**Universitatea Națională de Știință și Tehnologie Politehnica București**Facultatea de Electronică, Telecomunicații și Tehnologia Informației

Expertiză criminalistică pentru imagini şi secvenţe video

**Tema 6**

**Detectarea și recunoașterea fețelor în imagini de slabă calitate**

Profesor coordonator, Student,

Prof. Ioan VOICU Vintilescu Andreea-Alexandra

**București 2024**

**Cuprins**

[1. Introducere 3](#_Toc187092969)

[2. Provocările detectării fețelor în imagini de slabă calitate 4](#_Toc187092970)

[2.1. Variațiile de iluminare 4](#_Toc187092971)

[2.2. Rezoluția redusă 4](#_Toc187092972)

[2.3. Zgomotul digital 5](#_Toc187092973)

[2.4. Orientarea și expresiile faciale 5](#_Toc187092974)

[2.5. Ocluzia feței 6](#_Toc187092975)

[3. Soluția propusă 6](#_Toc187092976)

[3.1 Preprocesare 6](#_Toc187092977)

[3.1.1. Filtrarea zgomotului 6](#_Toc187092978)

[3.1.2. Normalizarea iluminării 7](#_Toc187092979)

[3.2. Detectarea trăsăturilor principale 7](#_Toc187092980)

[3.2.1. Algoritmi tradiționali 7](#_Toc187092981)

[3.2.2. Algoritmi bazați pe deep learning 7](#_Toc187092982)

[3.3. Recunoașterea fețelor 8](#_Toc187092983)

[3.3.1. Modele CNN 8](#_Toc187092984)

[3.3.2. Metrici de similaritate 8](#_Toc187092985)

[3.4. Evaluare manuală 8](#_Toc187092986)

[4. Testarea practică a soluției propuse 9](#_Toc187092987)

[4.1. Setul de date utilizat 9](#_Toc187092988)

[4.2. Metodologia de testare 9](#_Toc187092989)

[4.3. Rezultate și discuții 10](#_Toc187092990)

[5. Concluzii 13](#_Toc187092991)

[6. Bibliografie 14](#_Toc187092992)

# Introducere

Detectarea și recunoașterea fețelor joacă un rol important în domenii precum biometria și expertiza criminalistică, fiind indispensabile în aplicații legate de securitate, supraveghere video și identificarea persoanelor. Aceste procese presupun utilizarea unor tehnologii complexe pentru a analiza și compara imagini faciale, chiar și în condiții nefavorabile. Biometria se concentrează pe identificarea unei persoane pe baza caracteristicilor biologice, cum ar fi trăsăturile anatomice, fiziologice sau comportamentale, în timp ce identificarea criminalistică se bazează pe analiza trăsăturilor vizibile ale unei persoane, pentru a stabili identitatea acesteia. În cadrul criminalisticii, identificarea poate fi realizată fie prin metode automate, fie prin examinarea manuală a trăsăturilor observabile, cum ar fi fața sau alte semne distinctive.

Recunoașterea facială automatizată reprezintă compararea de imagini faciale cu imagini dintr-o bază de date, folosind algoritmi specializați. În schimb, în procesul de identificare criminalistică, tehnicile manuale implică analiza atentă a diferențelor și asemănărilor dintre imagini pentru a verifica identitatea unei persoane.

În ceea ce privește imaginile de calitate redusă, acestea complică semnificativ aceste procese, din cauza limitărilor tehnologice și a variabilității condițiilor de captură. Problemele frecvente includ zgomotul digital, iluminarea slabă, rezoluția scăzută, expresiile faciale variabile, ocluziile sau diferențele de orientare. Pentru a aborda aceste dificultăți, cercetătorii utilizează diverse metode de analiză, cum ar fi măsurarea distanțelor și unghiurilor între trăsături (analiza antropometrică), compararea formei și texturii feței (analiza morfologică), analiza semnelor particulare și alte tehnici care permit evidențierea detaliilor unice.

