

Sub-proiect 4: Dezvoltarea experimentală a instrumentelor din domeniul realității augmentate la nivelul rețelelor sociale online și studiul impactului acestora la nivelul utilizatorilor” (AR Media)

RAPORT PRIVIND

DESFĂȘURAREA ETAPELOR 4-3, 4-4 și 4.5

Etapa a 3-a 2020

Etapa 4-3: Dezvoltarea și implementarea de opțiuni și servicii (subsisteme) la nivelul rețelei sociale online ce utilizează realitatea augmentată AR Media

Activitatea 4-3-3: Dezvoltarea serviciului de recunoaștere a obiectelor cu ajutorul camerelor foto (object recognition) și afișarea a diverse informații despre acestea, precum traducerea numelor obiectelor în limbi străine în funcție de locația geografică a utilizatorului AR Media

STUDIUL 1: Recunoașterea obiectelor cu ajutorul camerelor foto (object recognition) și afișarea a diverse informații despre acestea, precum traducerea numelor obiectelor în limbi străine în funcție de locația geografică a utilizatorului AR Media

În cadrul acesti activități echipa de cercetare a studiat stadiul actual al cunoașterii din domeniul Computer Vision și a încercat să caute cele mai bune metode pentru realizarea aplicației propuse. Computer Vision este un domeniu interdisciplinar care tratează modul în care o platformă computațională poate fi construită și utilizată pentru a obține o analiză a informației extrase din imagini sau videoclipuri digitale. Din perspectiva ingineriei, Computer Vision automatizează sarcinile pe care sistemul vizual uman le poate efectua. (1)

Sarcinile incluse în această categorie sunt:

- metode de achiziție,
- metode de prelucrare,
- metode de analiză și înțelegere a imaginilor digitale din lumea reală pentru a produce informații

Problema clasică apărută o constituie procesarea imaginilor și capacitatea algoritmilor de a determina dacă datele imaginii conțin sau nu un anumit obiect, caracteristică sau activitate specifică.

În literatură sunt descrise diferite tipuri ale problemei de recunoaștere: recunoașterea obiectelor - poate fi recunoscut unul sau mai multe obiecte preferențiale sau învățate sau clase de obiecte, de obicei împreună cu pozițiile 2D ale acestora în imagine sau poziția 3D în scenă. Blippar, Google Goggles și LikeThat oferă programe de sine stătătoare care ilustrează această funcționalitate; identificare - este recunoscută o instanță individuală a unui obiect (exemple: identificarea feței sau amprentei digitale a unei persoane, identificarea cifrelor scrise de mână sau identificarea unui vehicul specific); detectare - datele imaginii sunt scanate pentru o anumită condiție (exemplele includ detectarea posibilelor celule sau țesuturi anormale în imagini medicale sau detectarea unui vehicul într-un sistem automat de taxare rutieră, detecția bazată pe calcule relativ simple și rapide este uneori folosită pentru găsirea unor regiuni de interes mai mici de date de imagine, care pot fi analizate în continuare prin tehnici mai exigente din punct de vedere computerizat pentru a produce o interpretare corectă). (1)

Pe lângă probleme clasice din computer vision, cum ar fi diferențele între clase sau diferitele unghiuri din care poate fi fotografiat obiectul, localizarea obiectelor trebuie să țină cont de rotirea și scalarea obiectului (obiecte mici), localizarea precisă a obiectului, viteza de detectare etc.

