

UNIVERSITATEA POLITEHNICA DIN BUCUREȘTI
FACULTATEA DE ELECTRONICĂ, TELECOMUNICAȚII ȘI TEHNOLOGIA
INFORMAȚIEI

Program de studii: Ingineria Informației (INF)

PROIECT DE DIPLOMĂ

Titlul proiectului:

**ÎNLĂTURAREA DEFOCALIZĂRII DIN IMAGINI FOLOSIND REȚELE
NEURONALE PROFUNDE**

Coordonator științific:

Ș.L. dr. ing. Cristian Constantin DAMIAN

Absolvent:

RĂDUCU M. Alexandru, 444A

An universitar: 2024 - 2025

Fișa de înscriere a temei

Titlul temei:

Înlăturarea defocalizării din imagini folosind rețele neuronale profunde

Descrierea temei:

Proiectul constă în realizarea unei rețele neuronale cu arhitectura U-net pentru a înlătura efectul de defocalizare din imagini fotografice.

Contribuția originală:

Studentul va face antrenarea și testarea rețelei utilizând un simulator al defocalizării și un set de date cu imagini multimedia și va compara performanțele pentru diferite arhitecturi, funcții cost și tipuri de antrenare.

Data înregistrării temei: 2025-02-23 16:15:23

Conducător lucrare: Ș.L. dr. ing. Cristian Constantin DAMIAN

Student: RĂDUCU M. Alexandru

Cod validare: ed4cf0faf1

Capitolul 1 – Introducere

1.1 Context general

În era digitală, imaginile au devenit o componentă esențială în numeroase aplicații precum supravegherea video, medicina imagistică, fotografia computațională, recunoașterea facială, realitatea augmentată și multe altele. O imagine de calitate slabă poate compromite performanțele unor sisteme automate care se bazează pe prelucrarea și înțelegerea vizuală a scenei.

Unul dintre cele mai comune tipuri de degradare a imaginii este defocalizarea, cunoscută și sub numele de blur optic. Aceasta apare atunci când lentilele unui dispozitiv de captură (precum o cameră foto sau video) nu focalizează corect pe planul dorit. Rezultatul este o imagine lipsită de claritate, cu margini estompate și detalii pierdute. În multe cazuri, cauza defocalizării nu este una intenționată artistic, ci o eroare de captare ce afectează semnificativ utilitatea imaginii.

Tradițional, această problemă era abordată folosind tehnici precum deconvoluția clasică, filtre inverse sau metode bazate pe estimarea kernelului de blur. Totuși, aceste abordări sunt sensibile la zgomot, presupun cunoașterea exactă a modelului de degradare și au rezultate limitate în contextul imaginilor reale.

Odată cu apariția și maturizarea tehnologiilor de **deep learning**, în special a rețelelor neuronale convoluționale (CNN), s-au deschis noi perspective în prelucrarea imaginilor degradate. Modele antrenate pe seturi de date mari pot învăța automat relații complexe între imaginea blurată și versiunea sa restaurată, fără a necesita modelarea explicită a fenomenului fizic de defocalizare.

1.2 Motivația lucrării

Necesitatea de a dezvolta un sistem robust, capabil să elimine defocalizarea din imagini reale, vine din mai multe direcții:

- **Practică:** în aplicații industriale sau medicale, imaginea trebuie să fie clară pentru a permite interpretarea automată sau manuală. De exemplu, într-un sistem de inspecție vizuală automată, o imagine blurată poate conduce la detecții greșite.
- **Socială:** pe rețele sociale sau în aplicații mobile, utilizatorii sunt interesați de îmbunătățirea fotografiilor capturate în condiții necorespunzătoare.
- **Științifică:** domeniul restaurării imaginilor este un subiect de cercetare activ, iar explorarea de noi arhitecturi sau tehnici contribuie la avansul general al inteligenței artificiale.

În acest context, proiectul de față își propune dezvoltarea unui sistem complet, autonom, capabil să îmbunătățească imaginile defocalizate folosind rețele neuronale adânci.

