Зборник радова Факултета техничких наука, Нови Сад

UDK: 4.633
DOI: https://doi.org/10.24867/33BE11Knezevic

ПРОШИРЕЊЕ АЛАТА AUTOPSY СА МОДУЛОМ ЗА ДЕТЕКЦИЈУ ОБЈЕКАТА НА ФОТОГРАФИЈАМА

EXTENDING AUTOPSY TOOL WITH OBJECT DETECTION MODULE

Исидора Кнежевић, *Факултет техничких наука, Нови Сад*

Област – ЕЛЕКТРОТЕХНИЧКО И РАЧУНАРСКО ИНЖЕЊЕРСТВО

Кратак садржај – *Дигитална форензика се суочава са изазовом анализе великих колекција фотографија које могу садржати кључне доказе у криминалистичкој истрази. У овом раду је представљено решење модул за Autopsy алат, „Детектор оружја и новца“, који користи YOLOv8 модел дубоког учења за детекцију објеката попут пиштоља, ножева, новчаница и кредитних картица, чиме се значајно убрзава процес анализе. Предложено решење показује знатно боље перформансе у детекцији објеката у односу на постојећа решења.*

Кључне речи: *дигитална форензика, детекција објеката, YOLOv8, Autopsy*

Abstract – *Digital forensics faces the challenge of analyzing large collections of photographs that may contain key evidence in a criminal investigation. This paper presents a solution, module for the Autopsy tool, “Weapon and Money Detector”, which uses the YOLOv8 deep learning model to detect objects such as guns, knives, banknotes and credit cards, significantly speeding up the analysis process. The proposed solution shows significantly better performance in object detection compared to existing solutions.*

Keywords: *digital forensics, object detection, YOLOv8, Autopsy*

1. УВОД

Дигитална форензика је научна дисциплина чији предмет су идентификација, прикупљање, чување, прегледање, анализа и презентација дигиталних доказа коришћењем научно и правно ваљаних метода и алата [1]. У савременој дигиталној форензици, фотографије представљају важан извор доказа који могу садржати информације од суштинског значаја за истрагу. Пораст броја рачунара и других рачунарских система које константно користимо у свакодневном животу, довео је до стварања великих количина слика. Међутим, ручна анализа великих колекција фотографија може бити изузетно временски захтевна за дигиталне форензичаре.

У овом раду акценат је стављен на проналажењу фотографија на којима се налазе објекти као што су пиштољи, ножеви, новчанице и кредитне картице. Њихова аутоматска детекција значајно смањује време потребно за анализу слика и пружа истражитељима могућност да се фокусирају на друге аспекте истраге.

Да би се ова аутоматизација постигла, коришћен је напредни модел дубоког учења YOLOv8 за детекцију наведених објеката. Модел је трениран и евалуиран на скупу података који садржи фотографије са објектима од интереса. Развијен је додатак за форензички алат Autopsy верзије 4.21.0, назван Детектор оружја и новца (енг. Weapons and Money Detector), који омогућава примену овог тренираног модела за аутомаско проналажење фотографија на којима се налазе наведени објекти.

2. ПРЕГЛЕД СЛИЧНИХ СОФТВЕРА

Главни критеријуми за избор сличних апликација били су њихова примена у форензичким истрагама, могућност аутоматске анализе мултимедијалног садржаја и подршка за идентификацију одређених објеката од интереса. Посебна пажња је посвећена алатима који укључују методе машинског учења за детекцију објеката. На основу ових критеријума, идентификовани су и анализирани алати као што су Autopsy, верзије 4.21.0, у поглављу 2.1, и Magnet Axiom, верзије 6.8, у поглављу 2.2.

