Site – Etapa I

Proiectul IDENEO "Identificarea comportamentului neobișnuit al persoanelor în fluxuri video" are ca scop principal dezvoltarea nivelului curent de cunoaștere și competențe în domeniul depistării comportamentului suspect/neobișnuit al persoanelor și mulțimilor în scopul combaterii unor acțiuni ilegale (e.g., violente, teroriste) în anumite circumstanțe. Se permite astfel forțelor de ordine să acționeze preventiv și nu reactiv.

Prima etapă a proiectului a avut în vedere studiul și evaluarea tehnologiilor State-of-theArt pentru dezvoltarea algoritmilor de detecție precum și identificarea limitărilor/constrângerilor de natura algoritmica/tehnica și a posibilităților de îmbunătățire a tehnologiei ce va fi adoptată în implementările viitoare corelată cu cerințele specifice. În cadrul acestei etape au fost prevăzute 7 activități care au fost realizate cu informarea și în parteneriat cu beneficiarul

Rezumat Livrabil I

Activitatea I.1 - *Colectarea şi analiza cerinţelor preliminare pentru soluția dezvoltată, de la beneficiar şi potenţiali utilizatori relevanţi* a avut ca obiective colectarea și analiza cerințelor preliminare provenite de la beneficiarul principal - Serviciul Român de Informații și de la alți potențiali utilizatori relevanți, în vederea proiectării, implementării, testării şioperaţionalizării unei platforme moderne, care să asigure servicii software pentru recunoaşterea automată a comportamentului anormal al persoanelor în fluxuri şi în înregistrări video. Îndeplinirea cerințelor tehnice specifice proiectului determină dezvoltarea unor algoritm software, bazați pe inteligență artificială, pentru identificarea comportamentului anormal individual și de acțiuni. De asemenea, platforma va implementa algoritmi pentru identificarea comportamentului anormal de grup.

Componenta hardware a sistemul informatic dezvoltat va asigura functionalitatea asumată in proiect. Astfel, sistemul va fi dat în exploatare instalat într-un rack independent, care include toate echipamentele de comunicație necesare interconectării la o rețea de camere IP 4K, va dispune de cel puțin un server capabil de procesare paralela, va include un post operator care conține cel puțin o stație PC dedicată, echipată cu mai multe monitoare care să asigure funcționalitățile de client ale platformei, va avea o arhitectură scalabilă și modulară, cu interfață grafică pentru selecția surselor video și a evenimentelor de interes, cu semnalarea acestora în timp real prin mecanisme de alertare și vizualizare georeferențiată, cu înregistrare și analiză post-eveniment după timp și/sau categorii de evenimente.

Rezumat Livrabil I.2

Activitatea I.2 - *Studiul şi documentarea (performanţe, limitări, dependinţe şi cerinţe hardware, etc.) algoritmilor de interes existenţi de procesare a datelor video. Analiză posibilităţi de îmbunătăţire şi nevoi neacoperite în scopul identificării persoanelor de interes* a avut ca obiectiv principal studiul, identificarea și documentarea principalelor direcții de dezvoltare și cercetare, metode și algoritmi de interes pentru beneficiar. Au fost analizate astfel metode de detectare a comportamentului anormal atât la nivel individual cât și în cadrul grupurilor de persoane, și au fost identificate principalele sub-domenii de interes și metode de procesare bazate pe rețele neuronale adânci pentru acele domenii.

O primă secțiune de interes a constat din explorarea avansului tehnologic pentru detecția unui set de acțiuni, cum ar fi alergarea, escaladarea, săritul, etc., metode ce pot fi adaptate și rafinate pentru a detecta un set specific de acțiuni.

