



**UNIVERSITATEA NAȚIONALĂ
DE ȘTIINȚĂ ȘI TEHNOLOGIE
POLITEHNICA BUCUREȘTI**



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

Decizie nr. 47 din 20-05-2024

**REZUMAT TEZĂ
DE DOCTORAT**

Ing. Nicolae-Cătălin RISTEA

**ÎNVĂȚARE AUTOMATĂ BAZATĂ PE MODELE FIZICE ÎN
TELEDETECȚIE**

PHYSICS-AWARE MACHINE LEARNING IN REMOTE SENSING

COMISIA DE DOCTORAT

Prof. Dr. Ing. Mihai CIUC NUST Politehnica București	Președinte
Prof. Dr. Ing. Andrei ANGHEL NUST Politehnica București	Coordonator PhD
Prof. Dr. Ing. Cosmin ANCUTI University Politehnica Timișoara	Referent
Dr. Bertrand CHAPRON Ifremer	Referent
Prof. Dr. Ing. Mihai DATCU NUST Politehnica București	Referent

BUCUREȘTI 2024

Cuprins

1	Introducere	1
1.1	Prezentarea domeniului tezei de doctorat	1
1.2	Domeniul tezei de doctorat	2
1.3	Conținutul tezei de doctorat	3
2	Concepte și Fundamente Teoretice	4
2.1	Concepte Bazate pe Modele Fizice	4
2.1.1	Radar cu Undă Continuă Modulată în Frecvență	4
2.1.2	Radar cu Apertură Sintetică	4
2.1.3	Descompunere în Subaperturi	5
2.1.4	Estimarea Centrozilor Doppler	5
2.2	Tehnici de Învățare Profundă	6
2.2.1	Rețele Neuronale Convoluționale	6
2.2.2	Arhitecturi Transformer	6
3	Reducerea Interferențelor Radar Auto	8
3.1	Introducere	8
3.2	Metode Propuse	9
3.2.1	Modelul de Semnal Radar	9
3.2.2	Preprocesarea Datelor	9
3.2.3	Reteaua Neuronală Propusă	10
3.2.4	Constrangerea Ponderilor	11
3.3	Seturi de Date	11
3.4	Rezultate	12
3.4.1	Metrici de Evaluare	12
3.4.2	Alegerea Hiper-Parametrilor	12
3.4.3	Rezultate pe ARIM-v2	13
3.4.4	Generalizare pe Date Reale	13
3.5	Discutii	13
4	Analiza Trasaturilor Oceanice	14
4.1	Introducere	14

4.2	Metode Propuse	14
4.2.1	Descompunerea în Subaperturi	14
4.2.2	Estimarea Centroizilor Doppler	15
4.2.3	Rețele Neuronale Nesupervizate	15
4.2.4	Căutarea Bazată pe Conținut a Imaginilor	16
4.3	Rezultate	17
4.3.1	Setul de Date	17
4.3.2	Alegerea Hiper-Parametrilor	17
4.3.3	Metrici de Evaluare	17
4.3.4	Rezultate Clasificare	17
4.3.5	Rezultate pentru Antrenare Nesupervizată	17
4.3.6	Rezultate Tehnici Cautare	19
4.4	Discutii	19
5	Segmentarea Gheții Oceanice din Imagini SAR	20
5.1	Introducere	20
5.2	Metoda Propusa	20
5.2.1	Preprocesarea Datelor	20
5.2.2	Arhitectura Reziduala Bazata pe CyTran	21
5.2.3	Transformer Bazat pe Atenție Transpusă Paralelă	21
5.2.4	Arhitectura UT-MHTA	21
5.3	Rezultate	23
5.3.1	Setul de Date	23
5.3.2	Alegerea Hiper-Parametrilor	23
5.3.3	Metrici de Evaluare	23
5.3.4	Rezultate Experimentale	24
5.3.5	Generalizare Către Date IW	24
5.4	Discutii	25
6	Concluzii	27
6.1	Reducerea Interferentelor Radar Auto	27
6.2	Analiza Trăsăturilor Oceanice	27
6.3	Segmentarea Gheții Oceanice din Imagini SAR	28
6.4	Lista de Publicații	28
6.5	Lista de Publicații în Afara Scopului Principal al Tezei	29
6.6	Direcții Viitoare	30
	Bibliografie	31

