

Universitatea POLITEHNICA din București

Facultatea de Automatică și Calculatoare,
Departamentul de Calculatoare



TEZĂ DE DOCTORAT REZUMAT

Selectarea Celei Mai Bune Abordări:
De la Analiza și Prelucrarea Imaginilor
la Dezvoltarea de Software

Conducător Științific:

Prof. Dr. Ing. Costin-Anton Boiangiu

Autor:

Cristian-Dumitru Avatavului

București, 2023

Cuprins

1	Introducere.....	1
1.1	Organizarea și Structura Tezei	1
2	Analiza de Imagine prin Combinarea Eficientă a Rezultatelor a Mai Multor Algoritmi ...	2
2.1	Recunoașterea Optică a Caracterelor (OCR).....	2
2.2	Estimarea și Analiza Mișcării.....	5
2.3	Strategii de Super-Rezoluție pentru Imagini	12
2.4	Corectarea și Adaptarea Înclinării Paginii.....	14
2.5	Evaluarea și Interpretarea Contrastului	17
3	Metodă Ierarhică de Clusterizare pentru Înțelegerea și Analiza Structurii Paginilor	20
4	Predicția Interacțiunilor Fizice Utilizând Rețele Neuronale.....	24
5	Investiții Eficiente în Dezvoltarea de Software	28
5.1	Analiza Jocurilor: Indie vs. AAA	29
5.2	Aspectele Open vs. Closed în Dezvoltarea de Software	32
6	Concluzii și Contribuții	33
6.1	Lucrări și Publicații	34
	Bibliografie.....	37

1 Introducere

În contextul transformărilor digitale din societatea contemporană, analiza de imagine și predicția interacțiunilor fizice au devenit domenii de cercetare tot mai relevante și complexe [1]. Acestea reprezintă arii esențiale pentru dezvoltarea de tehnologii avansate și îmbunătățirea interacțiunii om-mașină în diverse aplicații, de la prelucrarea datelor vizuale până la simulările complexe ale interacțiunilor fizice [2].

1.1 Organizarea și Structura Tezei

Structura acestei teze de doctorat a fost concepută cu scopul de a furniza o analiză sistematică și coerentă asupra subiectelor abordate, având ca obiectiv facilitarea înțelegerii și urmării progresului cercetării.

Capitolul 1, oferă un context amplu al cercetării, evidențiind relevanța transformărilor digitale din societatea contemporană, importanța analizei de imagine, predicției interacțiunilor fizice în dezvoltarea tehnologiilor avansate și îmbunătățirea interacțiunii om-mașină. Scopul și obiectivele lucrării sunt delimitate clar în această secțiune, conturând obiectivele majore ale tezei.

Capitolul 2, constituie primul pilon al cercetării și explorează în detaliu aspecte cheie ale analizei de imagine, cum ar fi recunoașterea optică a caracterelor (OCR), analiza mișcării, super-rezoluția, corectarea înclinării paginilor și evaluarea contrastului.

Capitolul 3, are ca scop identificarea și gruparea entităților relevante într-o structură ierarhică, furnizând astfel o perspectivă cuprinzătoare asupra organizării paginilor document.

Capitolul 4, este dedicat explorării potențialului rețelelor neuronale în predicția interacțiunilor fizice dinamice. Prin aplicarea acestor tehnologii avansate, se investighează și se înțeleg mai bine fenomenele complexe care apar în diferite contexte, contribuind la o mai profundă înțelegere și simulare a interacțiunilor fizice.

Capitolul 5, abordează aspectele esențiale, comparând paradigmele de dezvoltare „Indie vs AAA” și „Open-Source vs Closed-Source”. Prin analiza comparativă, se identifică cele mai eficiente strategii de dezvoltare software, luând în considerare costurile, calitatea și impactul în industrie.

Capitolul 6, reprezintă o sinteză a rezultatelor obținute în cadrul cercetării, subliniind contribuțiile aduse în domeniul analizei de imagine, predicției interacțiunilor fizice și dezvoltării eficiente de software.

2 Analiza de Imagine prin Combinarea Eficientă a Rezultatelor a Mai Multor Algoritmi

În acest capitol, ne concentrăm pe analiza de imagine și pe metodele inovatoare de a combina rezultatele obținute din mai mulți algoritmi pentru a îmbunătăți performanța și acuratețea în diferite aspecte ale procesării imaginilor. Subcapitolele următoare detaliază subiecte specifice ale analizei de imagine, precum recunoașterea optică a caracterelor (OCR), analiza mișcării, super-rezoluția, corectarea înclinării paginilor și evaluarea contrastului.

2.1 Recunoașterea Optică a Caracterelor (OCR)

Trăim într-o eră tehnologizată, în care un număr mare de documente fizice au început sau trebuie să fie digitalizate pentru a le face accesibile oriunde, pentru un număr mare de oameni. Recunoașterea optică a caracterelor (OCR) este una dintre tehnicile utilizate pe scară largă pentru a recunoaște caractere din imagini specifice obținute după scanare. Au fost dezvoltate diferite tipuri de sisteme pentru a efectua recunoașterea optică a caracterelor în diverse tipuri de documente, dar sarcina nu este ușoară, deoarece documentele diferă nu doar în termeni de conținut, ci și în formate, fonturi, vechime sau deteriorare. După revizuirea sistemelor existente, vom propune unul care utilizează două motoare OCR bine cunoscute și un principiu de votare bazat pe ponderi. Sunt analizate și rezultatele tehnicii noastre combinate, în opoziție cu abordarea individuală a celor două motoare alese [3].

Cele mai utilizate motoare OCR:

Ocropy¹ (cunoscut și sub numele de Ocropus) este un motor OCR bazat pe LSTM (Long Short-Term Memory - Memorie pe Termen Lung și Scurt), care abordează provocările complexe ale recunoașterii caracterelor în imagini cu diverse fonturi și calitate de imagine variabilă [4].

Ocrad² (cunoscut și sub denumirea de GNU OCR) reprezintă o tehnica de recunoaștere optică a caracterelor dezvoltată în cadrul proiectului GNU (GNU's Not Unix). Acesta oferă o soluție open-source pentru transformarea imaginilor cu text în conținut digital, permițând utilizatorilor să extragă și să prelucreze automat informații din imagini [5].

¹ <https://en.wikipedia.org/wiki/OCROpus>

² https://www.gnu.org/software/ocrad/manual/ocrad_manual.html

SwiftOCR¹ este o bibliotecă de recunoaștere optică a caracterelor (OCR) specializată pentru platformele Apple, cum ar fi iOS și macOS. Dezvoltată în limbajul de programare Swift, biblioteca se mulează perfect pe ecosistemul Apple, oferind o integrare nativă cu servicii și tehnologii precum Core Image și Vision Framework [6].

Attention-OCR² reprezintă o abordare modernă în domeniul recunoașterii optice a caracterelor bazată pe tehnici de învățare automată și mai precis pe modelele de rețele neuronale cu mecanism de atenție [7].

Tesseract³ - este un proiect OCR dezvoltat de Google începând cu 2006 [8]. A evoluat mult de-a lungul anilor, pornind de la un simplu cititor de text bazat pe rețele neuronale NN (Neural Network), fără suport pentru analiza template-ului, până la un sistem complet, care recunoaște machetele comune și oferă suport atât pentru recunoașterea NN, cât și pentru cea LSTM. Versiunile ulterioare ale Tesseract suportă diferite formate de ieșire, inclusiv hOCR (HTML - based OCR) cu informații despre machetă și formatare, sau pot fi integrate cu interfețe utilizator precum Ocropus [9] [10].

Asprise⁴ - este puternic și fiabil, dezvoltat de Asprise. Acesta oferă capacități avansate de recunoaștere a caracterelor și permite utilizatorilor să acceseze funcționalități OCR printr-o interfață simplă de programare API (Interfață de Programare a Aplicației). Asprise OCR se remarcă prin abilitatea sa de a manipula imagini cu diverse calități și formate, precum și prin performanțele sale în recunoașterea caracterelor în texte complexe [5].

Am prezentat anterior o serie de motoare OCR moderne și performante și cu toate acestea, Tesseract și Asprise au fost selectate datorită unor factori cheie: maturitate, comunitate, diversitatea funcționalităților, documentație, suport, performanță și precizie.

2.1.1 Metoda de combinare a rezultatelor Tesseract și Asprise

În cadrul acestei lucrări, s-a efectuat o analiză a motoarelor OCR Tesseract și Asprise, urmând propunerea un nou sistem care combină ambele tehnologii, implementând un mecanism de votare bazat pe ponderi pentru a obține cel mai precis rezultat posibil. Procesul de prelucrare

¹ <https://www.swift.org/>

² <https://github.com/emedvedev/attention-ocr>

³ <https://tesseract-ocr.github.io/tessdoc>

⁴ <http://asprise.com>

al fișierelor de intrare este împărțit într-o serie de etape, precum se poate observa în figura 2.1 [3] [11].

Prin aplicarea acestui mecanism de votare bazat pe ponderi, noul sistem asigură o abordare mai precisă și robustă în procesul de recunoaștere optică a caracterelor, ducând astfel la îmbunătățirea calității și acurateței rezultatelor finale. Rezultatele obținute prin această nouă abordare vor fi examinate și comparate în detaliu pentru a evalua performanța și eficiența sistemului propus în comparație cu metodele tradiționale. În final, acest studiu va contribui la avansarea și îmbunătățirea tehnologiilor OCR, aducând beneficii semnificative în diverse domenii de aplicare [12].

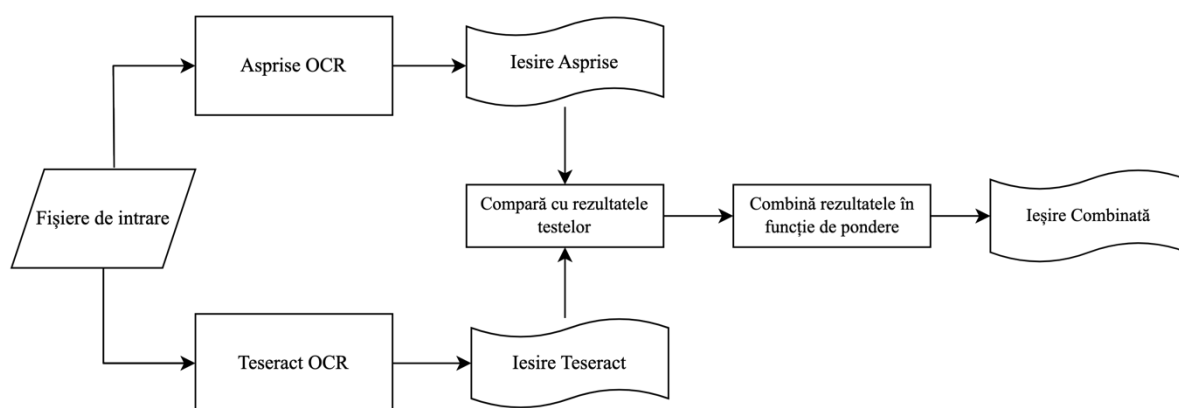


Figura Error! No text of specified style in document..1: Etapele de prelucrare a fișierelor de intrare [3].

După primirea fișierului de intrare, fiecare OCR a creat propria interpretare vizuală a textului pe baza nivelului de încredere.

Fișierul de intrare este reprezentat în prima imagine din figura 2.3, iar ieșirile OCR sunt delimitate în următoarele imagini prin culori în funcție de nivelul de încredere specific tehnicii.



Figura Error! No text of specified style in document..2: Imaginea de testare și rezultatul vizual generat de Tesseract și Asprise [3].

În ultima etapă, am combinat documentele pe baza ponderilor atribuite și am obținut rezultatul final, care poate fi observat în figura 2.6.

ENGLISH ROMAN. @ufque tandem abutere, Catilina, patientia nof'h'a? quamdiu nos etiam furor if'ce tuus eludet? quem a fincm
 fefe effremta jaefabit audacia? nihilne tc nofiumum praefidium palatii, nihilurbis vigiliae, nihil timor populi, nihil
 confou-|{us bonorum omnium, nihil hic munitifimus| ABCDEFGHIJKLMNOPBSTVUW

Figura Error! No text of specified style in document..3: Rezultatele după combinarea (Tesseract roșu și Asprise albastru) [3].

Au fost efectuate o serie de experimente pentru a evalua ponderile cele mai potrivite. În imaginile din figura 2.7 sunt prezentate câteva rezultate în funcție de nivelul de încredere al fiecărui motor OCR. Intensitatea culorii reprezintă nivelul de încredere: verde înseamnă încredere ridicată, roșu înseamnă încredere scăzută.

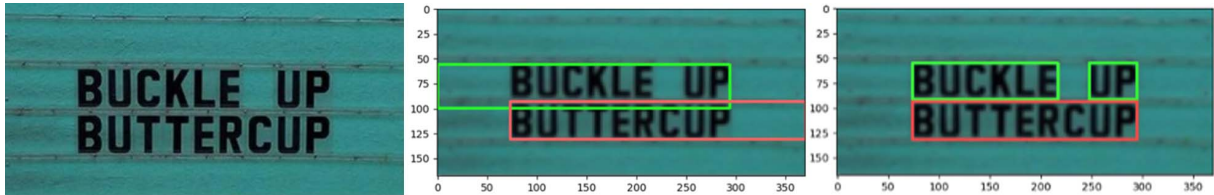


Figura Error! No text of specified style in document..4: Imaginea de testare, rezultatul Tesseract și Asprise [3].