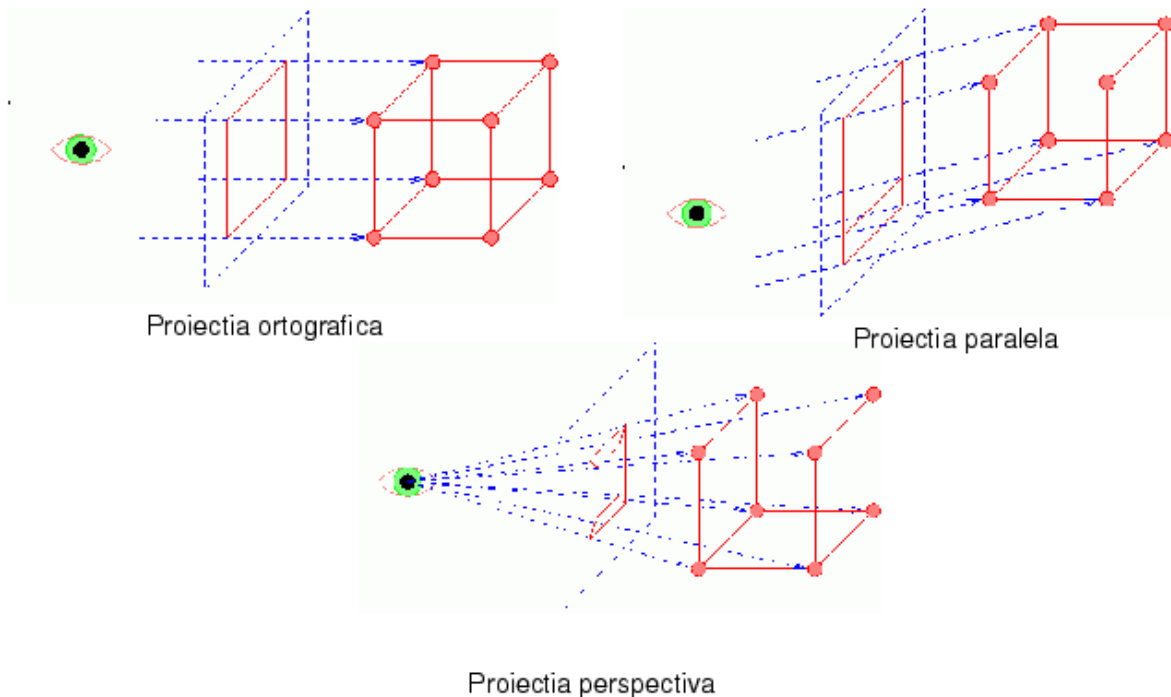
Modelul cel mai frecvent utilizat pentru geometria formării imaginii este cel al camerei obscure, în care imaginea se formează prin trecerea razelor luminoase printr-un orificiu infinezimal și proiectarea lor pe un plan. Într-un aparat tradițional, acest plan este un material fotosensibil; Camerele digitale au planul este compus din elemente electronice sensibile la radiația din spectrul de interes (vizibil, infraroșu, etc.). (2)

Transformarea geometrică în modelul camerei obscure este intuitivă și ușor de exprimat matematic - tehnologia este de *proiecția perspectivă*. Marea problemă a acestei transformări este însă faptul că relația dintre coordonatele punctelor din imagine și coordonatele tridimensionale ale lumii fotografiate este inerent neliniară.

Din această cauză, adesea se folosește un alt tip de proiecție, cea *ortografică*, în care un obiect în spațiul 3D este întâi proiectat perpendicular pe planul imaginii, ignorând *profundimea* (adică un punct (x, y, z) în 3 dimensiuni devine un punct (x, y) în planul imaginii). Apoi această proiecție este redusă la scară, în funcție de distanța medie a obiectului față de cameră. (2)

Modelul acesta, care folosește o proiecție ortografică urmată de o reducere la scară este o aproximare acceptabilă a modelului din perspectivă doar în cazurile în care obiectul este aproape de axa optică a camerei și relativ mic, în comparație cu distanța dintre obiect și aparatul optic. (2)

O altă transformare lineară, este *proiecția paralelă*. Și în acest caz obiectul este proiectat pe planul imaginii și apoi redus la scară, însă, de data aceasta, proiecția este realizată prin raze paralele orientate în direcția "medie" a razelor ce alcătuiesc proiecția prin perspectivă. Această metodă dă rezultate bune chiar dacă obiectul nu este aproape de axa optică. În cazul în care obiectul este chiar pe axa optică, proiecția paralelă și cea ortografică sunt echivalente. Cele trei proiecții sunt ilustrate în figura următoare (2):



Cele trei tipuri de proiecție folosite în modelarea aparatelor care captează imaginea.

