

ДЕТЕКЦИЈА ИНДУСТРИЈСКИХ КОМПОНЕНАТА ПРИМЕНОМ YOLOv3 МОДЕЛА DETECTION OF INDUSTRIAL COMPONENTS WITH YOLOv3 MODEL

Филип Брујић, Факултет техничких наука, Нови Сад

Област – МЕХАТРОНИКА

Кратак садржај – У овом раду описани су општи појмови машинског учења, конволутивних неуронских мрежа и машинске визије. У оквиру рада креиран је скуп података за тренирање, валидацију и тестирање модела вештачке интелигенције. На основу скупа решава се проблем детекције индустријских компоненти, применом конволутивне неуронске мреже. Као објекти за детекцију коришћени су пнеуматски цилиндар, разводни вентил и тастер.

Кључне речи: Машинско учење, конволутивне неуронске мреже, пнеуматске компоненте, машинска визија

Abstract – This paper describes the general principles of machine learning, convolutional neural networks, and computer vision. It also outlines the process of creating a dataset for training, validation, and testing of artificial intelligence models. Based on this dataset, the problem of industrial component detection is addressed using a convolutional neural network. The objects for detection include pneumatic cylinders, distribution valves, and push buttons.

Keywords: Machine learning, convolutional neural networks, pneumatic components, computer vision

1. УВОД

Данас, у 21. веку, са развојем технологије, индустрија постаје све напреднија. Постоје различити аутоматски и полуаутоматски системи који омогућавају ефикасну производњу готовог производа, спремног за тржиште. Дobar пример интеграције паметних фабрика је сама Индустрија 4.0. Овај појам карактеришу технолошка решења која су омогућила веће управљање производним системима, флексибилност производње и савремени начин дистрибуције производа. Неки од очекиваних исхода овакве индустријске трансформације, пре свега, јесу смањење трошкова залиха, повећање продуктивности, смањење периода престанка рада машина и повећање тачности предикција.

Машинско учење је подручје вештачке интелигенције које се бави развојем алгоритама и модела који омогућавају рачунарима да "уче".

НАПОМЕНА:

Овај рад проистекао је из мастер рада чији ментор је била др Ивана Шенк, ванр. проф.

Уместо да програмери напишу експлицитна правила за решавање одређених задатака, машинско учење омогућава рачунарима да идентификују обрасце у великим количинама података и да користе те обрасце за доношење одлука. Прикупљање података је први корак у машинском учењу. Подаци могу бити слике, текст, бројеви или било какви други подаци који су релевантни за конкретан проблем који је потребно решити. Након прикупљања података, рачунар користи те податке како би обучио, односно истренирао одређени модел за решавање одређеног задатка. Модел је математички алгоритам који се прилагођава подацима како би идентификовао обрасце.

Тренирање модела значи подешавање параметара тог алгоритма тако да што прецизније предвиђа жељене излазне вредности на основу улазних података. На крају тренирања, тестира се тачност и ефикасност модела. Ово се обично ради на посебним скуповима података који нису коришћени током тренинга.

Детекција објеката, с друге стране, представља специфичну примену машинског учења у области компјутерске визије. Овде рачунарима није само потребно рећи шта се налази на слици, већ и где се то налази. Детекција објеката идентификује присуство и позицију различитих објеката унутар слика или видео записа.

2. ТЕОРИЈСКЕ ОСНОВЕ

2.1 Машинско учење

Машинско учење [1, 2] има многе примене у различитим областима, укључујући препознавање облика, анализу текста, управљање ресурсима, медицинску дијагностику, финансијску анализу, препоручивање производа, аутономне аутомобиле и многе друге. Постоје различите врсте учења, а три основне категорије су надгледано, ненадгледано и појачано учење.

Надгледано учење је једна од најчешћих врста машинског учења. У овој категорији, постоји јасан скуп података који садржи улазне податке и одговарајуће излазне ознаке. Основна сврха је научити модел како да створи неку повезаност или пресликавање између улаза и излаза. Ненадгледано учење долази у игру када не постоје одговарајуће излазне ознаке за улазне податке.

У овом случају, циљ је пронаћи обрасце, односно различита груписања података. Кључна примена ненадгледаног учења је кластеровање. Појачано учење се користи за решавање проблема доношења одлука у интерактивним окружењима.

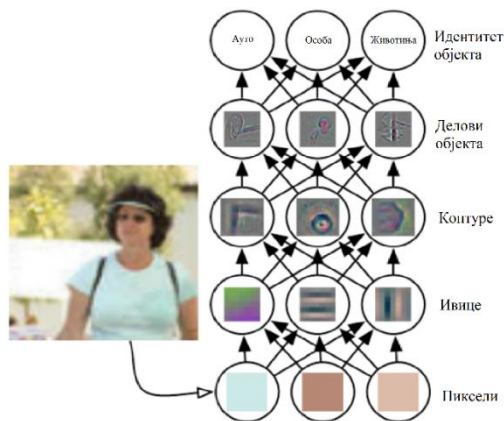
У овом типу учења, агент учи како да предузме одређене акције у окружењу како би максимизовао награду или минимизовао казну.