Capitolul 1

Introducere

Domeniul procesării semnalelor are o istorie vastă care se întinde pe parcursul mai multor decenii. Procesarea semnalelor implică manipularea, analiza și interpretarea semnalelor, acestea fiind reprezentări ale datelor în diverse forme. Istoria procesării semnalelor a cunoscut evoluții remarcabile, atât din punct de vedere teoretic, cât și în ceea ce privește aplicabilitatea practică. Metodele de procesare a semnalelor au avut un rol esențial în a extrage informații valoroase din date neprelucrate, în a îmbunătăți calitatea semnalelor și în a optimiza sistemele de comunicații. Domeniul a evoluat și mai mult odată cu apariția seturilor de date de mari dimensiuni în diferite sectoare, prin adoptarea algoritmilor de învățare automată și învățare profundă. Aceste abordări inovative au marcat o transformare majoră în studiul și aplicarea procesării semnalelor.

1.1 Prezentarea domeniului tezei de doctorat

Interferența Radarului Auto. Abordarea condusului autonom și siguranța rutieră reprezintă un subiect crucial pentru reducerea frecvenței accidentelor în trafic și a deceselor. Companiile auto se ocupă activ de acest subiect prin avansarea tehnologiilor vehiculelor autonome. O abordare predominantă implică utilizarea senzorilor radar pentru scanarea meticuloasă a mediului înconjurător. În industria auto, cea mai întâlnită alegere pentru senzorii radar este implementarea radarilor cu undă continuă modulată în frecvență (FMCW) sau a radarilor cu secvențe de chirp (CS), care transmit secvențe de semnale chirp liniare. Semnalele transmise și recepționate de astfel de senzori furnizează mijloacele necesare pentru estimarea distanței și a vitezei țintelor din apropiere (de exemplu, vehicule, pietoni sau alte obstacole). Cu toate acestea, creșterea utilizării senzorilor radar [17] crește probabilitatea de interferență între senzori din diferite vehicule, generând semnale corupte și inutilizabile. Interferența radiofrecvență ar putea crește semnificativ nivelul de zgomot, până în punctul în care deosebirea potențialelor ținte devine foarte dificilă sau chiar imposibilă. În astfel de cazuri, interferența poate ascunde complet țintele, diminuând sensibilitatea metodelor de detectare a acestora [3]. Pentru a face posibilă identificarea lor, este imperativ să se reducă interferența radarului.

Observarea Pământului. Observarea Pământului implică integrarea sistemelor fizice, chimice și biologice ale planetei noastre. Aceasta este realizată prin intermediul tehnologiilor de teledetecție (RS) utilizate în metodele de monitorizare a Pământului. Procesul implică colectarea, analiza și prezentarea datelor pentru a îmbunătăți înțelegerea diversității sistemului terestru [13]. Având în vedere că oceanul reprezintă aproximativ 71% din suprafața Pământului, observațiile oceanice au atras tot mai mult atenția comunității științifice în ultimele decenii. Înainte de 1978, cunoștințele noastre despre ocean erau limitate, până când a fost lansat Seasat, primul satelit conceput pentru a orbita Pământul și pentru a efectua teledetecția asupra oceanului planetar [43]. Deși Seasat a funcționat doar timp de aproximativ 100 de zile, misiunea a obținut mai multe date despre ocean decât toți senzorii utilizați anterior combinați. Acest eveniment a stimulat dezvoltarea rapidă a sateliților oceanici, conducând la lansarea unui număr tot mai mare de sateliți care transportă diferiți senzori (de exemplu, cu microunde, vizibili, infraroșu) pentru a îmbunătăți înțelegerea noastră despre ocean.

Pe lângă observarea oceanului, un alt aspect important care a atras o atenție semnificativă este scăderea zonelor marine cu gheață, în special în zona Arctică. Acest efect reprezintă o consecință a schimbărilor climatice globale. Prin urmare, gradul de acoperire și concentrația de gheață pe mări și oceane sunt factori esențiali atât în studiile privind schimbările climatice, cât și în navigația în regiunile polare. Abordarea cerințelor logistice ale industriei de transport în aceste zone necesită date detaliate, cu rezoluție mare, privind condițiile marine locale din zonele Arctice. Astfel de informații sunt indispensabile pentru planificarea operațională, determinarea rutelor și promovarea dezvoltării durabile a regiunilor polare [6]. Încă o dată, având în vedere volumul mare de date acumulate zilnic, necesitatea sistemelor automate are o importanță crucială.