Pentru detecția acțiunilor, au fost astfel explorați algoritmi de modelare temporala pentru detecția acțiunilor în fluxuri continue cu sursă deschisă, algoritmi ce exploatează mecanisme de memorie pe termen lung și scurt pentru a modela secvențe de durată continuă. De asemenea, s-au identificat metode ce presupun învățarea unui spațiu compact invariant la translație folosind doar coordonatele spațiale 2D ale părților corporale. Alte abordări au presupus extragerea unei stive de hărți de activări 3D ca reprezentări pentru detecția acțiunilor pe baza scheletonului uman. Pentru recunoașterea acțiunilor bazate pe skeleton, s-au studiat metode pentru a învăța dinamic diferite topologii și a agrega eficient trăsături ale părților corporale.

Detecția staționării îndelungate este în general considerată ca o clasă inclusă în categoria evenimentelor de tip anomalie. In acest context, au fost explorate metode MIL pentru detecția anomaliilor prin exploatarea de etichete artificiale în fluxuri video si metode multi-ramură care extrag reprezentări spațiale și de prim plan din secvențe video.

Detectarea evenimentelor violente, a infracțiunilor, a amenințărilor sau a activităților ilegale este abordată în literatura de specialitate prin dezvoltarea unor algoritmi ce extrag trăsături prin intermediul unor rețele convoluționale și pe care le procesează ulterior prin intermediul unor rețele de tip LSTM. In același context, detecția amenințării cu folosirea sau folosirea de arme albe sau arme de foc se bazează fie pe analiza și procesarea unor descriptori clasici, fie pe soluții bazate pe rețele neuronale adânci.

Detectarea comportamentului disimulat – loitering nu are o definiție exactă în literatură, însă majoritatea lucrărilor se concentrează asupra situației în care o persoană petrece un timp îndelungat într-o anumită zonă, de obicei interzisă. Provocarea majoră în detectarea loitering-ului constă în discriminarea eficientă a modelului comportamental de loitering ofensiv față de cel inofensiv. Urmărirea persoanelor este realizată și prin analiza caracteristicilor biometrice pe termen scurt în cadre succesive (culoarea și textura îmbrăcămintei). Din analiza traiectoriei, când sunt depășite pragurile de timp și deplasări, persoana este tratată ca loiterer.

O altă problemă studiată se referă la agregarea coordonatelor spațiale corespunzătoare membrelor detectate pentru a obține coordonatele globale corespunzătoare alurii sau feței persoanei. Utilizarea unor tehnici de urmărire de coordonate spațiale pot prezice traiectoria, distanțarea socială și viteza de deplasare a persoanelor din scenă iar utilizarea unor metode de modelare a fundalului pe termen scurt, mediu și lung permit extragerea obiectelor staționare (și astfel sunt utilizate pentru detecția bagajelor abandonate sau înstrăinate, sau a altor obiecte de interes ce își schimbă relațiile temporale față de purtător).

Detecția comportamentului anormal de grup în mod coerent este esențială pentru sistemele de securitate moderne, datorită frecvenței tot mai mari a problemelor de siguranță publică. Oamenii din scenele cu mulțime manifestă întotdeauna comportamente consistente și formează mișcări colective. Analiza mișcării colective a motivat o creștere a interesului pentru domeniul de computer vision; in acest caz efortul este îngreunat de natura complexă a mișcărilor colective. Detecția și clusterizarea grupurilor în mulțimi după densitate, energie și viteza de deplasare permit analiza fenomenului de deviație a mișcării care, coroborate cu informații oferite de rețele care permit generarea și clasificarea hărților de densitate permit detecția de anomalii a grupurilor în mulțime precum manifestări de panică sau de agresiune de grup, persoane și grupuri care staționează sau se deplasează în direcții diferite față de mulțime. Este propusă

utilizarea hărților de densitate coroborate cu rețele de atenție pentru estimarea numărului de persoane din scene complexe achiziționate printr-un singur canal video sau mai multe.

Rezumat Livrabil I.3

Activitatea I.3.*Identificarea și analiza senzorilor și surselor posibile de date video de la sol. Analiza modelelor de date, posibilități de interconectare* are ca scop studiul și descrierea surselor existente la momentul actual, în vederea alegerii celor mai potrivite surse pentru acest proiect precum și pentru asigurarea compatibilității soluției dezvoltate cu o gama cat mai mare de potențiale surse de date video. Sursele video utilizate au un rol fundamental în succesul proiectului, contribuind la performanțele finale ale sistemului implementat.