2.1 Конволуционе неуронске мреже

Конволуционе неуронске мреже [1] (енгл. Convolutional Neural Networks - *CNN*) су изузетно корисне за обраду различитих врста сигнала, укључујући звук, слику и текст. Њихова снага лежи у способности да аутоматски извлаче релевантне карактеристике из сирових сигнала, без потребе за унапред дефинисаним или ручно обликованим карактеристикама. Називају се "конволутивним" зато што уче филтре који откривају одређене карактеристике у сигнаlima.

У контексту обраде сигнала, конволуционе мреже омогућавају аутоматско откривање кључних особина сигнала. Ово је значајно јер су у традиционалном приступу инжењери морали ручно дизајнирати филтре за издвајање одређених карактеристика сигнала, на пример, филтре за откривање ивица на сликама. Једна од кључних предности конволутивних мрежа јесте њихова флексибилност.

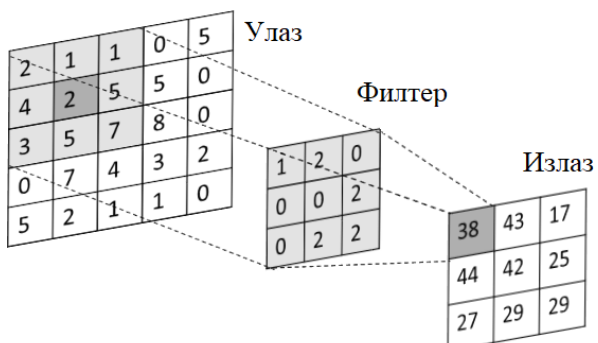
Принцип рада конволутивне неуронске мреже се може видети на слици 1.



Слика 1. Пример рада конволутивне неуронске мреже [1]

2.2 Конволуција

Конволуција је математичка операција која се често користи у обради слика и анализи података. У контексту дводимензионалних матрица, као што су слике, конволуција је начин за комбиновање две матрице, што се може видети на слици 2.



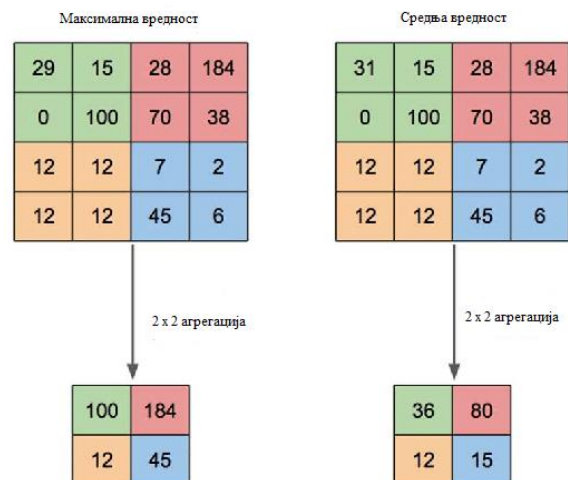
Слика 2. Конволуција две матрице

Да би се одржале димензије матрице улаза након конволуције, често се користи техника позната као проширивање (енгл. *padding*). Проширивање подразумева додавање пиксела око матрице улаза, обично нула или вредности са руба улазне матрице [1].

2.3 Сажимање

Слој сажимања у конволутивним неуронским мрежама има улогу управљања подацима који долазе из претходног конволутивног слоја. Ово се обично постиже применом једноставних функција сажимања на суседне јединице претходног слоја, као што су максимум или средња вредност, које се могу видети на слици 3. Сваки канал конволуције има одговарајући канал сажимања.

На пример, ако се користи сажимање на 3×3 пиксела, број излазних канала овог слоја биће 9 пута мањи од броја излазних канала претходног слоја. Максимална вредност је често популаран избор функције сажимања. Када се 3×3 јединице претходног слоја "проследе" једној јединици слоја сажимања, то значи да ће се сачувати информација о томе да је одређено својство (на пример, усправна линија) пронађено, али се неће сачувати информација о тачном положају тог својства на слици. Овакво занемаривање прецизног положаја често није проблем ако је циљ само препознати присуство одређених карактеристика (на пример, препознати лице). Информација о тачном положају можда није битна у таквим ситуацијама [1].