1.2 Domeniul tezei de doctorat

Având în vedere provocările prezentate în monitorizarea Pământului și în industria radarului auto, evoluția rapidă a senzorilor și volumul substanțial de date disponibile, algoritmi de procesare abia îndeplinesc cerințele actuale. În acest sens, obiectivul nostru este de a dezvolta rețele de învățare profundă robuste și precise, bazate pe modele fizice.

Principalele direcții sunt următoarele:

1. Identificarea, dezvoltarea și analiza sistemelor optime bazate pe rețele neuronale profunde pentru reducerea interferenței radarului auto în scenarii complexe (de exemplu, multiple surse de interferență).
2. Testarea capacității de generalizare a algoritmilor propuși pentru reducerea interferenței radarului pe mai mulți senzori.
3. Identificarea, dezvoltarea și analiza sistemelor optime bazate pe rețele neuronale profunde pentru analiza și clasificarea fenomenelor oceanice pe imagini SAR (Synthetic Aperture Radar).

4. Identificarea, dezvoltarea și analiza sistemelor optime bazate pe rețele neuronale profunde pentru segmentarea zonelor maritime de gheață în zonele arctice.
5. Furnizarea de motivații bazate pe modele fizice și metode de interpretare pentru arhitecturile dezvoltate.

1.3 Conținutul tezei de doctorat

În conformitate cu principalele direcții de cercetare urmate în timpul doctoratului, această teză este structurată în următoarele capitole:

Capitolul 2 (Fundamente teoretice și concepte) introduce principalele concepte care au fost explorate și folosite în cadrul lucrării curente, cum ar fi conceptele bazate pe fizică prezintă soluțiile FMCW, sistemele SAR și rețelele neurale.

Capitolul 3 (State of the Art) oferă o privire de ansamblu asupra direcțiilor noastre de cercetare. Atenția noastră s-a concentrat pe trei domenii principale: reducerea interferențelor radarului auto, algoritmi folosiți în analiza imaginilor SAR ale oceanului și metode pentru segmentarea zonelor marine de gheață în regiunile Arctice.

Capitolul 4 (Reducerea Interferenței Radarului Auto) introduce succint domeniul radarului auto și necesitatea algoritmilor de reducere a interferenței radarului. Se continuă cu explorarea în detaliu a descrierii, conceptului și rezultatelor arhitecturii noastre propuse FCN. Capitolul se încheie prin detalierea și discutarea avantajelor și dezavantajelor oferite de framework-ul nostru.

Capitolul 5 (Analiza Trăsăturilor Specifice Oceanelor) introduce domeniul imaginilor SAR oceanice, concentrându-se în special pe clasificarea vignete-urilor și obținerea imaginilor din zonele oceanice. Discuția se continuă cu descrierea și conceptele algoritmilor dezvoltați, subliniind motivațiile bazate pe modele fizice. Capitolul se încheie printr-o analiză detaliată a elementelor cheie introduse de framework-ul nostru, accentuând potențialul său de implementare în scenarii din lumea reală.

Capitolul 6 (Segmentarea Zonelor Marine De Gheață pentru Imagini SAR) introduce domeniul segmentării zonelor marine de gheață în regiunile arctice, folosind date SAR. Ulterior, descrie dezvoltarea arhitecturilor de rețele neuronale profunde cu scopul de a obține un echilibru optim între viteză și performanță. În plus, se descrie adaptarea datelor SAR IW-EW, împreună cu importanța practică și algoritmi utilizați. Capitolul se încheie printr-o analiză comprehensivă și discuție a elementelor inovatoare introduse în lucrarea noastră.

Capitolul 7 (Concluzii) cuprinde mai multe secțiuni care detaliază rezultatele obținute și contribuțiile aduse de teză. În plus, include o listă de articole publicate, urmată de perspective potențiale pentru cercetări viitoare.