În prezent, tehnologiile de detectare și recunoaștere facială au înregistrat progrese importante datorită integrării algoritmilor avansați de învățare automată și a rețelelor neuronale profunde. Cu toate acestea, performanța lor este semnificativ afectată în cazul imaginilor de calitate scăzută. Prin urmare, elaborarea unor metode robuste și eficiente pentru detectarea și recunoașterea fețelor în astfel de condiții reprezintă o direcție de cercetare de actualitate.

# Provocările detectării fețelor în imagini de slabă calitate

## Variațiile de iluminare

Iluminarea joacă un rol crucial în capturarea detaliilor faciale. Imaginile surprinse în condiții de iluminare scăzută sau neuniformă pot diminua claritatea trăsăturilor faciale, afectând negativ performanța algoritmilor de detectare și recunoaștere. Studiile arată că variațiile de iluminare pot reduce acuratețea sistemelor de recunoaștere facială cu până la 30% [1].

O imagine care conține îmbrăcăminte, persoană, încălțăminte, sacou

Descriere generată automatO imagine care conține îmbrăcăminte, persoană, încălțăminte, sacou

Descriere generată automat

Fig. 1. Omiterea detectării unei fețe cauzată de variațiile de iluminare

## Rezoluția redusă

Imaginile cu rezoluție scăzută nu conțin detaliile necesare pentru identificarea precisă a trăsăturilor faciale. De exemplu, în imagini cu o rezoluție sub 100x100 pixeli, algoritmii tradiționali de recunoaștere facială înregistrează o scădere semnificativă a performanței, cu rate de eroare de peste 40% [1].

O imagine care conține persoană, îmbrăcăminte, om, în aer liber

Descriere generată automat

Fig. 2. Omiterea detectării unei fețe cauzată de rezoluția redusă a imaginii

## Zgomotul digital

Zgomotul digital, introdus în timpul capturării sau procesării imaginilor, poate distorsiona trăsăturile faciale, complicând detectarea și recunoașterea. Prezența zgomotului poate duce la o scădere a acurateței cu până la 25% în sistemele automate de recunoaștere facială [2].

O imagine care conține îmbrăcăminte, persoană, sacou, încălțăminte

Descriere generată automatO imagine care conține îmbrăcăminte, încălțăminte, clădire, persoană

Descriere generată automat

Fig. 3. Omiterea detectării unei fețe cauzată de existența zgomotului digital

## 2.4. Orientarea și expresiile faciale

Variabilitatea unghiurilor de captură și diversitatea expresiilor faciale reprezintă provocări suplimentare. Algoritmii trebuie să fie suficient de robusti pentru a recunoaște fețele din multiple perspective și cu diferite expresii. Studiile indică faptul că variațiile de orientare pot reduce performanța sistemelor de recunoaștere facială cu până la 20% [4].

O imagine care conține vehicul, Vehicul de teren, în aer liber, stradă

Descriere generată automatO imagine care conține vehicul, Vehicul de teren, în aer liber, clădire

Descriere generată automat

Fig. 4. Omiterea detectării unei fețe cauzată de variabilitatea unghiurilor de captură

## 2.5. Ocluzia feței

Ocluzia parțială a feței, cauzată de obiecte precum ochelari, măști sau păr, poate împiedica detectarea completă a trăsăturilor faciale. Algoritmii tradiționali întâmpină dificultăți în prezența occluziunilor, înregistrând scăderi ale acurateței de până la 35% [3].





Fig. 5. Omiterea detectării unei fețe cauzată de ocluzia parțială a feței

# Soluția propusă

Pentru a aborda aceste provocări, lucrarea prezentă propune o metodologie integrată care combină tehnici de preprocesare a imaginii și utilizarea rețelelor neuronale convoluționale (CNN), însă pot fi efectuate și evaluări antropometrice manuale. Soluția propusă se desfășoară în următoarele etape:

## 3.1 Preprocesare

### 3.1.1. Filtrarea zgomotului

Filtrarea imaginilor digitale reprezintă o operație ce se aplică local fiecărui pixel din imagine, înlocuind valoarea intensității sau culorii pixelului curent cu o valoare ce depinde de valorile de intensitate/culoare ale pixelilor vecini.

