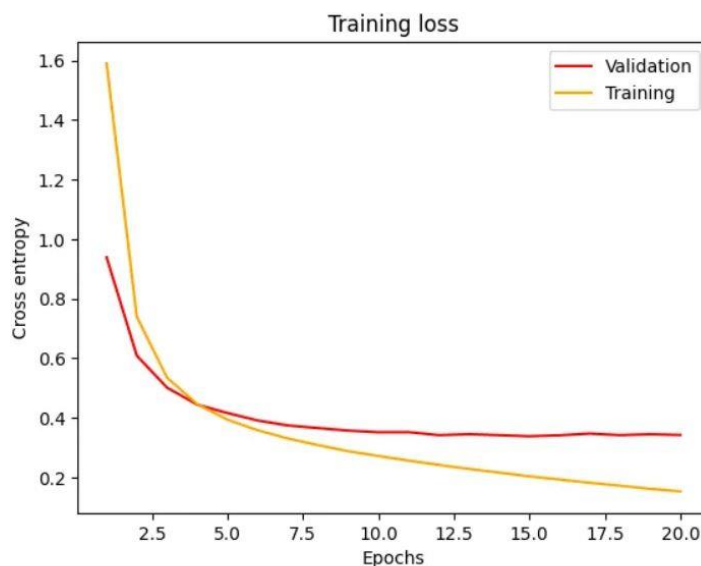
III. РЕЗУЛТАТИ

Моделот OCR постигна прецизност од 92.32%, покажувајќи ја неговата почетна ефикасност. Кога се споредува со перформансната метрика на EMNIST податочното множество, оваа прецизност се покажува како конкурентна за првата итерација на моделот. Меѓутоа постои простор за подобрување, а со искористување на алатките на Keras се овозможуваат директни модификации за подобрување на перформансите.



Сл. 12: Вредност на прецизност кај OCR за време на тренирање на моделот

За дополнително усовршување на моделот, направена е dropout регуларизација на третиот скриен слој во мрежата, воведувајќи веројатност на задржување од 0.5. Истотака направени се прилагодувања на бројот на неврони во скриените слоеви, кои се конфигурирани како 64, 3136, 512, 84, соодветно. Ова зголемување на бројот на неврони донесува дополнителна сложеност во процесот на донесување на одлуки. Овие модификации доведоа до забележително подобрување во точноста достигнувајќи 92.46%.

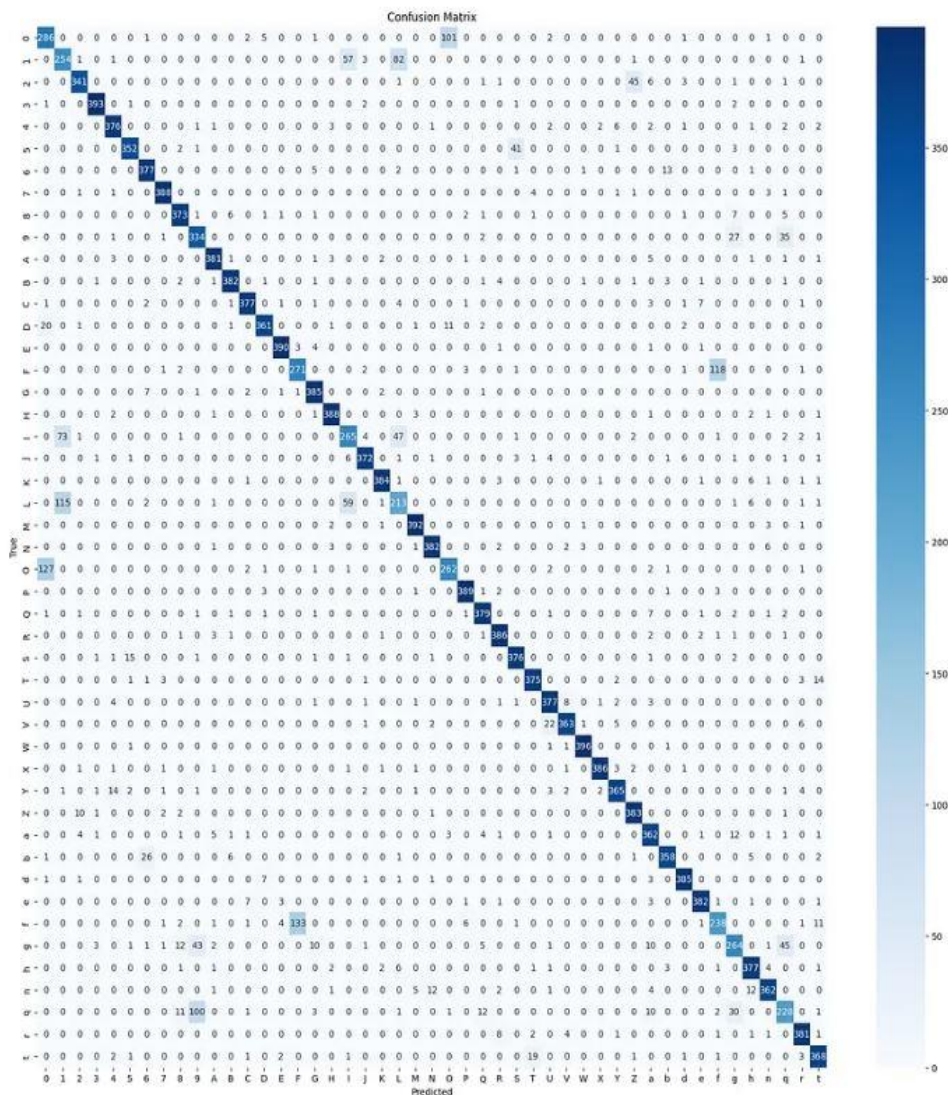


Сл. 13: Функција на чинење над тренирачко и валидациско податочно множество.