2.1. Autopsy

Autopsy [2] је алат за дигиталну форензику отвореног кода, који омогућава анализу дигиталних доказа са чврстих дискова рачунара и мобилних уређаја. Дизајниран и имплементиран је модуларно, што значи да постоје модули који проширују функције алата. Један од модула јесте и Object Detection модул који служи за детекцију објеката на сликама. Модул користи каскадни класификатор (енг. Cascade Classifier) алгоритам машинског учења из библиотеке за рачунарски вид и машинско учење OpenCV. Пошто библиотека OpenCV, верије 3.4.16, има само класификаторе за детекцију лица, очију, тела, мачака и регистарских таблица, а и Autopsy не садржи своје класификаторе уз модул, било је потребно тренирати и евалуирати класификаторе за објекте које желимо да детектујемо, а то су: пиштољи, ножеви, новчанице и кредитне картице. Метрике које су коришћене за евалуацију су прецизност, одзив и F1 мера. На табели 1 се могу видети вредности евалуације за сваки класификатор.

НАПОМЕНА:
Овај рад проистекао је из мастер рада чији ментор је био др Стеван Гостојић, ред. проф.

Табела 1. *Резултати евалуације класификатора*

Објекти	Прецизност	Одзив	F1 мера
Пиштољи	0.114	0.472	0.184
Новчанице	0.048	0.245	0.08
Ножеви	0.007	0.087	0.013
Кред. карт.	0.036	0.193	0.06

Резултати евалуације показују да каскадни класификатори имају ниске вредности прецизности и одзива, што указује на велики број лажних позитивних резултата и пропуштање стварних објеката.

2.2. Magnet Axiom

Magnet Axiom [3] је комерцијални софтвер за дигиталну форензику, који је специјализован за анализу података са различитих извора као што су чврсти дискови, мобилни уређаји, cloud сервиси и друштвене мреже. Једна од функционалности је и напредна функција за категоризацију текста у разговорима, слика и видео записа помоћу Magnet AI. Magnet AI је интегрисани систем који користи машинско учење за аутоматску анализу и категоризацију података. Може да идентификује слике које садрже специфичне објекте или сцене, као што су оружје, документа, новац, возила, неподобан садржај и разне друге објекте, У раду [4] Василарас и сарадници су извршили евалуацију функционалности категоризације слика, алата Magnet Axiom, верзије 6.8. Метрике које су коришћене за евалуацију су тачност, прецизност, одзив и F1 мера. Категорије које су евалуиране су: дрога, оружје, голотиња, новац и возила. У табели 2 су приказани резултати евалуације.

Табела 2. *Резултати евалуације алата Magnet Axiom*

Објекти	Тачност	Прецизност	Одзив	F1 мера
Дрога	0.988	0.066	0.279	0.107
Оружје	0.993	0.156	0.052	0.078
Голотиња	0.992	0.231	0.146	0.179
Новац	0.980	0.021	0.333	0.040
Возила	0.991	0.810	0.545	0.652

Резултати евалуације показују да Magnet AI има ограничене перформансе у неким категоријама, попут оружја, дроге и новца, где су ниске вредности прецизности и одзива. Ово указује на повећан број лажних позитивних и пропуштених објеката у тим категоријама. Функционалност категоризације слика помоћу Magnet AI је најпоузданија за категорију возила.

3. AUTOPSY

Овај одељак се бави анализом Autopsy алата за дигиталну форензику. У одељку 3.1 ће бити објашњене основне карактеристике алата и могућностима које нуди корисницима, док ће у одељку 3.2 бити објашњено како се могу проширити функционалности овог алата.

3.1. Основне карактеристике алата

Autopsy се базира на следећим кључним концептима:

- Случај (енг. case): дефинисан је као контејнер који садржи један или више извора података.
- Извор података (енг. data source): могу бити форензичке копије диска, локални дискови и логички фајлови и фолдери.
- Модули за обраду података (енг. ingest modules): обављају анализе фајлова из извора података и парсирају њихов садржај.
- Црна табла (енг. blackboard): представља начин комуникације између модула. Модули постављају своје резултате на црну таблу у облику артефаката, а кориснички интерфејс их приказује.

Неки од модула за обраду података који постоје у Autopsy-у су:

- File Type Identification – идентификује датотеке на основу њихових интерних потписа и не ослања се на екстензије датотека.
- Picture Analyzer – овај модул издваја Exchangeable Image File Format (EXIF) информације из слика, тј. метаподатке. Ове информације могу да садрже податке о геолокацији слике, времену и датуму настанка, моделу камере и њеним подешавањима као и друге информације.
- Object Detection – користи OpenCV да детектује објекте на сликама.