Слика 3. Два начина сажимања

3. ПОСТАВКА ПРОБЛЕМА

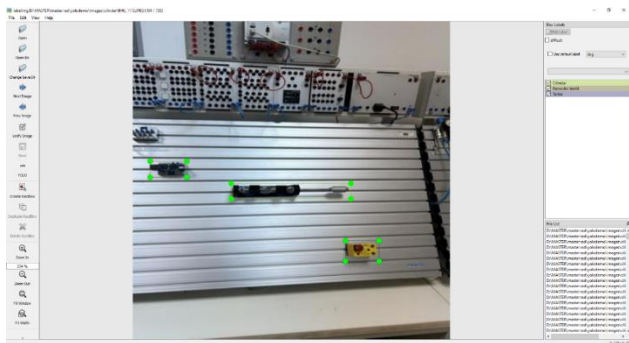
Главни циљ овог рада јесте детекција три индустријске пнеуматске компоненте (пнеуматски цилиндар, разводни вентил и пнеуматски тастер) уз помоћ YOLOv3 (енгл. You Only Look Once) модела машинског учења.

На једној слици могуће је видети један, два или сва три објекта потребна за детекцију. С обзиром на специфичност проблема, потребно је направити одговарајући сет слика за обуку, валидацију и тестирање.

3.1 Припреma сета података за тестирање, валидацију и обуку

Пошто је неуронска мрежа обучавана од почетка потребно је било скупити што више слика компонената. Укупно је искоришћено 1765 слика. За поставку граничних оквира (енгл. Bounding box) коришћен је софтвер LabelImg, чији интерфејс је приказан на слици 4. са сликом са све три компоненте.

Такође, да би се омогућио што већи број узорака, коришћене су различите технике аугментације слике, као што је, на пример, генерисање слика које се разликују по осветљењу, контрасту или слике које су насумично заротиране или исечене.



Слика 4. LabelImg софтвер

3.2 YOLOv3 модел

YOLO [3] је популаран алгоритам за детекцију објеката у реалном времену. YOLOv3 је једна од итерација овог алгоритма и представља унапређену верзију оригиналног модела. Познат је по својој брзини, јер може брзо детектовати објекте у стварном времену, што га чини одличним избором за апликације као што су праћење објеката, аутономна вожња и видео надзор.

Користи три различите скале за детекцију објеката. Свака скала омогућава моделу да детектује објекте различитих величина, што побољшава прецизност детекције.

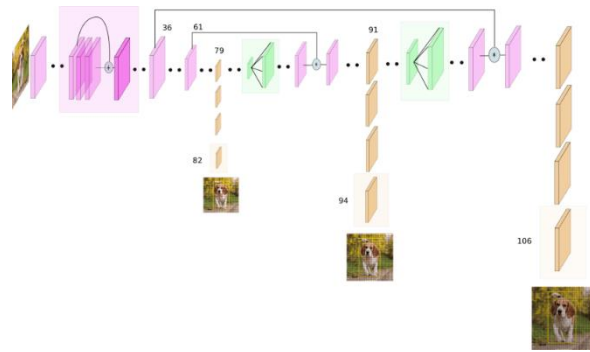
YOLOv3 користи конволуциону неуронску мрежу (CNN) за детекцију објеката на сликама. Мрежа је подељена на три дела: основу, врат и главу. Основа је низ конволутивних слојева који извлаче карактеристике из улазне слике.

Врат комбинује карактеристике различитих скала како би побољшао детекцију објеката.

Глава се састоји од скупа потпуно повезаних слојева који предвиђају локацију и класу сваког објекта на слици. Сама архитектура мреже се може видети на слици 5.

4. РЕЗУЛТАТИ И МЕРЕЊА

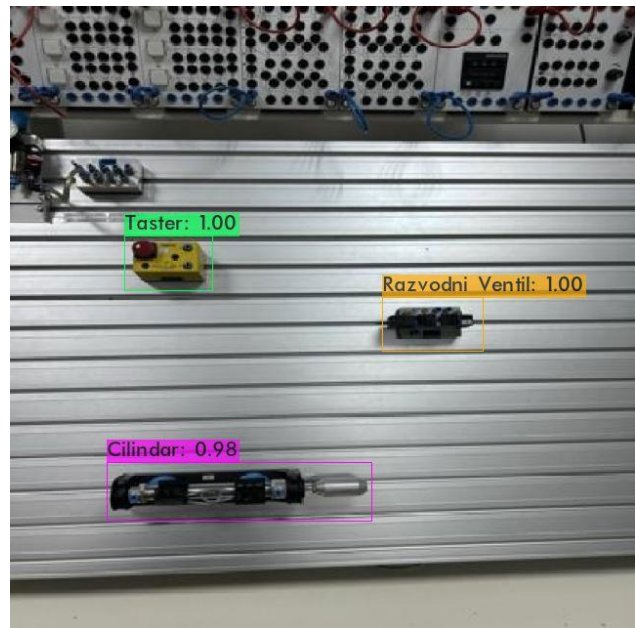
У оквиру овог поглавља представљени су резултати и понашање система током различитих тест ситуација. Првенствено ће се анализирати резултати где се модел налази у различитим срединама.