După analizarea mai multor situații, se observă că la orice nivel de neclaritate a imaginii, Asprise tinde să performeze mai bine decât Tesseract, diferența ce se vede în figura 2.8.

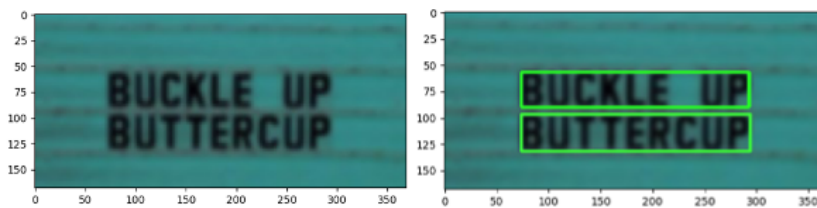


Figura Error! No text of specified style in document..5: Ieșire Tesseract (stanga) și Asprise (dreapta) [3]

Comparând rezultatele, se observă că fiecare dintre ele are propriile deficiențe în tratarea degradării fișierelor de intrare. Pentru a minimiza posibilele rezultate eronate oferite de fișierele de calitate scăzută, cu care majoritatea motoarelor se vor confrunta, am decis să creăm un mecanism de votare bazat pe ponderi, care analizează ambele rezultate și generează o ieșire bazată pe încrederea rezultatelor individuale [3].

2.2 Estimarea și Analiza Mișcării

Estimarea mișcării reprezintă un subiect esențial în domeniul viziunii computerizate, având o multitudine de aplicații, precum urmărirea și compresia video. Principiul de bază al estimării mișcării este, având la dispoziție o secvență de imagini (în general, un video), să identificăm mișcarea, adică ce a mișcat în secvență și cum sau cât de mult s-a mișcat (direcția și magnitudinea) pe măsură ce parcurgem imaginile din secvență. De obicei, dorim să

identificăm mișcarea dintre fiecare două cadre consecutive în același punct al spațiului imaginii [13].

În ultimii ani, domeniul estimării mișcării a evoluat considerabil [14,15]. În literatură sunt menționate două categorii majore de metode pentru estimarea mișcării: metode directe și metode indirecte [16].

În prima categorie, avem metodele prezentate în lumea științifică:

- metode dense/directe: estimează mișcarea în fiecare pixel pe baza variației temporale și spațiale a intensității în acel punct [17] [18] [19];
- metode de potrivire a blocurilor: împart spațiul imaginii în blocuri de dimensiuni egale și estimează mișcarea pentru fiecare bloc pe baza corelației dintre intensitățile pixelilor din acel bloc și cele din alt cadru. Acestea pot fi generalizate în așa-numitele metode de "potrivire a regiunilor" care folosesc regiuni de forme arbitrare în loc de blocuri dreptunghiulare [20–22];
- corelația de fază: estimează mișcarea globală dintre două cadre pe baza transformării lor Fourier. În unele materiale de cercetare, acestea sunt menționate ca metode în domeniul frecvenței [23];
- fluxul optic: calculează un vector de mișcare între două cadre consecutive [24].

A doua categorie este reprezentată de metodele indirecte, care se bazează pe caracteristici (detectarea colțurilor):

- metode bazate pe caracteristici: aceste metode identifică puncte caracteristice și le urmăresc pe mai multe imagini; estimează mișcarea doar a punctelor de caracteristică pe baza locațiilor în care sunt găsite în imagini [25];
- metodele dense se bazează pe o aproximare a seriei Taylor care este valabilă doar pentru deplasări mici ale aceluiași punct al obiectului în cele două cadre. Astfel, acestea sunt potrivite pentru videoclipuri cu o "mică" cantitate de mișcare. Pe de altă parte, metodele de potrivire a blocurilor pot detecta mișcări mai semnificative în funcție de dimensiunea regiunii de căutare; totuși, acestea sunt mai solicitante din punct de vedere computațional.

Metoda propusă calculează câmpurile vectorului de mișcare la același "nivel de raritate" folosind cele două metode descrise mai sus și apoi combină rezultatele lor într-o singură ieșire.

Flux optic

Această metodă [26,27] ia în considerare o secvență de imagini reprezentată printr-o funcție $I = I(x, y, t)$, unde x și y sunt coordonatele spațiale (locurile pixelilor) și t reprezintă

timpul (sau numărul cadrului). Scopul final este de a estima vectorul de mișcare $[u, v]$ pentru fiecare punct de eșantionare (x, y) în spațiul imaginii și între fiecare două cadre consecutive (t și $t + 1$).

S-a luat în considerare un anumit punct (punctul albastru din figura 2.9) la coordonatele (x, y) în cadru t , mișcându-se de-a lungul vectorului de mișcare $[u, v]$, astfel încât în cadru $t + 1$, punctul albastru va fi la locația $(x + u, y + v)$. O presupunere rezonabilă este că punctul va apărea în aceeași culoare (sau cu aceeași valoare de intensitate) în cadru $t + 1$ la noua locație, ceea ce oferă constrângerea de constanță a luminozității (ecuația (1)).

$$I(x, y, t) = I(x + u, y + v, t + 1) \quad (1)$$

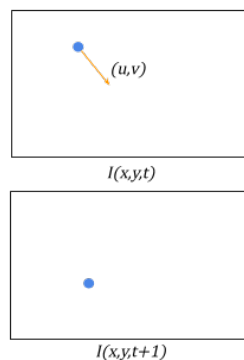


Figura **Error! No text of specified style in document..6**: Exemplu de Flux Optic [13].

Acum, dacă vectorul de mișcare $[u, v]$ este suficient de mic, se aplică următoarea expansiune a seriei Taylor:

$$I(x + u, y + v, t + 1) \approx I(x, y, t) + \frac{\partial I}{\partial x} \cdot u + \frac{\partial I}{\partial y} \cdot v + \frac{\partial I}{\partial t} \quad (2)$$

Din ecuațiile (1) și (2) se poate obține:

$$I_x \cdot u + I_y \cdot v + I_t = 0 \quad (3)$$

În ecuațiile (2) și (3), $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$, $\frac{\partial I}{\partial t}$, I_x , I_y , I_t reprezintă, estimările derivatei funcției de imagine I (în raport cu x , y , și t).

Vectorul de mișcare la (x, y) poate fi calculat rezolvând ecuația (3) pentru u și v . În acest caz, o soluție unică nu este disponibilă iar pentru a depăși această problemă, există unele abordări:

Horn & Schunck [19]– impun o constrângere de netezire asupra câmpului vectorului de mișcare pe întreaga imagine; împreună cu (II.3) aceasta devine o problemă de minimizare:

trebuie să găsim câmpul vectorului de mișcare $u(x, y), v(x, y)$ astfel încât următoarea expresie să fie minimizată:

$$\iint (I_x u + I_y v + I_t)^2 + \lambda(u_x^2 + u_y^2 + v_x^2 + v_y^2) dx dy \quad (4)$$

unde u_x, u_y, v_x, v_y reprezintă derivatele componentelor de mișcare și λ este un parametru ajustabil.

Lucas & Kanade [26] – presupun o netezire locală a câmpului vectorului de mișcare în vecinătatea punctului (x, y) ; mai precis, presupunerea este că vectorul de mișcare $[u, v]$ este același pentru fiecare pixel dintr-o mică fereastră în jurul aceluși punct; dimensiunea acestei ferestre este un parametru ce poate fi modificat. Aceasta este metoda utilizată în algoritmul bazat pe votare prezentat.

Scriind (3) pentru fiecare pixel i din fereastră, rezultă un sistem supra-constrâns de

N ecuații (numărul de pixeli din fereastră) cu doar 2 necunoscute, u și v :

$$\underbrace{\begin{bmatrix} I_{x_1} & I_{y_1} \\ I_{x_2} & I_{y_2} \\ \vdots & \vdots \\ I_{x_N} & I_{y_N} \end{bmatrix}}_A \begin{bmatrix} u \\ v \end{bmatrix} = - \underbrace{\begin{bmatrix} I_{t_1} \\ I_{t_2} \\ \vdots \\ I_{t_N} \end{bmatrix}}_b \quad (5)$$

unde indicele reprezintă numărul pixelului. Făcând notațiile de mai sus, rezultă:

$$A \cdot \begin{bmatrix} u \\ v \end{bmatrix} = b \quad (6)$$

Mai sus avem un sistem supra-constrâns de N ecuații și 2 necunoscute și, în general, nu are soluție. Cu toate acestea, putem găsi "cea mai apropiată soluție", cea care minimizează eroarea:

$$E = \left\| A \begin{bmatrix} u \\ v \end{bmatrix} - b \right\|^2 \quad (7)$$

rezolvând ecuația normală:

$$(A^T A) \begin{bmatrix} u \\ v \end{bmatrix} = A^T b \quad (8)$$

Dacă $A^T A$ este inversabilă, avem:

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b \quad (9)$$

care oferă următoarele expresii finale:

$$\begin{cases} u = \frac{\sum I_y I_t \cdot \sum I_x I_y - \sum I_x I_t \cdot \sum I_y^2}{\sum I_x^2 \cdot \sum I_y^2 - (\sum I_x I_y)^2} \\ v = \frac{\sum I_x I_t \cdot \sum I_x I_y - \sum I_y I_t \cdot \sum I_x^2}{\sum I_x^2 \cdot \sum I_y^2 - (\sum I_x I_y)^2} \end{cases} \quad (10)$$

unde sumele sunt efectuate peste întreaga fereastră în jurul punctului (x, y) .

Dacă punctul (x, y) corespunde unei regiuni plate, atunci vectorul de mișcare nu poate fi calculat, deoarece matricea

$$M = A^T A = \begin{bmatrix} I_x^2(x, y) & I_x(x, y)I_y(x, y) \\ I_x(x, y)I_y(x, y) & I_y^2(x, y) \end{bmatrix} \quad (11)$$

Se poate observa din (10) complexitatea calculării valorii unui vector de mișcare pentru 1 cadru este $O(W^2)$, unde W este dimensiunea ferestrei peste care se fac sumările.

Potrivirea blocurilor

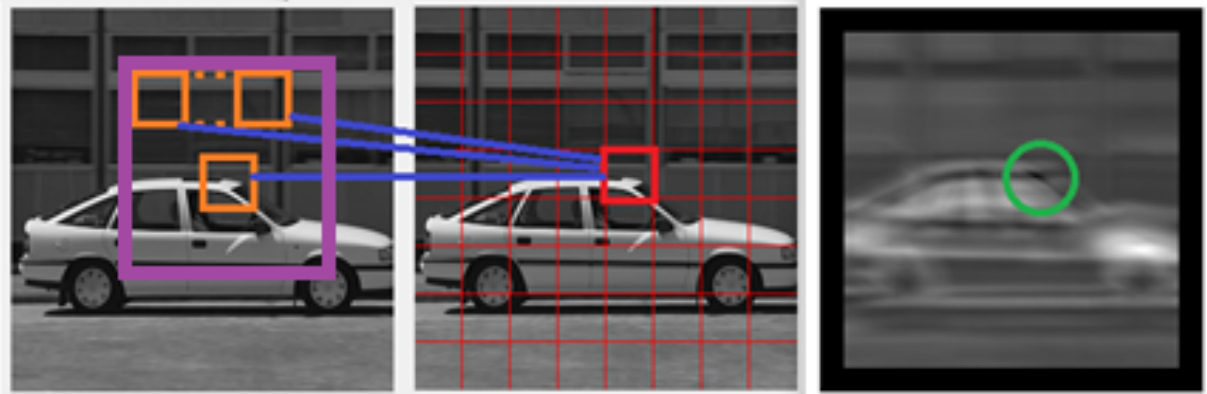


Figura Error! No text of specified style in document..7: De la stânga la dreapta: a. cadru de referință; b. cadru curent; c. funcție de (dis)similaritate. Unde: roșu – bloc curent, portocaliu – fereastră glisantă, mov – regiune de căutare, verde – minimumul funcției de disimilaritate [13].

Imaginea este împărțită la momentul t (cadru curent) într-o grilă dreptunghiulară de blocuri, așa cum se vede în figura 2.10. Apoi, pentru fiecare bloc, scopul este de a afla locația în spațiul imaginii unde se afla acel bloc în cadru anterior $t - 1$ (cadru de referință) – adică locația care se potrivește cel mai bine cu blocul curent [28]. Pentru a face acest lucru, s-a considerat o fereastră glisantă (având aceeași dimensiune ca blocul) care mătură cadru de referință și, pentru fiecare locație în cadru de referință, s-a calculat cât de similară este regiunea în fereastra glisantă cu blocul curent. Pentru asta, s-a folosit o funcție de (dis)similaritate precum eroarea pătratică medie.

