Autoritatea contractanta: Unitatea Executivă pentru Finanțarea Învățământului Superior, a Cercetării, Dezvoltării și Inovării (UEFISCDI)

Codul proiectului: PN-III-P1-1.1-TE-2019-1111

Tip proiect: Proiecte de cercetare pentru stimularea tinerelor echipe independente (TE) **Titlul proiectului:** Baza de date si tehnici de eliminare a ceții pentru scene afectate de

ceața densă

Acronimul proiectului: Data - Dehaze

Data inceperii proiectului: 01.11.2020 Data incheierii proiectului: 31.10.2022 Valoarea totală a contractului: 431.900 lei Durata: 24 luni

Director proiect:

Conf. Dr. habil. Ing. Codruta Ancuti Universitatea Politehnica Timisoara, Facultatea de Electronica si Telecomunicatii, B312 Blvd. Vasile Parvan 2, 300223, Timisoara, Romania Tel: (+40)-0256-403363 Fax: (+40)-0256-403295

Institutia coordonatoare:

Universitatea Politehnica Timisoara, Facultatea de Electronică, Telecomunicații și Tehnologii Informaționale, Departamentul de Masurari si Electronica Optica (MEO)

Rezumatul etapei II

Etapa a II-a. In aceasta etapa am realizat urmatoarele actiuni:

 Extinderea bazei de date cu imagini realiste afectate de ceata cu distributie neuniforma si evaluarea modelului optic din etapa I.

- Utilizarea acestei baze de imagini in competitia de image dehazing in cadrul conferintei NTIRE 2021 in conjunctie cu CVPR 2021.
- Evaluarea principaliilor algoritmi de image dehazing construiti pe baza modelului optic.
- Evaluarea algoritmilor de *image dehazing* bazati pe CNN si studiul mai multor modele de retele recente.
- In aceasta etapa echipa de cercetare a publicat 2 articole la conferinte internationale si 2 articole in reviste (toate indexate ISI/WOS).

A) Extiderea bazei de date (setului de imagini cu ceata) inceputa in etapa I. Acest set include perechi de imagini cu ceata (ceață reală, densă şi neomogenă) şi fără ceață (imagini referinta);

Am continuat sa inregistram setul de imagini, in special in zilele noroase, predominant in timpul diminetii sau spre seara, cand intensitatea vantului nu depasea mai mult de 3km/h (pentru a limita ca ceata sa se imprastie foarte repede in scena sau sa apara modificari vizibile in continutul scenei, precum miscari de frunze, iarba, etc). Cel mai greu criteriu de indeplinit a fost absenta vantului, si acesta este motivul pentru care etapa de inregistrare a fost destul de dificila si a durat destul de mult.

Similar cu etapa I, hardware-ul utilizat a fost compus din: camera Sony Alpha 7III, pe care am actionat-o de la distanta cu ajutorul unei telecomenzi Sony RMT-DSLR2 precum si doua masini profesionala de generat ceata Antari Z1200 si LSM1500 PRO 1500 W.

Am inregistrat imagini cu rezolutie de 4000x6000 si adancime de culoare 24 biti. La inceputul fiecarei achizitii de scena s-a realizat ajustarea manuala a parametriilor camerei. Acesti parametrii au ramas nemodificati pentru fotografierea imaginii de referinta fara ceata si imaginii pereche cu ceata. Acesti parametrii includ, setarea de *whitebalance*, timpul de expunere (*shutter speed/ exposure time*), apertura (F-Stop) si factorul ISO. Acest lucru permite ca zonele din cele doua imagini pereche care nu sunt afectate de ceata sa arate identic. Pentru operatia de *white-balance* am utilizat card-ul SpyderCHECKR. Acesta tehnica este utilizata foarte des in fotografie si presupune plasarea card-ului referinta in fata camerei.

Dupa ce toate conditiile de calibrare au fost indeplinite, in toate perechile de imagini am amplasat cardul de culoare Macbeth care permite o post-procesare a imaginilor inregistrate.