Слика 5. Архитектура YOLOv3 модела [3]

4.1 Детекција слика из тест скупа

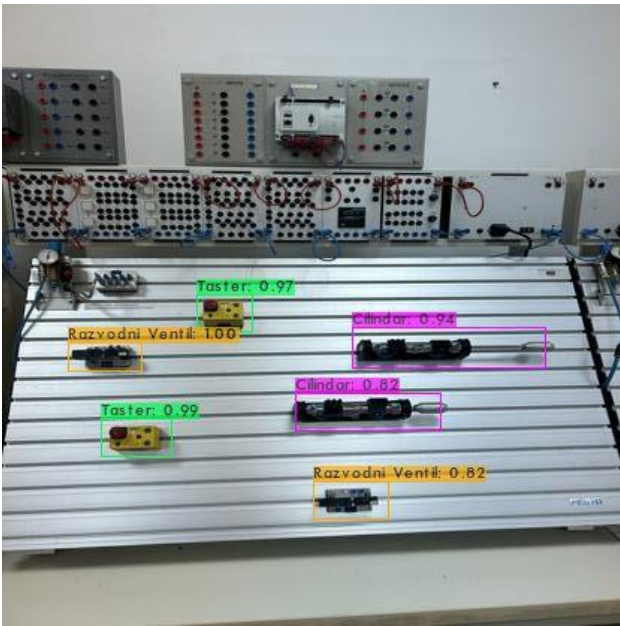
Идеја прављења самог тест скупа јесте приказивање података моделу које до сада није видео, односно на основу којих није био обучен. Гледајући слику 6 може се видети да модел правилно поставља граничне оквире и постиже веома високе проценте тачности за све три компоненте. Потребно је напоменути да су узорци за тренирање прављени тако да се имају у виду различите поставке компонената на шине.



Слика 6. Резултати првог теста

4.2 Детекција када се на слици налази више објеката

Овај експеримент подразумева детекцију објеката, када се на слици налази више од једне компоненте. На свакој слици се налазио по један пнеуматски цилиндар, један вентил и један тастер. Циљ овог теста је видети да ли ће модел успети да се снађе и детектује све објекте за које је обучен. Такође, ниједна слика, са које је модел учио, не поседује више од једног објекта за детекцију. Узимајући слике и видео снимак у обзир, модел најбоље резултате постиже са тастером, затим разводним вентилом и на крају са цилиндром, што се и може видети на слици 7.



Слика 7. Резултати другог теста

4.3 Тестирање детекције када је више објеката спојено

Идеја овог теста јесте да се види да ли је могуће „збунити“ модел да детектује само једну компоненту иако се на слици налазе две. Иако ће објекти бити постављени на већ познату околину, односно шине, и даље то не значи да ће модел моћи са лакоћом да изврши овај задатак. Као што је претходно већ речено, модел нема ниједну сличну ситуацију у тест скупу налик на ову. Посматрањем резултата можемо доћи до закључка да у овом случају модел није успео да изврши тражени задатак. На слици 8 види се да је цилиндар и разводни вентил детектовао као једну компоненту са веома великим процентом тачности. Како би се постигло да модел успешно детектује и овако детаљне случајеве, било би потребно додатно оптимизовати и проширити скуп за тренирање.



Слика 8. Резултати трећег теста

5. ЗАКЉУЧАК

Циљ овог рада био је прављење скупа података и обучавање YOLOv3 модела за детекцију индустријских пнеуматских компоненти. На основу добијених резултата, из различитих експеримената, може се закључити да је модел доста осетљив на промене. Од све три компоненте, тастер се показао најлакшим за детекцију, затим разводни вентил, па пнеуматски цилиндар. Ово истраживање је рађено на основу вишекласне класификације, где постоје три класе. Додатним проширивањем самог скупа података, где би били урачунати и закључци на основу урађених експеримената, постигао би се оптималнији и прецизнији модел.

6. ЛИТЕРАТУРА

- [1] Младен Николић, Анђелка Зачевић, Машинско учење, Београд, 2019.
- [2] D. Michie, D.J. Spiegelhalter, C.C. Taylor, Machine Learning, Neural and Statistical Classification, February 17, 1994.
- [3] <https://machinelearning.space.com/yolov3-tensorflow-2-part-1/> (приступљено у Септембру 2023.)

Кратка биографија



Филип Брујић рођен је у Новом Саду 1999. године. Основне академске студије завршава на Факултету Техничких Наука у Новом Саду, усмерење Мехатроника, 2022. године. Мастер академске студије уписује исте године на смеру Мехатроника. Контакт: filip.brujic@gmail.com