Astfel, pentru un bloc la coordonatele (x, y) , și o anumită poziție a ferestrei glisante, deplasată de față de (d_x, d_y) de la (x, y) , funcția de disimilaritate/cost a celor două regiuni este:

$$\epsilon(d_x, d_y) = \sum_{i,j} \left(I_t(x+i, y+j) - I_{t-1}(x+d_x+i, y+d_y+j) \right)^2 \quad (12)$$

unde sumarea se efectuează pe o regiune de dimensiunea blocului în jurul punctului (x, y) . Pe măsură ce diferența dintre intensitățile celor două regiuni la locațiile corespunzătoare este mai mare, funcția de cost este mai mare. Astfel, provocarea este de a găsi deplasarea (d_x, d_y) care minimizează funcția de cost [29]. Pentru a face acest lucru, vectorul de mișcare (u, v) este ales să fie:

$$(u, v) = -\operatorname{argmin}_{(d_x, d_y)} \epsilon(d_x, d_y) \quad (13)$$

Alternativ, se poate utiliza o funcție de similaritate, precum corelația încrucișată, și să încercăm să o maximizăm:

$$\epsilon(d_x, d_y) = \sum_{i,j} I_t(x+i, y+j) I_{t-1}(x+d_x+i, y+d_y+j) \quad (14)$$

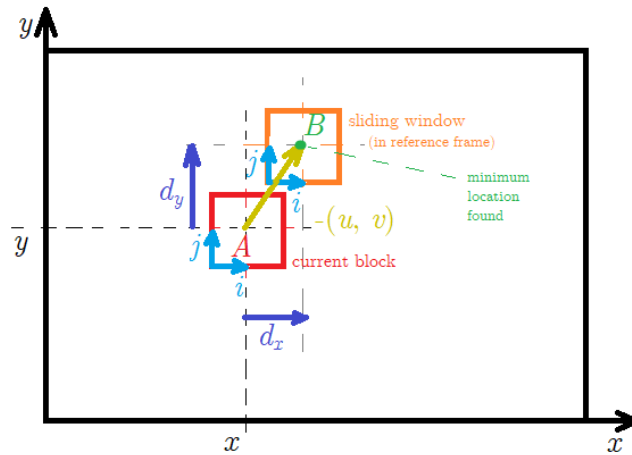


Figura **Error! No text of specified style in document.**8: Vectorul de compensare a mișcării \overline{AB} [13].

Dacă se denumește cu A centrul blocului curent și B locația care dă costul minim, se poate observa că punctul s-a mutat de la B la A , astfel încât vectorul de mișcare la punctul A este \overline{BA} (figura 2.11):

$$\begin{bmatrix} u \\ v \end{bmatrix} = \overline{BA} = \begin{bmatrix} x_A - x_B \\ y_A - y_B \end{bmatrix} \quad (15)$$

Dacă se știe că mișcarea este mai mică decât o valoare specifică u_{max} pixeli pe direcția x , $d_x x$ poate fi restricționat să varieze în intervalul $[-u_{max}, u_{max}]$ (similar pentru direcția y ; astfel, restricționarea este căutarea unui bloc într-o regiune de căutare în jurul (x, y)).

Complexitatea metodei este $O(S^2 B^2)$ pentru 1 vector de mișcare și pentru 1 singur cadru, unde S este dimensiunea regiunii de căutare și B este dimensiunea blocului.

Metoda votării

Această secțiune prezintă rezultatul celor două metode anterioare cu (u_1, v_1) și (u_2, v_2) . Apoi, metoda de votare consideră rezultatul metodei de votare ca fiind media aritmetică a celor două:

$$(u_{voting1}, v_{voting1}) = \left(\frac{u_1 + u_2}{2}, \frac{v_1 + v_2}{2} \right) \quad (16)$$

O altă posibilitate este să se ia media atât a magnitudinii, cât și a unghiului, așa cum se vede în figura 2.12.

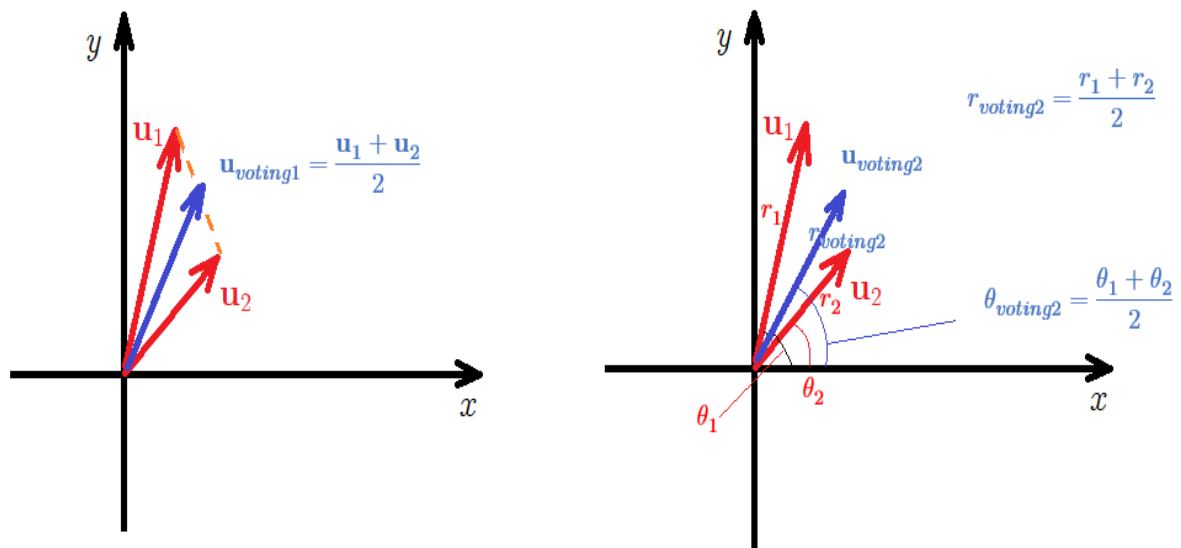


Figura Error! No text of specified style in document..9: Cele două cazuri pentru metoda de votare propusă [13].

Aplicația deschide un video și apoi parcurge fiecare cadru, generează și afișează un câmp de mișcare între cadru curent și cel anterior. În acest caz, această aplicație a considerat o grilă dreptunghiulară peste spațiul imaginii și efectuează estimarea mișcării în centrul fiecărei celule de grilă. Rezultatele abordării de votare sunt afișate în figura 2.13.



Figura Error! No text of specified style in document..10: Metoda propusă de votare. De la stânga la dreapta: a. imaginea originală, b. flux optic, c. potrivirea blocurilor, d. abordarea votului 1, e. abordarea votului 2 [13].

Metodele prezentate sunt intens consumatoare din punct de vedere computațional și au funcționat numai pentru imagini de intrare de dimensiune foarte mică. Metoda fluxului optic oferă adesea rezultate foarte imprecise, în special la limita obiectelor, unde presupunerea

câmpului constant de mișcare este încălcată. Metoda fluxului optic este instabilă în regiunile plate (numitorul din (10) tinde spre zero deoarece matricea din ecuația (11) este singulară). Mai mult, această metodă este susceptibilă la zgomot și funcționează numai pentru mișcări foarte mici (până la câțiva pixeli). Metoda de potrivire a blocurilor este mai robustă decât cea precedentă, dar din păcate, complexitatea este mai mare.

2.3 Strategii de Super-Rezoluție pentru Imagini

Super-rezoluția reprezintă o tehnică puternică în domeniul prelucrării imaginilor, având ca scop îmbunătățirea calității imaginilor, în special a celor cu rezoluție scăzută. Această tehnică are aplicabilitate într-o gamă variată de domenii, precum: fotografia digitală, video, imagistica medicală și în alte aplicații în care calitatea imaginilor joacă un rol foarte important [30] [31].

Procesul de super-rezoluție implică transformarea unei imagini de rezoluție redusă într-o imagine de rezoluție înaltă prin completarea acesteia cu informații ce trebuie approximate pe baza imaginii de calitate inferioară. Astfel, algoritmi de învățare automată sunt potriviți pentru această problemă, permițând obținerea de aproximări mai bune decât abordările euristice tradiționale. Tehnicile de învățare profundă reprezintă o alegere potrivită în acest context, deoarece pot extrage informații complexe și mai detaliate în procesul de învățare, în comparație cu rețelele mai puțin adânci, care de obicei extrag doar informații generale sau superficiale din imagini. Două dintre principalele clase de modele folosite pentru obținerea unor detalii excelente într-o imagine sunt CNN (rețele neuronale convoluționale) și GAN (rețele generative adversariale) [31].

În acest capitol, ne concentrăm pe selectarea rezultatelor obținute utilizând algoritmi bine cunoscuți din domeniul super-rezoluției pentru a obține rezultate mai bune în comparație cu utilizarea individuală a acestor algoritmi.

În subsecțiunile următoare, vom prezenta cele trei rețele neuronale profunde utilizate în experimentele noastre.

- RDN (Rețea Densă Reziduală);
- SRGAN (Rețea Generativă Adversarială pentru Super-Rezoluție);
- ESRGAN (Rețea Generativă Adversarială pentru Super-Rezoluție Îmbunătățită).

Algoritmii aleși pentru votare sunt următorii:

- Votarea Medie (pe bază de pixeli)
- Deplasare Maximă - se calculează distanțele dintre valorile pixelilor pentru fiecare algoritm, și se alege cel cu distanța maximă;

- Estimarea Zgomotului - imaginea este transformată în alb-negru, iar apoi se estimează cât de mult zgomot există în imagine, se alege imaginea cu cel mai mic zgomot (pe bază de fragmente);

Pentru fiecare imagine cu rezoluție înaltă, este oferită și o variantă a cărei calitate a fost redusă folosind una dintre următoarele proceduri:

- Interpolare Bicubică
- Operator Necunoscut – procedurile care au generat imaginea de calitate scăzută sunt păstrate ascunse pentru a evita găsirea unei metode care are rezultate doar pentru interpolarea bicubică.

Metode de evaluare

Pentru a evalua calitatea imaginilor, utilizăm metrica cea mai comună în publicațiile specializate și o comparăm cu rezultatele obținute prin rularea individuală a algoritmilor de învățare profundă.

Putem măsura succesul unei rețele prin aflarea în ce măsură reduce eroarea medie pătratică (MSE) între pixelii de ieșire și versiunea originală.

Cel mai bun rezultat este un MSE de 0, ceea ce înseamnă că imaginea originală de înaltă rezoluție și versiunea de înaltă rezoluție generată de rețea sunt identice.

$$MSE = \frac{1}{N} \sum_{i=1}^N (I_i - \hat{I}_{(i)})^2 \quad (17)$$

$$PSNR = 10 * \log_{10} \left(\frac{L^2}{MSE} \right) \quad (18)$$

Scopul ecuației PSNR (peak signal-to-noise ratio - raportul maxim semnal-zgomot) este de a calcula compensarea între MSE și valoarea maximă a pixelilor. Un PSNR mai mare reprezintă imagini generate de înaltă calitate.

În tabelul 2.3 prezentăm valorile PSNR obținute pentru cele trei metode individuale și apoi valorile obținute prin utilizarea strategiilor de votare propuse.

Algoritm	PSNR
RDN	19.17
SRGAN	19.36

ESRGAN	19.85
Votare medie	18.98
Cel mai îndepărtat	19.09
Estimarea zgomotului	19.37

Tabelul **Error! No text of specified style in document.**1: Valorile PSNR obținute prin utilizarea celor trei metode individuale și cele trei sisteme propuse de votare [13].

În acest capitol, am testat diverse modele de învățare profundă care rezolvă problema de super-rezoluție și am utilizat algoritmi de votare pentru a îmbunătăți performanța acestora. Deoarece o reconstituire perfectă a imaginii reduse este puțin probabilă, este necesar să se ia în considerare perspectiva umană asupra rezultatelor, deoarece din punct de vedere practic, este inutil ca o imagine să aibă PSNR ridicat și calitate subiectivă scăzută (așa cum este determinată de observatorii umani). Astfel, rezultatele ar trebui întotdeauna corelate cu observațiile umane.

2.4 Corectarea și Adaptarea Înclinării Paginii

Corecția înclinației paginii reprezintă o tehnică importantă în domeniul prelucrării imaginilor și al recunoașterii caracterelor. Ea implică detectarea și corectarea automată a înclinației sau a distorsiunilor de perspectivă dintr-o imagine a unei pagini, astfel încât textul să fie aliniat în mod corespunzător și să permită o recunoaștere precisă și eficientă a caracterelor [32].

Unul dintre principalele obiective ale algoritmilor de prelucrare a imaginilor este extragerea informațiilor. O clasă importantă de astfel de algoritmi este cea responsabilă pentru recunoașterea optică a caracterelor.

Această lucrare propune un sistem de votare care combină rezultatele din transformata Hough, profilarea proiecției și transformata Hough în domeniul frecvenței într-un singur răspuns. Din acest motiv, un nivel de încredere a fost adăugat la ieșirea fiecărui algoritm.

Pașii urmați pentru Algoritmul Transformării Hough sunt:

1. Prelucrați imaginea:
 - a. stabiliți un prag pentru imagine;
 - b. identificați contururile (litere/grupuri de litere) în imaginea cu prag;
 - c. generați o nouă imagine plasând puncte la aceleași coordonate ca și ancorele conturilor găsite în pasul anterior.

2. Căutați linii folosind funcția "HoughLines" din OpenCV ¹ V4.12 (funcția returnează o listă a liniilor cele mai votate, în ordine descrescătoare a voturilor).
3. Calculați nivelul de încredere:
 - d. liniile care au mai puține voturi decât cea mai votată, împărțit la o constantă, sunt eliminate (au fost date diferite valori constantei pentru a normaliza valorile de încredere ale algoritmilor);
 - e. calculați suma tuturor voturilor pentru liniile rămase;
 - f. calculați suma voturilor pentru liniile care au același unghi ca și linia dominantă;
 - g. împărțiți cele două sume și obțineți un număr între 0 și 1.

