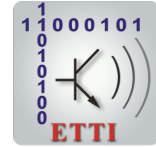




**UNIVERSITATEA POLITEHNICA
DIN BUCUREȘTI**



**Școala Doctorală de Electronică, Telecomunicații
și Tehnologia Informației**

Decizie nr. XXX din DD-MM-YYYY

**REZUMAT TEZĂ
DE DOCTORAT**

Ing. Mihai DOGARIU

**ÎNVĂȚARE ADÂNCĂ PENTRU APLICAȚII CU DATE
INSUFICIENTE**

DEEP LEARNING FOR SCARCE DATA APPLICATIONS

COMISIA DE DOCTORAT

Prof. Dr. Ing. Mihai CIUC Univ. Politehnica din București	Președinte
Prof. Dr. Ing. Bogdan IONESCU Univ. Politehnica din București	Conducător de doctorat
Prof. Dr. Eng. Constantin PALEOLOGU Univ. Politehnica din București	Referent
Cercetător Senior Dr. Ing. Hervé Le Borgne CEA LIST, Franța	Referent
Cercetător Senior Dr. Ing. Michael Riegler Arctic University of Norway, Norvegia	Referent

BUCUREȘTI 2021

Rezultatele prezentate în această teză au fost parțial obținute cu sprijinul
Ministerului Fondurilor Europene prin Programul Operational Capital
Uman, Contract nr. 51675/09.07.2019, cod SMIS 125125.

Cuprins

1	Introducere	1
1.1	Prezentarea domeniului tezei de doctorat	1
1.2	Scopul tezei de doctorat	2
1.3	Conținutul tezei de doctorat	2
2	Recuperarea Imaginilor	3
2.1	Aspecte Teoretice	3
2.1.1	Meta Căutarea Imaginilor	3
2.1.2	Recuperarea Imaginilor Bazată pe Conținut	3
2.1.3	Evaluare	4
2.2	Recuperarea Momentelor Lifelog	4
2.2.1	Prezentarea Literaturii Specifice	4
2.2.2	Abordarea Propusă	5
2.2.3	Concluzii	8
2.3	Recuperarea Bagajelor Abandonate	9
2.3.1	Prezentarea Literaturii Specifice	9
2.3.2	Abordarea Propusă	9
2.3.3	Concluzii	11
3	Generarea Nesupervizată de Date	12
3.1	Prezentarea Literaturii Specifice	12
3.1.1	Autoencoder	12
3.1.2	Autoencoder Variațional	12
3.1.3	Rețele Generative de Potrivire a Momentelor	13
3.1.4	Rețele Generative Adversariale	13
3.2	Generare de Logo-uri	13
3.2.1	Prezentarea Literaturii Specifice	13
3.2.2	Abordarea Propusă	14
3.2.3	Concluzii	15
3.3	Generarea Seriilor Temporale Financiare	15
3.3.1	Prezentarea Literaturii Specifice	15
3.3.2	Datele Financiare	16

3.3.3	Abordarea Propusă	16
3.3.4	Concluzii	21
4	Rezumatul Contribuțiilor și Lucrări Viitoare	22
4.1	Rezumatul Contribuțiilor	22
4.2	Contribuții Originale	23
4.3	Perspective de Viitor	24
4.4	Publicații	24
	Bibliografie	28

Capitolul 1

Introducere

1.1 Prezentarea domeniului tezei de doctorat

Datorită performanțelor superioare pe care le au metodele de învățare adâncă față de algoritmi clasici de învățare automată, învățarea adâncă a devenit abordarea de facto în domeniul multimedia. Aceasta a beneficiat atât de cercetarea teoretică, cât și de cea practică, iar acest aspect dublu poate fi considerat un criteriu de clasificare pentru majoritatea lucrărilor din domeniul multimedia. Progresele teoretice se concentrează pe găsirea celor mai înguste limite asupra modelelor matematice care pot caracteriza complet anumite sarcini, în timp ce cercetarea practică necesită atât inspirație, cât și intuiție pentru a proiecta noi sisteme care funcționează în sisteme reale. Aceste două trăsături ale învățării adânci sunt complementare și împărtășesc o complexitate nativă. Este un domeniu în care descoperirile majore într-una din cele două părți pot propulsa următoarea descoperire în cealaltă parte.

Indiferent de algoritmul pe care îl proiectăm sau îl implementăm, trebuie să îi demonstrăm aplicabilitatea printr-o validare extinsă. În mod tradițional, acest lucru a fost realizat prin raportarea rezultatelor pe seturi de date disponibile în mod deschis, care oferă premisele atât pentru antrenare, cât și pentru validare, dar acestea deseori nu pot ține pasul cu diversitatea aplicațiilor. Acest lucru ridică mai multe probleme din punct de vedere al procesului de colectare a seturilor de date, dintre care majoritatea sunt legate de dimensiunea setului de date. Este important să avem seturi de date suficient de mari, de preferință formate din exemple etichetate și adnotate cu atenție. Acest tip de date este de obicei dificil de găsit sau colectat, în plus față de necesitatea de a avea mai multe date disponibile pentru aplicații de învățare adâncă. Ne referim la seturi de date care nu au multe date etichetate pentru sarcini precise ca fiind insuficiente.

1.2 Scopul tezei de doctorat

În această teză abordăm problema învățării adânci pentru aplicații de date insuficiente. Astfel, aducem o contribuție laturii practice a domeniului multimedia, cu algoritmi care implică recuperarea imaginilor și generarea nesupervizată de date pentru aplicații cu date insuficiente. Identificăm două moduri de a face față unor astfel de probleme.

Prima abordare este de a proiecta algoritmi de învățare adâncă care funcționează cu foarte puține exemple. Ideea principală este să pornim de la modele care au capacități de generalizare foarte bune și să le rafinăm pe aplicații foarte precise. Acest lucru ne-a determinat să găsim abordări hibride între descriptorii clasici și învățarea adâncă.

A doua idee este de a extinde seturile de date insuficiente cu ajutorul modelelor generative. Este un fapt cunoscut că formarea și colectarea unui set de date este un proces complex care necesită o supervizare foarte strictă și resurse semnificative, iar rezultatul final ar putea face parte din cercetările moarte dacă este legat prea strict de o sarcină anume, fără nicio altă aplicabilitate pentru alți cercetători. Astfel, folosim seturi de date deja existente și le extindem prin generarea de eșantioane sintetice, pentru a îmbunătăți capacitățile predictive ale modelelor antrenate pe ele.

1.3 Conținutul tezei de doctorat

Teza conține 3 capitole. Primul este dedicat recuperării imaginilor în diverse aplicații care funcționează cu seturi de date insuficiente. Aceasta este compus din: o introducere teoretică privind recuperarea imaginii, o secțiune care prezintă algoritmi cu care am participat în cadrul unei competiții de benchmarking de recuperare a momentelor de tipul lifelog pe parcursul a 3 ani, respectiv o secțiune dedicată unui sistem de detecție automată a bagajelor, care a fost conceput pentru sistemele TV cu circuit închis.

Al doilea capitol abordează problema generării de date. Capitolul debutează cu un breviar teoretic despre modelele generative nesupervizate folosite. Ulterior, prezentăm progresul în domeniul generării de logo-uri. Scopul acestei părți este de a extinde seturile de date existente cu scopul de a îmbunătăți performanțele de detecție a logo-urilor și de a adăuga mai multă variabilitate seturilor de date prin utilizarea propagării înapoi a gradientului. A treia secțiune a acestui capitol se concentrează pe generarea de date financiare sub forma unor serii temporale. Această parte explorează un set mare de modele și metrici generative care pot fi utilizate pentru a evalua „realismul” probelor sintetizate. Deoarece nu există lucrări în literatura de specialitate care să acopere această ramură pentru seriile de timp financiare, considerăm că lucrarea noastră este un punct de plecare pentru cercetările viitoare.

În ultimul capitol al tezei prezentăm un rezumat al contribuțiilor originale și al rezultatelor pe care le-am obținut. Apoi oferim o perspectivă asupra perspectivelor noastre viitoare. În cele din urmă, enumerăm publicațiile care au validat această teză.

Capitolul 2

Recuperarea Imaginilor

Acest capitol conține lucrările dezvoltate în contextul recuperării imaginilor pentru aplicațiile de lifelog și de supraveghere.

2.1 Aspecte Teoretice

Recuperarea informațiilor este procesul de găsim a informațiilor dorite dintr-o colecție de date. Informațiile dorite se numesc de obicei “interogare de căutare” sau, pur și simplu, “interogare”. De obicei, interogarea nu identifică în mod unic un singur obiect din colectarea datelor. În schimb, corespunde unui set de astfel de obiecte, dar cu diferite grade de relevanță. Astfel, rezultatul se prezintă sub forma unei liste de entități clasificate în funcție de cât de bine corespund semnificației interogării de căutare.

2.1.1 Meta Căutarea Imaginilor

O modalitate de a efectua căutarea de imagini este bazată pe metainformații, cum ar fi etichete, cuvinte cheie sau fraze cheie. Fiecare imagine din setul de date este etichetată automat (sau manual) cu mai multe concepte, utilizate pentru a crea un sistem de indexare inversă. Apoi, interogarea introdusă de utilizator în motorul de căutare este tradusă în cuvinte cheie relevante și extrasă din indexul inversat, ordonat folosind o măsură de relevanță.

2.1.2 Recuperarea Imaginilor Bazată pe Conținut

O alternativă la meta căutarea imaginilor este recuperarea imaginilor pe baza conținutului furnizat de algoritmi de computer vision. Baza de date constă în descriptori extrași de algoritmi de computer vision din fiecare imagine individuală. În timpul procesului de căutare, aceiași descriptori sunt extrași din interogare și comparați cu cei stocați în baza de date. Imaginile care prezintă cea mai mare asemănare cu interogarea sunt returnate utilizatorului în ordinea descrescătoare a relevanței lor.

2.1.3 Evaluare

Deoarece recuperarea imaginilor este de așteptat să returneze un set de imagini care răspund la o interogare, este important să existe un mecanism care să cuantifice performanța unui astfel de sistem. Valorile utilizate în literatură pentru a evalua cât de relevante și complete sunt rezultatele obținute sunt **precizia** și **reamintirea**. Precizia reprezintă raportul dintre numărul de documente relevante recuperate și numărul total de rezultate recuperate. Reamintirea reprezintă raportul dintre numărul de documente relevante recuperate și numărul total de rezultate relevante.

Nici una dintre cele două valori menționate anterior nu este suficientă pentru a caracteriza complet rezultatele recuperării. Acesta este motivul pentru care aceste două valori sunt utilizate de obicei în tandem, fie exprimate ca două valori separate, fie combinate, sub forma măsurii F : $F_\beta = (1 + \beta^2) \frac{P \cdot R}{\beta^2 \cdot P + R}$. De obicei, $\beta = 1$ și măsura F devine măsura F_1 , reprezentând media armonică dintre precizie și reamintire: $F_1 = 2 \cdot \frac{P \cdot R}{P + R}$.

2.2 Recuperarea Momentelor Lifelog

Lifelogging este ansamblul de activități prin care o persoană își înregistrează activitățile zilnice într-un jurnal digital. De cele mai multe ori, aceste informații nu sunt etichetate, ceea ce face dificilă organizarea lor și navigarea, iar recuperarea evenimentelor specifice devine o sarcină complexă. Campania de benchmarking ImageCLEF și-a propus să atenueze aceste probleme prin introducerea sarcinii Lifelog Moment Retrieval (LMR). Următoarele subsecțiuni descriu progresul nostru în timpul competițiilor LMR 2017, 2018 și 2019.