Корисници да би покренули дигиталну форензичку истрагу помоћу Autopsy-а потребно је да направе случај и попуне све неопходне податке о њему. Затим додају извор података и бирају модуле за обраду података који ће се извршавати над тим извором. Након завршетка рада модула, корисници могу ручно да прегледају садржај фајлова и резултате модула како би идентификовали доказе. Када је завршена анализа корисник може да покрене генерисање финалног извештаја на основу одабраних ознака и резултата. Ово укључује креирање HTML или Excel извештаја.

3.2. Проширивост Autopsy-а

Autopsy је препознатљив као један од најприлагодљивијих алата за дигиталну форензику, због могућности проширења функционалности кроз различите модуле. Неке од врста модула које постоје у оквиру Autopsy-а су:

- Модули за обраду података (енг. Ingest Modules): ови модули се покрећу када се нови извор података дода у случај, и могу се поново покренути након тога. Ови модули постоје у два облика: Модули за обраду датотека (енг. File Ingest Modules) – позивају се за сваку датотеку у извору података, и модули за обраду извора података (енг. Data Source Ingest Modules) – позивају се једном за сваку слику диска или сет логичких датотека.
- Модули за извештаје (енг. Report Modules): ови модули се (типично) покрећу након што је корисник прегледао резултате и означио датотеке од интереса. Њихова сврха је креирање извештаја о резултатима, али могу се користити и за анализу.

Autopsy је писан у Java програмском језику, док модули могу бити писани у Java или Python језику (тј.

Jython језику, који претвара код писан у Python-у у Java код). Сваки Python модул, који се пише у једној Python скрипти, треба да буде у фолдеру за Python скрипте у оквиру Autopsy-а. У Autopsy GitHub репозиторијуму [5] се налазе примери за модул за обраду датотека, модул за обраду извора података и модул за извештаје. Сви примери имају TODO напомене које указују на то шта треба да се промени. Све детаљније информације о имплементацији модула, укључујући кораке, примере кода и додатне могућности, могу се пронаћи у званичној Autopsy документацији за девелопере [6].

4. ДЕТЕКЦИЈА ОБЕЈАКТА

Детекција објеката представља један од кључних задатака рачунарског вида и вештачке интелигенције, чији је циљ да препозна и лоцира објекте на дигиталним сликама или видео-снимцима [7]. У одељку 4.1 биће објашњено шта је детекција објеката и како функционише, док ће у одељку 4.2 бити представљени неки од модела за детекцију објеката.

4.1. Основни концепти детекције објеката

Детекција објеката је општи термин који описује скуп повезаних задатака у области рачунарског вида, а који подразумевају идентификацију објеката на дигиталним фотографијама. Општи принцип рада детекције објеката је следећи:

1. Улазна слика: Процес детекције објеката почиње анализом слике или видеа.

2. Предобрада: Слика се предобрађује како би се обезбедио одговарајући формат за модел који се користи.

3. Издвајање карактеристика: Модел разлаже слику на регионе и из сваког региона извлачи карактеристике како би детектовао обрасце различитих објеката.

4. Класификација: Сваки регион слике се класификује у категорије на основу извађених карактеристика. Задатак класификације може да се обавља помоћу SVM-а (support vector machines – тип алгоритма за надгледано учење који се може користити за задатке класификације или регресије) или неуронске мреже која рачуна вероватноћу сваке категорије присутне у региону.

5. Локализација: Истовремено са процесом класификације, модел одређује оквире за сваки детектовани објекат. Ово укључује израчунавање координата оквира који окружује сваки објекат, чиме се објекат прецизно лоцира унутар слике.

6. Сузбијање преклапања (енг. non-max suppression): Када модел идентификује више оквира за исти објекат, примењује се сузбијање преклапања. Ова техника задржава само оквир са највишим нивоом поузданости и уклања све остале који се преклапају.

7. Излаз: Процес се завршава оригиналном сликом на којој су означени оквири и називи класа, које илуструју детектоване објекте и њихове одговарајуће категорије.