Abstract: Optical Character Recognition (OCR) is an indispensable tool for technology users nowadays, as our natural language is presented through text. We live under the need of having information at hand in every circumstance and at the same time having machines understand visual content and thus enable the user to be able to search through large quantities of text.

To detect textual information and page layout in an image page, the latter must be properly oriented. This is the problem of the so-called document deskew, i.e. finding the skew angle and rotating by its opposite. This paper presents an original approach which combines various algorithms that solve the skew detection problem, with the purpose of always having at least one to compensate for the others' shortcomings, so that any type of input document can be processed with good precision and solid confidence in the output result.

Keywords: deskewing; skew angle detection; automatic document orientation; computer vision; OCR preprocessing; image document analysis

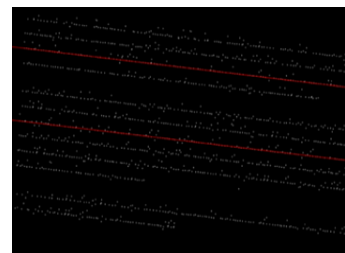
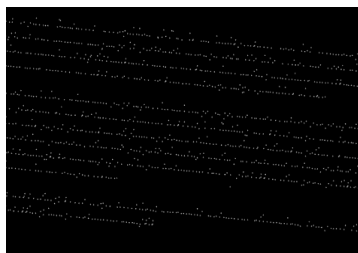


Figura **Error! No text of specified style in document..11**: De la stânga la dreapta: a. imaginea inițială; b. imagine cu puncte de ancorare; c. detectarea unghiului de înclinare [32].

Pașii pentru Algoritmul de Profilare prin proiecție sunt:

1. Prelucrarea imaginii
 - a. Aplicarea unui prag imaginii [11]
 - b. Identificarea conturilor (litere/grupuri de litere) în imaginea cu prag aplicat
 - c. Generarea unei noi imagini prin plasarea punctelor la aceleași coordonate ca și ancorele conturilor găsite în pasul anterior
2. Rotirea imaginii și găsirea unghiului pentru care proiecția verticală are o variație maximă.
3. Calculați nivelul de încredere:
 - a. împărțiți variația maximă la suma tuturor variațiilor (care a fost găsită, empiric, să aibă valori între 0 și 0,5);
 - b. scădeți 0,5;
 - c. luați valoarea absolută;
 - d. înmulțiți cu o constantă (pentru a normaliza valorile de încredere ale algoritmilor);
 - e. scădeți rezultatul din 1;

¹ OpenCV, Open source Computer Vision library, URL: <https://opencv.org>

Pașii pentru algoritmul Transformării Hough în Domeniul Frecvenței sunt:

1. Prelucrați imaginea:
2. calculați transformata Fourier
3. creați imaginea amplitudinilor
4. stabiliți un prag pentru imagine [1]
5. căutați linii folosind opencv houghlines
6. calculați nivelul de încredere (pașii sunt aceiași ca la transformata Hough obișnuită)

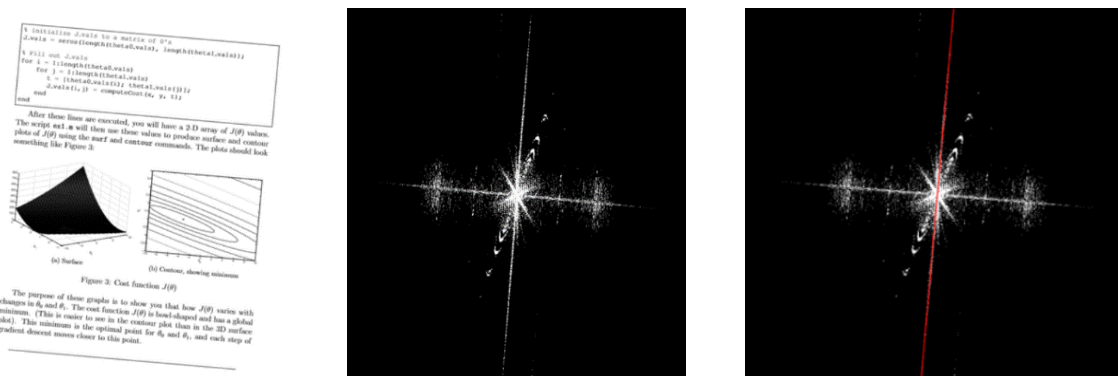


Figura 2.12: De la stânga la dreapta: a. imaginea inițială; b. FFT procesat; c. detectarea unghiului de înclinare [32].

Sistemul de votare combină unghiurile și încrederea rezultate din toți acești algoritmi și oferă o soluție optimă. Trei mecanisme de votare au fost abordate: primul cel mai bun, votul unanim și votul ponderat. Prima metodă alege rezultatul algoritmului cu cea mai mare încredere, neglijând complet celelalte. Votul unanim calculează unghiul de înclinare ca fiind media fiecărui rezultat al algoritmului. Ultima metodă, votul ponderat, funcționează în același mod ca votul unanim, dar de data aceasta se calculează o medie ponderată. Rezultate mai bune pot fi obținute dacă rezultatul cu o încredere mai mică decât jumătate din maxim este respins.

Votarea "Best First" produce cele mai bune rezultate, așa cum se poate observa în tabelul 2.4.

Metoda de corectare a înclinării	Acuratețe	Medie Eroare	Confidență medie	Timp de execuție (ms)
Linii Hough	99.150%	0.085°	0.632	4.473
Profilare prin proiecție	99.523%	0.048°	0.545	131.019
Linii hough pe frecvențe	99.363%	0.064°	0.481	51.512

Best First Voting (votare primul cel mai bun)	99.568%	0.043°	0.752	187.004
Votare unanimă	99.541%	0.046°	0.616	187.004
Votare ponderată	99.553%	0.045°	0.651	187.004

Tabelul **Error! No text of specified style in document..2**: Măsurători de Performanță [32].

Rezultatele sistemelor de vot sunt ușor mai bune decât rezultatele individuale ale votanților. Pe un set de date de 1414 imagini, acuratețea a crescut cu 0,04%, ceea ce poate părea o creștere foarte mică. Cu toate acestea, acuratețea inițială era deja peste 99%, astfel încât creșterea acurateței este semnificativă.

2.5 Evaluarea și Interpretarea Contrastului

Contrastul joacă un rol esențial în domeniul imaginilor digitale și al viziunii artificiale, influențând în mod semnificativ calitatea vizuală și interpretarea ulterioară a acestora. Evaluarea și cuantificarea acestui contrast au devenit o nevoie critică într-un spectru larg de aplicații, inclusiv imagistica medicală, teledetecția și fotografia digitală. Acest studiu oferă o analiză comprehensivă a diferitelor metrici disponibile pentru evaluarea contrastului imaginilor, concentrându-se pe formulele lor subiacente, interpretări, avantaje, dezavantaje și scenarii de utilizare relevante. Sunt comparate metrici din perspective bazate pe histograme, frecvență spațială și statistice. Se explorează complexitatea computațională, precizia și robustețea acestor metrici, inclusiv sensibilitatea lor la zgomot și alte degradări ale imaginilor [34].

Având în vedere rolul crucial al contrastului, capacitatea de a-l măsura și cuantifica cu precizie are o importanță deosebită. Această necesitate nu este valabilă doar pentru tehnicile de îmbunătățire a imaginilor, ci și pentru evaluarea calității imaginilor și performanța algoritmilor de procesare a imaginilor [35] [36]. Metricile precise de contrast servesc ca repere pentru evaluarea eficacității algoritmilor de îmbunătățire a imaginilor și pot ghida dezvoltarea de noi metodologii [37] [38].

În contextul importanței evaluărilor precise ale contrastului, acest capitol prezintă o analiză detaliată a mai multor metrici de contrast utilizate pe scară largă. Acestea includ atât metrici bazate pe histograme globale și locale [39], metrici bazate pe frecvențe spațiale, precum contrastul bazat pe transformata Fourier (FTC) și contrastul bazat pe transformata Wavelet

(WTC), cât și metrici statistice, cum ar fi contrastul mediu pătratic, contrastul Michelson și contrastul lui Weber [40] [41].

Fiecare dintre aceste metrici are caracteristici, avantaje și dezavantaje unice, care le fac potrivite pentru aplicații și scenarii specifice [42].

Metrici bazate pe histogramă

1. Contrast global al histogramelor (*GHC*):

$$GHC = Max(I) - Min(I) \quad (22)$$

unde $Max(I)$ și $Min(I)$ reprezintă valorile maxime și minime ale intensităților pixelilor în imaginea I [39].

2. Contrastul local al histogramelor (*LHC*):

$$LHC(x, y) = Max(I(x, y)) - Min(I(x, y)) \quad (23)$$

unde $Max(I(x, y))$ și $Min(I(x, y))$ reprezintă nivelurile maxime și minime ale intensităților pixelilor în imaginea I din blocul situat la coordonatele (x, y) .

Metrici bazate pe frecvența spațială

1. Contrastul bazat pe transformata Fourier (*FTC*):

$$FTC = \sum |FT(I(u, v))|^2 \quad (24)$$

unde $FT(I(u, v))$ reprezintă transformata Fourier a imaginii I la coordonatele frecvenței (u, v) [43]. Pătratul magnitudinii (notat cu $|\cdot|^2$) este însumat peste toate frecvențele pentru a obține FTC.

2. Contrastul bazat pe transformata Wavelet (*WTC*):

$$WTC = \sum |WT(I(x, y))| \quad (25)$$

unde $WT(I(x, y))$ reprezintă transformata Wavelet a imaginii I la coordonatele spațiale (x, y) , iar suma este luată peste toate imaginile de sub-benzi [43].

Metrici de contrast statistic

1. Contrastul radacina medie patrată (RMS):

$$RMS\ Contrast = \sqrt{\frac{1}{M \cdot N} \sum (I(x, y) - \mu)^2} \quad (26)$$

unde $I(x, y)$ indică intensitatea pixelului la poziția (x, y) în imagine, μ reprezintă intensitatea medie a imaginii, iar M și N sunt dimensiunile imaginii [40].

2. Contrastul Michelson:

$$MC = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (27)$$

unde I_{max} reprezintă cea mai mare intensitate a imaginii, iar I_{min} reprezintă intensitatea minimă.

3. Contrastul lui Weber

$$WC = \frac{I - I_b}{I_b} \quad (28)$$

unde I reprezintă intensitatea obiectului, iar I_b reprezintă intensitatea fundalului [38].

Comparativ: metrici de evaluare

Metricile de contrast au fost utilizate în mod extensiv în prelucrarea imaginilor, viziunea artificială și domenii conexe pentru a evalua și îmbunătăți calitatea imaginilor. Cu toate acestea, eficacitatea fiecărei metrici poate varia în funcție de mai mulți factori, inclusiv complexitatea computațională, precizia, robustețea și sensibilitatea la zgomot și alte degradări ale imaginii. În această parte, comparăm măsurile de contrast menționate în acest studiu pe diferite dimensiuni.

În ceea ce privește complexitatea computațională, metricile globale cum ar fi contrastul global al Histogramelor, contrastul RMS, contrastul Michelson și contrastul lui Weber au, în general, o încărcătură computațională mai mică. Acestea necesită fie calcule statistice de bază, fie operații simple pe intensitățile pixelilor [40,44,45].

Pe de altă parte, metricile bazate pe frecvență spațială, cum ar fi contrastul bazat pe transformata fourier (FTC) și contrastul bazat pe transformata Wavelet (WTC), sunt mai exigente din punct de vedere computațional. Ele necesită efectuarea fie a unei transformate Fourier, fie a unei transformate Wavelet asupra imaginii, operații care pot fi computațional intensive, în special pentru imagini mai mari [43,46].

Precizia și robustețea unei metrici de contrast depind în mare măsură de tipul și complexitatea imaginilor evaluate. Metricile globale pot oferi estimări precise ale contrastului [35] pentru imagini mai simple, dar pot neglija variațiile de contrast locale în imagini mai complexe.

În schimb, FTC și WTC pot oferi estimări ale contrastului mai precise și mai robuste pentru imagini complexe, deoarece surprind variațiile locale de contrast. Cu toate acestea, alegerea funcției Fourier sau a funcției Wavelet și a altor parametri poate afecta semnificativ precizia acestor metrici [36,43].

În ceea ce privește sensibilitatea la zgomot și alte degradări ale imaginii, metricile bazate pe histogramă și statistică tind să fie sensibile la valorile extreme și la zgomot, care pot supraevalua estimarea contrastului [37].

Alegerea metricii de contrast depinde de cerințele specifice ale aplicației, inclusiv resursele computaționale disponibile, complexitatea imaginilor și nivelul de zgomot și alte degradări prezente în imaginile respective. În timp ce nicio metrică singulară nu este universal superioară, o combinație de metrici poate oferi adesea o evaluare cuprinzătoare și precisă a contrastului imaginii [43].