În cazul de față, pentru filtrarea zgomotului, sunt aplicate mai multe filtre, precum filtrul median sau cel gaussian, ce îmbunătățesc astfel claritatea trăsăturilor faciale. Filtrul median este eficient în eliminarea zgomotului de tip sare și piper, în timp ce filtrul gaussian este adecvat pentru reducerea zgomotului gaussian.

O imagine care conține îmbrăcăminte, accesoriu de cap, pălărie, Accesoriu vestimentar

Descriere generată automatO imagine care conține îmbrăcăminte, accesoriu de cap, pălărie, Chip de om

Descriere generată automatO imagine care conține accesoriu de cap, îmbrăcăminte, pălărie, Chip de om

Descriere generată automat

1. Imagine originală b) Imagine afectată de c) Imagine afectată de

zgomot median zgomot gaussian

Fig. 6. Imagini afectate de zgomotul median și gaussian

### 3.1.2. Normalizarea iluminării

Pentru a corecta variațiile de iluminare, au fost utilizate tehnici precum egalizarea histogramei sau algoritmi mai avansați, cum ar fi Retinex. Aceste metode redistribuie intensitățile pixelilor pentru a accentua detaliile din zonele mai puțin iluminate ale imaginii.

## 3.2. Detectarea trăsăturilor principale

### 3.2.1. Algoritmi tradiționali

Detectarea fețelor prin metode tradiționale, precum clasificatorii Haar sau metode bazate pe histograme ale gradientului orientat (HOG), oferă o abordare rapidă și eficientă pentru imagini de calitate medie. Totuși, performanța lor scade semnificativ în prezența imaginilor de slabă calitate .

### 3.2.2. Algoritmi bazați pe deep learning

Soluțiile moderne se bazează pe rețele neuronale convoluționale (CNN), cum ar fi Multi-task Cascaded Convolutional Networks (MTCNN) și YOLO (You Only Look Once). Acestea identifică fețele și punctele de referință (landmarks) cu o acuratețe superioară, chiar și în condiții dificile.

## 3.3. Recunoașterea fețelor

### 3.3.1. Modele CNN

Pentru recunoașterea fețelor, rețelele precum FaceNet, VGGFace sau ResNet extrag caracteristici relevante, transformând fețele într-un spațiu vectorial. Vectorii pot fi apoi comparați utilizând măsuri de similaritate, precum distanța cosinusului.

### 3.3.2. Metrici de similaritate

Compararea trăsăturilor extrase se face utilizând metrici precum similaritatea cosinusului, care cuantifică gradul de apropiere între vectorii de trăsături faciale. Această metodă este robustă în fața variațiilor mici de iluminare sau poziționare .

Așadar, similaritatea cosinusului este utilizată pentru:

* Măsurarea similarității: Dacă doi vectori faciali au o similaritate cosinus ridicată (de exemplu, dacă valoarea > 0.8), ei sunt considerați ca aparținând aceleiași persoane.
* Clasificare: Această metrică ajută la compararea fețelor necunoscute cu cele dintr-o bază de date pentru a identifica persoanele.

### 3.4. Evaluare manuală

În situațiile în care metodele automate eșuează, evaluarea manuală devine esențială și implică utilizarea unor tehnici variate. Acest proces ține cont de faptul că pot exista persoane cu proporții faciale similare sau cu trăsături morfologice apropiate. Totodată, identificarea unei diferențe clare, care nu poate fi justificată prin calitatea imaginii, orientarea sau expresia facială, poate conduce la excluderea unui suspect și la formularea unei concluzii negative certe.