În ceea ce privește sursele de date video în timp real, au fost studiate toate tipurile de surse, atât cele analogice, cât și cele bazate pe IP (Internet Protocol). În plus, s-a avut în vedere și faptul că sistemul dezvoltat va putea să utilizeze și alte tipuri de camere.

In urma evaluării tipurilor de camere existente s-a determinat faptul ca exista avantaje certe ale camerelor video de tip IP față de cele analogice. Astfel, camerele IP oferă în general o calitate video mai mare decât camerele analogice (camerele digitale oferă rezoluții de 6 până la 20 de ori mai mari decât camerele analogice), oferind un câmp vizual mai larg, și capabilități de zoom-are (mărire) mai performante. În plus, sunt oferite mai multe detalii video, ceea ce le face mai potrivite pentru procesarea video. De asemenea, camerele analogice sunt mult mai vulnerabile la breșele de securitate, deoarece fluxurile pot fi interceptate fizic, iar casetele și dispozitivele de înregistrare pot fi furate. Fluxurile video analogice nu sunt criptate. Pe de alta parte, camerele IP fac datele dificil de interceptat. Acestea criptează și comprimă datele înainte de a le transporta prin Internet pe server, și au suport VPN (Virtual Private Network). Este de precizat însă că în ceea ce privește fiabilitatea, ambele tipuri de camere performează bine.

Proiectul actual poate funcționa atât cu surse video live, cât și cu surse video înregistrate. In acest context au fost documentate principalele formate utilizate în stocarea fișierelor video (MP4, AVI, MOV, MKV)

De asemenea a fost studiat si documentat standardul ONVIF, care va fi utilizat în cadrul acestui proiect, pentru a asigura compatibilitatea dintre mai multe dispozitive de rețea. În general, în momentul în care se implementează un sistem în care se dorește configurarea unor dispozitive care aparțin unor producători diferiți, pot apărea probleme în ceea ce privește compatibilitatea dintre camere și înregistratoarele IP. Rolul standardului ONVIF este de a furniza și de a promova interfețe standardizate care să asigure interoperabilitatea eficientă a produselor de securitate bazate pe IP, eliminând în acest mod orice problemă de compatibilitate ar putea apărea. În prezent, acest standard este o asociație de câteva sute de companii, fiind susținut de producătorii de top din industria CCTV.

Ulterior, ONVIF a început să dezvolte și să introducă profiluri standard, care sunt utilizate pentru a îmbunătăți compatibilitatea dispozitivelor. În cadrul acestui sistem se potrivesc 4 dintre profilurile ONVIF: S, T, G, și M.

RezumatLivrabil I.4

ActivitateaI.4 - *Analiza vectorilor UAV și surselor video aeropurtate - comunicații, modele de date, etc*. obiectivele principale ale acesteia (de analiză a vectorilor UAV și a surselor video aeropurtate) a avut ca obiective principale dezvoltarea și operaționalizarea a unor servicii software dedicate de analiză video, pentru detecția și recunoașterea automată în timp real a comportamentului anormal individual, de grup și de acțiuni, în fluxuri video provenite de la

camere de supraveghere și de la sisteme UAV. Astfel, activitatea are ca scop analiza caracteristicilor tehnice ale semnalelor video generate de sistemele UAV actuale și a modalităților de transmitere a acestora către centrul de comandă, în vederea alegerii unei soluții optime, în concordanță cu cerințele de proiectare.

Pornind de la cerințele proiectului, zonele de interes ale câmpului tactic supravegheat vor fi filmate cu camere video Full HD 1080p, existente la bordul sistemelor UAV. Semnalul video, astfel creat, este transmis stației de la sol prin intermediul unui canal unidirecțional de radiocomunicații, de înaltă frecvență, rezistent la interferențe și bruiaj. Fluxul video provenit de la dronă, va fi redirecționat de stația de la sol către platforma VMS, care realizează integrarea, managementul și procesarea specifică a acestuia.