2.2.1 Prezentarea Literaturii Specifice

LMR este un concept destul de nou, care a fost întâlnit în primul rând la campanii de benchmarking precum ImageCLEF și NTCIR. Prezentăm pe scurt cele mai relevante progrese în aceste competiții. O lucrare abordează un lanț care efectuează segmentarea setului de date de evaluare pe baza introducerii manuale a etichetelor temporale și a metadatelor constând din concepte. O abordare similară implică rularea fiecărei imagini printr-un detector de obiecte, numărarea persoanelor și detectarea locurilor, după care se aplică o măsură de similaritate pe descriptor. Una dintre abordările noastre [5] a folosit o distanță de similaritate bazată pe WordNet. De asemenea, a fost propusă o abordare în care obiectele, caracteristicile scenei și acțiunile au fost extrase și combinate împreună cu descrieri textuale într-un index inversat. Abordările noastre [6, 7] propun o tehnică similară, aplicând un sistem de prag pentru încetșoare ca prim pas, apoi extragând mai multe caracteristici precum locuri, concepte, obiecte și combinându-le cu informații textuale.

2.2.2 Abordarea Propusă

A. Seturile de Date O descriere exactă a activităților de lifelogging implică captarea cât mai multor aspecte senzoriale posibile, cum ar fi mișcarea, locația, acțiunile etc. Considerăm că este esențial să înțelegem bine datele din registrul de viață disponibile pentru a realiza o recuperare exactă. Un punct forte al competiției ImageCLEF LMR este că organizatorii au lansat în fiecare an un set de date nou, de fiecare dată mai divers. Rezumăm în tabelul 2.1 principalele aspecte ale seturilor de date care au fost utilizate pentru dezvoltarea algoritmilor de recuperare de lifelogs.

Tabel 2.1 Rezumatul seturilor de date LMR care au fost utilizate din 2017 până în 2019. “yes” / “no” marchează prezența/absența caracteristicii respective din setul de date de dezvoltare.

	LMR_2017	LMR_2018	LMR_2019
users	3	1	2
images	88k	80k	81k
image concepts	ImageNet	Microsoft	Microsoft
locations	130	135	61
activity	yes	yes	yes
biometrics	no	yes	yes
music	no	yes	yes
places	no	no	yes
objects	no	no	yes
dev set topics	5	10	10
test set topics	10	10	10

B. Recuperarea Momentelor Lifelog la ImageCLEF 2017 Lucrarea noastră [5] în timpul ImageCLEF 2017 Lifelog Moment Retrieval parcurge diagrama prezentată în Fig. 2.3.

Fiecărei imagini din setul de date al competiției i se atribuie un set de 1000 de concepte cu încrederi diferite. Păstrăm doar cele mai relevante (cea mai mare încredere)

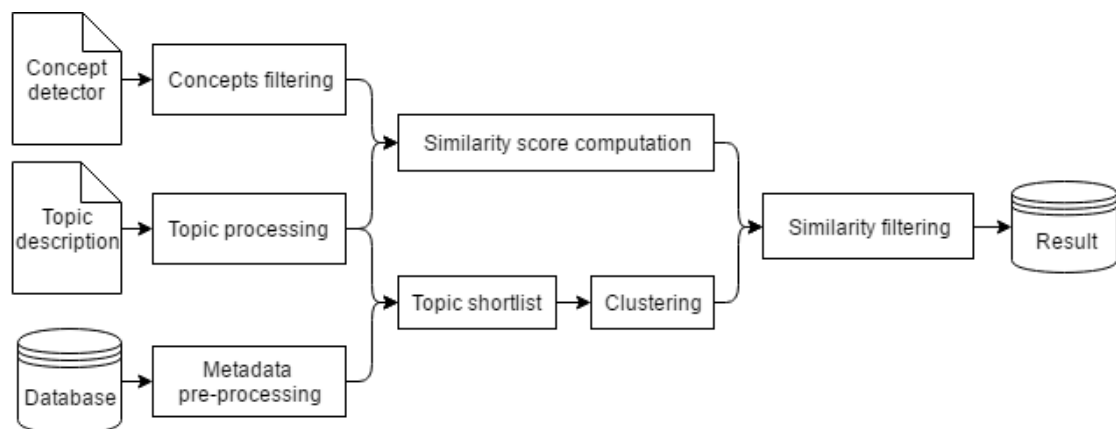


Figura 2.3 Lanțul de procesare LMR 2017 [5].

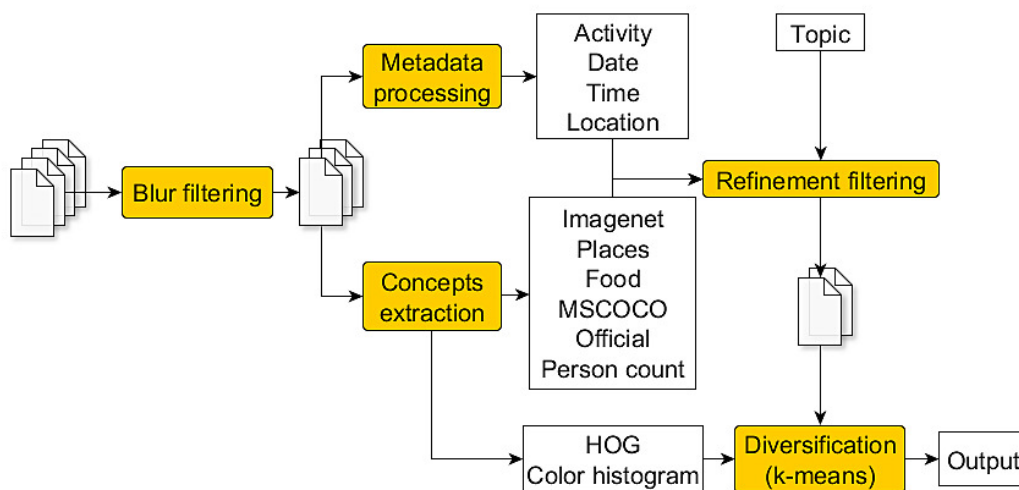


Figura 2.8 Lanțul de procesare LMR 2018 [6].

concepte pentru fiecare imagine. Apoi, propunem o abordare care utilizează instrumente WordNet și păstrează doar substantivele și verbele din descrierea textuală care constituie interogarea noastră. În cele din urmă, pentru fiecare subiect obținem niște informații grosiere, dar fundamentale, legate de acea interogare specifică. Apoi am eliminat imaginile care sunt în disonanță cu metadatele interogării (numărul utilizatorului, locația, activitatea) reducând drastic dimensiunea setului de date relevant. Ulterior, am implementat un algoritm ierarhic de clusterizare bazat pe histograma gradientilor orientați (HOG) extras din fiecare imagine și am oprit algoritmul ierarhic de clusterizare când s-au format 30 de cluster. Măsura de similaritate pe care am propus-o a fost calculată între un set de concepte și un set de cuvinte care corespund unei anumite descrieri a subiectului. Pentru aceasta, am folosit măsura de similitudine Wu-Palmer a WordNet. Fiecare distanță este ponderată cu încrederea conceptului și adăugăm toate distanțele dintre perechile (*concept detectat, cuvânt de descriere*). Odată ce obținem clusterele și scorurile de similitudine dintre fiecare imagine din cluster și descrierea subiectului, selectăm cei mai buni candidați pentru trimiterea unui test. Sortăm clusterele în ordine descrescătoare pe baza valorii medii a scorurilor de similaritate ale imaginilor pe care le conține. În final, am selectat cele mai bine clasate 2 imagini din 25 de cluster. Valoarea oficială pentru rularea noastră a fost **F1@10=0.132**.

B. Recuperarea Momentelor Lifelog la ImageCLEF 2018 În această subsecțiune prezentăm algoritmul cu care am participat la campania de benchmarking ImageCLEF 2018 [6]. Lanțul de procesare al sistemului nostru este prezentată în Fig. 2.8.

Mai întâi eliminăm imaginile neclare care au o măsură de focalizare sub un prag impus. Apoi, rulăm fiecare imagine prin mai multe clasificatoare și un detector: clasificatoare Imagenet, Places365, Food101 și un detector de obiecte Faster R-CNN [24]. Clasa “persoană” a detectorului de obiecte a fost de asemenea folosită pentru numărarea

Tabel 2.3 Rezultatele oficiale LMR 2018 pentru rulările trimise.

Run	$F1@5$	$F1@10$	$F1@20$	$F1@30$	$F1@40$	$F1@50$
Run 1	0.235	0.216	0.224	0.218	0.203	0.199
Run 2	0.154	0.169	0.215	0.21	0.207	0.199
Run 3	0.158	0.168	0.217	0.214	0.199	0.206
Run 4	0.129	0.166	0.184	0.184	0.178	0.188
Run 5	0.412	0.443	0.446	0.438	0.419	0.405

persoanelor. În plus, am folosit conceptele oficiale, data, ora, activitatea și metadatele de locație asociate fiecărei imagini, furnizate de organizatori. Am implementat două tipuri de rafinare. În primul rând, interpretăm manual întregul text al subiectului și extragem constrângeri semnificative asupra metadatelor asociate fiecărei imagini și le eliminăm pe cele care nu satisfac constrângerile date. În mod similar, eliminăm imaginile care conțin anumite obiecte/concepte. Apoi calculăm scorul de relevanță ca o sumă ponderată între toate confidențele conceptului detectat și vectorii de referință corespunzători. Fiecare imagine este apoi reprezentată prin concatenarea a doi vectori normalizați: un vector 1536-D reprezentând histograma gradientilor orientați (HOG) și un vector 512-D reprezentând descriptorul histogramă de culoare. Executăm algoritmul K-means cu 5, 10, 25 sau 50 de clustere. Pentru lista finală a imaginilor propuse, selectăm din fiecare cluster imaginea cu cel mai mare scor de relevanță într-un mod round-robin.

Am trimis o rulare în timpul competiției și alte 4 rulări după încheierea competiției. În Tabelul 2.3 vă prezentăm rezultatele finale $F1@X$ pe care le-am obținut pentru fiecare rulare cu cele mai bune valori cu text îngroșat. Ultima noastră rulare a obținut cele mai bune rezultate, deoarece a implicat o abordare intens supravegheată. **Run 1** urmează lanțul descris mai sus, pentru **Run 2** am exclus imaginile nou adăugate, care sunt prea asemănătoare vizual cu cele deja din listă, pentru **Run 3** am construit vectorii de referință cu aceeași tehnică pe care am folosit-o în [5], **Run 4** run a fost similar cu **Run 3**, cu singura diferență fiind că toate ponderile au fost setate la 1. În cele din urmă, **Run 5** a fost realizat cu aceeași abordare ca și **Run 1**, de data aceasta efectuând o rafinare a tuturor parametrilor de sistem pentru subiectele care aveau probleme prin încercare și eroare. Rezultatul oficial pe care l-am obținut a fost cel al primei runde, **$F1@10=0.216$** .

D. Lifelog Moment Retrieval at ImageCLEF 2019 Pentru ediția din 2019 de recuperare a momentelor Lifelog, abordarea noastră [7] s-a axat pe excluderea imaginilor neinformative ca prim pas, și apoi calculează un scor de relevanță pentru subsetul de imagini rămas. Din experiența anterioară, am observat că a fi mai strict cu criteriile de excludere a imaginilor neinformative duce la rezultate mai bune. Arhitectura sistemului nostru poate fi văzută în Fig. 2.10.

Lanțul de procesare începe rulând un sistem de detectare a neclarității, calculând varianța kernel-ului Laplacian pentru fiecare imagine. Imaginile care nu îndeplinesc un

anumit prag sunt eliminate. Apoi, metadatele imaginilor sunt verificate pentru a fi în conformitate cu restricțiile impuse de subiectul interogat. Informațiile despre numărul utilizatorului, locația, ora, acțiunea, fusul orar sunt apoi utilizate pentru a elimina o altă parte a imaginilor rămase. În unele cazuri, această selecție de metadate poate suferi modificări de la un subiect la altul. În acest moment, setul de imagini a scăzut drastic în comparație cu setul de date original. Apoi, trecem setul de imagini rămase prin procesul de calcul al scorului de relevanță.