La mijlocul anului 2021 am participat si organizat workshop-ul NTIRE 2021 din cadrul prestigioasei conferinte CVPR 2021 – etapa de image dehazing. La aceasta competitie s-a utilizat baza de date construita in etapa 1 (O1) - cu imagini realiste afectate de ceata cu distributie neuniforma. Aceasta baza de imagini include perechi de imagini cu ceata (ceață reală, *densă și neomogenă) și fără ceață (referinta) obtinute in conditii similare de iluminare.*



Figura 1. Exemple de perechi de imagini (cu vizibilitate – coloana din stanga si afectate de ceata neomogena – coloana din dreapta) inregistrate in conditii similare de iluminare

Competitia NTIRE 2021 a înregistrat 327 de participanți și 23 de echipe au fost clasate în faza finală. Participantii au experimentat diverse arhitecturi de retele neuronale și au propus câteva soluții noi care au îmbunătățit semnificativ rezultatele existente.

Clasamentul final a fost făcut în raport cu Media Scorul de Opinie rezultat din studiul nostru asupra utilizatorilor și soluțiile au fost împărțite în trei categorii în ceea ce privește percepția acestora.

DWT dehaze - best PSNR value



Figura 2. Rezultate obtinute de cei mai performanti algoritmi submisi la competitia de image dehazing, organizata in cadrul NTIRE-CVPR2021

B) Cercetarea modului de estimare a parametrilor modelului optic şi identificarea limitărilor acestuia prin analizarea influenţei ceţei asupra distribuţiei informaţiei şi a atenuării optice şi distribuţia informaţiei.

O prima observație este legată de influența constantei de airlight care este direct influentata de tipul de atenuare al culorilor si/sau de iluminarea artificiala, care este responsabila in mare parte de distorsiunile introduse de foarte multe metode existente.

Aceasta observatie importanta am investigat-o din perspectiva estimarii transmisiei (pe baza DCP- Dark Channel Prior [13]) care presupune ca utilizand doar informatia de intensitate nu se poate estima corect transmisia. Acest aspect era vazut ca o cauza principala pentru care algoritmii de eliminare a cetii din imagini nu genereaza rezultate satisfacatoare.

Pentru completarea acestor afirmatii ne-am bazat pe modelul optic de imbunatatire a imaginiilor afectate de ceata (Koschmieder [1]), descris mai jos, precum si pe metoda de estimare a transmisie pe baza informatiei priori numite DCP (Dark Channel Prior) [13].

Model optic care descrie formarea imaginilor cu ceata acceptat in general este modelul introdus de Koschmieder [1]. Pentru a intelege mai bine acest model, ecuatia si parametrii principali sunt reprezentati cu imagini in figura urmatoare.



Figura 3. Reprezentarea modelului optic (Koschmieder [1]) cu ajutorul imaginilor.

Dupa cum se poate observa I_h este imaginea afectata de ceata (sau imaginea de intrare), I este imaginea clara (ideala, fara ceata) care se doreste sa fie estimata, factorul t este transmisia si A_{∞} este culoarea atmosferica care se urmareste sa se estimeze la o distanta infinita (constanta de airlight). De asemenea un aspect important este ca transmisia depinde de distanta in mod exponential pe baza coeficientului β care este coeficientul de absorbtie al mediului, iar acest aspect este exprimat prin ecuatia:

$$t = \exp(-\beta d)$$

Putem spune ca in estenta, datorită impurităților aerosolului sau a fenomenelor naturale precum ceata, lumina reflectată de suprafețele obiectului este disipata de-a lungul directiei de propagare. Mai mult, coeficientul de absorbție prezintă un alt factor important care atenuează lumina. Pentru a rezolva ecuatia modelului optic este necesara estimarea a doi parametrii: transmisia *t* si respectiv a constantei de airlight A_{∞} .