La nivelul legăturii de control se utilizarea intensiv banda ISM 2,4GHz, tehnicile de bandă largă cu spectru împrăștiat, cu salt de frecvență FHSS (FrequencyHoppingSpreadSpectrum) sau cu secvență directă DSSS (Direct SequenceSpreadSpectrum) și unele standarde/protocoale proprietare. Pentru transmiterea fluxului video se utilizează benzile ISM 2,4GHz și 5,8GHz, tehnica digitală OFDM și standardele Wi-Fi 802.11. Legătura de date telemetrică (prin care se trimit către stația de la sol informații despre dronă, legate de coordonatele GPS sau nivelul de încărcare al bateriei acesteia) aceasta este realizata fie tot prin canalul video, fie printr-un canal independent pe frecvențe de 433MHz, 868MHz sau 2,4GHz.

Este de precizat ca in cadrul unui UAS, pentru transmiterea semnalului video de la UAV la centrul (punctul) de comandă îndepărtat este necesară existența unui sistem de radiocomunicații mobil de bandă largă, precum cele din cadrul standardelor Wi-Fi, 4G și 5G.

În urma studiilor efectuate în cadrul acestei activități au fost stabilite caracteristicile principale ale camerei UAV: obiectiv videofocal, explorare întrețesută/progresivă, rezoluție înaltă și dimensiuni mari ale senzorului video CCD/CMOS, raport de aspect 16/9, codare PAL, certificare IP67, funcții PTZ.

În ceea ce privește analiza metodelor de transmitere a fluxului video, care este caracterizat de un debit binar de mare viteză - Mbps, putem sintetiza următoarele: se vor utiliza canale de radiofrecvență de bandă largă (MHz), modulații digitale rapide, rezistente la zgomot - QAM, QPSK, tehnică de transmisie OFDM, fie în cadrul standardelor ISM fără licență, precum Wi-Fi 2,4GHz și 5,8GHz, fie cu licență în cadrul sistemelor de radiocomunicații celulare 4G sau 5G, codare de canal, (criptare flux binar), iar schimbul de mesaje se va realiza într-o arhitectură de tip client-server, cu adresare de pachete IP, transportate utilizând protocoalele TCP sau UDP.

Rezumat Livrabil I.5

ActivitateaI.5 - *Studiul și documentarea reglementărilor legate de GDPR, siguranța și securitatea datelor, managementul identităților*este dedicată investigăriiprincipalelor reglementări referitoare la aplicarea normelor GDPR pe parcursul activităților ce vor fi desfășurate în cadrul proiectului. A fost detaliată modalitatea în care constrângerile GDPR sunt respectate la achiziția seturilor de date necesare antrenării algoritmilor de AI capabili să recunoască comportamente neobișnuite ale persoanelor în fluxuri video, la stocarea și accesul la aceste de date, dar și la managementul identităților și al credențialelor.

Pentru implementarea activităților din proiectul IDENEO, se ridică problema respectării normelor GDPR. Identificarea comportamentului neobișnuit al persoanelor în fluxuri video înseamnă, în foarte multe situații, prelucrarea datelor (sub o anumită formă) extrase din aceste fluxuri. În primul rând, trebuie să stabilim contextul în care sursele acestor fluxuri pot fi

configurate într-o anumită poziție geografică de interes. Există o serie de prevederi legale ce stabilesc condițiile de legitimitate pentru prelucrarea datelor cu caracter personal prin utilizarea sistemelor video, în special în cazul firmelor de pază sau a autorităților, scopul fiind protejarea obiectivelor de interes. Instalarea și utilizarea sistemelor de supraveghere video în astfel de situații se face în conformitate cu prevederile legale din domeniu,în livrabil fiind precizate care sunt aceste prevederi.