Setul de date de dezvoltare conține informații despre atributele, categoriile și conceptele detectate în fiecare imagine. Atributele se referă la diferite aspecte ale peisajului, obiectelor și clasificării imaginilor. De asemenea, am ținut evidența numărului de detecții ale fiecărui obiect din fiecare imagine, deoarece fiecare obiect ar putea declanșa mai multe detecții în aceeași imagine. Calculăm scorul de relevanță ca suma confidențelor caracteristicilor care trebuie detectate în imagine. Spre deosebire de abordarea noastră din 2018 [6], nu folosim o sumă ponderată, deoarece ponderile trebuie reglate manual pentru fiecare interogare individuală și s-ar abate prea mult de la ideea procesării automate. În cele din urmă, am trimis cele mai bune 50 de imagini clasate per subiect, în funcție de scorul de relevanță.

Facem observația că utilizatorul trebuie să selecteze manual parametrii atât pentru restricțiile de metadate, cât și pentru lista de elemente care determină scorul de relevanță, iar aceasta este singura intrare manuală solicitată de la utilizator. Din câte știm, nu a fost încă dezvoltată o metodă clară privind modul în care acest proces de reglare a parametrilor ar putea fi complet automatizat. Rezultatul oficial al rulării noastre a fost **F1@10=0.127**.

2.2.3 Concluzii

Am participat la 3 ediții consecutive ale campaniei de benchmarking ImageCLEF și am propus 3 abordări diferite pentru a rezolva recuperarea momentelor Lifelog. În tabelul 2.8 prezentăm rezultatele oficiale ale sistemelor noastre de-a lungul anilor, împreună cu cel

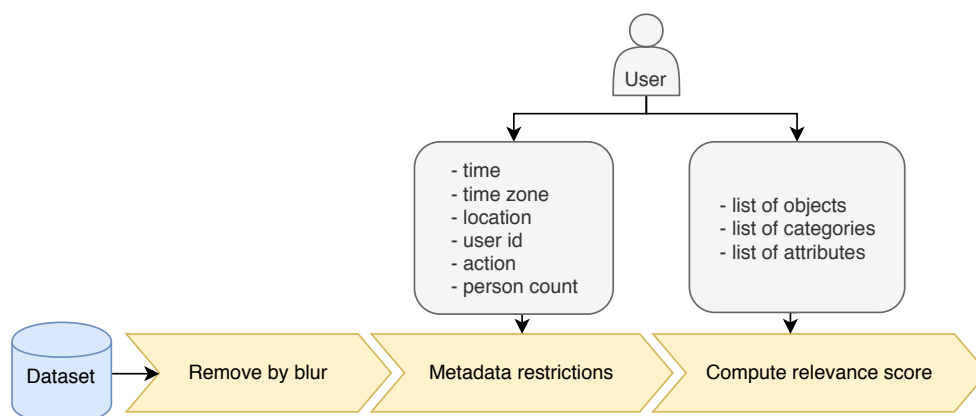


Figura 2.10 Lanțul de procesare LMR 2019 [7].

mai bun scor și baseline și tragem mai multe concluzii. În primul rând, proiectarea un sistem care să răspundă la întrebări de recuperare cu grade nelimitate de libertate este extrem de dificilă. În al doilea rând, rulările cu adevărat automate vor rămâne în urma acelor încercări care implică intervenția umană. În cele din urmă, întrucât seturile de date lifelog acoperă un interval de timp mic, majoritatea evenimentelor vor avea loc doar de un număr foarte limitat de ori (posibil doar o singură dată), transformând această sarcină într-o problemă de manipulare a datelor insuficiente.

Tabel 2.8 Rezultatele oficiale $F1@10$ LMR din 2017 până în 2019. Comparatie între Comparison baseline, rularea noastră și cel mai bun rezultat al competiției.

Year	2017	2018	2019
Baseline	0.103	0.131	0.289
Our result	0.132	0.216	0.127
Best result	0.497	0.545	0.61

2.3 Recuperarea Bagajelor Abandonate

Această secțiune acoperă un alt tip de recuperare a imaginii pe care o abordează această teză, și anume recuperarea bagajelor abandonate. Această problemă a devenit mai importantă odată cu creșterea numărului de cazuri de explozivi ascunși în pachete obișnuite, cum ar fi rucsaci, valize sau pungi și pot fi plasate în locuri aglomerate, neobservate. Întrucât raportul dintre imaginile în care sunt prezente bagaje abandonate și numărul total de imagini captate de un sistem de supraveghere este aproape de zero, ne confruntăm cu o problemă de date insuficiente.

2.3.1 Prezentarea Literaturii Specifice

Cele mai multe abordări din literatura de detecție a bagajelor abandonate se concentrează pe separarea semantică a fundalului de prim-plan și apoi pe urmărirea atât a obiectelor statice, cât și a celor în mișcare. Spre deosebire de acestea, propunem un sistem compus din trei module: o componentă de detectare a obiectelor, un subsistem de detectare a suspecților și o componentă de reidentificare a persoanei [9]. Folosim Mask R-CNN [14] pentru a efectua detectarea obiectelor și pentru a extrage toate caracteristicile relevante. O lucrare similară cu a noastră a fost realizată de Intel [2], dar abordarea lor este limitată doar la detecția obiectelor.

2.3.2 Abordarea Propusă

Sistemul nostru folosește arhitectura Mask R-CNN, antrenată pe setul de date MS-COCO. Din clasele de obiecte le selectăm doar pe cele care au impact asupra sistemului nostru

direct și le eliminăm pe restul. Vom descrie acum setarea pe care am folosit-o pentru fiecare dintre cele 3 module și felul în care acestea interacționează.

Detecția Bagajelor Abandonate Algoritmul de detectare a bagajelor abandonate folosește mecanismul Mask R-CNN. Ne propunem să detectăm doar un subset de clase: “persoană”, “rucsac”, “geantă de mână” și “valiză”. Ultimele 3 clase au fost grupate într-o clasă generală considerată bagaj. Am etichetat un obiect ca fiind abandonat atunci când caseta de delimitare a obiectului nu intersectează caseta de delimitare a unei persoane detectate.

Detecția Suspecților Folosim descriptorul generat de Rețeaua de Propuneri Regionale (RPN) a Mask R-CNN și căutăm bagajul abandonat prin toate imaginile disponibile și clasăm aceste imagini în ordinea descrescătoare a similarității cu bagajul, cu condiția să existe în imagine o persoană a cărei casetă de delimitare intersectează caseta de delimitare a bagajelor. Testăm similitudinea utilizând distanța euclidiană. Ulterior, efectuăm un clasament al acestor distanțe și afișăm imaginile unde este cel mai probabil ca bagajul abandonat să fi fost detectat în prezența unei persoane și considerăm că această persoană este suspectă.

Re-identificarea Suspecților În pasul următor, sistemul nostru pornește de la suspectul care tocmai a fost detectat și îl caută în fluxurile camerei. Am folosit aceeași procedură pentru a căuta persoana pe care am folosit-o pentru bagaje. Totuși, de data aceasta, am efectuat un clasament per cameră și am obținut pentru fiecare cameră un set de imagini în care a fost detectat suspectul. Acest lucru este motivat de faptul că vrem să urmărim traseul acestei persoane pe tot perimetrul supravegheat.

Setul de Date În plus, am creat un mic set de date în scopuri demonstrative. Am strâns 1 oră de imagini înregistrate de sistemul CCTV al centrului nostru de cercetare. Am restricționat zona observată la subsol, parter și exteriorul clădirii. Am stabilit un scenariu în care o persoană ar abandona un rucsac pe hol și ar pleca. Mai multe alte persoane care purtau rucsaci au fost capturate în acest set de date. În plus, am creat o interfață grafică cu scop demonstrativ pentru a ajuta un operator uman.

Validare și Rezultate Am efectuat teste pe mai multe arhitecturi ale miezului sistemului de detecție de obiecte și am raportat precizia medie pentru scorul Intersecție peste Uniune (IoU) mai mare de 0,75¹ împreună cu timpul necesar fiecărei arhitecturi pentru a procesa o imagine în tabelul 2.9. Arhitecturile miezului trebuie citite după cum urmează:

¹Rezultatele performanței detecției pe baza de date MSCOCO raportate de Facebook: https://github.com/facebookresearch/detectron2/blob/master/MODEL_ZOO.md

R/X (ResNet sau ResNeXt), 50/101 (numărul de straturi), C4/DC5/FPN (ResNet conv4, ResNet conv5 sau ResNet + FPN, respectiv), 1x/3x (multiplicatorul numărului de epoci).

Tabel 2.9 Performanța diferitelor modele de detecție de obiecte.

Backbone	Bbox AP@IoU=0.75	Inference time (s/image)
R50-C4_1x	35.7	0.392
R50-DC5_1x	37.3	0.408
R50-FPN_1x	37.9	0.228
R50-C4_3x	38.4	0.398
R50-DC5_3x	39.0	0.396
R50-FPN_3x	40.2	0.231
R101-C4_3x	41.1	0.482
R101-DC5_3x	40.6	0.474
R101-FPN_3x	42.0	0.308
X101-FPN_3x	43.0	0.591

Am obținut timpii de inferență prezentați pe un singur GPU NVIDIA QUADRO M4000. Considerăm că este mai bine să optăm pentru un model mai rapid în defavoarea unui model mai precis. Pierderea preciziei de detecție poate fi compensată prin stabilirea unui prag de detecție mai mic pentru a forța propuneri suplimentare și a reduce rata falsului negativ. Cu toate acestea, scăderea timpului de inferență este mult mai dificilă. În cazul nostru de utilizare, un răspuns rapid este un aspect critic al sistemului. Prin urmare, selectăm modelul R50-FPN_3x pentru sistemul propus.

Componenta de reidentificare a persoanei a fost testată pe setul de date CUHK03 și a obținut o precizie top-1 de 70,8%. Aceeași tehnică a fost utilizată de Xiao et al. [29]. În timpul demonstrației noastre am reușit să surprindem toate evenimentele care sunt de interes: să detectăm bagajul abandonat, persoana care l-a lăsat acolo și apoi să detectăm prezența acelei persoane pe camerele individuale. Mai mult, am reușit să extragem momente importante, cum ar fi când persoana a intrat în clădire în timp ce transporta rucsacul și când persoana a părăsit clădirea, fără rucsac.

2.3.3 Concluzii

Am prezentat un detector de bagaje abandonate care poate fi implementat pe sisteme CCTV. Abordarea noastră este compusă din 3 module, fiecare conceput pentru a efectua o sarcină diferită, dar din același descriptor: detecția bagajelor abandonate, detecția proprietarului bagajului și găsirea prezenței acelei persoane pe camerele CCTV. Am folosit Mask R-CNN pentru a efectua detecția și am calculat asemănările pe baza descriptorilor generați de modulul RPN al rețelei. Am adunat un mic set de date pentru experimentul nostru, simulând un caz de utilizare real, am construit o interfață cu utilizatorul pentru operatori și am demonstrat că sistemul funcționează conform așteptărilor din scenariul propus.

Capitolul 3

Generarea Nesupervizată de Date

Acest capitol conține lucrările în scopul augmentării seturilor de date insuficiente, folosind tehnici de învățare adâncă nesupravegheate. Debutează cu o scurtă recenzie a algoritmilor generativi existenți nesupervizați. Apoi, prezentăm avansul în domeniul generării de imagini prin prezentarea unei aplicații de generare a logo-urilor. În continuare, ne concentrăm pe generarea de serii temporale pentru analiza pieței bursiere.