3 Metodă Ierarhică de Clusterizare pentru Înțelegerea și Analiza Structurii Paginilor

Această cercetare introduce o tehnică robustă și fiabilă pentru structurarea ierarhică a imaginilor document prin exploatarea puterii triangulației Delaunay. Element central al abordării noastre îl constituie formarea unui arbore de clusterizare, care încapsulează conținutul paginii prin exploatarea strategică a aranjamentelor elementelor de structură și a distanțelor relative dintre acestea. Prin aplicarea tehnicilor noastre, categorisim eficient pagina în clustere distincte care cuprind imagini, titluri și paragrafe. Cadrul ierarhic consecvent, întemeiat pe arborele de clusterizare, stabilește o schemă durabilă și de încredere a structurii documentului, accelerând astfel înțelegerea și examinarea acestuia [47] [47].

Capitolul de față se bazează pe lucrarea [48], adăugând unele explicații noi, scenarii de testare mai bune, un set de date mai adecvat și o nouă fază de postprocesare menită să îmbunătățească și mai mult rezultatele metodei propuse.

Primul pas implică prelucrarea datelor de intrare pentru a satisface cerințele algoritmului. Importantă pentru această strategie este utilizarea documentelor alb-negru. Prin urmare, indiferent de schema de culoare inițială, fiecare document este convertit în alb și negru [38] [49].

Generarea conturilor și segmentelor

După selecția de intrare, procesul implică generarea segmentelor de imagine, denumite "componente conectate" în cadrul imaginii. Un grup de pixeli negri conectați constituie o componentă conectată, o determinare care poate fi realizată printr-un algoritm simplu (figura 3.1). Pornind de la un pixel negru, algoritmul traversează toți pixelii negri până când rămân doar pixeli albi vecini [48].

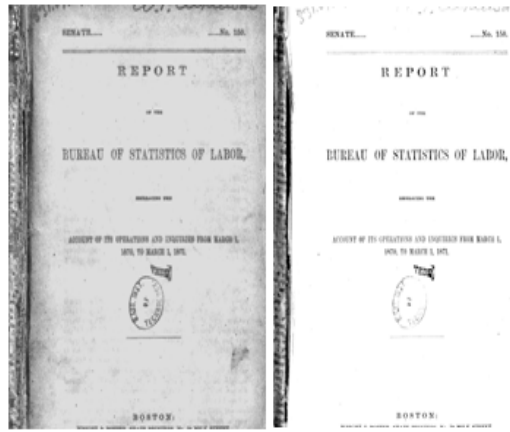


Figura Error! No text of specified style in document..13: Convertirea imaginii gri inițiale în format alb-negru [48].

Procedura este aplicată iterativ tuturor pixelilor negri care nu au fost vizitați, rezultând în dobândirea componentelor conectate. După cum este prezentat în figura 3.2 și 3.3, o colecție de pixeli negri poate fi închisă într-o varietate de configurații, cu dreptunghiul fiind cel mai comun.

Die Bundesregierung sagt voraus, daß es 1998 ein Wirtschaftswachstum von bis zu drei Prozent und am Jahresende weniger Arbeitslose geben wird. So steht es im Jahreswirtschaftsbericht, den das Kabinett am Mittwoch verabschiedete. Die Opposition sprach von „Schönfärberei“, Gewerkschaften vom „Prinzip Hoffnung“.

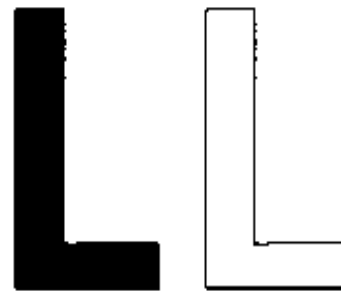


Figura Error! No text of specified style in document..14: Dreptunghiurile de înconjurare ale componentelor conectate [48].

Figura Error! No text of specified style in document..15: Rezultatul algoritmului de detectare a conturului [48].

Triangulația Delaunay

După selectarea conturilor, este folosit algoritmul de triangulație Delaunay constrânsă pentru a conecta toate componentele conectate. Acest lucru duce la un număr excesiv de conexiuni. Prin urmare, se efectuează o fază de prelucrare pentru a elimina triunghiurile care conectează mai mult de două componente conectate, lăsând doar componente conectate în perechi.

Ca urmare, acest capitol stabilește convențional două categorii distincte de puncte, și anume "puncte curente" și "puncte destinație". Aceste denumiri sunt derivate din asocierea sau lipsa de asociere cu o componentă conectată. Utilizând triangulația Delaunay, mai multe triunghiuri emană din fiecare componentă conectată și se extind către o altă componentă conectată. În cadrul acestor triunghiuri, punctele situate pe componenta conectată curentă sunt

denumite puncte curente, iar cele care aparțin triunghiurilor situate pe o componentă conectată diferită sunt denumite puncte destinație.

Proximitatea este relația dintre două sau mai multe componente conectate. Prin procesarea iterativă a triunghiurilor prezente în triangulația Delaunay constrânsă și excluderea triunghiurilor care conectează două componente conectate distincte (triunghiuri interne), proximitatea este determinată. Sunt excluse din procesare triunghiurile generate în cadrul unei singure componente conectate (triunghiuri interne) sau care implică trei componente conectate distincte.

Conceptul de arbore de clusterizare

Metodologia noastră implică construirea unui model ierarhic utilizând un arbore de clusterizare, o formă specializată a unui arbore cu mai multe ramuri. Componentele conectate servesc ca noduri frunză în această structură de arbore, în timp ce nodurile interne reprezintă clustere de componente conectate. Diametrul clusterului reprezintă distanța maximă între oricare două componente interconectate din același cluster sau între componente conectate învecinate care pot forma o secvență pentru a conecta oricare două astfel de componente (vezi figura 3.4). Obiectivul principal al acestui arbore este să categorisească componente conectate în clustere caracterizate prin diametre în creștere. Rădăcina arborelui semnifică clusterul cu cel mai mare diametru.

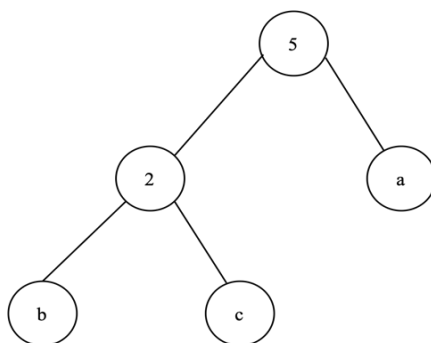


Figura **Error! No text of specified style in document.**16: Arbore de clusterizare simplu în care nodurile interne sunt etichetate cu diametrele clusterelor [48].

Algoritmul pentru construirea arborelui începe prin calcularea lungimii minime a laturii triunghiului Delaunay care conectează fiecare pereche de componente conectate. Arborele este construit de jos în sus. Se creează un cluster în jurul unei componente conectate aleator. Apoi, se determină și se adaugă în cluster componenta conectată cea mai apropiată de componenta inițială. Ulterior, se identifică și se adaugă în cluster componenta conectată cea mai apropiată de una dintre cele două componente conectate existente, iar dacă distanța până la această componentă conectată este comparabilă cu distanța dintre primele două, a treia componentă

conectată este adăugată în cluster. Acest procedeu se repetă până când cea mai apropiată componenta conectată cea mai apropiată are o magnitudine atât de mare încât nu poate fi inclusă în clusterul inițial. Clusterelor rămase sunt generate folosind o abordare similară, deși algoritmul ar putea introduce în loc de cea mai apropiată componentă conectată, cel mai apropiat cluster. Rezultatul algoritmului este structura de arbore căutată, care reflectă cu exactitate ierarhia paginii.

Die Strafe r nicht bewiesen

medizinischen Attesten
1. Sie sprach während des
lich von „schlechter Be-
strafte werden müsse.
nun, daß nach dem Frei-
sten die Jugendlichen er-
gestellt werden könnten.
egen elf der Schüler we-
ft in einer terroristischen
r ausgesetzt worden, weil
ie möglicherweise unter
gekommen waren. Die
ren zu jeweils 15 Jahren
orden.

Öffentlicher Dienst

Schlichter zeigen sich „gemäßigt optimistisch“

stg BREMEN, 11. März. Die beiden
Vorsitzenden der Schlichtungskommission für die schwierige Tarifrunde im Öffentlichen Dienst sind nach eigenen Aussagen „gemäßigt optimistisch“, daß sie ein Kompromißpaket schnüren können. Die Tarifparteien seien einigungswillig, sagten Bremens Ex-Bürgermeister Hans Koschnick (SPD) und der frühere Regierungschef von Rheinland-Pfalz, Carl-Ludwig Wagner (CDU), am Mittwoch nach dem ersten Tag der Schlichtungsgespräche in Bremen. Die Verhandlungen sollen am Montag fortgesetzt werden.



Figura Error! No text of specified style in document..17: Imaginea inițială și rezultatul triangulației Delaunay [48].

A doua abordare pentru construirea ierarhiei implică utilizarea coordonatelor derivate din formele de înconjurare ale componentelor conectate și utilizarea distanței dintre aceste forme de înconjurare ca metrică. Mecanismul de înveliș convex este o opțiune adecvată pentru forma de înconjurare. În acest caz, învelișul convex al fiecărei componente conectate este calculat folosind punctele sale de contur, iar distanța minimă dintre structurile de înconjurare este determinată și folosită, așa cum s-a făcut și înainte, ca distanță minimă între componentele conectate. Similar cu metoda anterioară, algoritmul începe cu un set gol și dezvoltă un cluster prin adăugarea de componente conectate cu cea mai scurtă distanță minimă dintre structurile de înconjurare. Clusterelor sunt construite în mod similar, cu posibilitatea de a include alte clusterelor dacă distanțele lor sunt de aceeași magnitudine.

Figura 3.5 ilustrează două versiuni ale aceleiași imagini: una înainte de triangulația Delaunay și cealaltă după aplicarea sa pentru a oferi o reprezentare vizuală și a facilita înțelegerea. În aceste imagini, componentele conectate sunt vizibil legate prin numeroase muchii, accentuând interconexiunea lor. Fiecare culoare reprezintă conexiunile dintre punctele de pe marginea a două componente conectate, oferind o indicație distinctă a distanțelor lor respective. Ca intrare pentru algoritm, doar conexiunea cu cea mai scurtă distanță este selectată din această colecție.

Metoda de analiză a structurii prezentată în acest document reprezintă o evoluție naturală a procedurilor de clusterizare ierarhică. Ea simulează retragerea progresivă a privitorului dintr-

un document, rezultând într-o imagine care devine din ce în ce mai "neclară". Cu toate că detaliile precise se pierd, structura generală a documentului este încă vizibilă, inclusiv poziționarea paragrafelor, titlurilor, tabelelor și imaginilor.

Utilizarea structurilor de triangulare Delaunay asigură menținerea preciziei pe parcursul mai multor niveluri de eșantionare, capturând cu fidelitate fenomenul de grupare (clusterizare) perceput de ochiul uman pe măsură ce distanța de vizualizare se mărește.

În plus, sistemul vizual uman manifestă o sensibilitate mai mare față de structurile caracterizate de o formă rectangulară. Prioritatea este acordată reconstrucției rectangulară a clusterelor din document prin utilizarea funcțiilor de arie a clusterelor care încurajează maximizarea locală.

După ce analiza imaginii documentului a generat o structură preliminară, este necesară o etapă de postprocesare vitală pentru a rafina rezultatul, eliminând în mod eficient elementele potențial suprapuse și asigurând cea mai bună reprezentare a elementelor structurii. Acest proces utilizează Suprimarea Valorilor Non-Maxime (NMS), o tehnică des utilizată în sarcinile de viziune artificială pentru a reduce numărul de dreptunghiuri de încadrare suprapuse și a le reține doar pe cele mai potrivite [50,51].

4 Predicția Interacțiunilor Fizice Utilizând Rețele Neuronale

Obiectivul principal este de a proiecta, dezvolta și evalua o arhitectură de rețea neuronală artificială capabilă să emuleze și să prezică tiparele dinamice de interacțiune manifestate în timpul întâlnirii dintre două entități distincte. Acest capitol se concentrează în principal pe învățarea și înțelegerea computațională a impulsurilor fizice asociate care apar atunci când aceste obiecte intră în contact, elucidând interacțiunile fizice complexe implicate. Acest proces include utilizarea strategică a unui motor fizic existent pentru a genera seturile de date necesare antrenării, oferind astfel o bază solidă și completă pentru antrenarea și evaluarea ulterioară a performanței rețelei neuronale [52].

Pentru a examina și a valida eficacitatea modelului de rețea neuronală artificială propus, acest studio include și o analiză comparativă riguroasă. Principalul focus al acestei comparații este de a juxtapune rezultatele obținute de rețeaua neuronală antrenată față de cele produse de motorul fizic original. Scopul aici este de a evalua precizia, fiabilitatea și aplicabilitatea practică a modelului antrenat în prezicerea precisă a impulsurilor fizice, demonstrând astfel potențialul său de a servi ca o alternativă viabilă la motorul fizic tradițional.

Un aspect crucial în predicția interacțiunilor fizice este selectarea și pregătirea adecvată a setului de date folosit pentru antrenarea rețelelor neuronale. Acest set de date trebuie să conțină informații relevante despre interacțiunile fizice dintre obiecte sau particule, astfel încât rețelele neuronale să poată învăța să realizeze predicții precise. Procesul de pregătire a setului de date implica: colectarea datelor experimentale, curățarea datelor, etichetarea datelor și divizarea setului de date.

În această secțiune, sunt prezentate diferite arhitecturi și modele de rețele neuronale utilizate pentru a realiza predicții cu privire la interacțiunile fizice.