Principalii pași în crearea unei aplicații pentru recunoașterea obiectelor (2):

- Captarea imaginii
- Comparația imaginilor
- Crearea bazei de date multimedia
- Recunoașterea obiectelor
- Segmentarea
- Stereoscoopia
- Învățarea

Captarea imaginii - O imagine este o funcție, definită în două dimensiuni a cărei valoare este culoarea (sau intensitatea luminii, în cazul imaginilor alb-negru). Pentru că fiecare element fotosensibil are dimensiuni

finite, în procesul de captare a imaginii domeniul este discretizat într-o grilă; fiecare celulă a grilei corespunde unui element fotosensibil (pixel). Această transformare din spațiul continuu în cel discret se numește eșantionare (sampling). Valoarea culorii sau a intensității luminii este de asemenea într-un spațiu continuu; aparatele vor discerne însă un număr finit de valori diferite. (2)

Comparația imaginilor - O operație simplă este compararea a 2 imagini: căutăm imagini pe Internet sau în biblioteci digitale cu recunoașterea obiectelor (prin compararea cu o imagine-prototip). Algebra ne pune la dispoziție unelte foarte simple cu care putem rezolva în mod foarte eficace măcar o parte din problema comparației. Pentru fiecare imagine "măsurăm" într-un anumit fel o serie de parametri care ni se par importanți pentru clasificare. Pentru fiecare astfel de parametru obținem o valoare. (2)

Iată câteva exemple de parametri:

- mărimea imaginii în pixeli;
- numărul de culori sau nuanțe de gri diferite;
- dimensiunea celei mai mari "pete" de culoare mată din imagine;
- numărul de puncte dintr-o anumită culoare;
- numărul mediu de puncte consecutive dintr-o anumită culoare pe o direcție dată;
- media culorilor din imagine;
- intensitatea maximă și minimă a punctelor din imagine.

Asamblăm pentru fiecare imagine colecția aceasta de valori într-un vector. Dacă avem n valori diferite pentru o imagine, obținem un vector într-un spațiu n -dimensional. În caz extrem, putem vedea întreaga imagine ca pe un vector într-un spațiu $x * y$ -dimensional, unde x și y sunt dimensiunile în pixeli, iar valoarea în fiecare punct din imagine este intensitatea culorii din acel punct.

În general această soluție extremă nu este practică, din două motive (2):

1. Toate imaginile cu care operăm vrem să fie reprezentate de vectori cu același număr de dimensiuni, pentru că aceasta este crucial pentru a le compara.
2. Numărul de pixeli dintr-o imagine este de obicei mult prea mare, comparat cu numărul de trăsături care ne interesează. Adesea vrem să comprimăm descrierea imaginii într-un număr relativ mic de trăsături (de la câteva la câteva sute).

Odată ce avem o descriere a unei imagini printr-un vector, putem folosi două metrici simple pentru a compara vectorii (2):

- Putem calcula distanța dintre doi vectori;

- Putem calcula unghiul plan dintre doi vectori.

A doua metodă este recomandabilă, pentru că este insensitivă la scalarea vectorului (de exemplu, pentru parametrii exemplificați mai sus, unghiul dintre doi vectori nu se schimbă dacă dublăm dimensiunea uneia dintre ele). Ambele metrice se pot calcula foarte ușor: prima necesită n scăderi, iar a doua n înmulțiri și adunări plus un radical.

Baze de date multimedia – Distincția între procesul de căutare a unei imagini după conținut și cel de recunoaștere a imaginii este destul de fină. Numeroase sisteme de recunoaștere pot fi folosite pentru a căuta imagini într-o bază de date. Distincția constă în faptul că, în cazul bazelor de date, imaginile sunt disponibile înainte de a începe căutarea deci ele pot fi pre-procesate. Putem astfel crea structuri de date off-line cu algoritmi prea costisitori pentru recunoașterea interactivă a obiectelor. (2)

Recunoașterea obiectelor - Pentru a recunoaște obiecte, se pornește de la un set de modele ale obiectelor; când o imagine este prezentată sistemului pentru analiză, modelele disponibile sunt folosite pentru a determina care sunt obiectele din imagine. Problemele de rezolvat sunt însă foarte dificile: pentru a recunoaște un obiect într-o imagine, ar trebui să știm cum arată văzut din orice poziție, și cu orice sursă de iluminare. În practică această cerință este imposibil de satisfăcut. În cazul în care avem un model complet al obiectului, algoritmi de recunoaștere încearcă să determine care este poziția și iluminarea cea mai plauzibilă pentru a "explica" imaginea. (2)