3.1 Prezentarea Literaturii Specifice

În viitor, tot mai mult din noul conținut multimedia nu va fi etichetat, deoarece este imposibil ca etichetarea datelor să țină pasul cu crearea lor. În consecință, înțelegerea modului de a trata și de a învăța din datele neetichetate a devenit mult mai interesantă decât în anii precedenți. Există mai multe tipuri de modele care beneficiază de date neetichetate pe care le-am folosit în munca noastră, pe care le vom prezenta în continuare.

3.1.1 Autoencoder

Autoencoder-ele [20] sunt rețele neuronale care sunt antrenate pentru a reproduce intrarea la ieșirea lor. Acestea sunt compuse din 2 subsisteme în cascadă, un codor și un decodor, conectate într-o rețea feed-forward. Intrarea x este codată în vectorul latent h , care, la rândul său, este decodat în ieșirea x' . Un dezavantaj al autoencoder-elor este că acestea nu pot fi utilizate pentru generarea eșantioanelor. Autoencoder-ele învață doar reprezentarea codului latent a datelor de antrenare, neputând generaliza bine. Acest spațiu de cod latent este aproape de fiecare dată discret, făcând autoencoder-ul nesigur pentru generare.

3.1.2 Autoencoder Variațional

Autoencoder-ele Variaționale (VAE) [19] sunt modele generative dirijate probabilistic. Similar cu autoencoder-ele, utilizează un codor, un decodor și o formulare vectorială latentă, dar procedura lor de antrenare este foarte diferită. În cazul VAE, proce-

sul de generare începe cu un eșantion z extras din distribuția codului latent $p_{model}(z)$. Acesta este transmis prin generator, $g(z)$. În sfârșit, x este eșantionat din distribuția $p_{model}(x; g(z)) = p_{model}(x|z)$. În etapa de antrenare, z se obține prin rețeaua de codare $q(z|x)$. La rândul lui, $p_{model}(x|z)$ este considerat decodorul. În practică, caracterul aleatoriu al eșantioanelor extrase din $p(z)$ face imposibilă efectuarea de propagare înapoi de către algoritmi de optimizare, așa că se folosește în schimb un truc de reparametrizare.

3.1.3 Rețele Generative de Potrivire a Momentelor

GMMN-urile [21] utilizează un cadru de testare a ipotezelor statistice numit discrepanță medie maximă (MMD). În primul rând, un autoencoder este antrenat pe un anumit set de date. Apoi, codorul este utilizat pentru a transforma datele de intrare în spațiul de cod latent, formând distribuția discretă a codului latent. Generatorul (decodorul autoencoder-ului) este apoi antrenat pentru a prelua date din distribuția codului latent și a le transforma în eșantioane noi. După cum s-a menționat anterior, autoencoder-ele au o distribuție discretă a codului latent, ceea ce le face inutilizabile pentru generare. Cu toate acestea, GMMN învață o distribuție continuă a datelor prin spațiul de cod latent.

3.1.4 Rețele Generative Adversariale

Rețelele Generative Adversarial (GAN) [13] se bazează pe un scenariu al teoriei jocului în care sunt implicate 2 părți (generator și discriminator) care concurează unul împotriva celuilalt într-un joc cu sumă zero în care o funcție $v(\theta^{(g)}, \theta^{(d)})$ determină remunerarea discriminatorului. În mod similar, generatorul este recompensat cu $-v(\theta^{(g)}, \theta^{(d)})$. Cu fiecare dintre cele două părți care încearcă să-și maximizeze propria recompensă, convergența este atinsă la $g^* = \arg \min_g \max_d v(g, d)$ [12]. În acel moment, discriminatorul nu va putea face diferența între eșantioanele reale și cele sintetizate de generator și va genera $\frac{1}{2}$ peste tot, urmând ca doar generatorul să fie utilizat în continuare.

3.2 Generare de Logo-uri

Generarea de logo-uri implică crearea de noi logo-uri ale companiilor pentru a augmenta seturile de date existente. Principala aplicație privind logo-urile este detecția automată, deoarece companiile doresc să știe cât de vizibil este produsul lor în diferite media. În continuare, vom prezenta abordarea noastră [8] în generarea de logo-uri.

3.2.1 Prezentarea Literaturii Specifice

Detecția logo-urilor este, de fapt, o aplicație de detecție a obiectelor, domeniu care și-a atins maturitatea. Detectoarele în două etape [14, 24] au obținut în mod tradițional rezultate mai precise, dar în detrimentul unor algoritmi mai complecși care necesită

timp de calcul mai mare, spre deosebire de omologii lor cu o singură etapă. Un alt aspect important al detecției logo-urilor este setul de date care este utilizat pentru antrenarea detectorului. Detecția logo-urilor, spre deosebire de detecția obiectelor, este slab reprezentată, cu doar câteva seturi de date disponibile pentru detecție, cum ar fi FlickrLogos-47 [25]. În plus, mai există Large Logo Dataset (LLD) [26], un set de date care constă numai din logo-uri reprezentate digital, fără exemple in-the-wild.

3.2.2 Abordarea Propusă

Propunem atenuarea problemei lipsei de date prin augmentarea cu logo-uri generate artificial a seturilor de date deja existente. Algoritmul nostru este structurat în 3 pași: 1. antrenarea unui model de generare a logo-urilor, 2. recuperarea vectorului latent pentru un logo dat prin propagarea înapoi a gradientului și alterarea minoră a acestuia, astfel încât să se obțină instanțe diferite ale unui eșantion dat, utilizat pentru augmentare și 3. validarea abordării cu un sistem de detecție automată a logo-urilor.

Sinteza Logo-urilor Partea de sinteză a logo-urilor se realizează cu ajutorul unei rețele generative adversariale, și anume modelul DCGAN [23], antrenat pe LLD. Am antrenat mai multe modele DCGAN pe LLD sub diferite setări de parametri. Deoarece evaluarea calității imaginilor generate este cunoscută a fi dificilă, am examinat manual rezultatele tuturor modelelor noastre pentru a decide cea mai bună configurație.

Propagarea Înapoi a Gradientului GAN-urile transformă vectorii latenți z în imagini I prin intermediul generatoarelor lor, $I = G(z)$. Ne interesează găsirea unei aproximări a vectorului latent z' $approxz = G^{-1}(I)$ care poate fi folosit pentru a genera un logo dat. Maparea inversă, de la imagini la vectori, se poate face prin propagarea înapoi a gradientilor funcțiilor de cost [22]. Pornind de la un vector de zgomot de dimensiunea 100, distribuit uniform între -1 și 1 , $z \sim U([-1, 1]^{100})$, generatorul unui GAN va produce o imagine $G(z)$. Vom genera un alt vector de zgomot aleator z' și imaginea corespunzătoare a acestuia $G(z')$. Apoi, vrem să forțăm z' să fie cât mai aproape de z pentru a obține $G(z')$ cât mai aproape de $G(z)$. Acest lucru se face prin minimizarea normei L_2 . Deoarece există, de asemenea, o constrângere suplimentară conform căreia vectorul de zgomot ar trebui să se situeze în interiorul hipercubului $[-1, 1]^{100}$, se propune o optimizare modificată, unde fiecare valoare care se încadrează în afara acestui interval va fi înlocuită cu o valoare aleatorie eșantionată uniform din intervalul $[-1, 1]$.

Detecția Logo-urilor După cum a fost menționat anterior, evaluarea calității probelor generate de un GAN este un subiect complicat în sine. Pentru a determina cât de bună este reconstrucția, configurăm o un detector de logo-uri. Aceasta este, în esență, o problemă de detecție a obiectelor. Prin urmare, am aplicat o arhitectură Faster R-CNN [24] și am

adaptat-o la logo-uri. Fiecare logo din setul de date Flickr_47 a fost extras, reconstruit cu metoda de propagare înapoi și inserat înapoi în imaginea originală. Acest set de date nou format a fost adăugat la setul de date de antrenare deja existent, obținându-se astfel două versiuni pentru fiecare imagine din setul de date de antrenare: una cu logo-urile originale și una cu logo-urile reconstituite. Ne referim în continuare la acest set de date drept versiunea extinsă a Flickr_47. Am pre-antrenat algoritmul de detecție pe setul de date MS-COCO, apoi l-am rafinat pe versiunea extinsă a Flickr_47 și l-am testat pe setul de date de testare Flickr_47. Am obținut $\text{mAP}@0.5 = 0.6019$, ceea ce este foarte promițător. De asemenea, am antrenat un alt model de detecție a logo-urilor pe setul de date de antrenament Flickr_47 fără imagini reconstituite și l-am testat pe setul de date de testare Flickr_47 și am obținut $\text{mAP}@0.5 = 0.6725$.

3.2.3 Concluzii

Am prezentat un algoritm de generare a logo-urilor bazat pe GAN-uri și propagarea înapoi a gradientului. Am antrenat un model DCGAN pentru a genera logo-uri aleatorii pornind de la setul de date LLD și am extras logo-uri in-the-wild din setul de date Flickr_47. Din aceste logo-uri am extras reprezentările codului latent și apoi am folosit acești vectori de zgomot pentru a ghida generatorul DCGAN să producă logo-uri similare. Aceste noi logo-uri au fost utilizate pentru a înlocui omoloagele lor originale și am rulat un sistem de detecție a logo-urilor pentru a valida procesarea noastră.

3.3 Generarea Seriilor Temporale Financiare

O altă problemă se referă la disponibilitatea unor seturi de date mai puțin frecvente. În această secțiune, prezentăm abordarea noastră¹ în domeniul generării automate de serii temporale financiare realiste. Pentru a surprinde proprietățile care stau la baza datelor, exploatăm diverse modele generative nesupervizate. Introducem problema evaluării calității datelor generate artificial și propunem mai multe soluții. De asemenea, propunem o metodă de generare a unor serii temporale financiare de lungime arbitrară.

3.3.1 Prezentarea Literaturii Specifice

Există două direcții principale pe care cercetătorii le urmează pentru seriile temporale financiare. O abordare este de a se concentra asupra prezicerii următoarelor eșantioane într-o secvență bazată pe istoricul recent disponibil (regresie). Cealaltă abordare este de a genera un număr fix de eșantioane consecutive, cu scopul de augmentare a setului de date. Această procedură necesită extragerea ferestrelor cu o lungime dată L_w atunci când se creează setul de date de antrenare pentru GAN. Apoi, eșantioanele sintetizate sunt de

¹Cercetarea a fost finanțată de Hana TI [4, 10]

obicei validate de un algoritm de predicție bursieră. De asemenea, aceasta reprezintă principala aplicație pentru algoritmul nostru de generare a seriilor temporale financiare.

3.3.2 Datele Financiare

Seria temporală financiară a unei companii reprezintă evoluția cronologică a mai multor indicatori, denumiți OHLC (Open, High, Low, Close). Pentru a capta diferența relativă între zilele consecutive, vor fi utilizate log returns:

$$r_i = \log \frac{C_i}{C_{i-1}}, \quad (3.7)$$

unde C_i reprezintă prețul de închidere al zilei i și r_i prețul de închidere din ziua i . Acest lucru reduce atât intra-variația seriei temporale, cât și inter-variația dintre companii.

Se știe că seriile temporale financiare au vârful mai ascuțit decât distribuțiile normale și prezintă un comportament cu “coadă grasă”. De asemenea, modificările mari ale prețurilor tind să se clusterizeze. În cele din urmă, randamentele empirice ale activelor sunt necorelate pentru orice valoare a decalajului mai mare decât 1, dar nu independente. Modelele generative ar trebui să se ocupe în special de aceste aspecte.

3.3.3 Abordarea Propusă

Această secțiune prezintă soluțiile noastre [4, 10] pentru a genera serii temporale financiare sintetice realiste de lungime arbitrară. Ne comparăm rezultatele cu cele ale Takahashi et al. [27], deoarece principiul pe care îl urmează este cel mai apropiat de munca noastră.