1.3 Aplicații practice

Metodele automate de deblurare au aplicații directe în:

- **Supraveghere video** – îmbunătățirea clarității înregistrărilor de securitate;
- **Recunoaștere facială** – clarificarea trăsăturilor faciale pentru identificare;
- **Industrie auto** – imagini de la camere montate pe vehicule autonome;
- **Medicină** – restaurarea imaginilor ecografice sau RMN defocalizate;
- **Realitate augmentată (AR) și virtuală (VR)** – îmbunătățirea capturilor live pentru integrare în scenarii 3D;
- **Editare foto/video** – software de post-procesare care aplică deblurare automată.

Prin implementarea unui sistem ce poate procesa imagini în mod automat, cu intervenție umană minimă, se pot economisi resurse și îmbunătăți calitatea serviciilor în toate aceste domenii.

1.4 Scopul lucrării

Scopul principal al lucrării este **realizarea unui sistem software autonom** care să înlăture defocalizarea din imagini folosind **rețele neuronale profunde**, mai exact o arhitectură inspirată din U-Net, modificată pentru reconstrucția imaginilor degradate.

Pentru atingerea acestui scop, sunt definite următoarele obiective operaționale:

- Proiectarea și implementarea unei rețele neuronale convoluționale pentru restaurarea imaginilor;
- Crearea unui sistem complet de **generare automată de date** (imagini sharp și blur);
- Antrenarea modelului pe un set diversificat de imagini și optimizarea parametrilor rețelei;
- Dezvoltarea unei interfețe grafice simple care să permită încărcarea și procesarea imaginilor;
- Compararea rezultatelor obținute cu metode alternative existente în literatură;
- Analiza calitativă și cantitativă a performanței modelului propus (PSNR, SSIM);
- Redactarea unei documentații complete conform cerințelor facultății.

Prin realizarea acestei lucrări, se dorește nu doar demonstrarea capacităților rețelelor neuronale în restaurarea imaginilor, ci și oferirea unui instrument concret, ușor de utilizat, pentru procesarea defocalizării în diverse scenarii.

Capitolul 2 – Obiectivele lucrării

2.1 Defocalizarea: cauze și modele fizice

Defocalizarea este o formă comună de degradare a imaginilor care apare atunci când sistemul optic al camerei nu focalizează corect planul de interes. Această eroare duce la pierderea detaliilor fine și la estomparea marginilor obiectelor din imagine. Fenomenul poate fi descris matematic prin convoluția imaginii reale cu o funcție de răspuns a punctului (Point Spread Function – PSF), rezultând:

$$I_{\text{blur}} = I_{\text{sharp}} * h + n$$

unde:

- I_{blur} este imaginea defocalizată,
- I_{sharp} este imaginea ideală (focalizată),
- h este PSF-ul (modelul blurului),
- n este zgomotul aditiv.

Cauzele defocalizării pot include:

- Mișcarea camerei sau a obiectului în timpul expunerii;
- Focalizare greșită a lentilei;
- Limitări ale adâncimii câmpului;
- Calitatea slabă a obiectivului optic;
- Vibrări în timpul captării.

PSF-ul poate fi de tip:

- **Gaussian** – simulează o defocalizare uniformă;
- **Disc (circular)** – model frecvent pentru defocalizare produsă de diafragme;
- **Complex, necunoscut** – în cazul imaginilor reale.

Estimarea și inversarea PSF-ului este o sarcină dificilă, de unde și necesitatea metodelor de tip învățare automată.

2.2 Tehnici tradiționale vs. metode deep learning

Metode tradiționale:

1. **Deconvoluție clasică** – presupune cunoașterea exactă a PSF-ului, ceea ce rareori este cazul.
2. **Filtru Wiener** – atenuează zgomotul, dar produce rezultate mediocre în defocalizare severă.
3. **Metode variationale (TV, MAP)** – impun constrângeri asupra regularității imaginii (ex: linii drepte, margini clare).
4. **Estimarea kernelului + deconvoluție** – presupun modelarea PSF-ului, dar sunt lente și instabile.

Metode bazate pe învățare profundă:

1. **Rețele CNN antrenate end-to-end** – învață direct relația blur→sharp fără cunoașterea PSF-ului.
2. **Autoencodere și U-Net** – arhitecturi encoder-decoder pentru restaurare.
3. **GAN-uri (ex: DeblurGAN)** – învață să genereze imagini realiste ca ieșire, dar necesită multă regularizare.

Avantajele deep learning:

- Scalabilitate la seturi de date mari;
- Eliminarea nevoii de model explicit;
- Rezultate stabile, clare și rapid reproduse.