Proceduri utilizate:

* Analiza fotogrammetrică (sau antropometrică): constă în compararea relațiilor spațiale între diverse elemente ale fizionomiei, prin măsurarea distanțelor și a unghiurilor dintre punctele caracteristice și semnele particulare ale feței;
* Analiza morfologică: implică evaluarea formei și aspectului diferitelor elemente fizionomice;
* Analiza prin superpoziție: se bazează pe suprapunerea parțială a figurilor comparate pentru a verifica continuitatea trăsăturilor în imaginile disponibile;
* Analiza semnelor particulare [5].

# Testarea practică a soluției propuse

## 4.1. Setul de date utilizat

Am utilizat două seturi de date publice: **Labeled Faces in the Wild (LFW)** și **CelebA**, care conțin imagini cu fețe etichetate, capturate în condiții variate de iluminare și rezoluție .

**Labeled Faces in the Wild (LFW)** conține aproximativ 13.000 de imagini ale peste 5.700 de persoane, fiecare persoană fiind reprezentată de una sau mai multe fotografii. LFW se remarcă prin faptul că imaginile provin din viața reală și sunt realizate în contexte naturale, ceea ce oferă o provocare semnificativă pentru modelele de recunoaștere facială. Setul de date include variabilități în ceea ce privește poziția capului, expresiile faciale, iluminarea și calitatea imaginii, făcându-l ideal pentru evaluarea performanței sistemelor de recunoaștere în condiții variate.

Pe de altă parte, **CelebA** este un set de date cu un număr semnificativ mai mare de imagini, depășind 200.000, și conținând aproximativ 10.000 de celebrități. Acesta este folosit pentru sarcini complexe privind recunoașterea facială și clasificarea atributelor faciale. Fiecare imagine din CelebA este etichetată cu 40 de atribute faciale binare precum și informații suplimentare despre locația punctelor-cheie faciale, cum ar fi ochii, nasul și gura.

## 4.2. Metodologia de testare

Pentru testarea soluției propuse m-am inspirat din codul pus la dispozitie de către Iván de Paz Centeno pe GitHub: <https://github.com/ipazc/mtcnn> unde am adăugat mai multe filtre de reducere a zgomotului, tehnici de egalizare a histogramei, ș.a.m.d. astfel încât aplicația să se muleze mai mult pe cerințele lucrării de față. Așadar, aplicația cuprinde următoarele etape:

* **Preprocesare**

Primul pas a constat în aplicarea unor tehnici de filtrare și normalizare a iluminării, menite să îmbunătățească calitatea imaginilor preluate din cele două seturi de date. Aceasta a fost o etapă critică, având în vedere variabilitatea iluminării în diferite condiții de captură a imaginilor. În cadrul acestei faze, am utilizat algoritmi de ajustare a contrastului și operații de filtrare pentru a reduce zgomotul și a uniformiza distribuția iluminării pe întreaga imagine. Scopul principal al acestui pas a fost de a mă asigura că modelele de recunoaștere facială nu sunt influențate negativ de variațiile de iluminare, iar fețele erau extrase într-un mod mai clar și mai consistent. În plus, a fost realizată normalizarea culorii pentru a compensa diferențele de tonuri și luminozitate între imagini, îmbunătățind astfel stabilitatea și fiabilitatea algoritmilor de recunoaștere în condiții variate de iluminare.

* **Detectare**

După preprocesarea imaginilor, am trecut la detecția fețelor utilizând modele pre-antrenate, MTCNN (Multi-task Cascaded Convolutional Networks), care s-au dovedit a fi eficiente în localizarea fețelor într-o gamă largă de imagini. MTCNN este un algoritm bazat pe rețele neuronale convoluționale care detectează fețele printr-un proces în mai multe etape, asigurându-se că sunt identificate corect, indiferent de poziție, dimensiune sau expresie. Acest model este capabil să îndeplinească simultan sarcini de detectare a feței și aliniere a punctelor-cheie faciale, ceea ce îl face deosebit de eficient pentru imagini variate. Utilizarea unui model pre-antrenat a permis accelerarea procesului, evitând necesitatea antrenării unui model de la zero, și a asigurat o acuratețe mare în detecția feței chiar și în condiții dificile, cum ar fi unghiuri neobișnuite sau iluminare slabă.