Modele Generative Primul pas al algoritmului constă în generarea unui vector de eșantioane de lungime fixă cu ajutorul mai multor modele generative. Am efectuat un studiu aprofundat asupra unui număr mare de arhitecturi generative. Am investigat 3 clase majore de modele generative: GAN [13], VAE [19] și GMMNs [21].

Dorim să generăm vectori de lungime n și apoi să îi combinăm în serii temporale arbitrar de lungi. Lungimea vectorului sintetizat a fost setată la 250 pentru toate modelele, echivalentul unui an de lucru în domeniul financiar. Toate arhitecturile sunt prezentate în versiunile lor optimizate, realizate după studii aprofundate de ablație. Am propus 4 perceptroni multistrat diferiți (MLP_1 la MLP_4), 6 GAN complet convoluționale ($FCGAN_1$ la $FCGAN_4$, $snFCGAN$ și $FCGANmc$), 2 modele VAE și 2 arhitecturi GMMN.

Crearea Setului de Date Sistemul nostru este conceput pentru a genera serii temporale financiare sintetice de lungime fixă. Prin urmare, antrenăm generatorul să producă tablouri 1-dimensionale cu lungime fixă sau, în cazul arhitecturilor multicanal, tablouri 4-dimensionale, cu dimensiunea L_w . Ne antrenăm modelele generative pe setul de date S&P furnizat de Hana Institute of Technology. Acest set de date este format din 1506

companii cu înregistrări zilnice OHLC din 1 ianuarie 2000 până în 31 martie 2020. Seriile temporale financiare încep la date diferite, dar se termină la aceeași dată. Am împărțit fiecare serie temporală disponibilă în segmente de lungime fixă folosind o fereastră glisantă. Pornind de la cea mai timpurie poziție, începem să extragem segmente de 250 de probe pentru fiecare companie. Am ales să renunțăm complet la segmentele incomplete (datorită listării companiei pe piață în interiorul intervalului temporale al ferestrei). Procesăm restul setului de date glisând o fereastră, cu un pas de rulare de 30 de eşantioane, echivalent cu 6 săptămâni lucrătoare, și le adăugăm la setul nostru de date de antrenare. Denumim setul de antrenare obținut ca $D = \text{bigcup}_{\text{substack}i=1}^{\text{textstep}30} W_i$, unde W_i este setul tuturor segmentelor complete începând cu ziua i . Toate datele au fost transformate în log returns, așa cum se explică în ecuația 3.7. Pentru arhitecturile multicanal, am generat Open, High, Low, Close folosind patru canale în loc de unul în configurarea complet convoluțională. Codăm relația dintre aceste valori și prețul de închidere în loc să folosim valorile lor absolute. Astfel, folosim $\frac{\text{fracHigh}_n - \text{Close}_n}{\text{Close}_{n-1}}$, $\frac{\text{fracClose}_n - \text{Low}_n}{\text{Close}_{n-1}}$ și $\frac{\text{fracClose}_n - \text{Open}_n}{\text{Close}_{n-1}}$.

Pregătirea Setului de Date Setul nostru de date final, D , conține mai mult de 200k intrări unidimensionale fiecare cu lungimea L_w . Propunem 2 moduri de eşantionare a datelor din acest set de date. În primul rând, considerăm că D este un amestec omogen de ferestre și eşantionăm aleatoriu loturi de date din acesta. Notăm această configurație $Data_1$. Considerăm că aceasta face parte dintr-un studiu preliminar pe care l-am folosit pentru aprofundare. A doua abordare constă în păstrarea fiecărui set de segmente W_i în forma lor originală și tratarea fiecărui astfel de subset ca un întreg lot în procesul nostru de formare. Această metodă alimentează modelul cu date corelate încrucișat, ajutându-l să învețe în mod inerent această proprietate. Notăm această configurație $Data_2$.

Împărțirea în Regimuri Bursele sunt în general “în creștere”, ceea ce înseamnă că valorile prețurilor de închidere cresc pe termen lung. În termeni de log return, asta înseamnă că valorile pozitive le depășesc pe cele negative. Dacă diferența dintre media secțiunii transversale și media de rulare a secțiunii transversale este pozitivă, atunci regimul este considerat ca fiind ascendent. Altfel, este descendent.

Împărțirea pe baza regimurilor are ca rezultat etichetarea a 68,07% a zilelor ca fiind ascendente, iar restul de 31,93% ca fiind descendente. Am urmat 2 strategii pentru antrenarea fiecăruia dintre modelele noastre: (1) am antrenat fiecare model cu întregul set de date D ; (2) am calculat regimurile pentru fiecare zi pe setul de date original, am împărțit setul de date conform celor două regimuri și apoi am aplicat mecanismul de fereastră pe fiecare dintre cele două regimuri, rezultând în 3 modele pentru fiecare arhitectură: unul complet, antrenat pe ansamblul întregului set de date, unul ascendent, antrenat doar cu zilele ascendente și unul descendent, antrenat numai cu zilele cu tendințe descendente. Notăm aceste 3 versiuni ca “complete”, “up” și “jos”, respectiv.

Formarea Sintetică a Seriilor Temporale Odată ce generatoarele sunt instruite să producă o serie temporală cu lungime fixă, trebuie să le combinăm astfel încât să obținem serii temporale de dimensiuni arbitrare. În cazul modelelor “complete”, generăm mai multe loturi de lungime fixă și le concatenăm împreună. Cu toate acestea, pentru abordarea regimurilor “mixte”, aplicăm următoarea procedură. Ne bazăm pe faptul că regimurile ascendente și descendente apar în rafale de 20 până la 120 (determinate statistic) eșantioane consecutive. Mai mult, știm cota finală a fiecăruia dintre cele două regimuri. Prin urmare, eșantionăm bucăți de lungime aleatorie cuprinse între 20 și 120 de eșantioane cu o probabilitate de 68% ca acestea să provină din generatorul de model “up” și 32% dintre acestea să provină din modelul “down”. Concatenăm aceste bucăți până atingem lungimea dorită a seriei temporale. De asemenea, ajustarea dimensiunii lotului pentru generator este echivalentă cu setarea numărului de stocuri pe care dorim să le generăm pentru o anumită perioadă, adică dimensiunea universului financiar.

Evaluare Evaluarea datelor generate artificial este încă o problemă deschisă, deoarece nu există metode standardizate de evaluare a performanței unui GAN în cadrul seriilor temporale financiare. Prin urmare, analizăm și propunem o serie de valori care au fost inspirate din problemele de generare a semnalelor multimodale. Majoritatea acestor valori sunt încă experimentale în ceea ce privește cât de precis pot descrie performanța unui GAN de date financiare, dar pot ajuta în continuare la stabilirea unei ierarhii între diferite modele. Abordăm evaluarea la trei niveluri: (i) calitativ, (ii) cantitativ și (iii) test de precizie predictivă, așa cum este prezentat în cele ce urmează.

1. **Analiza calitativă** — primele 4 momente centrale, autocorelație, distribuție cu coadă grea, grupare de volatilitate, t-SNE pentru vizualizare 2D, sumă cumulată a randamentelor și raporturi de tendință. Am investigat manual dacă probele reale și sintetice au același comportament pentru toate valorile de mai sus.
2. **Analiza cantitativă** — divergența Kullback-Leibler, divergența Jensen-Shanon, testul Kolmogorov-Smirnov și distanța Earth Mover. Valorile mici ale măsurilor menționate anterior indică distribuții similare. Am ales un lot aleatoriu de probe reale ca referință acolo unde era cazul.
3. **Test de precizie predictivă** — prezicem mișcarea pieței bursiere încadrând predicția ca o problemă de clasificare, discernând dacă stocurile sunt sau nu performante într-un moment predefinit. Metoda noastră propusă constă în patru etape: (i) gruparea statistică a stocurilor pe baza randamentelor normalizate ale acestora, (ii) etichetarea statistică a stocurilor în performanțe ascendente sau descendente, (iii) reducerea dimensiunilor și a dimensionalității și (iv) formarea de modele predictive pentru a genera ieșirea cu un pas înainte. În etapa de predicție, rulăm simulările noastre pe o perioadă de 20 de ani (începând din 2000 până în 2020), folosind un protocol divizat în două părți. În partea de antrenare, folosim datele

din ultimii șapte ani pentru a antrena modelele folosind un amestec de date reale și sintetice, iar în partea de testare, utilizăm anul următor folosind strict doar date reale pentru a prezice performanța.

Rezultate și Discuții

Am creat instantanee ale fiecărei setări de rețea ori de câte ori aceasta întâlnește o nouă valoare optimă pentru oricare dintre valorile cantitative propuse. Ulterior, am parcurs toate aceste instantanee și am inspectat manual toate valorile calitative. Rezultatele empirice arată că divergența Jensen-Shannon este cel mai bun indicator pentru care model are o performanță generală mai bună, așa că prezentăm rezultatele pentru instantaneele care au obținut cel mai bun JSD pentru fiecare model de rețea.

Realizăm o etapă experimentală preliminară pe datele din implementarea setului de date $Data_1$. Apoi, instruiem modelele prezentate în tabelul 3.3 (marcate cu citația [4]) și le folosim pentru a genera numai probe sintetice cu lungime fixă. A doua fază a experimentelor noastre implică aplicarea lanțului de procesare constând în implementarea $Data_2$ și întreaga configurare explicată în secțiunea “Date Financiare”. Toate experimentele din această parte au fost rulate cu Wasserstein GAN cu setarea de penalizare în gradient. Tabelul 3.3 sintetizează rezultatele obținute.

Procedura de Antrenament Mai întâi, sub formularea GAN, antrenamentul Wasserstein (cu penalizare de gradient) îi depășește pe omologii săi simpli. Alegerea dintre modele mixte sau complete nu are o influență majoră asupra preciziei predicției tendințelor, ambele tehnici oferind rezultate similare. În sfârșit, straturile de normalizare a loturilor nu mai afectează antrenamentul modelului, așa cum sa raportat în [4, 27].

Metrici Valorile ridicate ale JSD înseamnă că modelul nu funcționează bine. Cu toate acestea, valorile scăzute nu indică neapărat modele bune, deoarece această valoare poate fi minimizată atunci când modelul se prăbușește într-un singur eșantion, marcat cu “*” în Tabelul 3.3. Mai mult, eșantioanele care urmează exact aceeași cale în graficul sumă cumulativă indică faptul că modelul generatorului s-a prăbușit pe un singur set de valori. Am observat că modelele care generează probe a căror funcție de densitate de probabilitate (PDF) se potrivește cu funcția de densitate de probabilitate a probelor reale tind să ofere rezultate bune. De asemenea, am observat că modelele care au reușit să potrivească momentul central de ordinul 4, adică kurtosis, se comportă bine la multe niveluri. În plus, tehnica de pregătire a setului de date propusă a ajutat la capturarea autocorelației prin alimentarea cu probe corelate încrucișat la fiecare iterație. Dintre modelele propuse, câștigătorul clar este $VAE_{FC\ complete}$, care depășește orice alt model în aproape fiecare aspect.

Tabel 3.3 Rezultate calitative și cantitative (✓ înseamnă că proprietatea a fost realizată, pe când ✗ contrariul).