Capitolul 2

Concepte și Fundamente Teoretice

2.1 Concepte Bazate pe Modele Fizice

2.1.1 Radar cu Undă Continuă Modulată în Frecvență

În soluțiile radarului FMCW, semnalul transmis $s_{TX}(t)$ este o secvență de semnale chirp, a cărui frecvență urmează de obicei un model de tip dinte de fierăstrău. În prezența interferențelor reciproce, antena de recepție colectează un amestec din două semnale, semnalul reflectat și semnalul de interferență. Prin urmare, semnalul primit este definit formal mai jos.

$$s_{RX}(t) = \sum_{i=0}^{N_t-1} \underline{A}_i \cdot s_{TX}(t - \tau_i) + \sum_{l=0}^{N_{int}-1} s_{RFI,l}(t), \quad (2.1)$$

unde $\underline{A}_i = A_i \cdot e^{j\phi_i}$ este amplitudinea complexă, τ_i este întârzierea de propagare a țintei i , N_t este numărul țintelor, iar N_{int} este numărul de interferențe. Antena de recepție colectează semnalul reflectat $s_{RX}(t)$, care este ulterior amestecat cu semnalul transmis și filtrat trece jos, rezultând semnalul de bătai $s_b(t)$. La combinarea semnalului reflectat de o țintă punctiformă cu semnalul transmis, se obține un semnal cu frecvență constantă. Astfel, amestecarea unei interferențe necorelate cu semnalul chirp transmis rezultă într-un semnal chirp în banda de bază.

2.1.2 Radar cu Apertură Sintetică

Radarul cu apertură sintetică (SAR) este o tehnologie radar modernă care depășește limitele tradiționale și deține capacități de imagistică fără precedent. Spre deosebire de radarele convenționale care se bazează pe antene de mari dimensiuni pentru a obține o rezoluție înaltă, SAR folosește aperturi virtuale obținute din exploatarea mișcării platformei radar. Prin utilizarea acestei mișcări și procesarea semnalelor radar reflectate, SAR poate obține rezoluții în azimut de ordinul metrilor sau chiar a centimetriilor, depășind cu mult capacitățile radarelor tradiționale.

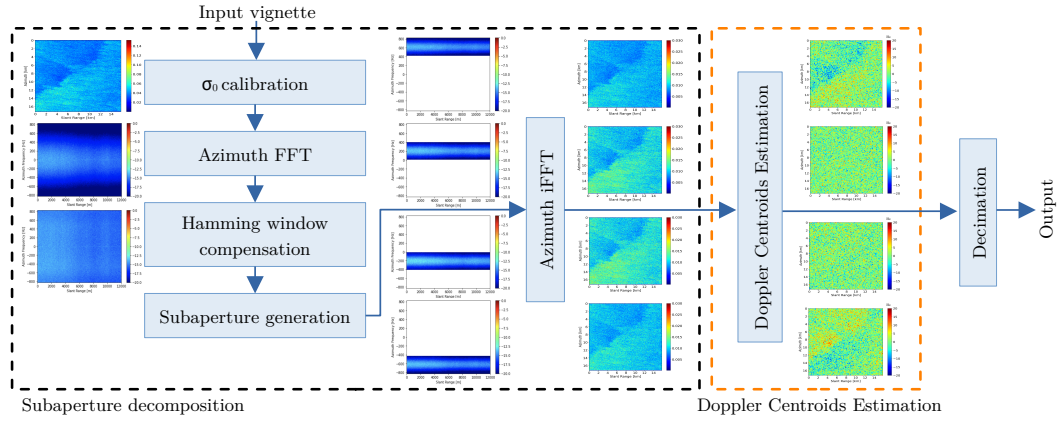


Figura 2.1 Schema de preprocesare a descompunerii pe subaperturi. Blocul DCE este un algoritm suplimentar, care poate fi omis în conformitate cu experimentul efectuat.

La bază, SAR funcționează pe principiul imaginii radar coerente, în care o serie de impulsuri radar sunt transmise către zona țintă în timp ce platforma radar se deplasează pe o traiectorie predefinită. Pe măsură ce undele radar interacționează cu zona terestră, acestea sunt dispersate și reflectate înapoi către senzorul radar. Măsurând faza și amplitudinea acestor semnale reflectate pe întreaga apertură, SAR creează o imagine cu rezoluție înaltă a scenei țintă. Prin exploatarea informațiilor de fază conținute în ecurile radar, SAR poate atenua eficient contribuțiile nedorite, rezultând imagini clare și precise chiar și în condiții de mediu adverse.