* **Recunoaștere**

Odată ce fețele au fost detectate și aliniate, am continuat cu recunoașterea propriu-zisă. Trăsăturile faciale au fost extrase prin metode de învățare profundă ce sunt capabile să extragă caracteristici unice. Compararea acestor trăsături s-a realizat utilizând măsuri de similaritate, precum distanța cosinus pentru a determina gradul de asemănare dintre fețele detectate și cele stocate în baza de date. Aceste măsuri au permis stabilirea unui scor de similaritate, iar pe baza acestuia, sistemul a putut să identifice sau să verifice identitatea persoanei respective. Procesul a fost optimizat pentru a reduce numărul de erori și a asigura o precizie ridicată în cadrul aplicațiilor practice, cum ar fi sistemele de securitate sau autentificare.

## 4.3. Rezultate și discuții

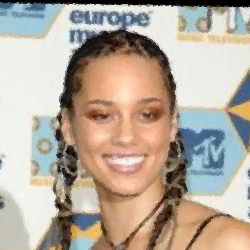
Am atașat mai jos exemplul oferit de către autor cu privire la functionalitatea aplicației. Algoritmul a citit imaginea originală ( în cazul de față ivan.jpg ), a redimensionat-o pentru a putea fi detectate fețe de diferite dimensiuni, iar mai apoi, în cadrul primei rețele P-Net a avut loc detecția inițială unde s-a realizat acel chenar portocaliu ( în esență coordonatele unor potențiale fețe) si scorul de incredere pentru fiecare zonă propusă ca posibilă față, eliminând suprapunerile și păstrând doar cadrele cu cele mai mari probabilități. Mai apoi, regiunile candidate de la rețeaua anterioară au fost transmise către a doua rețea R-Net, ce are ca scop îmbunătățirea preciziei pozitiei chenarului portocaliu propus și eliminarea zonelor ce nu prezintă fețe reale. Detecția finală a fost realizată în ultima rețea O-Net, ce primește bounding box-urile deja rafinate, urmând mai apoi să ajusteze și mai precis coordonatele și să estimeze locația punctelor cheie: poziția ochilor, vârful nasului, colțurile gurii. În plus, aceasta păstrează doar detecțiile cu grad înalt de încredere și elimină duplicatele.

O imagine care conține text, captură de ecran, persoană, îmbrăcăminte

Descriere generată automat

Fig. 7. Captură de ecran cu rezultatul obținut la simpla rulare a proiectului

Așadar, pentru testarea soluției propuse, m-am folosit de 6 imagini JPG din setul de date **Labeled Faces in the Wild** preluat de pe <https://www.kaggle.com/datasets/jessicali9530/lfw-dataset?resource=download> . De menționat faptul că nu am păstrat estimatul punctelor cheie în afișarea noilor rezultate. Prin urmare, în cadrul pozelor nu vor fi observate locatiile exacte ale poziției ochilor, nasului sau colțurile gurii.

O imagine care conține Chip de om, persoană, cravată, îmbrăcăminte

Descriere generată automat

a) Angela Bassett b) Adrian Năstase C) Alicia Keys

Fig. 8a. Imagini folosite în validarea codului

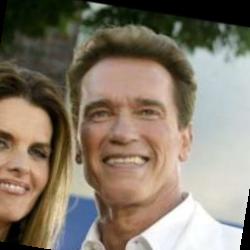
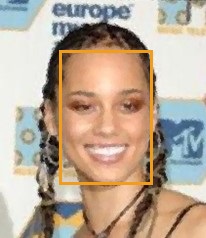
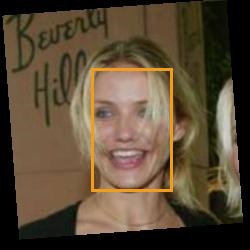
 d) Arnold Schwarze e) Cameron Diaz f) Ben Affleck