Model	Regime	Mean	Var	Skew	Kurtosis	Autocorr	Heavy tail	Cluster Volatility	Cum Sum	Trend Ratio	JS	KL	K-S	EM
<i>MLP</i> ₁	complete	✓	✓	✓	✗	✓	✗	✗	✓	✓	0.4511	237.8	0.2759	0.1254
<i>MLP</i> ₁	up	✗	✗	✓	✗	✗	✗	✗	✓	✓	0.575	698.4	0.3546	0.1053
<i>MLP</i> ₁	down	✗	✗	✓	✗	✓	✗	✗	✓	✓	0.4782	394	0.302	0.1103
<i>MLP</i> ₁ [27, 4]	N/A	✗	✗	✓	✗	✗	✗	✗	N/A	N/A	0.5757	721.5	0.5757	0.114
<i>MLP</i> ₂	complete*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.4845	276.9	0.3204	0.1196
<i>MLP</i> ₂	up*	✓	✗	✗	✗	✗	✗	✗	✗	✗	0.6263	961.6	0.4019	0.1424
<i>MLP</i> ₂	down*	✓	✗	✗	✗	✗	✗	✗	✗	✗	0.6841	677.4	0.4184	0.4029
<i>MLP</i> ₂ [27, 4]	N/A	✓	✓	✓	✗	✓	✓	✗	N/A	N/A	0.0897	29.81	0.0301	0.0028
<i>MLP</i> ₃	complete*	✓	✗	✗	✗	✗	✗	✗	✗	✗	0.5836	621.4	0.3774	0.2012
<i>MLP</i> ₃	up*	✓	✗	✗	✗	✗	✗	✗	✗	✗	0.7216	1359	0.445	0.3872
<i>MLP</i> ₃	down*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.6552	845.8	0.4115	0.413
<i>MLP</i> ₃ [4]	N/A	✓	✓	✓	✗	✗	✓	✗	N/A	N/A	0.1235	128	0.0465	0.0088
<i>WMLP</i> ₃ [4]	N/A	✓	✓	✓	✓	✓	✓	✗	N/A	N/A	0.1031	39.08	0.0323	0.003
<i>MLP</i> ₄	complete	✓	✓	✓	✗	✓	✓	✗	✓	✓	0.1225	35.69	0.07936	0.00697
<i>MLP</i> ₄	up	✓	✓	✗	✗	✓	✓	✗	✓	✓	0.2692	186.4	0.1641	0.01232
<i>MLP</i> ₄	down	✓	✗	✗	✗	✓	✗	✗	✓	✓	0.6354	536.4	0.3845	0.281
<i>MLP</i> ₄ [4]	N/A	✓	✓	✓	✗	✗	✗	✗	N/A	N/A	0.2095	99.89	0.131	0.0106
<i>FCCGAN</i> ₁	complete	✓	✗	✗	✗	✓	✗	✓	✗	✓	0.2947	140.1	0.1818	0.06226
<i>FCCGAN</i> ₁	up	✓	✓	✓	✗	✗	✗	✗	✓	✓	0.4541	395.2	0.2745	0.07718
<i>FCCGAN</i> ₁	down	✓	✓	✗	✗	✓	✗	✗	✓	✓	0.4595	287.7	0.3001	0.2058
<i>FCCGAN</i> ₁ [27, 4]	N/A	✗	✗	✗	✗	✗	✗	✗	N/A	N/A	0.2315	197.5	0.1709	0.0115
<i>WFCCGAN</i> ₁ [4]	N/A	✓	✓	✓	✗	✓	✓	✗	N/A	N/A	0.057	18.23	0.0359	0.0018
<i>FCCGAN</i> ₂	complete	✗	✓	✓	✗	✗	✓	✗	✓	✓	0.1592	42.44	0.1385	0.01779
<i>FCCGAN</i> ₂	up	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.3776	256.9	0.2438	0.06838
<i>FCCGAN</i> ₂	down	✗	✗	✓	✗	✓	✗	✗	✓	✓	0.5569	381.8	0.3451	0.2074
<i>FCCGAN</i> ₂ [27, 4]	N/A	✓	✓	✓	✓	✗	✓	✗	N/A	N/A	0.0454	13.26	0.0178	0.0011
<i>WFCCGAN</i> ₂ [27, 4]	N/A	✓	✓	✗	✗	✗	✓	✓	N/A	N/A	0.0825	26.06	0.0387	0.0034
<i>FCCGAN</i> ₃	complete*	✓	✗	✗	✗	✗	✓	✗	✗	✗	0.1198	83.27	0.05527	0.00156
<i>FCCGAN</i> ₃	up*	✓	✗	✗	✗	✗	✓	✗	✗	✗	0.1576	267.4	0.07755	0.00376
<i>FCCGAN</i> ₃	down*	✓	✗	✗	✗	✗	✓	✗	✗	✗	0.2508	92.44	0.1136	0.08977
<i>FCCGAN</i> ₃ [4]	N/A	✗	✗	✗	✗	✗	✗	✗	N/A	N/A	0.5341	673.8	0.4007	0.1026
<i>FCCGAN</i> ₄	complete	✓	✓	✗	✗	✗	✓	✗	✓	✓	0.1324	82.87	0.0525	0.002
<i>FCCGAN</i> ₄	up	✓	✗	✓	✓	✗	✓	✗	✓	✓	0.1739	92.61	0.1051	0.0056
<i>FCCGAN</i> ₄	down	✓	✓	✗	✗	✗	✓	✗	✗	✗	0.1294	46.47	0.0902	0.0107
<i>snFCCGAN</i>	complete*	✓	✓	✓	✗	✗	✓	✗	✗	✗	0.0666	5.613	0.0383	0.0039
<i>snFCCGAN</i>	up*	✓	✓	✗	✗	✗	✗	✗	✗	✗	0.4882	719.4	0.3216	0.0525
<i>snFCCGAN</i>	down*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.2985	140.1	0.2476	0.0475
<i>snFCCGAN</i> [4]	N/A	✗	✗	✗	✗	✗	✓	✗	N/A	N/A	0.0953	23.53	0.0408	0.0031
<i>FCCGAN</i> _{mc} [4]	N/A	✓	✓	✓	✓	✗	✓	✗	N/A	N/A	0.1101	78.16	0.0797	0.0032
<i>GMMN</i> _{AE} _{FC}	complete*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.1211	155.9	0.03031	0.00153
<i>GMMN</i> _{AE} _{FC}	up*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.1926	436.9	0.1069	0.00358
<i>GMMN</i> _{AE} _{FC}	down*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.2783	985.3	0.141	0.00951
<i>GMMN</i> _{AE} _{MLP}	complete*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.3671	1349	0.1937	0.0165
<i>GMMN</i> _{AE} _{MLP}	up*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.2576	1417	0.1237	0.0052
<i>GMMN</i> _{AE} _{MLP}	down*	✗	✗	✗	✗	✗	✗	✗	✗	✗	0.2278	822.4	0.0941	0.0072
<i>VAE</i> _{FC}	complete	✓	✓	✓	✓	✓	✓	✗	✓	✓	0.0929	14.26	0.0553	0.00566
<i>VAE</i> _{FC}	up	✓	✓	✓	✓	✓	✓	✗	✓	✓	0.07917	42.5	0.0273	0.00156
<i>VAE</i> _{FC}	down	✓	✓	✓	✓	✓	✓	✗	✓	✓	0.09543	30.97	0.0343	0.00357
<i>VAE</i> _{MLP}	complete	✓	✓	✓	✓	✗	✓	✗	✓	✓	0.1058	54.4	0.07054	0.00312
<i>VAE</i> _{MLP}	up	✓	✗	✓	✓	✓	✓	✗	✓	✓	0.06591	21.67	0.03615	0.00182
<i>VAE</i> _{MLP}	down	✓	✗	✓	✗	✗	✓	✗	✓	✓	0.09136	19.79	0.04364	0.00331

Mecanismul de Predicție Am testat 3 rețele diferite în etapa de predicție, și anume un LSTM bidirecțional cu 2 straturi, ResNet-50 și un MLP cu 4 straturi. Am augmentat setul de date de antrenament cu date sintetice obținute cu fiecare dintre modelele prezentate în Tabelul 3.4 și am testat algoritmul de predicție pe date reale. Cele mai mari acurateți au fost obținute cu modelul LSTM, deci le raportăm doar pe acestea. Raportăm acuratețea medie și maximă obținută pentru cele 10 perechi antrenare-testare. Acuratețile maxime sunt în medie cu 0,11% mai mari decât acuratețile medii care arată o variație importantă a performanței, confirmând natura volatilității piețelor bursiere. Valoarea baseline a fost obținută prin antrenarea și testarea algoritmului de predicție numai pe date reale.

Tabel 3.4 Acuratețea predicției tendinței pe seturile de date augmentate.

Model	Mean accuracy	Max accuracy
<i>MLP₁ complete</i>	50.12%	50.26%
<i>MLP₁ mixed</i>	50.30%	50.58%
<i>MLP₂ complete</i>	50.41%	50.56%
<i>MLP₂ mixed</i>	50.40%	50.48%
<i>MLP₃ complete</i>	50.38%	50.43%
<i>MLP₃ mixed</i>	50.40%	50.50%
<i>MLP₄ complete</i>	50.34%	50.45%
<i>MLP₄ mixed</i>	50.27%	50.38%
<i>F CGAN₁ complete</i>	50.27%	50.40%
<i>F CGAN₁ mixed</i>	50.35%	50.45%
<i>F CGAN₂ complete</i>	50.40%	50.54%
<i>F CGAN₂ mixed</i>	50.11%	50.27%
<i>F CGAN₃ complete</i>	50.37%	50.43%
<i>F CGAN₃ mixed</i>	50.28%	50.37%
<i>F CGAN₄ complete</i>	50.39%	50.46%
<i>F CGAN₄ mixed</i>	50.32%	50.41%
<i>sn_FCGAN complete</i>	50.31%	50.48%
<i>sn_FCGAN mixed</i>	50.39%	50.44%
<i>GMMN_AE_FC complete</i>	50.40%	50.44%
<i>GMMN_AE_FC mixed</i>	50.42%	50.52%
<i>GMMN_AE_MLP complete</i>	50.40%	50.48%
<i>GMMN_AE_MLP mixed</i>	50.41%	50.57%
<i>VAE_FC complete</i>	50.31%	50.41%
<i>VAE_FC mixed</i>	50.29%	50.42%
<i>VAE_MLP complete</i>	50.18%	50.28%
<i>VAE_MLP mixed</i>	50.22%	50.32%
Baseline	50.04%	

3.3.4 Concluzii

Am propus un cadru complex pentru generarea unor serii temporale financiare realiste. Am propus un nou mod de extragere a loturilor de date din setul de antrenare, adaptat la seriile temporale financiare. Am investigat 3 clase majore de modele generative cu diferite compoziții de modele, setări, hiperparametri, cadre de antrenare și regimuri de date. Am examinat diferite valori calitative și cantitative și am testat capacitatea de augmentare a setului de date pe date reale, într-un scenariu complex de predicție.

Am identificat autoencoder-ul variațional ca fiind cel mai potrivit model pentru sarcina curentă. În cele din urmă, subliniem necesitatea unui cadru metric sau de validare care să valorifice atât proprietățile obiective, cât și cele subiective, sub o singură valoare cuantificabilă.

Capitolul 4

Rezumatul Contribuțiilor și Lucrări Viitoare

4.1 Rezumatul Contribuțiilor

În această secțiune vom prezenta rezultatele obținute, per capitol. Deoarece aceste rezultate au fost discutate pe larg în secțiunile corespunzătoare, le vom introduce doar pe scurt.

În capitolul 2 am discutat mai multe abordări pentru recuperarea imaginilor pe seturi de date insuficiente. Această lucrare a fost împărțită în două părți după cum urmează.