In teorie odata ce acesti parametrii sunt estimati, prin simpla inversare a modelului optic vom putea sa recuperam valoare lui *I*, imaginea clara. Pentru estimarea transmisiei multi algoritmi pornesc de la informatia oferita de DCP care este o generalizare a dark object. Pe scurt, DCP porneste de la observatia ca in imaginile care nu contin cerul, exista in fiecare regiune cel putin un pixel pe un canal de culoare care este intunecat. Pentru estimarea transmisie, este necesar sa estimam acele valori minime, si sa filtram aceste minime locale in functie de informatia din imagine pentru reducerea artifactelor de intensitate sau culoare.

Precum se poate observa in exemplele generate mai jos, imaginile imbunatatite cu algoritmul si transmisia pe baza DCP amplifica erorile de culoare. Comparativ cu tehnica de DCP, o simpla operatie de imbunatatire a culorii inainte de prelucrare a imaginii genereaza rezultate satisfacatoare chiar daca se estimeaza transmisia doar pe valoarea de intensitate.

Dupa cum se poate observa mai sus imaginile prezentate sunt obtinute fara aplicarea unor filtre avansate si au ca scop demonstrativ impactul modificarii distributiei informatiei pe canalul de culoare asupra rezultatului final. Acest rezultat este important deoarece prelucrarea separata a intensitatii si respectiv a culorii permite o mai buna manipulare a informatiei vizuale si permite reducerea artifactelor care apar ineret dupa procesul de image dehazing precum cresterea zgomotului sau schimbarea culorii (color shifting) in mod artificial.



Figura 4. De la stanga la dreapta: imaginea originala cu ceata, imaginea imbunatatita pe baza canalului dark channel fara modificare de culoare, imaginea imbunatatita cu modificare de culoare pe baza luminantei.

Pe baza rezultatelor si a observatilor obtinute legate de modelul optic, suntem in etapa de scriere a unei publicatii pe acest subiect.

O2. Deep-learning image dehazing (dezvoltarea și antrenarea rețelelor neuronale pentru imbunatatirea vizibilitatii in imaginile cu ceata)

a) Am abordat acest objectiv urmarind studiul si extinderea unor algoritmi eficienti de image dehazing construiti cu ajutorul CNN (*convolutional neural networks*).

In aceasta etapa am analizat algoritmi care estimau parametrii modelului optic precum transmisia si airlight-ul dar si algorimi care generau rezultate imbunatatite direct din imaginea de intrare. Am observat in plus ca modelele CNN implemetate de tipul encoderdecoder genereaza rezultate rezonabil vizuale pentru imagini cu ceata omogena. Pentru imaginile neomogene exista inca loc de imbunatatire. Pentru training-ul retelelor am uitilizat baza de imagini nou achizitionata cu imagini cu ceata neomogena NH-HAZE prezentata in prima parte.

In urmatoarea sectiune sunt prezentate pe scurt 5 dintre modele mai avansate care au fost analizate pe baza performantelor obtinute utilizand baza de imagini introdusa NTIRE 2021.

Adaptive Dehazing Network (ADN) [2,3] este un model care este compus din doua ramuri WaveletNet si AttentionNet, care urmareste sa se adapteze la regiunile afectate de ceata subtire sau ceata densa. Prima ramura gestioneaza regiunile cu ceata mai putin densa, iar ramura a doua este dedicata pentru imbunatatirea regiunilor grav contaminate de ceata. Rezultatele generate de cele doua ramuri sunt proiectate in generatorul ghidat de hartile de impact care genereaza doua harti asociate fiecarui canal de intrare ce determina in ultima etapa contributia fiecarei ramuri la rezultatul final.



Figura 5. Structura modelului **Adaptive Dehazing Network (ADN)** compus din doua ramuri WaveletNet (pentru ceata subtire) si AttentionNet (pentru ceata densa)

Modelul DWT dehaze [3,4,5,6] este o solutie construita pe doua ramuri de retele adversiale, numite DW-GAN. Pe prima ramură au propus ideea de încorporarea cunoștințelor din domeniul frecvenței în reteaua de dehazing. Pentru aceasta s-a derivat arhitectura U-NET [7] astfel incat s-a construit aceasta prima ramura ca o retea wavelet.