Legalitatea prelucrării poate fi stabilită în contextul următoarelor situații: interes legitim, necesitatea prelucrării datelor personale, necesitatea de a îndeplini o sarcină în interesul public sau în exercitarea funcției, punerea la dispoziția organelor de cercetare a imaginilor video și procesarea categoriilor speciale de date.

Un operator sau o terță persoană poate utiliza supravegherea video dacă este necesară pentru îndeplinirea interesului legitim, în situația foarte clară, ca acest interes să fie unul real și de actualitate (ex. prevenirea acțiunilor de vandalism sau distrugere în unele locații publice sau private aflate într-o zonă cu criminalitate crescută).

In contextul surselor publice ne-am orientat către licențele *Creative Commons* (CC) care sunt utilizate pentru a indica drepturile altor persoane în raport cu sursele respective. Fiecare astfel de sursă este protejată automat de drepturi de autor si astfel fiecare utilizator al acesteia trebuie sa ceară permisiunea proprietarului drepturilor de autor.

Deoarece există mai multe licențe *Creative Commons* (CC BY, CC BY-SA, CC BY-NC, CC BY-ND, CC BY-NC-SA, CC BY-NC-ND) prin intermediul cărora de specifică condițiile de utilizare, în cadrul proiectului IDENEO ne-am orientat către sursele publicate sub licențele CC BY (licență care permite utilizatorului să redistribuie, să creeze derivate, să utilizeze publicația pentru activități comerciale, cu condiția să se acorde credit corespunzător autorului (BY) și ca utilizatorul să indice dacă publicația a fost modificată) și respectiv CC BY-SA (SA - sharealike indică faptul că sursa ajustată ar trebui să fie partajată sub aceleași drepturi de reutilizare, deci cu aceeași licență CC). De asemenea, s-au utilizat sursele cu licența CC-0 care implică faptul că autorul a renunțat la toate drepturile de autor asupra lucrării (i.e., lucrarea este dedicată domeniului public și reprezintă faptul că oricine o poate utiliza în orice fel). Este de precizat că nu s-au utilizat sursele conținând acronimele NC (utilizare necomercială) și ND (fără lucrări derivate) - surse care impun constrângeri asupra produselor dezvoltate.

Pentru realizarea videoclipurilor ce conțin scenele corespunzătoare acțiunilor ce se doresc a fi detectate s-a obținut consimțământul liber exprimat al participanților, în deplină cunoștință de cauză a contextului pentru care se efectuează prelucrarea și procesarea datelor personale. Astfel, participanții au fost informați că prelucrarea și procesarea datelor cu caracter personal survenite prin monitorizarea cu camere video sunt realizate cu respectarea legislației în vigoare și sunt necesare în vederea detectării comportamentului neobișnuit al persoanelor în fluxuri video, în cadrul proiectului IDENEO, PN-III-P2-2.1-SOL-2021-0024. De asemenea, participanților li s-a comunicat că activitățile desfășurate vor presupune detectarea persoanelor (incluzând detecția fețelor persoanelor și ale membrelor acestora) precum și a urmăririi acestora în cadrul fluxurilor video, precizându-se totodată ca procesările efectuate nu vor conduce la identificarea unică a persoanelor și vor fi utilizate doar în scopurile anterior menționate.

Pentru asigurarea respectării condițiilor de funcționare în conformitate cu specificațiile GDPR, implementarea algoritmilor în platformă trebuie completată de module GDPR, ce permit ajustarea modului de lucru pentru diferite condiții.

Implementarea activităților din proiectul IDENEO, ridică problema respectării normelor GDPR datorită situațiilor implicate în procesul de prelucrare al fluxurilor video, și anume

identificarea comportamentului neobișnuit din aceste fluxuri video, ce înseamnă, în foarte multe cazuri utilizarea algoritmilor de prelucrare a datelor personale. În activitatea prezentă s-au stabilit, într-o formă parțială, condițiile de utilizare ale operațiilor de prelucrare a datelor ce urmează a fi implementate. Ulterior, selecția și identificarea soluțiilor eficiente pentru implementarea acestor algoritmi va duce la clarificarea situațiilor prezente, și astfel analiza critică legată de conformitatea aplicației cu regulile GDPR va fi finalizată.