На Слика 13, може да се забележи дека се користи cross-entropy како функција на загуба, а загубите за тренирање и валидација започнуваат на релативно високи нивоа. Меѓутоа, како што тренирањето на моделот напредува низ следните епохи, забележливото намалување на двете загуби станива очигледно. Ова намалување ја означува подобрената способност на моделот прецизно да ги погодува излезните цели. Конвергенцијата на функцијата на чинење кај тренирачкото и валидациското податочно множество означува ефективно учење на моделот, каде моделот ги усовршува параметрите за поблиску да се усогласи со вистинската дистрибуција на податоците. Надолниот тек на функциите на чинење покажува дека невронската мрежа успешно се адаптира и ги минимизира грешките помеѓу предвидените и вистинските излези за време на текот на тренирачките епохи. Оваа позитивна патека го нагласува капацитетот на моделот да препознае комплексни задачи и да добро да генерализира и за тренирачкото податочно множество и за невиденото валидациско множество.

Во следното визуелно претставување, на Слика 14 е претставена confusion matrix, каде се нуди преглед на перформансите на класификацискиот модел при препознавање на знаци. Оваа матрица визуелно ги прикажува точните предвидувања по должината на дијагоналата

и погрешните класификации во сите останати елементи кои не се по должина на дијагонала, обезбедувајќи знаење за силните страни на моделот и потенцијалните области за подобрување.



Сл. 14: Confusion matrix на предвидените лабели и вистинските вредности.

Ќелиите на матрицата ја прикажуваат фреквенцијата со која се предвидуваат инстанци од вистинската класа како секоја можна класа. Дијагоналните елементи претставуваат точни предвидувања, формирајќи ја главната дијагонала на матрицата. Елементите кои не се на главната дијагонала покажуваат грешни класификации, нудејќи увид во тенденциите на моделот и области каде што може да има проблеми. Резултантната мапа овозможува јасна визуелизација од перформансите на моделот над различни класи, олеснувајќи ја проценката на неговата точност. Како што може на се види на горната матрица, не е лесно да се направи разлика помеѓу голема буква “O” и бројот “0”, голема буква “L” и бројот “1”, мала буква “r” со голема буква “F”.

IV. ЗАКЛУЧОК

Во ова истражување се запознавме со целосен систем за препознавање на регистарски таблички приспособен за најразлични сценарија. Нашите пронајдоци го демонстрираат моќниот перформанс на предложениот пристап. Клучен пристап е воведување на мрежа која го олеснува препознавањето и откривањето на регистарските таблички од најразлични агли, со генерирање на матрица за трансформација за секоја ќелија за препознавање. Овој чекор ја олеснува задачата на мрежата за препознавање на оптички знаци (OCR), на тој начин што ја намалува дисторзијата. Дополнително, видовме како овие мрежи се совршено прилагодени за препознавање на слики, покажувајќи ја значителната ефективност во снимањето на сложени задачи и репрезентација на слики со високи карактеристики за препознавање.

РЕФЕРЕНЦИ

1. License Plate Detection and Recognition in Unconstrained Scenarios, Institute of Informatics - Federal University of Rio Grande do Sul Porto Alegre, Brazil - https://openaccess.thecvf.com/content_ECCV_2018/papers/Sergio_Silva_License_Plate_Detection_ECCV_2018_paper.pdf
2. A New Implementation of Deep Neural Networks for Optical Character Recognition and Face Recognition - Proceedings of the New Trends in Information Technology, The University of Jordan Computer Engineering Department, Khaled Younis, - https://www.researchgate.net/publication/324877673_A_New_Implementation_of_Deep_Neural_Networks_for_Optical_Character_Recognition_and_Face_Recognition
3. Deep Learning and EMNIST: How to use a Convolutional Neural Network for image recognition, Approaching Image Recognition with a Convolutional Neural Network using Keras - <https://medium.com/@mspeciale46/deep-learning-and-emnist-how-to-use-a-convolutional-neural-network-for-image-recognition-81acbcfa99eb>
4. Detect and Recognize Vehicle's License Plate with Machine Learning and Python - Part 1: Detection License Plate with Wpod-Net, Quang Nguyen - <https://medium.com/@quangnhatnguyenle/detect-and-recognize-vehicles-license-plate-with-machine-learning-and-python-part-1-detection-795fda47e922>
5. An automated license plate detection and recognition system based on wavelet decomposition and CNN - Ibtiissam Slimani, Abdelmoghith Zaarane, Wahban Al Okaishi - <https://www.sciencedirect.com/science/article/pii/S2590005620300254>
6. Improving Warped Planar Object Detection Network For Automatic License Plate Recognition - Nguyen Dinh Tra, Nguyen Cong Tri, Phan Duy Hung - <https://arxiv.org/abs/2212.07066>
7. License Plate Detection and Recognition in Unconstrained Scenarios - Sergio Montazzolli Silva - <http://sergiomontazzolli.com/pubs/alpr-unconstrained/>