4.2. Методе и алгоритми детекције објеката

Детекција објеката ослања се на различите методе и алгоритме који се углавном деле на традиционалне (пре 2014. године) и модерне методе (после 2014. године), засноване на дубоком учењу. Традиционалне методе које се издвајају су: Виола-Џонс детектори (тј. каскадни класификатори), HOG (histogram of oriented gradients) детектор и DPM (deformable part-based model) модел. Модерне методе се деле на две категорије: „двостепена детекција“ и „једностепена детекција“. Код „двостепених детекција“ издвајају се методе: RCNN, SPPNet, Fast RCNN, Faster RCNN, док се код „једностепених детекција“ издвајају методе: you only look once (YOLO) и single shot multibox detector (SSD).

4.2.1. YOLOv8

YOLOv8 је развијен од стране Ultralytics тима. Архитектура YOLOv8 се састоји од два главна дела, кичме и главе, која чине две конволуционе неуронске мреже.

У раду [8] Билоус и сарадници су поредили перформансе модела, као што су YOLOv4-v8, Faster RCNN, SSD и EfficientDet, за детекцију људи и техничких објеката. Главне метрике за поређење биле су тачност (mAP), прецизност, одзив, F1 мера и брзина обраде (FPS). На табели 3 су приказани резултати евалуације модела. YOLOv8 се показао као најефикаснији модел захваљујући својој високој тачности, брзини обраде и способности прилагођавања разноврсним сценаријима. И због тога је YOLOv8 модел изабран за решавање проблема овог рада.

Табела 3. *Резултати евалуације модела*

Модел	Прецизност	Одзив	mAP	F1 мера	FPS
YOLOv4	0.81	0.83	0.82	0.82	40
YOLOv5	0.73	0.73	0.73	0.73	45
YOLOv6	0.62	0.62	0.62	0.62	42
YOLOv7	0.68	0.68	0.68	0.68	43
YOLOv8	**0.87**	**0.89**	**0.88**	**0.88**	**48**
Faster RCNN	0.84	0.82	0.83	0.83	10
SSD	0.76	0.75	0.75	0.75	35
EfficientDet	0.82	0.80	0.81	0.81	30

5. ДЕТЕКТОР ОРУЖЈА И НОВЦА НА ДИГИТАЛНИМ СЛИКАМА

Детектор оружја и новца (енг. Weapons and Money Detector) је модул развијен за форензички алат Autopsy, верзије 4.21.0, који анализира дигиталне слике из извора података и издваја слике на којима се налазе пиштољи, ножеви, новчанице и кредитне картице. Модул је имплементиран у Python скрипти и за детекцију наведених објеката користи YOLOv8 модел, који је трениран и евалуиран на скупу података који садржи фотографије са објектима од интереса. Скуп података над којим је трениран модел преузет је из рада [9] Перез-Ернандеза и сарадника. На табели 4 је приказан скуп података који је коришћен за тренирање и евалуацију модела.

Табела 4. *Скуп података*

	Пиштољи	Новчанице	Ножеви	Кредитне картице
Тренирање	855	425	621	222
Тест	214	102	172	57

Модел је трениран у 25 епоха. Након тренирања је урађена евалуација. Метрике које су коришћене за евалуацију су прецизност, одзив и F1 мера. На табели 5 се могу видети резултати евалуације модела.

Табела 5. *Резултати евалуације YOLOv8 модела*

Објекти	**Прецизност**	**Одзив**	**F1 мера**
Пиштољи	0.913	0.937	0.925
Новчанице	0.911	0.931	0.921
Ножеви	0.930	0.820	0.872
Кред. карт.	0.882	0.474	0.617

Модул Weapons and Money Detector у Autopsy алату садржи фолдер utils, у коме се налазе два фолдера model и dist. У фолдеру model се налази фајл best.pt, који представља YOLOv8 модел, док се у фолдеру dist налази weaponsMoneyDetector.exe датотека.
Корисник када започиње дигиталну форензичку истрагу потребно је прво да направи нови случај и дода извор података. Након тога покреће Weapons and Money Detector модул, који издваја све слике са подржаном екстензијом и копира их у привремени фолдер. Модул затим позива weaponsMoneyDetector.exe датотеку која на сликама детектује објекте од интереса (пиштоље, ножеве, новчанице и кредитне картице) и резултате записује у текстуалну датотеку report.txt. Модул обрађује резултате из текстуалне датотеке и слике са детектованим објектима додаје на Blackboard у секцији Interesting Files груписане у листу под називом Weapons and Money in Images. Након завршетка рада модула, модул обавештава корисника да је процес обраде завршен и корисник може да генерише финални HTML или Excel извештај.