Exemple de arhitecturi de rețele neuronale folosite în predicția interacțiunilor fizice:

- Rețele Neuronale Convoluționale (CNN - Convolutional Neural Networks)
- Rețele Neuronale Recurente (RNN - Recurrent Neural Networks):
- Rețele Neuronale bazate pe Atenție (Attention-based Neural Networks)

Arhitectura aplicației experimentale

Pentru a antrena eficient rețeaua neuronală, este important să furnizăm date de intrare și de ieșire care reflectă în mod precis logica utilizată de motorul fizic original în determinarea rezultatului unei coliziuni între două obiecte și vitezele ulterioare ale acestor corpuri după impact.

În acest scop, datele au fost colectate din motorul fizic funcțional în timpul executării diverselor teste demonstrative sau "scenarii". Aceste scenarii au implicat obiecte care trec prin coliziuni aleatoare, iar datele poziționale ale celor două obiecte care intra în coliziune, precum și vitezele și vitezele angulare ale acestora la punctul de impact, au fost colectate ca date de intrare. Corespunzător, datele de ieșire au constat în viteza rezultată și viteza unghiulară a obiectelor în urma contactului lor.

Motorul fizic Box2D [53] a fost instrumentul-cheie în procesul de generare a datelor. Într-un scenariu cu mai multe obiecte care ar putea intra în coliziune, motorul separă convenabil scena în perechi de obiecte care intra în coliziune și actualizează vitezele rezultate pentru fiecare pereche în mod separat. Această abordare asigură rețelei neuronale date de înaltă calitate și reprezentative, consolidând capacitatea sa de a prezice cu precizie rezultatele coliziunilor.

Optimizarea rețelei neuronale

Pentru a obține o funcție de pierdere diminuată, a fost necesară o rafinare meticuloasă a setului de date de antrenament, datorită influenței sale profunde asupra acurateții rețelei neuronale. Inițial, s-a folosit un set de date generat aleator, trecând ulterior la un set de date

definit, bine organizat, pentru o eficiență îmbunătățită. În iterația inițială, seturile de puncte (x , y) au fost generate aleatoriu în intervalul $[15, 15]$ pentru axa Ox . Punctul de ordonată a fost fixat în mod constant la 15, în timp ce viteza unghiulară și rotația au fost de asemenea obținute aleatoriu.

Deși această metodă a produs funcții de pierdere apropiate de ținta noastră, au apărut anumite situații în timpul simulării realiste a rețelei neuronale în motor (unde rețeaua neuronală a înlocuit logica tradițională utilizată pentru calcularea vitezei derivate la contact), care nu au funcționat conform așteptărilor. Cauza principală a acestui comportament inconsistent poate fi atribuită naturii aleatorii a generării setului de date, care ar fi putut să conducă la o supra-reprezentare a anumitor scenarii și la o sub-reprezentare a altora.

Conștientizând acest model, s-a decis să renunțăm la seturile de date generate aleatoriu în favoarea unei metode iterative care să asigure o acoperire egală a unui set predefinit de cazuri. Acest lucru a fost realizat prin iterarea prin intervalul Ox inițial de $[-15, 15]$ cu un pas granular de 0,05. Pentru a ajuta rețeaua, acest interval a fost restrâns la $[-5, 5]$, cu includerea a două scenarii distincte: unul cu o singură cutie lovită de cutia cu bombă, iar celălalt implicând o stivă de 10 cutii supusă aceluiași impact al cutiei cu bombă (ultimul implicând că bomba a fost lansată din aceeași poziție pentru ambele scenarii).

Această cercetare deliniează o adaptare îmbunătățită a studiului prezentat în [54], realizată prin rafinări minuțioase pe mai multe dimensiuni.

În primul rând, structura setului de date de antrenament și de evaluare a suferit modificări semnificative pentru a se potrivi mai bine cerințelor de calcul. Acest lucru a fost realizat printr-un amestec complex de date, cu accent pe îmbogățirea combinației dintre seturile de date generate aleatoriu și cele sintetice. Datele sintetice utilizate în această cercetare simulează mai aproape datele reale, generând un set de date mai robust și versatil. Acest amestec favorizează un mediu de învățare mai cuprinzător pentru rețeaua neuronală, permițându-i să extrapoleze eficient pe o gamă mai largă de scenarii și să-și sporească capacitățile de generalizare.

În al doilea rând, s-a produs o schimbare semnificativă către o explorare mai cuprinzătoare în spațiul hiperparametrilor, ceea ce a condus inevitabil la îmbunătățiri ale performanței rețelei neuronale. Calibrarea subtilă a hiperparametrilor este crucială pentru optimizarea procesului de învățare, influențând în mod decisiv performanța generală a modelului de rețea neuronală. Prin urmare, o căutare minuțioasă și ajustarea acestor parametri au fost instrumentale în identificarea setului optim care conferă dinamică superioară de învățare.

După implementarea acestor constrângeri, generarea datelor de antrenament în conformitate cu aceste constrângeri și antrenarea rețelei neuronale cu setul de date rafinat, funcții de pierdere în intervalul 10-4 au fost obținute după 5000 de iterații de antrenament. Tabelul 4.1 ilustrează funcția de pierdere rezultată într-un ciclu de antrenament care cuprinde 2000 de iterații.

Iterații	Scorul Testului	Scorul de antrenament
0	0.105798	0.106044
10	0.051244	0.051222
50	0.015246	0.015388
100	0.007767	0.007789
1000	0.001212	0.001225
2000	0.000728	0.000730

Tabelul **Error! No text of specified style in document..3**: Rezultatul funcției de pierdere, după un ciclu de antrenament care cuprinde 2000 de iterații, cu constrângerile impuse pentru datele de antrenament [52].

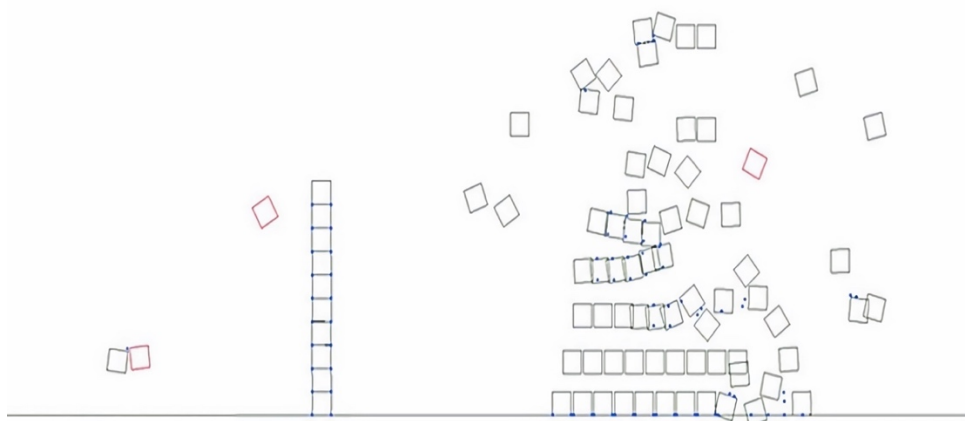


Figura **Error! No text of specified style in document..18**: Scenarii ilustrate utilizate pentru generarea datelor de antrenament [52].

Scenariul 1 (figura 4.4) demonstrează interacțiunea unui singur obiect, Scenariul 2 demonstrează interacțiunea unui singur obiect cu o stivă de zece obiecte suplimentare, iar Scenariul 3 evidențiază o interacțiune complexă care implică o piramidă compusă din șaiszeci

și șase de obiecte. Generarea datelor de antrenament implică lansarea unui obiect (pătratul mov, denumit bomba) într-un alt obiect sau o stivă de zece obiecte. Bomba începe de la poziția (-5, 15) și iterează până la poziția finală (5, 15) cu un pas de 0,05, colectând date din ambele scenarii cu aceeași poziție a obiectului bombă. Datele de intrare colectate includ pozițiile (în coordonate carteziane), vitezele, vitezele unghiulare și rotațiile celor două obiecte care urmează să intre în coliziune. Datele de ieșire reprezintă derivatele (în coordonate carteziane), care sunt apoi aplicate la vitezele inițiale înainte de contact pentru a obține vitezele finale (și direcțiile) ale obiectelor după contact.

Funcția de pierdere utilizată este funcția de eroare medie pătratică (MSE), cu o valoare de aproximativ 10^{-4} pentru rețea. Cu această valoare a funcției de pierdere, există încă erori, așa cum era de așteptat, deoarece această valoare nu reflectă o acuratețe de 100%, ceea ce ar implica o replicare perfectă a motorului fizic utilizat pentru antrenament.

Rețeaua are, de asemenea, un impact asupra performanței motorului fizic. Acest lucru este de așteptat, deoarece rețeaua neuronală necesită mai multă computație în comparație cu motorul fizic original, care utilizează o ecuație newtoniană simplă pentru a calcula derivata vitezei. Pentru scenarii simple, cum ar fi scenariul 1 și scenariul 2, acest impact nu este observat, deoarece atât motorul fizic, cât și rețeaua neuronală prezintă performanțe similare. Cu toate acestea, pentru scenarii mai complexe cu multe obiecte care interacționează (scenariul 3), computația necesară pentru predicția rețelei neuronale este semnificativ mai mare decât cea pe care motorul tradițional o va calcula, ceea ce duce la un impact semnificativ asupra performanței grafice.

În funcție de scenariul de testare, rețeaua neuronală atinge o rată de predicție de succes între 60% și 91%. Aceasta nu este încă o rată de predicție ideală, mai ales că nu există o rată consistentă în toate scenariile de testare.

Utilizarea unei rețele neuronale pentru a replica un motor fizic în anumite scenarii este cu siguranță posibilă și poate oferi o acuratețe acceptabilă. Această acuratețe ar putea fi îmbunătățită potențial prin generarea unui set de date mai mare, extinderea timpului de antrenament și ajustarea hiperparametrilor. Cu toate acestea, abordarea convențională, în special utilizarea funcțiilor de activare clasice, nu poate înlocui complet motorul fizic.

5 Investiții Eficiente în Dezvoltarea de Software

În acest capitol, vom explora în detaliu dezvoltarea de software, un domeniu esențial și dinamic care joacă un rol central în era tehnologiei și informației. Dezvoltarea de software

implică procesul complex de creare, implementare și gestionare a aplicațiilor software, care sunt utilizate într-o varietate de domenii și industrii.

Primul studiu este o comparație cuprinzătoare între producția de jocuri pentru calculator de tip AAA și Indie. Deși ambele abordări își propun să ofere cea mai bună experiență posibilă publicului lor, calitatea produsului poate varia în funcție de diferiți factori, inclusiv bugetul și tehnologia utilizată. Pentru a obține succes, dezvoltatorii de jocuri trebuie să aibă o înțelegere clară a scopului care poate fi realizat într-un interval de timp rezonabil. Acest studiu examinează rentabilitatea potențială a metodelor de dezvoltare a jocurilor [1] AAA și Indie. Deși nu există o singură metodă pentru succes, postulăm că dezvoltarea de jocuri AAA aduce în general un câștig mai mare decât dezvoltarea de jocuri Indie. Analiza noastră compară titluri de jocuri selectate în ceea ce privește bugetul de dezvoltare, efortul de marketing, dimensiunea echipei, tehnologia și oportunitățile de franciză. Descoperirile noastre sugerează că, în ciuda costului mai mic al dezvoltării jocurilor Indie, jocurile AAA oferă o metodă mai sigură de obținere a veniturilor [55].

Al doilea studiu oferă o comparație cuprinzătoare între proiectele open-source și proiectele closed-source, examinând mai multe aspecte-cheie, inclusiv numărul de contribuitori sau angajați, numărul de funcționalități introduse pe proiect, numărul de vulnerabilități prezente în software, veniturile sau profiturile și tehnicile de gestionare a proiectelor. Comparând aceste aspecte într-o serie de proiecte open-source și closed-source, încercăm să evaluăm potențialul ca aceste modele de distribuție să se completeze reciproc și să identificăm contextele în care un model ar putea fi mai eficient decât celălalt. Această analiză își propune să furnizeze informații valoroase pentru părțile interesate implicate în dezvoltarea software-ului, inclusiv dezvoltatori, manageri de proiect și factori de decizie [56].

5.1 Analiza Jocurilor: Indie vs. AAA

În acest subcapitol oferim o perspectivă asupra caracteristicilor și diferențelor dintre dezvoltarea jocurilor video "Triple-A" (AAA) și cele independente (Indie). Prima categorie este alcătuită din studiouri mari și editori care investesc resurse financiare substanțiale în crearea și promovarea francizelor de înaltă tehnologie pentru consolele de jocuri acasă, în timp ce dezvoltarea Indie implică în mod obișnuit echipe mici de până la 15 dezvoltatori care acordă prioritate independenței creative și autonomiei financiare [57].

Ecosistemul jocurilor AAA

Se pare că principalul stimul pentru dezvoltatorii de jocuri de a căuta angajare în companiile de jocuri AAA este perspectiva de a participa la un efort de echipă pe scară largă și

de a contribui semnificativ la dezvoltarea jocului. În contextul angajării în companiile de jocuri AAA, stabilitatea financiară este adesea menționată ca un avantaj-cheie pentru angajați. Lucrul pentru un studio mare poate oferi beneficii precum acoperire medicală și program flexibil [58]. În dezvoltarea jocurilor AAA, un buget și o echipă mare nu garantează o dezvoltare de joc de succes. Presiunea de a livra rezultate pentru o echipă și o bază de jucători numeroasă poate fi semnificativă și necesită ca indivizii să înțeleagă planul de dezvoltare și programul înainte de a se angaja într-un proiect [59].