Conformitatea cu regulile GDPR a fost sintetizată pentru algoritmii ce urmează a fi dezvoltați în trei categorii importante, prima fiind categoria algoritmilor ce se ocupă de distanțarea socială, traiectoria persoanelor și obiectele abandonate. A doua categorie se referă la mulțimi și în final categoria algoritmilor de identificare a comportamentelor anormale individuale. Fiecare categorie implică o atenție sporită, fiind subliniate situațiile specifice, pentru implementarea algoritmilor astfel încât să fie în conformitate cu regulile GDPR.

De asemenea, este prezentată soluția implementată conformă cu regulile GDPR pentru activitatea de culegere a seturilor de date ce vor fi utilizate în dezvoltarea algoritmilor AI.

Rezumat Livrabil I.6

Activitatea I.6*Analiza și documentarea topologiilor posibile de lucru - ex. procesare pt extragere descriptori Server-side sau on-the-edge, în Cloud Public și utilizare/interpretare Metadate în mediu "confidențial"* este dedicata analizei si documentarii arhitecturilor disponibile pentru dezvoltarea unui sistem de procesare video care are ca scop principal identificarea comportamentului neobișnuit al persoanelor, atât în fluxuri video în timp real, cât și în fluxuri video înregistrate. Ca în cazul oricărui sistem video, sunt produse volume mari de date video. Una din provocările principale ale unui astfel de sistem este transportul datelor produse, în timp real, și într-un mod cât mai sigur, care să asigure confidențialitatea datelor și conformitatea sistemului cu reglementările GDPR.

Analiza video implică generarea automată a metadatelor care descriu ceea ce se întâmplă într-un flux video. Acestea pot fi folosite pentru a detecta comportamentul suspect al persoanelor. Este important ca un sistem de prelucrare video să asigure confidențialitatea acestor metadate. Există diverse moduri de procesare a informației video. Raportul prezent are ca scop principal documentarea și alegerea celei mai bune topologii de lucru, pentru ca acest proiect să fie conform cu cerințele specifice. Sunt prezentate topologiileprincipale ale sistemelor de prelucrare video (arhitectura centralizată/ distribuită) pentru alegerea celei mai bune variante.

În prima parte, am analizat și descris în raport arhitectura, caracteristicile și dezavantajele fiecărei din cele două topologii, cât și scenariile în care fiecare este mai potrivită. Astfel, s-a demonstrat că, în timp ce o arhitectură distribuită este mai potrivită în cazul sistemelor complexe, de sute de camera, în timp ce o arhitectură centralizată este mai utilizată în sistemele de dimensiuni mici și mijlocii.

A fost realizată o comparație între cele două arhitecturi, împreună cu documentarea arhitecturii hibride, care urmărește să unifice cele două arhitecturi prezentate anterior, prin combinarea analizei distribuite cu procesarea pe server și gestionarea bazelor de date.

Alegerea tipologiei de lucru este foarte importantă la implementarea unui sistem de prelucrare video, afectând în mod direct performanțele acestuia, viteza de procesare, cât și securitatea datelor provenite de la sursele din sistem.

Acest raport are ca scop principal documentarea topologiilor de lucru, în vederea alegerii celei mai potrivite soluții pentru acest proiect, pentru a fi conform cu cerințele specific. Astfel, s-au analizat cele mai populare și utilizate topologii, precum procesarea centralizată, de tip server-

side, și procesarea distribuită. Au fost analizate avantajele și, respectiv, limitările fiecărei soluții, și cele mai potrivite scenario de lucru pentru fiecare în parte, dar și o soluție hibridă, care are ca scop minimizarea dezavantajelor unei topologii, prin combinarea tuturor caracteristicilor oferite de cele două topologii principale.