2.3 Rețele neuronale convoluționale (CNN)

Rețelele neuronale convoluționale sunt modele inspirate biologic care procesează date vizuale folosind filtre învățate automat. Structura lor generală conține:

- **Straturi convoluționale** – filtre de dimensiuni mici (ex: 3x3, 5x5) care extrag caracteristici locale;
- **Straturi de activare (ReLU)** – aplică nelinearități pentru a permite învățarea de funcții complexe;
- **Pooling (max/avg)** – reduce dimensiunile și ajută la generalizare;
- **Straturi complet conectate** – opționale în sarcinile de clasificare;
- **Dropout, BatchNorm** – tehnici pentru regularizare și stabilizare.

CNN-urile sunt ideale pentru restaurarea imaginilor datorită:

- Localizării spațiale a filtrelor;
- Posibilității de a învăța filtre specializate (edge detector, textură etc.);
- Abilității de a reconstrui imaginea strat cu strat.

2.4 Arhitectura U-Net și variante

U-Net este o arhitectură propusă inițial pentru segmentare biomedicală, dar adaptată extensiv pentru sarcini de restaurare a imaginilor.

Componentele U-Net:

- **Encoder** – reduce treptat dimensiunea imaginii, extrăgând caracteristici (feature maps).
- **Bottleneck** – strat de mijloc cu reprezentare latentă.
- **Decoder** – reconstruiește imaginea prin upsampling + concatenare (skip connections).
- **Skip connections** – păstrează detaliile locale de la encoder și le injectează în decoder.

Avantajele U-Net în deblurare:

- Îmbină contextul global cu detaliile locale;
- Permite propagarea directă a gradientului (skip);
- Se poate antrena relativ rapid, cu puține date.

În cadrul acestei lucrări, s-a folosit o variantă ușor modificată de U-Net cu blocuri reziduale și normalizare batch pentru stabilitate și acuratețe îmbunătățite.

2.5 Metrice de evaluare a calității imaginilor (PSNR, SSIM)

Evaluarea performanței unui model de deblurare necesită metrice cantitative obiective.

PSNR – Peak Signal-to-Noise Ratio:

- Compară imaginea restaurată cu cea de referință.
- Se exprimă în decibeli (dB).
- Valori tipice peste 30 dB indică o restaurare bună.

Formula:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right)$$

unde MAX este valoarea maximă a pixelului (255), iar MSE este eroarea pătratică medie.

SSIM – Structural Similarity Index:

- Măsoară similaritatea structurală între două imagini.
- Ia în considerare luminozitatea, contrastul și structura.
- Valori între -1 și 1, unde 1 înseamnă identitate perfectă.

Este o metrică mai apropiată de percepția vizuală umană decât PSNR.

Capitolul 3 – Metodologie

3.1 Arhitectura aleasă: U-Net modificat

Rețeaua este inspirată din U-Net clasic, dar adaptată cu blocuri reziduale și batch normalization. Codul sursă definește un encoder–decoder cu concatenare de feature-uri între nivele pentru a păstra detalii spațiale.

3.2 Implementare tehnică

Codul este scris în Python, folosind PyTorch. Se folosește un Dataset personalizat pentru încărcarea imaginilor blurate și clare, cu augmentare simplă (blur gaussian).

3.3 Generator dataset

Sistemul automat descarcă imagini random, aplică blur, și creează perechi blur/sharp pentru antrenare și testare.

3.4 Antrenare

Modelul este antrenat cu L1Loss, optimizat cu Adam, pe 500 de epoci, cu salvarea automată a modelului optim.

3.5 Vizualizare rezultate

Aplicația include o funcție de comparație între imaginile blurate și cele procesate, afișate una lângă alta.

Capitolul 4 – Progres actual

- Codul de bază este complet implementat (arhitectură + trainer + generator de date)
- S-au generat imagini blur/sharp și s-a testat pipeline-ul
- Modelul a fost antrenat pe date inițiale și produce rezultate vizual acceptabile
- Urmează ajustări pe dataset-uri reale și compararea cu metode din literatură

Capitolul 5 – Bibliografie

1. Zhang, Kaihao et al. Deep Image Deblurring: A Survey, IJCV 2022
2. <https://github.com/binorchen/AIFNET>
3. <https://pytorch.org/docs/stable/index.html>

4. Ronneberger, O. et al. U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015