Segmentarea - O alta problemă fundamentală este cea de a separa într-o imagine diferitele obiecte care apar. Dificultatea acestei probleme variază în funcție de imaginea cu pricina; cu cât iluminarea este mai uniformă, obiectele au culori mai contrastante și imaginea conține mai puțin zgomot, cu atât problema segmentării este mai simplă. (2)

Printr-o metodă oarecare, marcăm "centrele" obiectelor principale din imagine. Putem face asta fie cu informații de la utilizator, fie alegând punctele de contrast maxim, fie alte informații, depinzând de aplicație.

Calculăm apoi "distanța" între fiecare două puncte vecine în imagine (putem folosi fie patru fie opt vecini) astfel: cu cât culorile sunt mai diferite, cu atât distanța este mai mare. Rulăm apoi un proces iterativ de calcul în care fiecare punct este atribuit în aceeași regiune cu cel mai apropiat vecin. Algoritmul se termină când fiecare punct este atribuit uneia dintre regiunile inițiale. (2)

Stereoscopia- Stereoscopia este procedeul prin care se poate obține informație tridimensională pornind de la două sau mai multe imagini. Principiul este cel pe care se bazează percepția umană. (2)

Învățarea – Practic, orice algoritm care determină parametri de funcționare poate fi atribuit acestei categorii. Un exemplu este "antrenarea" unui sistem pentru segmentare de imagini. (2)

Serviciul de recunoaștere pentru detectarea obiectelor – proiect FUTUREWEB

Serviciul de recunoaștere pentru detectarea obiectelor cu ajutorul camerelor foto este realizat în python ce folosește pytorch și biblioteca de la facebook, Detectron pentru detectarea obiectelor.

PyTorch este o bibliotecă optimizată pentru învățarea profundă folosind GPU-uri și CPU.

Detectron este o bază de cod performantă pentru detectarea obiectelor, acoperind posibilitățile de identificare și segmentare pentru obiectele identificate. Detectron a fost construit de Facebook AI Research (FAIR) pentru a sprijini implementarea rapidă și evaluarea noilor cercetări de viziune computerizată. Acesta include implementări pentru următorii algoritmi de detectare a obiectelor: Mask R-CNN, RetinaNet, Faster R-CNN, RPN, Fast R-CNN, R-FCN.

Detectron poate fi utilizat imediat pentru detectarea obiectelor generale sau modificat pentru a antrena și a executa inferența pe propriile seturi de date. Este scris în Python și va fi alimentat de cadrul de învățare profundă PyTorch 1.0.

Un aspect foarte important, îl reprezintă baza de date ce conține obiectele "cunoscute" de aplicație. Fiecare obiect este descompus în trăsăturile sale esențiale, putând fi recunoscut din orice unghi 3D.

Aplicația realizată are un avantaj evident prin faptul că folosește Deep Learning, ceea ce permite "învățarea" de noi obiecte. Practic, numărul obiectelor din baza de date poate fi nelimitat și astfel gradul de recunoaștere al obiectelor din orice imagine crește semnificativ.

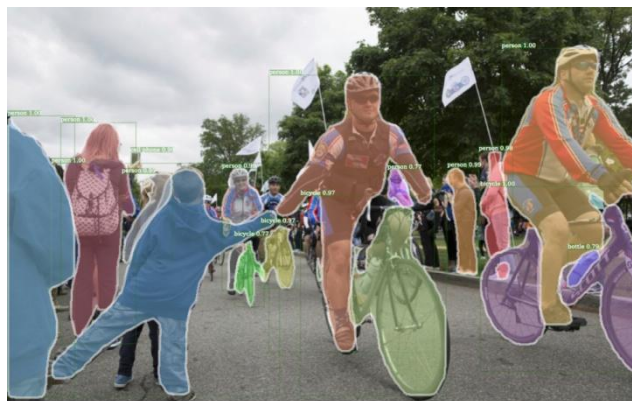


Fig.01 – Exemplu de detectare a obiectelor dintr-o fotografie utilizând cod Detectron

Activitatea 4-3-4: Dezvoltarea serviciului de etichetare (object tagging) prin realitatea augmentată a obiectelor reale din mediul înconjurător și datarea etichetei astfel încât aceste informații să fie disponibile spre vizualizare altor utilizatori înregistrați ai rețelei AR Media

STUDIUL 2.