În soluțiile centralizate, analiza video se face la nivelul serverului central, cu fluxul video complet transmis prin rețea. În soluțiile distribuite, software-ul de analiză video este încărcat pe un procesor de semnal digital (DSP - Digital Signal Processor) sau pe un procesor încorporat instalat în camera video la momentul fabricației. În unele cazuri, software-ul este încărcat direct în dispozitivul de supraveghere video ca plug-in, în timp ce dispozitivul se află pe teren.

Arhitectura distribuită are capacitatea de a analiza videoclipul înainte de a-l trimite prin rețea. În schimb, soluțiile bazate pe server oferă mai multă putere de procesare decât poate fi furnizată la dispozitivele din arhitectura distribuită. De aceea, aplicațiile care necesită o putere de procesare mai mare, precum detecția și recunoașterea a diverse obiecte și acțiuni, sunt în mod normal bazate pe server, în timp ce soluțiile distribuite sunt limitate la aplicații mai simple, cum ar fi numărarea persoanelor și tripwire.

În urma analizei, s-a ajuns la concluzia că, pentru un sistem cu un număr redus de camera, cea mai performantă varianta este arhitectura centralizată. Totuși, pentru acest proiect se are în vedere, pe lângă funcționarea centralizată, și o arhitectură distribuită, utilizând ca dispozitive dedicate mai multe module dee tip NVIDIA Jetson, pentru a prelua o parte din rolul serverului din topologia centralizată.

Rezumat Livrabil I.7

Activitatea I.7 - *Colectare baze de date publice adnotate și realizare baze de date video interne proiect adnotate reprezentative pentru proiect* este dedicată documentarii bazelor de date existente (publice/sintetice) si a modalității prin care se pot realiza astfel de baze de date din surse înregistrate.

Modelele de învățare profundă s-au dezvoltat in mod deosebit în ultimii ani, mai ales în cazuri ce au presupus sarcini de detecție, clasificare sau segmentare în cadrul domeniului vederii computaționale (Computer Vision). Această evoluție fulminantă a fost posibilă atât de progresul în dezvoltarea arhitecturilor de rețele profunde, de existența unei puteri de procesare disponibile din ce în ce mai mare, dar și de accesul la seturi mari de date. O opinie general acceptată în domeniul învățării automate este că seturile de date mari implică obținerea unor modele „mai bune" de învățare profundă. Cu toate acestea, există situații când seturile de date disponibile nu reflectă în mod corespunzător elementele ce se doresc a fi detectate/clasificate sau aceste seturi de date sunt supuse unor constrângeri legate de utilizare.

În cadrul proiectului IDENEO vom folosi fie seturi de date fie din surse înregistrate în cadrul proiectului cu acordul GDPR al participanților (Secțiunea 2), fie din surse publice (Secțiunea 3), fie din date sintetice (Secțiunea 4).

În cadrul acestei activități a fost realizat un set de date denumit IDENEO-UOC și colectate informații despre seturi de date internaționale (cu fluxuri video și/sau imagini reale și sintetice – modele de date, incluzând oameni, realizate în platforma Unity). Pentru fiecare caz în parte s-au specificat situațiile care ar putea intra sub incidența constrângerilor impuse de GDPR sau, acolo unde nu este cazul, tipul de licență care stipulează drepturile de folosire ale acestora.

Pentru seturile de date din Secțiunile 2 și 3 ne-am orientat către licențele Creative Commons (CC) care sunt utilizate pentru a indica drepturile altor persoane în raport cu sursele respective. Cele mai multe dintre aceste seturi de date au licența de tip CC BY, licență care

permite utilizatorului să redistribuie, să creeze derivate, să utilizeze publicația pentru activități comerciale, cu condiția să se acorde credit corespunzător autorului (BY) și ca utilizatorul să indice dacă publicația a fost modificată.