Pentru a îndeplini cerințele pentru extragerea informatiei din domeniului de frecvență sunt utilizate cinci module de subeșantionare DWT și șase straturi de downsampling convoluționale pentru a construi un codificator. Astfel, reprezentările spațiale și frecvențele sunt concatenate ca intrare a procesului de subeșantionare. Pe ramura a doua s-a utilizat reteaua Res2Net [7] ca si encoder. Pentru intializare au utilizat weightsurile preantrenate de la ImageNet. În modulul decodor, s-a utilizat stratul de pixel-shuffle pentru supra-eșantionare, ceea ce face ca dimensiunea caracteristicii recuperate sa creasca treptat la rezoluția originala. De asemenea s-au adaugat conexiuni de ignorare între codificator și decodor. In ultima parte de executie s-a adaugat un strat de convolutie 7x7 pentru operatia de fuziune a caracteristicilor dintre cele doua ramuri care genereaza imaginea finala.



Figura 6. Structura solutiei DWT-Dehaze este formata din doua ramuri de retele Wavelet Net si Attention Net. Aceeași culoare folosită în dreptunghiuri denotă aceeași operație. 'Conv', 'BN', 'TConv', 'MP', 'PS', 'AP', 'LReLu' sunt abrevieri (din engleza) pentru convolution, batch normalization, transpose-convolution, max-pooling, pixel-shuffle, averagepooling, si leakyReLu. 'B2N', 'C-PA' si 'DWT' sunt abrevieri (din engleza) pentru bottle2neck, channel and pixel-wise attention și discrete wavelet transform.

Mac-dehaze [3] este de asemenea o retea formata din doua ramuri astfel: prima ramura este construita pe ImageNet [8] si Res2Net [9,10] pretreinuita. In aceasta etapa s-au eliminat legaturile complete (care asigura detaliile fine din imagine), cu scopul de a forta reteaua sa extraga o reprezentare globala. In plus aceasta abordare are rolul de a minimiza problemele legate de insuficienta informatiei de training.

Ramura a doua este inspirata de *residual channel attention network* (RCAN) [11] si este constituita din cinci grupuri reziduale, unde fiecare grup este constituit din zece blocuri reziduale. Spre deosebire de implementare RCAN initiala, care realizeaza donsamplingul imaginii, aceasta abordare mentine rezolutia originala si urmareste mentinerea detaliilor fine. Aceasta ramura are rezultate foarte bune pentru imaginile din domeniul de antrenament. Imaginea de iesire este generata de un strat de fuziune care concateneaza caracteristicile din fiecare ramura intr-o imagine de iesire.



Figura 7. Arhitectua modelului Mac dehaze

Modelul Deep Multi-Patch Hierarchical Network (DMPHN) [11] este o arhitectura multistrat care a fost initial utilizata pentru *image deblurring*. Aceasta arhitectura include o pereche encoder-decoder pentru fiecare nivel (prezentate in figura 5). Solutia analizata DMPHN (1-2-4) utilizeaza 1,2 respectiv 4 zone (patches) pornind top-to-bottom. Astfel nivelul superior (top) considera o zona (patch) per imagine, nivelul 2 extinde numarul de zone la 2, iar nivelul inferior (bottom) imparte imaginea in 4 zone (patches).Arhitectura este de tip encoder-decoder, descrisa in Figura 8. leşirea decodor-ului de la nivelul inferior este adăugată cu patch-uri la nivelul următor și alimenteaza la encoder-ul. De asemenea hărțile caracteristice rezultate sunt concatenate spatial.



Figura 8. Arhitectura de functionare a modelului Deep Multi-Patch Hierarchical Network. Simbolurile '{' marcheaza concatenarea spatiala, respectiv ⊕ marcheaza adaosul rezidual.



Figura 9. Arhitectura modulelor encoder-decoder utilizate de modelul DMPHN



Figura 10. Modelul Deep Multi-Patch Hierarchical Network (DMPHN) se poate extine la un numar mai mare de nivele (4 in aceasta figura).