Etichetarea (object tagging) prin realitatea augmentată a obiectelor reale din mediul înconjurător și datarea etichetei astfel încât aceste informații să fie disponibile spre vizualizare altor utilizatori înregistrați ai rețelei AR Media

Serviciul de etichetare (object tagging) prin realitatea augmentată a obiectelor reale din mediul înconjurător se realizează pentru fiecare obiect care este recunoscut ca fiind existent în baza de date.

Practic, se parcurg următorii pași:

P1 – Se încarcă imaginea care se dorește analizată

P2 – Se analizează imaginea, identificându-se obiectele din imagine

P3 – Pentru fiecare obiect identificat din imagine se face o comparație a similarității cu obiectele din baza de date de obiecte

P4 – Este identificat obiectul din baza de date care are indicele de similaritate maxim cu cel al obiectului din imaginea analizată

P5 – Se aplică o etichetă prin realitate augmentată pentru obiectul identificat, ce conține informații ce se consideră relevante ca descriere a obiectului respectiv. (Aceste informații sunt stocate în baza de date alături de obiecte)



F1-1 - Imaginea inițială ce va fi analizată



F1-2 - Imaginea conținând etichete pt obiectele identificate (persoana 98%, persoana 99%, persoana 100%)



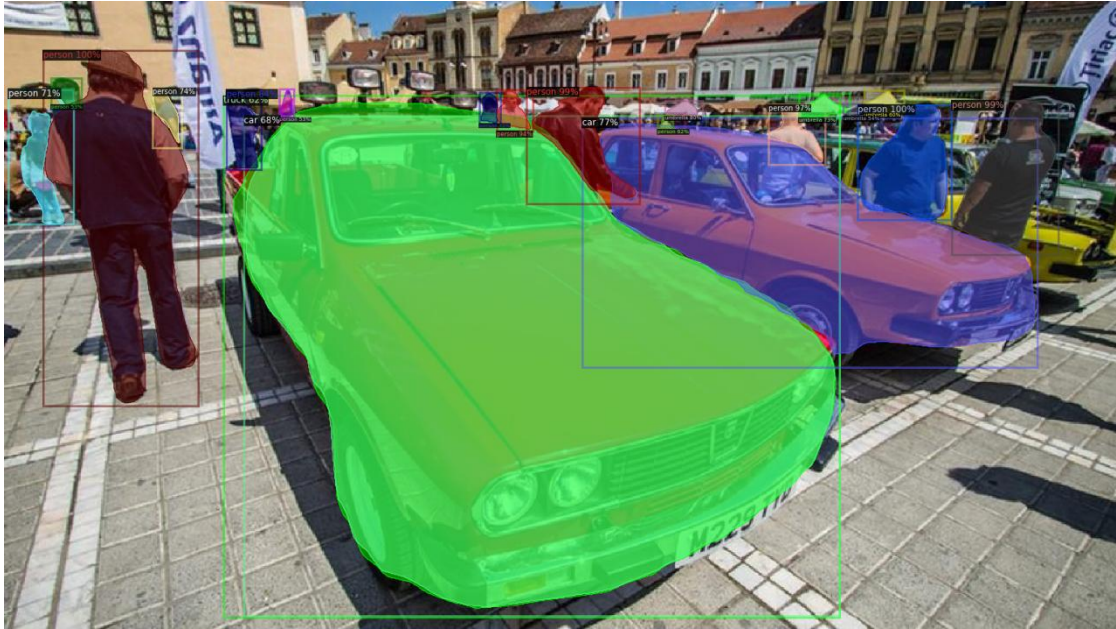
F2-1 - Imaginea inițială ce va fi analizată



F1-2 - Imaginea conținând etichete pt obiectele identificate (persoană 100%, bancă 100%, persoană 85%, bancă 84%, persoană 85%...)