- Secțiunea 2.2 acoperă progresele noastre pe parcursul a 3 ani (2017, 2018 și 2019) la o competiție de benchmarking de lifeloggin.
 - În 2017 [5] am propus o abordare nouă care analizează automat asemănarea dintre o interogare text și un set de imagini pentru care sunt disponibile concepte.
 - În 2018 [6] am propus o abordare mai sofisticată care presupunea analizarea interogării textului pentru informații critice privind activitatea, locația, ora și data. Coroborarea acestora cu mai mulți detectori are ca rezultat o listă scurtă de imagini relevante.
 - În 2019 [7] am optat pentru o filtrare mai agresivă a setului de date, excludând imaginile care nu îndeplineau un set de restricții legate de neclaritate, activitate, locație, oră, fus orar, număr de persoane etc. și am aplicat un mecanism de notare ponderat pe asemănările dintre conceptele relevante extrase din interogarea de căutare și cele extrase din imagini.
- Secțiunea 2.3 prezintă sistemul pe care l-am propus [9] pentru a rezolva detecția bagajelor abandonate. Am rulat un detector de obiecte pe fiecare imagine din setul de date și am detectat bagajele abandonate. Apoi, comparăm vectorul caracteristic

al obiectului izolat cu alte obiecte similare din fluxul CCTV pentru a determina cine l-a lăsat acolo. După ce am găsit un suspect, căutăm în fluxul CCTV pentru a afla unde a mai fost văzută această persoană.

În Capitolul ?? abordăm problema datelor insuficiente dintr-o perspectivă diferită. În loc să găsim modalități de adaptare a algoritmilor noștri la disponibilitatea limitată a datelor, propunem diferite metode de augmentare a setului de date folosind modele generative.

- Secțiunea 3.2 prezintă metoda pe care o propunem [8] pentru controlul logo-urilor pe care un DCGAN le poate halucina. Antrenăm un model DCGAN pentru a genera logo-uri, aplicăm un mecanism de propagare înapoi pentru a recupera o aproximare a codului latent care a generat logo-urile și mărim setul de date original cu versiuni reconstituite ale logo-urilor originale.
- Secțiunea 3.3 prezintă abordarea noastră [4] cu privire la problema generării de serii temporale financiare. Propunem mai multe arhitecturi generative (GAN, VAE, GMMN), cu un întreg set de variații pentru fiecare tip, pentru a genera serii temporale realiste în contextul piețelor financiare. Propunem noi modalități de generare și amestecare a probelor sintetice, astfel încât acestea să pară realiste. De asemenea, investigăm cum poate fi efectuată validarea acestui tip de date.

4.2 Contribuții Originale

În această secțiune prezint contribuțiile originale pe care le-am avut la fiecare publicație individuală.

În [Ch1], [Ch2], [Ch3], [Ch4], [C7], and [C9] am făcut parte din comitetul de organizare al competițiilor de benchmarking ImageCLEF, unde am fost implicat în organizarea competițiilor, însărcinat cu înregistrarea participanților.

În [J1] am proiectat și implementat cadrul generativ propus pentru seriile temporale financiare. Aceasta a constat în extinderea cercetării noastre anterioare [4] la noi modele generative (de exemplu, GMMN, VAE, WGAN etc.), adaptarea acestora la configurația financiară, implementarea de noi metrice, implementarea împărțirii regimului și amestecarea modelelor, și rularea întregului lanț de procesare. Ulterior, am participat și la etapa de predicție, unde am derulat o parte considerabilă din experimentele prezentate și am centralizat rezultatele finale.

În [J2] am propus adoptarea propagării înapoi la DCGAN pentru a controla generarea eșantioanelor. De asemenea, am implementat întregul sistem și am efectuat validarea necesară. Această lucrare a fost efectuată în timpul unui stagiu ERASMUS la CEA-List, Palaiseau, Franța.

În [C8] am proiectat și implementat cadrul generativ propus pentru seriile temporale financiare. Aceasta a constat în propunerea mai multor arhitecturi de rețea, adaptarea

lor la configurația financiară, găsirea unor metrici posibile pentru validarea acestora și execuția întregului algoritm.

În [C6] am propus și implementat ideea utilizării caracteristicilor detectorului de obiecte pentru rezolvarea tuturor celor 3 submodele implicate (detecția bagajelor abandonate, detecția suspectului și reidentificarea suspectului). Am fost, de asemenea, responsabil cu procesul de colectare a seturilor de date de validare care a constat în proiectarea și implementarea unei GUI specifice sarcinilor pentru adnotările persoanei, coordonarea adnotatorilor, centralizarea tuturor datelor și gestionarea întregului proces.

În [C5] am fost implicat în procesul de extragere a descriptorilor persoanelor.

În [C4] și [C2] am propus și implementat mai multe proceduri de filtrare care ar elimina majoritatea imaginilor neinformative dintr-un set de date lifelog. Apoi, am rulat mai multe detectoare conceptuale și am proiectat o regulă care să clasifice imaginile pe baza scorurilor obținute de detectoarele conceptuale.

În [C3] am fost implicat în proiectarea autoencoder-ului care a fost utilizat pentru extragerea descriptorilor semnalelor de electroencefalogramă.

În [C1] am propus și implementat o metodă care stabilește o conexiune între descrierea textuală a mai multor interogări de căutare și ieșirea unui detector de concepte rulat pe un set de date cu imagini lifelog. Această conexiune se bazează pe metode NLP care examinează similitudinea dintre diferite cuvinte care fac parte dintr-un dicționar mare. Au fost stabilite praguri empirice pentru a limita cantitatea mare de propuneri de imagini pentru fiecare interogare de căutare.

4.3 Perspective de Viitor

Această teză s-a axat pe modul în care mai multe aplicații pot face față unor seturi de date insuficiente. Bineînțeles, vom continua eforturile pe care le-am realizat în domeniul de detecție a obiectelor și lucrul cu date insuficiente, concentrându-ne în special pe învățarea cu câteva fotografii și detectarea obiectelor cu câteva fotografii. De asemenea, vom beneficia de experiența pe care am dobândit-o cu modele generative pentru augmentarea claselor slab reprezentate. Acest lucru va ajuta, de asemenea, cu problema dezechilibrului de clasă care este atât de des întâlnită în aplicațiile de învățare adâncă, în special în scenarii din lumea reală.

4.4 Publicații

Capitole de Carte

[Ch4] Ionescu, B., Müller, H., Péteri, R., ..., **Dogariu, M.**, ..., Deshayes, J. (2021). Overview of the ImageCLEF 2021: Multimedia Retrieval in Medical, Nature,

Internet and Social Media Applications. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, Cham. Springer International Publishing. [17]

[Ch3] Ionescu, B., Müller, H., Péteri, R., ..., **Dogariu, M.**, Ștefan, L. D., Constantin, M. G., Deshayes, J., and Popescu, A. (2021). The 2021 ImageCLEF benchmark: Multimedia retrieval in medical, nature, internet and social media applications. In *Advances in Information Retrieval*, pages 616–623, Cham. Springer International Publishing. [16]

[Ch2] Ionescu, B., Müller, H., Péteri, R., ..., **Dogariu, M.**, Ștefan, L. D., and Constantin, M. (2020). Overview of the ImageCLEF 2020: Multimedia retrieval in medical, lifelogging, nature, and internet applications. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, pages 311–341, Cham. Springer International Publishing. [15]

[Ch1] Ionescu, B., Müller, H., Péteri, R., ..., **Dogariu, M.**, Ștefan, L. D., and Constantin, M. G. (2020). ImageCLEF 2020: Multimedia retrieval in lifelogging, medical, nature, and internet applications. In *Advances in Information Retrieval*, pages 533–541, Cham. Springer International Publishing. [18]

Reviste

[J1] **Dogariu, M.**, Ștefan, L.-D., Boteanu, B. A., Lamba, C., Kim, B. and Ionescu, B. (2021). Generation of Realistic Synthetic Financial Time-Series. In *ACM Transactions on Multimedia Computing Communications and Applications*. Paper under review. [10]

[J2] **Dogariu, M.**, Le Borgne, H., and Ionescu, B. (2021b). Backpropagation aided logo generation using generative adversarial networks. In *University Politehnica of Bucharest Scientific Bulletin Series C-Electrical Engineering And Computer Science*, 83(2):59–70. [8]

Conferințe

[C9] Berari, R., Tauteanu, A., Fichou, D., Brie, P., **Dogariu, M.**, Ștefan, L.-D., Constantin, M. G. and Ionescu, B. (2021). Overview of ImageCLEFdrawnUI 2021: The Detection and Recognition of Hand Drawn and Digital Website UIs Task. In *Working Notes of CLEF 2021 - Conference and Labs of the Evaluation Forum.*, Bucharest, Romania, September 21-24, volume 2936. [1]

[C8] **Dogariu, M.**, Ștefan, L.-D., Boteanu, B. A., Lamba, C., and Ionescu, B. (2021). Towards realistic financial time series generation via generative adversarial learning. In *2021 29th European Signal Processing Conference (EUSIPCO)*. [4]

- [C7] Fichou, D., Berari, R., Brie, P., **Dogariu, M.**, Ștefan, L.-D., Constantin, M. G. and Ionescu, B. (2021). In *Working Notes of CLEF 2020 - Conference and Labs of the Evaluation Forum.*, Thessaloniki, Greece, September 22-25, volume 2696. [11]
- [C6] **Dogariu, M.**, Ștefan, L.-D., Constantin, M. G., and Ionescu, B. (2020). Human-object interaction: Application to abandoned luggage detection in video surveillance scenarios. In *2020 13th International Conference on Communications (COMM)*, pages 157–160. IEEE. WOS:000612723900028. [9]
- [C5] Ștefan, L.-D., Abdulamit, Ș., **Dogariu, M.**, Constantin, M. G., and Ionescu, B. (2020). Deep learning-based person search with visual attention embedding. In *2020 13th International Conference on Communications (COMM)*, pages 303-308. IEEE. WOS:000612723900053. [3]
- [C4] **Dogariu, M.** and Ionescu, B. (2019). Multimedia lab @ ImageCLEF 2019 lifelog moment retrieval task. In *Working Notes of CLEF 2019 - Conference and Labs of the Evaluation Forum*, Lugano, Switzerland, September 9-12, volume 2380. CEUR-WS.org. [7]
- [C3] Tăuțan, A.-M., **Dogariu, M.**, and Ionescu, B. (2019). Detection of Epileptic Seizures using Unsupervised Learning Techniques for Feature Extraction. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2377-2381). IEEE. WOS:000557295302184. [28]
- [C2] **Dogariu, M.** and Ionescu, B. (2018). Multimedia lab @ ImageCLEF 2018 lifelog moment retrieval task. In *Working Notes of CLEF 2018 - Conference and Labs of the Evaluation Forum*, Avignon, France, September 10-14, volume 2125. CEUR-WS.org. [6]
- [C1] **Dogariu, M.** and Ionescu, B. (2017). A textual filtering of HOG-based hierarchical clustering of lifelog data. In *Working Notes of CLEF 2017 - Conference and Labs of the Evaluation Forum*, Dublin, Ireland, September 11-14, volume 1866. CEUR-WS.org. [5]

Proiecte Internaționale de Cercetare

- [IntR2] *Octombrie 2020 - prezent: Student doctorand* la Universitatea Politehnica din București, în proiectul H2020 AI4Media "A European Excellence Centre for Media, Society and Democracy", coordonator CERTH, Grecia, axa H2020 ICT-48-2020 / Towards a vibrant European network of AI excellence centres.
- [IntR1] *Februarie 2017 - Aprilie 2019: Trainer* la Universitatea Politehnica din București, în proiectul UMETECH "University and Media Technology for Cultural Heritage", finanțat de Comisia Europeană, ID 574105-EPP-1-2016-1-IT-EPPKA2-CBHE-JP.