Setul de date IDENEO-UOC conține filmări create în contextul proiectului IDENEO de către echipa de la UOC, proiect al cărui scop este detectarea comportamentului neobișnuit în fluxuri video. Setul de date a fost realizat prin regizarea scenariilor specifice obiectivelor propuse în cadrul proiectului si au implicat persoane cărora li s-a prezentat motivația realizării filmărilor și li s-a cerut acordul GDPR. Aceste fluxuri video vor fi folosite pentru detectarea persoanelor (incluzând detecția fețelor persoanelor și ale membrelor acestora) precum și a urmăririi acestora, precizându-se participanților că procesările efectuate nu vor conduce la identificarea unică a persoanelor și vor fi utilizate doar în scopurile menționate și asumate prin proiect.Setul de date denumit IDENEO-UOC conține 72 de fluxuri video, organizate în 20 de scenarii distincte.

Legat de colectarea seturilor de date publice, ne-am orientat către seturi de date care sunt disponibile pentru scopul proiectului nostru, cu licența de tip CC BY (pentru marea majoritate). Pentru fiecare caz în parte, am precizat date statistice aferente, împreună cu un exemplu sugestiv. Toate conțin fluxuri video pentru diverse scenarii, mai puțin setul de date COCO care este format doar din imagini.

În livrabilul respectiv am prezentat și o grupare a acestor seturi de date in raport cu acțiunile de studiat și asumate prin proiect: detecția persoanelor, detecția acțiunilor, identificarea persoanelor, detecția si recunoașterea fețelor, detecția obiectelor, estimarea traiectoriei persoanelor, distanțare socială, interacțiune om-obiect, detecția violenței, detecția comportamentului disimulat – loitering, detecția si clusterizarea în mulțimi, estimarea numărului de persoane, traiectoria grupurilor, detecția deviației mișcării, detecția anomaliei de grup.

Algoritmii dezvoltați vor fi integrați în platforma *VMS KVision*. Clientul de vizualizare este conceput pentru a fi cât mai configurabil și poate fi setat pentru a vizualiza simultan un număr mare de fluxuri video cât și monitorul de alarme, făcând operarea extrem de ușoara.

Soluția propusă și asumată în proiect va conține mai mulți algoritmi care vor fi testați pe seturi de date dintre cele propuse în această activitate. Astfel, avem din punct de vedere al comportamentului persoanelor: detecția loitering - comportament disimulat, staționare îndelungată, alergare, escaladare, cădere, comportament agresiv, bagaje abandonate și evaluarea distanțării sociale; din punct de vedere al comportamentului mulțimilor: evaluarea densității, vitezei și direcției, estimarea numărului de persoane, detecția panicii sau violenței, staționării grupurilor și contra-flux; asigurarea conformității GDPR.

Algoritmii propuși pentru acest proiect sunt bazați pe învățare profundă, prin rețele neuronale optimizate pentru a răspunde cerințelor stringente actuale folosind tehnici care să permită înțelegerea raționamentului de decizie a modelelor.

Ținând cont de complexitatea domeniului și de numărul larg de tipuri de evenimente care intră în definiția conceptului de comportament agresiv propunem utilizarea unui număr mare de baze de date diferite pentru antrenare.

Diseminarea rezultatelor proiectului în aceasta etapa inițială a proiectului a avut în vedere dezvoltarea cadrului de cercetare-dezvoltare în acord cu tematica și obiectivele asumate la nivelul partenerilor. In acest sens s-au desfășurat activități ce au implicat studenți de la cele trei cicluri de studii (licența, masterat, doctorat). Astfel, s-a asigurat promovarea proiectului în rândul noii generații de cercetători.

Popularizarea activităților realizate s-a materializat în mediul online prin intermediul site-ului web dar si prin postări în social-media. De asemenea, s-au emis comunicate de presa, proiectul IDENEO având astfel și o expunere în presa.

O alta direcție importantă a fost constituită de organizarea conferinței *4th International Workshop on Research&Innovation for Secure Societies @ IEEE COMM 2022*, conferință ce va atrage un număr însemnat de cercetători în tematica generala a proiectului și care va constitui o platformă de popularizare a rezultatelor obținute.