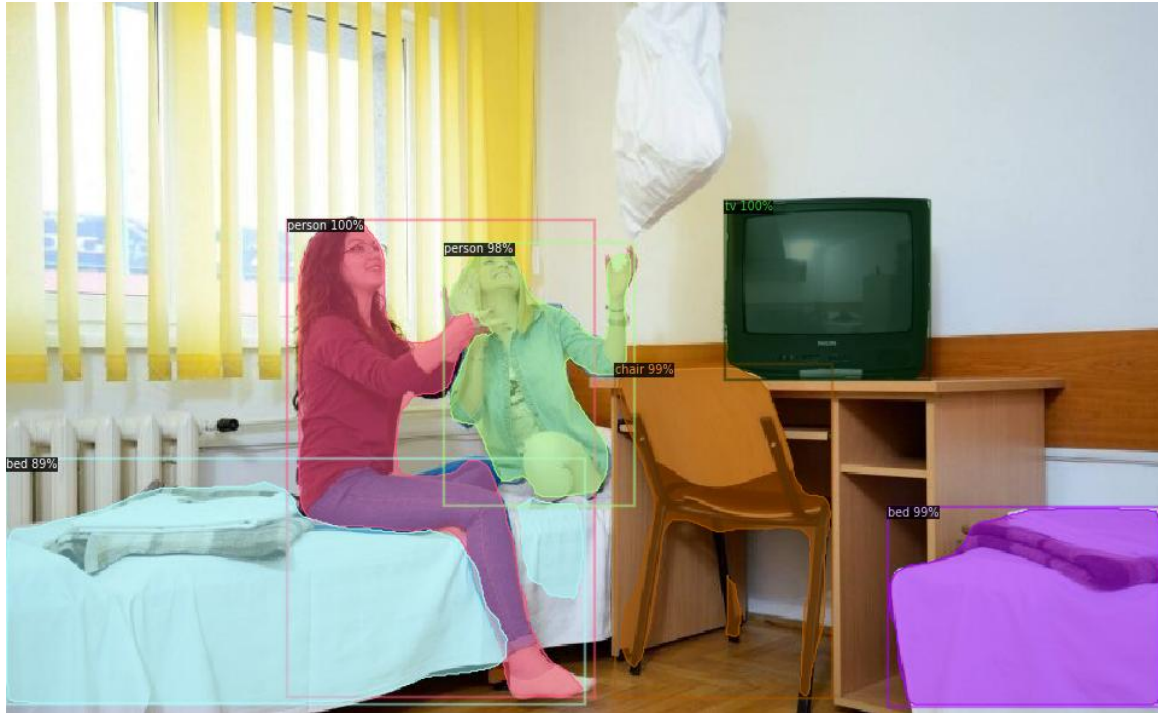
F3-1 - Imaginea inițială ce va fi analizată



F1-2 - Imaginea conținând etichete pt obiectele identificate (persoană 100%, persoană 99%, mașină 77%, mașină 68%...)



F4-1 - Imaginea inițială ce va fi analizată



F1-2 - Imaginea conținând etichete pt obiectele identificate (persoană 100%, persoană 98%, scaun 99%, TV 100%, pat 99%)



F5-1 - Imaginea inițială ce va fi analizată



F1-2 - Imaginea conținând etichete pt obiectele identificate (persoană 100%, persoană 99%, bicicletă 99%, bicicletă 100% ...)