Fig. 8b. Imagini folosite în validarea codului

Rezultate obținute pentru cele 6 imagini folosite în testarea soluției:

 a) b) c)

d) e) f)

Fig. 9. Rezultatele obținute pentru cele 6 poze folosite în testarea soluției propuse

După cum se poate observa în figura 9, algoritmul de detecție al fețelor identifică și delimitează chipurile persoanelor, însă fiecare imagine ilustrează o situație ușor diferită:

* În imaginea a) se pot observa două persoane, una a cărui chip apare parțial și o față vizibilă integral. În aest caz, detectia reușeste să surprindă conturul ambelor fețe, chiar dacă primul subiect este surprins la marginea cadrului.
* În imaginea b) fața bărbatului este aproape frontală, cu iluminare ușor neuniformă și deși părul subiectului este grizonat și fundalul întunecat, detecția este făcută în mod corect, semn că algoritmul este destul de robust la diferențe de culoare și contrast.
* În imaginea c), deși calitatea acesteia nu este foarte bună și prezintă o formă de artefact de compresie în care a fost degradată calitatea, chipul este detectat semn că detectarea feței rămâne robustă chiar și în aceste situații.
* În ceea ce privește imaginea d), este detectată o singură persoană, în comparație cu imaginea a) unde au fost detectate ambele persoane. Motivul principal poate fi legat de faptul că cealaltă față este prea puțin vizibilă, iar scorul de detecție se află cel mai probabil sub pragul impus.
* În imaginea e) apare o ușoară ocluzie cauzată de șuvițele de păr care acoperă parțial fruntea, sprâncenele și chiar și privirea, însă chiar și așa, detecția rămâne bună, încadrând corect trăsăturile principale ale feței.
* În imaginea d), algoritmul își focalizează atenția doar pe regiunea feței, realizând o detecție corectă, deși imaginea prezintă un fundal aglomerat.

# Concluzii

Detectarea și recunoașterea fețelor în imagini de slabă calitate rămâne o provocare semnificativă, întrucât atât zgomotul, cât și rezoluția scăzută sau iluminarea neuniformă pot distorsiona semnificativ trăsăturile faciale. Totuși, soluțiile moderne, bazate pe preprocesare avansată și pe modele de tip CNN, au demonstrat că pot gestiona o parte din aceste dificultăți, obținând rezultate promițătoare.

Performanța aestora depinde în continuare de calitatea inițială a imaginilor, dar odată cu îmbunătățirea algoritmilor (prin tehnici de denoising, normalizare a iluminării, etc.) și cu extinderea bazelor de date care să includă o varietate mai mare de condiții și subiecți, se deschid perspective pentru creșterea semnificativă a acurateței în detecția și recunoașterea facială. Astfel, direcțiile de cercetare viitoare pot urmări integrarea mai strânsă între modele de deep learning, metode inteligente de preprocesare și resurse extinse de date, în scopul de a obține performanțe cât mai bune în scenariile din lumea reală.

# Bibliografie

1. S. Shan, W. Gao, D. Zhang, "Illumination Normalization for Robust Face Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003.
2. P. Viola, M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE CVPR, 2001.
3. K. Zhang, Z. Zhang, et al., "Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks," IEEE SPL, 2016.
4. F. Schroff, D. Kalenichenko, et al., "FaceNet: A Unified Embedding for Face Recognition and Clustering," CVPR, 2015.
5. V. Ioan, " Expertiza criminalistică pentru imagini şi secvenţe video : L5 - Biometrie din imagini şi secvenţe video"
6. Cod github: <https://github.com/ipazc/mtcnn>
7. <https://www.kaggle.com/datasets/jessicali9530/lfw-dataset?resource=download>
8. <https://www.kaggle.com/datasets/jessicali9530/celeba-dataset>