Modelul tradițional AAA

Modelul de dezvoltare a jocurilor AAA funcționează în mod similar cu alte modele de afaceri, având ca focalizare principală generarea de profituri. Această accentuare a profitului provine din investițiile financiare considerabile necesare pentru dezvoltarea, testarea și promovarea unui singur proiect, dat fiind amploarea mare a acestor jocuri.

Evoluția și impactul jocurilor Indie

În contrast cu jocurile AAA, care necesită milioane de dolari, jocurile Indie sunt dezvoltate în mod obișnuit cu bugete mult mai mici, în general în gama de mii de dolari. Ca urmare, majoritatea dezvoltatorilor de jocuri Indie operează sub constrângeri de resurse, adesea îndeplinind mai multe roluri, cum ar fi tester, programator și designer [60].

În ceea ce privește dezvoltarea jocurilor Indie, obținerea finanțării suficiente poate fi o sarcină dificilă. Cu toate acestea, există diverse strategii pe care dezvoltatorii le pot adopta pentru a aborda această problemă, inclusiv: autofinanțare, finanțare prin strângere de fonduri, finanțare prin editori, tehnologia grafică, promovare de marketing, francizare și dimensiunea echipei.

În primele săptămâni ale lunii ianuarie 2023, a fost efectuat un sondaj pentru a determina preferințele persoanelor cu privire la jocurile video. Sondajul a avut 115 participanți, constând din 34 de studenți (30%), 48 de persoane cu vârsta cuprinsă între 11 și 18 ani (48%) și 33 de adulți angajați cu vârste cuprinse între 30 și 60 de ani (22%).

Sondajul a inclus următoarele întrebări:

1. Câte ore pe săptămână petreceți jucând jocuri Indie?
2. Câte ore pe săptămână petreceți jucând jocuri AAA?
3. Ce tip de joc (Indie sau AAA) preferați să jucați și de ce?
4. Pe o scară de la 0 la 5, cât de importante sunt grafica și efectele vizuale pentru dumneavoastră când jucați un joc?

5. Pe o scară de la 0 la 5, cât de importante sunt povestea și personajele pentru dumneavoastră când jucați un joc?
6. Pe o scară de la 0 la 5, cât de importantă este componenta multiplayer a unui joc pentru dumneavoastră?
7. Ați jucat vreodată un joc de tip " Indie " care credeți că ar fi trebuit să fie la fel de popular ca un joc de tip AAA?
8. Este probabil să recomandați unui prieten un joc Indie sau AAA?

Descoperirile studiului nostru indică faptul că numărul mediu de ore pe săptămână petrecute de către studenți jucând jocuri video Indie este de aproximativ 12, în timp ce elevii de liceu tind să petreacă puțin mai puține ore, aproximativ 10, jucând astfel de jocuri [55].

A doua întrebare a sondajului a produs rezultate remarcabile, care s-au diferențiat de cele ale interogării anterioare. Media de ore săptămânale pe care studenții le petrec jucând jocuri video AAA s-a redus la aproximativ 10 ore, în timp ce persoanele cu vârsta cuprinsă între 11 și 18 ani au petrecut de trei ori mai mult timp, în medie 30 de ore pe săptămână.

Întrebarea 3 a avut ca scop investigarea preferinței definitive a celor trei grupuri de participanți, iar rezultatele obținute sunt în concordanță cu descoperirile întrebărilor 1 și 2. Grupul de studenți a manifestat o mai mare înclinație către jocurile Indie, așa cum indică numărul de respondenți care preferă această categorie în fața jocurilor AAA (24 versus 10). În schimb, elevii de liceu au preferat jocurile AAA într-o mai mare măsură decât jocurile Indie (30 versus 18). Grupul de adulți, pe de altă parte, a demonstrat o preferință relativ echilibrată, cu 19 respondenți indicând preferința pentru jocurile AAA și 13 exprimând preferința pentru jocurile Indie. Cu privire la partea cu răspuns deschis a întrebării, motivul cel mai frecvent pentru popularitatea jocurilor Indie a fost stilul lor artistic distinct, în timp ce pentru jocurile AAA, factorul principal a fost posibilitatea de a participa în multiple confruntări între jucători (PvP).

Întrebarea următoare a oferit un rezultat concluziv, deoarece aproximativ 40 de respondenți au indicat că grafica și efectele vizuale sunt factori semnificativi în experiența lor de joc, indiferent de tipul de joc. Cei 75 de respondenți rămași au furnizat predominant evaluări moderate între 2 și 3 pe scara de importanță, cu 45 de răspunsuri în acest interval.

Întrebarea 5 a produs răspunsuri mai uniform distribuite pe scara de evaluare, cu o ușoară îndreptare către partea inferioară a scării în comparație cu întrebarea 4. În mod specific, 35 de respondenți au indicat că povestea și personajele unui joc sunt de o importanță mare (evaluare de 4 sau 5 pe scară), în timp ce 28 de respondenți au exprimat incertitudine sau un nivel mai scăzut de importanță (evaluare de 2 sau 3 pe scară).

Întrebarea 6 a colectat răspunsuri mixte din partea participanților, similar cu întrebarea anterioară. Dintre cei 115 respondenți, 30 au indicat o preferință puternică pentru componenta multiplayer a jocurilor video, în timp ce 25 au arătat puțin interes pentru aceasta. Cei 60 de respondenți rămași se încadrează într-o zonă intermediară, sugerând că importanța funcționalității multiplayer este extrem de subiectivă și depinde de jucătorul individual.

Întrebarea 7 a obținut rezultate intrigante, cu răspunsurile populare prezentând deja jocuri Indie prominente precum "Stardew Valley" și "Hollow Knight". Respondenții doresc ca jocurile Indie să obțină mai multă recunoaștere în lumea jocurilor AAA, cu titluri precum "Celeste" concurând cu jocurile Ubisoft sau Rockstar.

Întrebarea 8 reprezintă o întrebare esențială, având în vedere că recomandările personale din partea cunoștințelor sunt influente în modelarea preferințelor de joc. Datele sondajului evidențiază faptul că 72 de respondenți ar recomanda un joc AAA, în timp ce 43 ar recomanda un joc Indie.

5.2 Aspectele Open vs. Closed în Dezvoltarea de Software

Obiectivul principal al acestui subcapitol este de a efectua o comparație exhaustivă între proiectele software open-source și cele closed-source. Comparația se bazează pe mai multe aspecte-cheie, incluzând numărul de contribuatori/angajați, numărul de funcționalități introduse pe proiect, numărul de vulnerabilități (bug-uri sau probleme mai grave) prezente în software, veniturile și tehnicile de management de proiect. Un metric important de luat în considerare este numărul de contribuatori, care este relevant datorită diferențelor fundamentale în modelele de dezvoltare subiacente celor două tipuri de proiecte. Proiectele open-source se bazează pe contribuțiile din partea unui grup larg și divers de dezvoltatori, care pot participa în mod part-time sau ca angajați plătiți ai altor organizații. În schimb, proiectele closed-source depind de un număr fix de angajați care lucrează full-time și cărora li se atribuie sarcini specifice.

Analiza funcționalităților devine semnificativă în evaluarea succesului unui proiect de software, ceea ce influențează direct acceptarea acestuia de către utilizatori sau investitori. Numărul de funcționalități și timpul de dezvoltare sunt indicatori critici care determină viteza și eficiența cu care software-ul evoluează și este distribuit către utilizatorii finali. În cazul proiectelor open-source, am obținut date de la 10 aplicații, provenind din TAWOS Dataset [61], pe care le-am analizat și procesat conform cerințelor noastre de cercetare.

Examinând mai multe proiecte open-source, am calculat raportul dintre contribuatori și numărul total de funcționalități pentru primele trei produse într-un interval de timp comparabil de doi ani (2018-2020): MongoDB Core Server - 0,37, Moodle - 0,51, Sonatype Nexus - 0,92.

Proiectele closed-source se concentrează pe funcționalități specifice solicitate de clienții lor în timp ce proiectele open-source tind să fie mai difuze în obiectivele lor de funcționalitate, unde o anumită funcționalitate poate fi introdusă de un contribuitor pentru utilizarea personală sau în beneficiul unui număr mic de membri ai comunității, prioritizând practicitatea în defavoarea servirii unui număr mare de utilizatori finali. Această distincție poate fi rezumată prin termenul de impact: proiectele closed-source se concentrează pe mai puține funcționalități cu un impact mai mare, în timp ce proiectele open-source se extind asupra unui număr mai mare de funcționalități cu un impact mai mic.

Securitatea este un aspect crucial al dezvoltării software și, în consecință, este esențial să o luăm în considerare atunci când scriem cod. Prin urmare, a treia noastră metrică se concentrează asupra securității proiectelor atât closed-source, cât și open-source.

Beneficiile și riscurile investiționale

Investițiile în dezvoltarea de software reprezintă un aspect critic în economia digitală actuală, deoarece tehnologia joacă un rol esențial în succesul și inovarea afacerilor. Acest capitol se concentrează asupra beneficiilor și riscurilor asociate investițiilor în dezvoltarea de software, subliniind importanța unei abordări echilibrate și informate în luarea deciziilor.

Beneficiile investițiilor în dezvoltarea de software: îmbunătățirea eficienței operaționale, creșterea competitivității, adaptabilitatea și scalabilitatea, controlul asupra securității datelor, economii pe termen lung,

Riscurile investițiilor în dezvoltarea de software: costuri ridicate, riscul de eșec, durata de dezvoltare, dependența de echipa de dezvoltare, necesitatea actualizării și întreținerii.

6 Concluzii și Contribuții

Pe baza analizei efectuate în această teză de doctorat, s-au obținut contribuții semnificative în diverse domenii tehnologice, de la optimizarea tehnicilor OCR și estimarea mișcării, până la dezvoltarea de software și analiza structurii paginilor. Un aspect cheie al acestei cercetări a fost adoptia unei abordări interdisciplinare, care a permis identificarea de soluții inovatoare pentru probleme complexe. De exemplu, mecanismul de votare ponderat a îmbunătățit acuratețea în tehnologiile OCR, în timp ce analiza contrastului și estimarea mișcării au deschis noi căi pentru îmbunătățirea procesării imaginilor. În plus, studiul a examinat diferențele între jocurile Indie vs AAA, subliniind că inovația și adaptabilitatea pot învinge adesea resursele financiare substanțiale în dezvoltarea software-ului de succes.

Avansările în domeniile studiate nu sunt doar teoretice, ci au aplicabilitate directă în industrii variate. Metricile de contrast discutate sunt vitale pentru domenii precum imagistica medicală, teledetecția și fotografie digitală. Totodată, înțelegerea și optimizarea interacțiunilor fizice prin rețele neuronale pot avea implicații semnificative în domenii precum simulațiile fizice sau robotică. Cu toate acestea, lucrarea și-a asumat și rolul de a identifica lacunele și limitele actuale, oferind astfel un cadru pentru cercetările viitoare. De exemplu, deși rețelele neuronale pot îmbunătăți simulațiile fizice, ele nu pot încă înlocui complet motoarele fizice tradiționale. Acest echilibru între progresul realizat și identificarea domeniilor care necesită îmbunătățiri ulterioare face ca această teză să fie nu doar un punct de referință academic, dar și un instrument practic pentru industrie.