Modelul MPRNet propus in [3,12] recupereaza progresiv imaginile in doua etape. In prima etapa se aplica trei sub-retele de tip encoder-decoder pentru fiecare canal rosu, verde si albastru ale imaginii cu ceata. Consecutiv pe fiecare canal prelucrat s-a aplicat un modul SAM (Supervised Attention Module)[12] iar mai apoi rezultatele sunt concatenate si reprocesate ca o intrare in ultima etapa cu scopul de minimizare a erorilor vizuale.



Figura 11. Reprezentarea functionarii separate pe fiecare canal de culoare a modelulului MPRNet

In mare parte, algoritmii studiati au abordat solutia de tip end-to-end care evita estimarea paramteriilor modelului optic si estimeaza direct imaginea dehazed direct din imaginea cu ceata. Din acest motiv, pentru ca se evita intelegerea modelului optic consideram ca acesti algoritmi pot sa fie imbunatatiti daca se include o etapa sau ramura de procesare care sa incoroporeze informatia despre modelul optic.

Pe baza studiului si a rezultatelor obtinute suntem in etapa de valorificare a experientei si de extindere a algoritmiilor. In urmatoarea etapa a acestui proiect ne propunem sa finalizam acest studiu prin publicarea unei metode originale care sa imbunatateasca rezultatele metodelor existente.

Reviste si articole publicate in Etapa II (2 reviste si 2 articole cotate ISI/WOS) :

1. C. O. Ancuti, C. Ancuti, F.-A. Vasluianu, R. Timofte, M. Fu, H. Liu et al., *"NonHomogeneous Dehazing Challenge Report*", NTIRE 2021, IEEE CVPR 2021, US

2. A. Kis, H. Balta, C. O. Ancuti, *"The Impact of Haze Non-Homogeneity on the Recent Image Dehazing Methods"*, IEEE ELMAR, Zadar, Croatia, septembrie, 2021

3. M. Sbert, C. Ancuti, C. O. Ancuti, J Poch, S Chen, M Vila, *"Histogram Ordering*", IEEE Access, 2021

4. Q Hao, Q Zhao, M Sbert, Q Feng, C. Ancuti, M Feixas, M Vila, J Zhang, *"Information-Theoretic Channel for Multi-exposure Image Fusion*", The Computer Journal 2021

Gradul de realizare al obiectivelor – 100%

Referinte:

[1] H. Koschmieder. Theorie der horizontalen sichtweite. In Beitrage zur Physik der freien Atmosphare, 1924.

[2] Codruta O. Ancuti, Cosmin Ancuti, Florin-Alexandru Vasluianu, Radu Timofte, Minghan Fu, Huan Liu et al., NonHomogeneous Dehazing Challenge Report, NTIRE 2021

[3] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Multiscale single image dehazing using perceptual pyramid deep network. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pages 902–911, 2018.

[4] Pengju Liu, Hongzhi Zhang, Kai Zhang, Liang Lin, and Wangmeng Zuo. Multi-level wavelet-cnn for image restoration. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pages 773–782, 2018.

[5] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion attention network for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, pages 11908–11915, 2020.

[6] Christian Ledig, Lucas Theis, Ferenc Husz´ar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photorealistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4681–4690, 2017.

[7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Unet: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015

[8] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei- Fei. Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 248–255, 2009.

[9] S. H. Gao, M. M. Cheng, K. Zhao, X. Y. Zhang, M. H. Yang, and P. Torr. Res2net: A new multi-scale backbone architecture. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(2):652–662, 2021.

[10] C.O. Ancuti and C. Ancuti. Single image dehazing by multi-scale fusion. IEEE Transactions on Image Processing, 22(8):3271–3282, 2013

[11] Hongguang Zhang, Yuchao Dai, Hongdong Li, Piotr Koniusz, Deep Stacked Hierarchical Multi-patch Network for Image Deblurring, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

[12] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In CVPR, 2021

[13] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," IEEE TPAMI, 2011