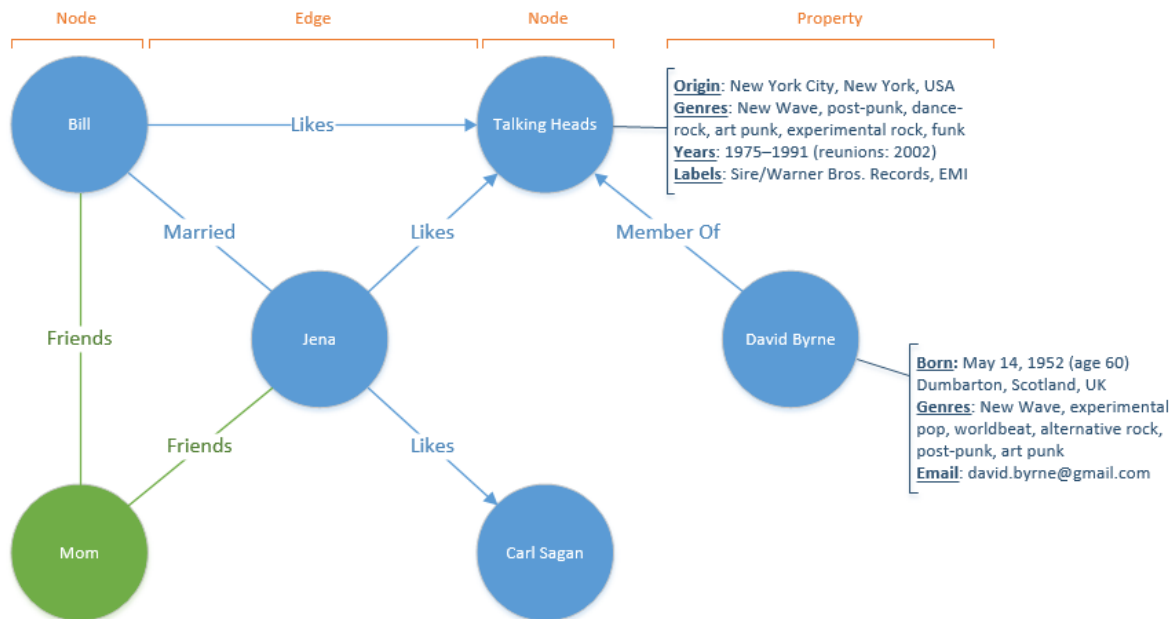
Activitatea 4-3-5: Dezvoltarea serviciului de conectare, preluare și analiză a activității utilizatorilor de pe alte rețele sociale specifice precum LinkedIn, Twitter, Flickr, Youtube, Spotify sau alte rețele în cadrul AR Media pentru a analiza activitatea recentă a acestora, imaginile și clipurile media favorite

STUDIUL 3. Conectare, a preluarea și analiza activității utilizatorilor de pe alte rețele sociale specifice precum LinkedIn, Twitter, Flickr, Youtube, Spotify sau alte rețele în cadrul AR Media pentru a analiza activitatea recentă a acestora, imaginile și clipurile media favorite

Serviciul de conectare, preluare și analiză a activității utilizatorilor de pe alte rețele sociale se bazează pe API-uri specifice fiecărei rețele interconectate. Graph API-ul este principala modalitate de a scoate date din platforma Facebook și, de asemenea, de a posta date pe ea. Este un API bazat pe HTTP pe care aplicațiile îl pot utiliza pentru interogarea datelor, postarea de articole noi, gestionarea anunțurilor,

încărcarea fotografiilor și efectuarea unei game largi de alte sarcini, prin intermediul programelor / scripturilor.

Structura unui Graph API utilizează noduri pentru a obține date despre obiecte individuale. Se folosesc elemente marginale pentru a obține colecția de obiecte asociate unui nod sau pentru a publica obiecte în acele colecții. Se folosesc apoi câmpuri pentru a specifica că toate datele dorite sunt incluse în răspunsuri.



LinkedIn oferă un serviciu care permite oamenilor să-și aducă profilurile și rețelele LinkedIn pe site-ul sau aplicația dvs. prin intermediul API-ului lor bazat pe OAuth.

Totuși, preluarea automată de date trebuie făcută ținând cont de portabilitatea datelor din GDPR - "noul drept la portabilitatea datelor are ca scop responsabilizarea persoanelor vizate în ceea ce privește propriile date cu caracter personal întrucât facilitează capacitatea lor de a muta, copia sau transmite datele cu caracter personal cu ușurință de la un mediu IT la altul"

Etapa 4-4: Dezvoltarea conținutului și promovarea utilizării rețelei sociale online cu realitatea augmentată AR Media și integrarea acesteia la nivelul platformei FutureWeb

Activitatea 4-4-1: Îmbunătățirea și optimizarea rețelei sociale online AR Media prin contribuția directă a membrilor consorțului de cercetare și a utilizatorilor aplicației și integrarea acesteia la nivelul platformei FutureWeb

STUDIUL 4.