6.1 Lucrări și Publicații

Articole științifice în reviste și comunicări în proceedings de conferințe

Jurnale Indexate ISI

1. Cristian Avatavului, Rares-Cristian Ifrim, Mihai Voncila "CAN NEURAL NETWORKS ENHANCE PHYSICS SIMULATIONS?", BRAIN. Broad Research in Artificial Intelligence and Neuroscience Vol.14 No.2 - 2023, e-ISSN: 2067 – 3957, pp. 76-92 <https://doi.org/10.18662/brain/14.2/445>
2. Cristian Avatavului, Costin-Anton BOIANGIU „A HIERARCHICAL CLUSTER TREE APPROACH LEVERAGING DELAUNAY TRIANGULATION?”, BRAIN. Broad Research in Artificial Intelligence and Neuroscience, e-ISSN: 2067 – 3957 [Acceptat în Vol 14(3), 2023, în curs de publicare]

Comunicări în proceedings de conferințe indexate ISI

3. Iulia-Cristina Stănică, Costin-Anton Boiangiu, Giorgiana Violeta Vlăsceanu, Marcel Prodan, Cristian Avatavului, Răzvan-Adrian Deaconescu, Codrin Tăut, "A Survey on History, Present and Perspectives of Document Image Analysis Systems", in New technologies and redesigning learning spaces Book of abstracts, page 36, 15th eLearning and Software for Education Conference, Bucharest, 2019, ISSN 2360-2198, doi: 10.12753/2066-026X-19-025
4. Giorgiana Violeta Vlăsceanu, Costin-Anton Boiangiu, Răzvan-Adrian Deaconescu, Marcel Prodan, Cristian Avatavului, Răzvan Rughiniș, Irina Mocanu, "Designing a Document Image Analysis System on 3 Axis: Education, Research and Performance", in New technologies and redesigning learning spaces Book of abstracts, page 37, 15th eLearning

and Software for Education Conference, Bucharest, 2019, ISSN 2360-2198, doi: 10.12753/2066-026X-19-027

Indexate în alte baze de date

5. Remus Petrescu, Sergiu Manolache, Costin-Anton Boiangiu, Giorgiana Violeta Vlăsceanu, Cristian Avatavului, Marcel Prodan, Ion Bucur; "Combining Tesseract and Asprise results to improve OCR text detection accuracy", The Journal of Information Systems & Operations Management, Vol.13(1), 2019, ISSN 1843-4711, pp. 57-64
6. Alexandru ILINU, Cristian AVATAVULUI, Giorgiana Violeta VLĂSCEANU, Costin-Anton BOIANGIU; "VOTING-BASED MOTION ESTIMATION ", The Journal of Information Systems & Operations Management, Vol.14(1), 2020, ISSN: 1843-4711, pp. 82-92
7. Nicolae Tarbă, Daniel Schmidt, Anda - Elena Popovici, Eduard Stăniloiu, Cristian Avatavului, Marcel Prodan, "On performing skew detection and correction using multiple experts' decision", The Journal of Information Systems & Operations Management, Vol.14 No.2 - 2020 pp. 188-195
8. Robert STANCA, Eduard-Marius COJOCEA, Cristian AVATAVULUI, Costin-Anton BOIANGIU "On How To Combine Single Image Super-resolution Algorithms", The Journal of Information Systems & Operations Management, Vol.14 No.1 - 2020, ISSN: 1843-4711, pp. 140-150
9. Cristian Avatavului, Marcel Prodan „Evaluating image contrast: a comprehensive review and comparison of metrics”, The Journal of Information Systems & Operations Management, Vol.17 No.1 – 2023, ISSN 1843-4711, pp. 143-160.
10. Giorgiana Violeta Vlăsceanu, Cristian Avatavului, Costin-Anton Boiangiu, "An extensive review of metrics for evaluating image binarization", Journal of Information Systems & Operations Management, Vol 17 No. 1, 2023, ISSN: 1843-4711, pp: 200-222
11. Cristian Avatavului, Ilie-Octavian Sandu, Nicu-Cătălin Ioviță, Bogdan Vasile, Ciprian Duță, Costin-Anton Boiangiu, Mihai-Lucian Voncilă and Nicolae Tarbă „Indie vs AAAs: A Fair Comparison”, The Journal of Information Systems & Operations Management, ISSN 1843-4711 [Acceptat în Vol 17 (2), 2023, în curs de publicare]
12. Cristian Avatavului, Andrei-Iulian Cucu, Alexandru-Mihai Gherghescu, Costin-Anton Boiangiu, Iulia-Cristina Stănica, Cătălin Tudose , Mihai-Lucian Voncilă, Daniel Rosner „Open-source and closed-source projects: A fair comparison”, The Journal of Information Systems & Operations Management, ISSN 1843-4711 [Acceptat în Vol 17 (2), 2023, în curs de publicare]

Postere

13. Cristian AVATAVULUI, “Combining Tesseract and Asprise results to improve OCR text detection accuracy” – in Semicentennial Anniversary of the Department of Computers, University Politehnica of Bucharest 2018.

Bibliografie

1. Mueller JP, Massaron L. Artificial Intelligence For Dummies. 1st ed. For Dummies; 2018.
2. Szeliski R. Computer Vision: Algorithms and Applications. Springer; 2010.
3. Petrescu R, Manolache S, Boiangiu CA, Avatavului C, Prodan M, Bucur I, et al. Combining Tesseract and Asprise results to improve ocr text detection accuracy. *Journal of Information Systems Management*. 2019 May 1;13:57.
4. Bhowmick A, Govindaraju V. A Framework for Evaluation of Uncertain Inputs in OCR Systems. In: *Proceedings of the 8th IAPR International Workshop on Document Analysis Systems*. 2008.
5. SDK AOCR. Asprise OCR and its applications. Asprise; 2020.
6. Wshah S, Kumar G, Govindaraju V. Script Independent Word Spotting in Offline Handwritten Documents". In: *International Conference on Document Analysis and Recognition*. 2011.
7. Graves A, Schmidhuber J. Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks". In: *Advances in Neural Information Processing Systems* 21. 2009.
8. GitHub - tesseract-ocr/tessdoc: Tesseract documentation [Internet]. [cited 2023 Feb 1]. Available from: <https://github.com/tesseract-ocr/tessdoc>
9. In: OCR Community Help Wiki, URL [Internet]. Available from: <https://help.ubuntu.com/community/OCR>,
10. Graves A, Jaitly N. Towards End-to-End Speech Recognition with Recurrent Neural Networks". In: *Proceedings of the 31st International Conference on Machine Learning*. 2014.
11. Park SC, Park MK, Kang MG. Super-Resolution Image Reconstruction: A Technical Overview". *IEEE Signal Processing Magazine*. 2003;
12. Canny J. A Computational Approach to Edge Detection. *IEEE Trans Pattern Analysis and Machine Intelligence*. 1986;8[6]:679–98.
13. Ilinu A, Avatavului CD, Vlăsceanu GV, Boiangiu CA. Voting-Based Motion Estimation. *Journal of Information Systems & Operations Management*. 2020;14.
14. Moeslund T, Granum E. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*. 2001 Mar 1;81:231–68.
15. J.R. B, P. A, K.J. H, R H. Hierarchical modelbased motion estimation. In: Sandini G, editor. *Computer Vision*. p. 92.
16. Cédras C, Shah M. Motion-based recognition a survey. *Image and Vision Computing* [Internet]. 1995 Mar 1;13[2]:129–55. Available from: <https://www.sciencedirect.com/science/article/pii/026288569593154K>

17. A. French Optical Flow. Computerphile, editor. 2019.
18. A IE. Bobick Introduction to Computer Vision, udacity.com. 2015.
19. Shah M. Optical Flow. UCF Computer Vision Video Lectures; 2012.
20. Zhu S, Ma KK. A new diamond search algorithm for fast block-matching motion estimation. *IEEE Transactions on Image Processing*. 2000 Feb;9[2]:287–90.
21. Coursera [Internet]. Fundamentals of Digital Image and Video Processing. Available from: <https://www.coursera.org/learn/digital>
22. de Haan G, Biezen PWAC, Huijgen H, Ojo OA. True-motion estimation with 3-D recursive search block matching. *IEEE Transactions on Circuits and Systems for Video Technology*. 1993 Oct;3[5]:368–79.
23. Min Li, Biswas M, Kumar S, Truong Nguyen. DCT-based phase correlation motion estimation. In: 2004 International Conference on Image Processing, 2004 ICIP '04 [Internet]. Singapore: IEEE; 2004. p. 445–8. Available from: <http://ieeexplore.ieee.org/document/1418786/>
24. Sun D, Roth S, Black M. Secrets of Optical Flow Estimation and Their Principles. In 2010. p. 2432–9.
25. Torr PHS, Zisserman A. Feature Based Methods for Structure and Motion Estimation. In: Triggs B, Zisserman A, Szeliski R, editors. *Vision Algorithms: Theory and Practice*. Berlin, Heidelberg: Springer; 2000. p. 278–94. [Lecture Notes in Computer Science].
26. Argyriou V, Vlachos T. A study of sub-pixel motion estimation using phase correlation. In 2006. p. 387–96.
27. Xu L, Jia J, Matsushita Y. Motion Detail Preserving Optical Flow Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2012 Sep;34[9]:1744–57.
28. Bruenig M, Niehsen W. Fast full-search block matching. *Circuits and Systems for Video Technology, IEEE Transactions on*. 2001 Mar 1;11:241–7.
29. Nie Y, Ma KK. Adaptive rood pattern search for fast block-matching motion estimation. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*. 2002 Feb 1;11:1442–9.
30. Freeman WT, Jones TR, Pasztor EC. Example-based super-resolution. *IEEE Computer Graphics and Applications*. 2002;22[2]:56-65,.
31. STANCA R, COJOCEA EM, AVATAVULUI C. Costin-Anton BOIANGIU “On How To Combine Single Image Super-resolution Algorithms.” *The Journal of Information Systems & Operations Management*. 14[1–2020, ISSN]:1843–4711, 140–50.
32. Tarbă N, Schmidt D, Popovici AE, Stăniloiu E, Avatavului CD, Prodan M. On Performing Skew Detection and Correction Using Multiple Experts' Decision. *Journal of Information Systems & Operations Management*. 2020;14.
33. Boiangiu CA, Dvornic AI. Dan Cristian Cananau, Binarization for Digitization Projects Using Hybrid Foreground- Reconstruction. In: *Proceedings of the 5th IEEE International*

- Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca. 2009. p. 141–4.
34. Avatavului CD, Prodan M. Evaluating Image Contrast: A Comprehensive Review and Comparison of Metrics. *Journal of Information Systems & Operations Management*. 2023;17.
 35. Li C, Bovik AC. Three-component weighted structural similarity index. In: *Proceedings of SPIE - The International Society for Optical Engineering*. 2009.
 36. Kovesei P. - Image features from phase congruency - Videre. *Journal of Computer Vision Research*. 1999;1[3]:1–26.
 37. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. - Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*. 2004;13[4]:600–12.
 38. Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*. 1979;9[1]:62–6.
 39. Zuiderveld K. - Contrast Limited Adaptive Histogram Equalization. In: Heckbert P, editor. *Graphics Gems IV*. Academic Press; 1994. p. 474–85.
 40. D. Shaked, Tastl I. Sharpness measure: Towards automatic image enhancement. In: *Proceedings of the IEEE International Conference on Image Processing*. 2005.
 41. Asamoah D, Oppong E, Oppong S, Danso J. Measuring the Performance of Image Contrast Enhancement Technique. *International Journal of Computer Applications*. 2018;181:6–13.
 42. PIELLA G, HEIJMANS HJAM. A new quality metric for image fusion. In: *Proceedings of the IEEE International Conference on Image Processing*. 2003.
 43. STARK JA. Adaptive image contrast enhancement using generalizations of histogram equalization. 889-896. *IEEE Transactions on Image Processing*. 2000;9[5].
 44. Gonzalez RC, Woods RE. - *Digital Image Processing*. 4th ed. Pearson; 2018.
 45. HARALICK RM, SHAPIRO LG. - *Computer and Robot Vision - Volume I*. Addison-Wesley; 1992.
 46. Matheron G. - *Random sets and integral geometry* - Wiley. 1975.
 47. Avatavului CD, Boiangiu CA. A hierarchical cluster tree approach leveraging Delaunay triangulation. *BRAIN Broad Research in Artificial Intelligence and Neuroscience*. 14[3].
 48. Boiangiu CA, Cananau DC, Raducanu B, Bucur I. A Hierarchical Clustering Method Aimed at Document Layout Understanding and Analysis. *International Journal of Mathematical Models and Methods in Applied Sciences*. 2008 Jan 1;2:413–22.
 49. Sauvola J, Pietikäinen M. Adaptive document image binarization. *Pattern Recognition [Internet]*. 2000 Feb 1;33[2]:225–36. Available from: <https://www.sciencedirect.com/science/article/pii/S0031320399000552>

50. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017 Jun;39[6]:1137–49.
51. Malisiewicz T, Gupta A, Efros AA. Ensemble of Exemplar-SVMs for Object Detection and Beyond. In: *2011 International Conference on Computer Vision*. 2011.
52. Avatavului CD, Ifrim RC, Voncila M. Can Neural Networks Enhance Physics Simulations? *BRAIN Broad Research in Artificial Intelligence and Neuroscience*. 2023 Jun 30;14[2]:76–92.
53. Catto E. Fast and Simple Physics using Sequential Impulses. In: *Proceedings of Game Developer Conference*. 2006.
54. Ifrim RC, Penariu P, Boiangiu CA. Replicating Impulse-Based Physics Engine Using Classic Neural Networks. *Journal of Information Systems & Operations Management*. 2021;15[2].
55. Avatavului CD, Prodan M. Indie vs AAAs: A Fair Comparison. *Journal of Information Systems & Operations Management*. 2023;18.
56. Avatavului CD. Open-Source and Closed-Source Projects: A Fair Comparison. *Journal of Information Systems & Operations Management*. 2023;18.
57. Karthikeya K. The Difference Between Working in Indie and AAA Game Development [Internet]. *Gameopedia*; 2022. Available from: <https://www.gameopedia.com/indie-aaa-aa-games-comparison>
58. Shoshanah WALL. Should I Work for a AAA or Indie Video Game Studio?, *CGSpectrum*, 2022 [Internet]. Available from: <https://www.cgspectrum.com/blog/difference-between-aaa-vs-indie-game-studio>
59. Campbell C. What aaa can learn from indies – according to indies [Internet]. 2023. Available from: <https://www.gamedeveloper.com/audio/what-aaa-can-learn-from-indies--according-to-indies>
60. Mathews CC. Nia WEARN - How Are Modern Video Games Marketed? *The Computer Games Journal*. 2016;5:23-37,.
61. Tawosi V, Al-Subaihin A, Moussa R, Sarro F. A versatile dataset of agile open source software projects. In: *Proceedings of the 19th International Conference on Mining Software Repositories* [Internet]. Pittsburgh Pennsylvania: ACM; 2022. p. 707–11. Available from: <https://dl.acm.org/doi/10.1145/3524842.3528029>