Proiecte Naționale de Cercetare

- [NR4] *Iunie 2020 - prezent*: **Student doctorand** la Universitatea Politehnica din București, în proiectul SMARTRetail "Enhancing and Improving Customer Experience and Services in Supermarkets via SMART Artificial Intelligence Powered Systems", coordonator Softrust Vision Analytics, ID PN-III-P2-2.1-PTE-2019-0055.
- [NR3] *Iunie 2020 - prezent*: **Student doctorand** la Universitatea Politehnica din București, în proiectul GRAVI "Virtual Guardian: Artificial Intelligence Powered Multi-Sensor System for Automatic Securing of Areas of Interest", coordonator Softrust Vision Analytics, ID PN-III-P2-2.1-PTE-2019-0570.
- [NR2] *Mai 2017 - Aprilie 2020*: **Student doctorand** la Universitatea Politehnica din București, în proiectul SPIA-VA("Intelligent Systems for Video and Audio Analysis - Technologies and Innovative Video Systems for Person Re-identification and Analysis of Dissimulated Behavior", coordonator Universitatea Politehnica din București, ID PN-III-P2-2.1-SOL-2016-02-000.
- [NR1] *Ianuarie 2017 - Iunie 2018*: **Student doctorand** la Universitatea Politehnica din București, în proiectul SPOTTER("Intelligent Real-time Surveillance System with Specific Regions Detection Integrated on IP Cameras"), finanțat de UEFISCDI, coordonator Universitatea Politehnica din București, ID PN-III-P2-P2.1-PED-2016-1065.

Proiecte Industriale de Cercetare

- [IndR3] *Martie 2020 - Octombrie 2020*: **Student doctorand** la Universitatea Politehnica din București, în proiectul Keysight 1 "Machine Learning Techniques for Generating Network Traffic Data", coordonator Universitatea Politehnica din București, Centrul de Cercetare CAMPUS, beneficiar Keysight Technologies Romania, ID Keysight 1/23-03-2020.
- [IndR2] *Aprilie 2020 - Iulie 2020*: **Student doctorand** la Universitatea Politehnica din București, în proiectul Hana 2 "Financial Data Augmentation and Forecasting Using Advanced AI Techniques", coordonator Institutul de Cercetare Dezvoltare și Inovare "Politehnica", beneficiar Hana Institute of Technology, Coreea de Sud, ID Hana 2/01-04-2020.
- [IndR1] *Iulie 2019 - Decembrie 2019*: **Asistent de cercetare** la Universitatea Politehnica din București, în proiectul Hana 1 "Machine Learning Techniques for the Processing and Analysis of Financial Data", coordonator Institutul de Cercetare Dezvoltare și Inovare "Politehnica", beneficiar Hana Institute of Technology, Coreea de Sud, ID Hana 1/22-07-2019.

Bibliografie

- [1] Berari, R., Tauteanu, A., Fichou, D., Brie, P., Dogariu, M., Ștefan, L. D., Constantin, M. G., and Ionescu, B. (2021). Overview of ImageCLEFdrawnUI 2021: The detection and recognition of hand drawn and digital website uis task. In *Working Notes of CLEF 2021 - Conference and Labs of the Evaluation Forum.*, volume 2936 of *CEUR Workshop Proceedings*, pages 1121—1132, Bucharest, Romania. CEUR-WS.org <<http://ceur-ws.org>>.
- [2] Builders, A. I. D. N. N. I. A. (2014). Unattended baggage detection using deep neural networks in intel® architecture. Technical report, Intel Corporation.
- [3] Ștefan, L.-D., Abdulamit, c., Dogariu, M., Constantin, M. G., and Ionescu, B. (2020). Deep learning-based person search with visual attention embedding. In *2020 13th International Conference on Communications (COMM)*, pages 303–308.
- [4] Dogariu, M., Ștefan, L.-D., Boteanu, B. A., Lamba, C., and Ionescu, B. (2021a). Towards realistic financial time series generation via generative adversarial learning. In *2021 29th European Signal Processing Conference (EUSIPCO)*.
- [5] Dogariu, M. and Ionescu, B. (2017). A textual filtering of hog-based hierarchical clustering of lifelog data. In Cappellato, L., Ferro, N., Goeuriot, L., and Mandl, T., editors, *Working Notes of CLEF 2017 - Conference and Labs of the Evaluation Forum, Dublin, Ireland, September 11-14, 2017.*, volume 1866. CEUR-WS.org.
- [6] Dogariu, M. and Ionescu, B. (2018). Multimedia lab @ imageclef 2018 lifelog moment retrieval task. In Cappellato, L., Ferro, N., Nie, J.-Y., and Soulier, L., editors, *Working Notes of CLEF 2018 - Conference and Labs of the Evaluation Forum, Avignon, France, September 10-14, 2018.*, volume 2125. CEUR-WS.org.
- [7] Dogariu, M. and Ionescu, B. (2019). Multimedia lab @ imageclef 2019 lifelog moment retrieval task. In Cappellato, L., Ferro, N., Losada, D. E., and Müller, H., editors, *Working Notes of CLEF 2019 - Conference and Labs of the Evaluation Forum, Lugano, Switzerland, September 9-12, 2019.*, volume 2380. CEUR-WS.org.
- [8] Dogariu, M., Le Borgne, H., and Ionescu, B. (2021b). Backpropagation aided logo generation using generative adversarial networks. *University Politehnica of Bucharest Scientific Bulletin Series C-Electrical Engineering And Computer Science*, 83(2):59–70.
- [9] Dogariu, M., Ștefan, L.-D., Constantin, M. G., and Ionescu, B. (2020). Human-object interaction: Application to abandoned luggage detection in video surveillance scenarios. In *2020 13th International Conference on Communications (COMM)*, pages 157–160. IEEE.
- [10] Dogariu, M., Ștefan, L.-D., Boteanu, B., Lamba, C., Kim, B., and Ionescu, B. (2021c). Generation of realistic synthetic financial time-series. *ACM Transactions on Multimedia Computing Communications and Applications*. Paper under review.

- [11] Fichou, D., Berari, R., Brie, P., Dogariu, M., Ștefan, L. D., Constantin, M. G., and Ionescu, B. (2020). Overview of ImageCLEFdrawnUI 2020: The detection and recognition of hand drawn website uis task. In *Working Notes of CLEF 2020 - Conference and Labs of the Evaluation Forum.*, volume 2696 of *CEUR Workshop Proceedings*, Thessaloniki, Greece. CEUR-WS.org <<http://ceur-ws.org>>.
- [12] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- [13] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc.
- [14] He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.
- [15] Ionescu, B., Müller, H., Péteri, R., Abacha, A. B., Datla, V., Hasan, S. A., Demner-Fushman, D., Kozlovski, S., Liauchuk, V., Cid, Y. D., Kovalev, V., Pelka, O., Friedrich, C. M., García Seco de Herrera, A., Ninh, V.-T., Le, T.-K., Zhou, L., Piras, L., Riegler, M., Halvorsen, P., Tran, M.-T., Lux, M., Gurrin, C., Dang-Nguyen, D.-T., Chamberlain, J., Clark, A., Campello, A., Fichou, D., Berari, R., Brie, P., Dogariu, M., Ștefan, L. D., and Constantin, M. G. (2020a). Overview of the imageclef 2020: Multimedia retrieval in medical, lifelogging, nature, and internet applications. In Arampatzis, A., Kanoulas, E., Tsirikla, T., Vrochidis, S., Joho, H., Lioma, C., Eickhoff, C., Névéol, A., Cappellato, L., and Ferro, N., editors, *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, pages 311–341, Cham. Springer International Publishing.
- [16] Ionescu, B., Müller, H., Péteri, R., Abacha, A. B., Demner-Fushman, D., Hasan, S. A., Sarrouti, M., Pelka, O., Friedrich, C. M., de Herrera, A. G. S., Jacutprakart, J., Kovalev, V., Kozlovski, S., Liauchuk, V., Cid, Y. D., Chamberlain, J., Clark, A., Campello, A., Moustahfid, H., Oliver, T., Schulz, A., Brie, P., Berari, R., Fichou, D., Tauteanu, A., Dogariu, M., Ștefan, L. D., Constantin, M. G., Deshayes, J., and Popescu, A. (2021a). The 2021 imageclef benchmark: Multimedia retrieval in medical, nature, internet and social media applications. In Hiemstra, D., Moens, M.-F., Mothe, J., Perego, R., Potthast, M., and Sebastiani, F., editors, *Advances in Information Retrieval*, pages 616–623, Cham. Springer International Publishing.
- [17] Ionescu, B., Müller, H., Péteri, R., Ben Abacha, A., Sarrouti, M., Demner-Fushman, D., Hasan, S. A., Kozlovski, S., Liauchuk, V., Dicente, Y., Kovalev, V., Pelka, O., de Herrera, A. G. S., Jacutprakart, J., Friedrich, C. M., Berari, R., Tauteanu, A., Fichou, D., Brie, P., Dogariu, M., Ștefan, L. D., Constantin, M. G., Chamberlain, J., Campello, A., Clark, A., Oliver, T. A., Moustahfid, H., Popescu, A., and Deshayes-Chossart, J. (2021b). Overview of the ImageCLEF 2021: Multimedia retrieval in medical, nature, internet and social media applications. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, Proceedings of the 12th International Conference of the CLEF Association (CLEF 2021), Bucharest, Romania. LNCS Lecture Notes in Computer Science, Springer.
- [18] Ionescu, B., Müller, H., Péteri, R., Dang-Nguyen, D.-T., Zhou, L., Piras, L., Riegler, M., Halvorsen, P., Tran, M.-T., Lux, M., Gurrin, C., Chamberlain, J., Clark, A., Campello, A., Seco de Herrera, A. G., Ben Abacha, A., Datla, V., Hasan, S. A., Liu, J., Demner-Fushman, D., Pelka, O., Friedrich, C. M., Dicente Cid, Y., Kozlovski, S., Liauchuk, V., Kovalev, V., Berari, R., Brie, P., Fichou, D., Dogariu, M., Ștefan,

- L. D., and Constantin, M. G. (2020b). Imageclef 2020: Multimedia retrieval in lifelogging, medical, nature, and internet applications. In Jose, J. M., Yilmaz, E., Magalhães, J., Castells, P., Ferro, N., Silva, M. J., and Martins, F., editors, *Advances in Information Retrieval*, pages 533–541, Cham. Springer International Publishing.
- [19] Kingma, D. P. and Welling, M. (2014). Auto-encoding variational bayes. In *International Conference on Learning Representation (ICLR)*.
- [20] Lecun, Y. (1987). *Modeles connexionnistes de l'apprentissage*. PhD thesis, These de Doctorat, Universite Paris.
- [21] Li, Y., Swersky, K., and Zemel, R. (2015). Generative moment matching networks. In *International Conference on Machine Learning*, pages 1718–1727. PMLR.
- [22] Lipton, Z. C. and Tripathi, S. (2017). Precise recovery of latent vectors from generative adversarial networks. In *International Conference on Learning Representation (ICLR) Workshop*.
- [23] Radford, A., Metz, L., and Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. In Bengio, Y. and LeCun, Y., editors, *4th International Conference on Learning Representations, ICLR*.
- [24] Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99.
- [25] Romberg, S., Pueyo, L. G., Lienhart, R., and Van Zwol, R. (2011). Scalable logo recognition in real-world images. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval*, pages 1–8.
- [26] Sage, A., Agustsson, E., Timofte, R., and Van Gool, L. (2018). Logo synthesis and manipulation with clustered generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5879–5888.
- [27] Takahashi, S., Chen, Y., and Tanaka-Ishii, K. (2019). Modeling financial time-series with generative adversarial networks. *Physica A: Statistical Mechanics and its Applications*, 527:121261.
- [28] Tăuțan, A.-M., Dogariu, M., and Ionescu, B. (2019). Detection of epileptic seizures using unsupervised learning techniques for feature extraction. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2377–2381.
- [29] Xiao, T., Li, S., Wang, B., Lin, L., and Wang, X. (2017). Joint detection and identification feature learning for person search. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3415–3424.