Îmbunătățirea și optimizarea rețelei sociale online AR Media prin contribuția directă a membrilor consorțului de cercetare și a utilizatorilor aplicației și integrarea acesteia la nivelul platformei FutureWeb

Un aspect important al aplicației dezvoltate în cadrul proiectului I-a reprezentat îmbunătățirea și optimizarea acesteia.

Pentru aceasta, un rol important I-a avut echipa proiectului de cercetare care a supus la teste atât funcționalitățile practice ale rețelei sociale realizate dar și stabilitate și fiabilitatea în timp.

Concret s-au urmărit dependențele față de alte aplicații online sau baze de date, conexiunea la Internet, gradul de încărcare a bazei de date, etc.

La nivelul serverului s-au analizat listele de log-uri, urmărindu-se eventualele probleme apărute precum și elementele de cauzalitate, încercându-se pentru viitor obținerea unei platforme cât mai stabile.

Etapa 4-5: Diseminarea pe scară largă a rezultatelor proiectului

Activitatea 4-5-1: Diseminarea pe scară largă a rezultatelor proiectului

În anul 2020, membrii echipei de cercetare au realizat și publicat 4 articole cu acknowledgement proiect FutureWeb și o broșură de prezentare a proiectului pe site.

Articole publicate/în curs de publicare:

PERSONALIZED ONLINE MARKETING USING FACIAL AND EMOTION RECOGNITION/2020/Journal of Smart Economic Growth/Radu Lixândriou, Cătălin Maican, Gheorghe Epuran, Gabriel Brătucu, Lavinia Dovleac/Publicat

<https://jseg.ro/index.php/jseg/article/view/109>

AUGMENTED REALITY TECHNOLOGIES IN EDUCATION - A LITERATURE REVIEW /
2020/ Bulletin of the Transilvania University of Brasov/ Daniela Roxana Vuță /în curs de
publicare

http://webbut.unitbv.ro/bulletin/Series%20V/Contents_V_1_2020.html

AUGMENTED REALITY AND FACIAL RECOGNITION TECHNOLOGIES. BUILDING
BRIDGES BETWEEN THE HOSPITALITY INDUSTRY AND TOURISTS DURING PANDEMIC/
2020/ Bulletin of the Transilvania University of Brasov/ Ioana-Simona Ivasciuc/ în curs de
publicare

http://webbut.unitbv.ro/bulletin/Series%20V/Contents_V_1_2020.html

OPPORTUNITIES OF USING NEW TECHNOLOGIES (VR/AR) IN ORDER TO FACILITATE
THE ACCESS OF PERSONS WITH DISABILITIES TO TOURISM PRODUCTS/ 2020/ Bulletin
of the Transilvania University of Braşov/ Gheorghe Epuran, Alina Simona Tecău, Cristinel
Petrişor Constantin, Bianca Tescaşiu, Ioana Bianca Chiţu/ în curs de publicare

http://webbut.unitbv.ro/bulletin/Series%20V/Contents_V_1_2020.html

Borşura de promovare

https://futureweb.unitbv.ro/images/brosura_aplicatie.pdf

Bibliografie:

(1) Samuel Sandu - Modelarea și simularea algoritmului de detecție a obiectelor în Python pentru Raspbery
Pi

(https://www.researchgate.net/publication/346003836_Modelarea_si_simularea_algoritmului_de_detec_tie_a_obiectelor_in_Python_pentru_Raspbery_Pi)

(2) <https://www.cs.cmu.edu/~mihai/articole/vedere/vedere-html.html>

(3) <https://analyticsindiamag.com/top-8-algorithms-for-object-detection/>

- (4) <https://www.analyticsvidhya.com/blog/2018/10/a-step-by-step-introduction-to-the-basic-object-detection-algorithms-part-1/>
- (5) <https://heartbeat.fritz.ai/introduction-to-basic-object-detection-algorithms-b77295a95a63>
- (6) <https://machinelearningknowledge.ai/different-types-of-object-detection-algorithms/>
- (7) <https://cv-tricks.com/object-detection/faster-r-cnn-yolo-ssd/>
- (8) <https://blog.netcetera.com/object-detection-and-tracking-in-2020-f10fb6ff9af3>