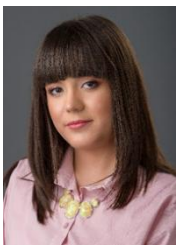
6. ЗАКЉУЧАК

Проблем који је анализиран у овом раду односи се на детекцију потенцијално опасних и значајних објеката, као што су пиштољи, новчанице, ножеви и кредитне картице, у дигиталним сликама. Као метод решавања проблема коришћен је YOLOv8 модел за дубоко учење и развијен је модул за форензички алат Autopsy верзије 4.21.0, Детектор оружја и новца (енг. Weapons and Money Detector), који омогућава примену овог модела за аутоматско проналажење фотографија на којима се налазе наведени објекти. Мотивација за овакав приступ лежи у потреби за ефикаснијом анализом дигиталних доказа у форензичким истрагама, као и у ограничењима постојећих решења, као што су Autopsy и Magnet Axiom, која показују недовољне резултате у прецизности и одзиву њихових модела. Иако YOLOv8 модел показује знатно боље резултате од конкуренције, његова детекција кредитних картица и даље остаје изазов. Правци даљег развоја обухватају побољшање детекције кредитних картица, кроз повећање скупа података за тренирање са већим бројем слика на којима се налазе кредитне картице. Такође, тренирање модела да детектује још више категорија објеката, што ће проширити примену решења у другим областима.

7. ЛИТЕРАТУРА

[1] A. Arnes, Digital Forensics: An Academic Introduction, John Wiley & Sons Ltd, 2018.
[2]"Autopsy User Documentation 4.21.0," [Online]. Available: https://sleuthkit.org/autopsy/docs/user-docs/4.21.0/index.html. [Accessed 7 December 2024].
[3] "Magnet Axiom," Magnet Forensics, [Online]. Available: https://www.magnetforensics.com/products/magnet-axiom/. [Accessed 9 December 2024].
[4] A. Vasilaras, N. Papadoudis and P. Rizomiliotis, "Artificial intelligence in mobile forensics: A survey of current status, a use case analysis and AI alignment objectives," Forensic Science International: Digital Investigation, vol. 49, no. 301737, 2024.
[5] https://github.com/sleuthkit/autopsy/tree/autopsy-4.21.0/pythonExamples [Accessed 9 December 2024].
[6]"Autopsy Forensic Browser Developer's Guide and API Reference," [Online]. Available: http://sleuthkit.org/autopsy/docs/api-docs/4.21.0//. [Accessed 10 December 2024].
[7] J. Murel and E. Kavlakoglu, "What is object detection?" IBM, 3 January 2024. [Online]. Available: https://www.ibm.com/think/topics/object-detection. [Accessed 12 December 2024].
[8] N. Bilous, V. Malko, M. Frohme and A. Nechyporenko, "Comparison of CNN-Based Architectures for Detection of Different Object Classes," AI, vol. 5, pp. 2300-2320, 2024.
[9] F. Pérez-Hernández, S. Tabik, A. Lamas, R. Olmos, H. Fujita and F. Herrera, "Object Detection Binary Classifiers methodology based on deep learning to identify small objects handled similarly: Application in video surveillance," Knowledge-Based Systems, vol. 194, no. 105590, 2020.

Кратка биографија:



Исидора Кнежевић рођена је у Новом Саду 2000. год. Дипломирала је на Факултету техничких наука у Новом Саду 2023. године и исте године уписала мастер студије на Факултету техничких наука, студијски програм Рачунарство и аутоматика, смер Примењене рачунарске науке и информатика – Електронско пословање.