2.1.3 Descompunere în Subaperturi

În sistemele SAR clasice, semnalul reflectat de obiectele iluminate este captat în diverse poziții și unghiuri azimutale de-a lungul traiectoriei radarului. Într-o scenă extrem de dinamică, cum ar fi suprafața oceanului, observarea sa din diferite unghiuri poate furniza informații suplimentare despre zona iluminată. Prin urmare, prin descompunerea imaginii SAR în subaperturi, scena ar putea fi văzută artificial din diferite unghiuri de observație, capturând informații suplimentare. Fiecare subapertură corespunde unei imagini formate folosind doar o parte din unghiul total al azimutului, așa cum este ilustrat în Figura 2.1 după blocul de generare a subaperturii.

2.1.4 Estimarea Centrozilor Doppler

Fie $X_i \in \mathbb{R}^{m \times n}$ subapertura i pentru o imagine SAR, unde $m, n \in \mathbb{N}$ sunt dimensiunile azimut și distanță. Fie $Y_i \in \mathbb{R}^{m \times n}$ versiunea întârziată cu un eșantion în azimutul axei X_i . Estimăm centrozii Doppler pentru subapertura i astfel:

$$D_i = -PRF \cdot \frac{\text{angle}(Z_i)}{2\pi}, \quad (2.2)$$

unde $Z_i = \text{filt}(X_i \cdot Y_i^*)$, Y_i^* este conjugata complexă a lui Y_i , PRF este frecvența de repetare a impulsurilor, $\text{angle}()$ returnează unghiul valorii complexe de intrare, iar $\text{filt}()$ este un filtru mediu bidimensional cu dimensiunea kernelului $d_1 \times d_2$. În partea dreaptă a Figurii 2.1 am ilustrat rezultatul DCE pentru o imagine de intrare SAR.

2.2 Tehnici de Învățare Profundă

2.2.1 Rețele Neuronale Convoluționale

Rețelele neuronale convoluționale (CNN-uri) sunt o clasă de modele de învățare profundă special concepute pentru procesarea datelor structurate sub formă de matrice, cum ar fi imagini, audio și serii temporale. Acestea au obținut rezultate remarcabile în diverse sarcini, cum ar fi clasificarea imaginilor classification [16, 29, 42, 24], detectarea obiectelor [7, 30, 5] și segmentarea semantică [22, 28, 12]. Blocurile de bază ale CNN-urilor includ straturi convoluționale, straturi de extragere și straturi complet conectate. În continuare, vom descrie pe scurt primele două blocuri.

Stratul convoluțional. Un strat convoluțional aplică un set de filtre învățabile pe datele de intrare pentru a produce hărți de trăsături. Fără a pierde generalitatea, vom descrie formal scenariul bidimensional, generalizarea fiind trivială. Fie $X \in \mathbb{R}^2$ tensorul de intrare, $F \in \mathbb{R}^2$ tensorul filtrului și $Y \in \mathbb{R}^2$ harta de trăsături de ieșire. Operația de convoluție este definită ca:

$$Y_{i,j} = (X * F)_{i,j} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \cdot F_{m,n}, \quad (2.3)$$

unde $M \in \mathbb{N}$ și $N \in \mathbb{N}$ sunt dimensiunile filtrului F , iar $*$ reprezintă operația de convoluție.

Stratul de extragere. Straturile de pooling reduc dimensiunile spațiale ale hărților de trăsături de intrare, menținând totodată informația cea mai importantă, conform unei operații matematice date, cum ar fi maximul, minimul sau media. Stratul este aplicat separat pe fiecare canal, astfel încât doar dimensiunea spațială a datelor de intrare este afectată, în timp ce numărul de canale rămâne neschimbat.

2.2.2 Arhitecturi Transformer

Blocul Transformer. În partea stângă a Figurii 2.2 este ilustrată arhitectura generală a Transformer-ului. Fie $X = [x_1, x_2, \dots, x_n]$ o secvență de intrare, unde $x_i \in \mathbb{R}^m$ este vectorul de caracteristici pentru a i -a intrare. Secvența este proiectată în descriptori $T = [t_1, t_2, \dots, t_n]$, unde fiecare descriptor $t_i \in \mathbb{R}^d$ corespunde datelor de intrare x_i și $d \in \mathbb{N}$ este dimensiunea de înglobare a descriptorului. Ulterior, sunt adăugate trăsăturile de poziție pentru toți descriptorii, astfel încât fiecare descriptor t_i să poată fi identificat în mod unic prin trăsătura de poziție dată. Pentru procesarea paralelă, toate secvențele sunt

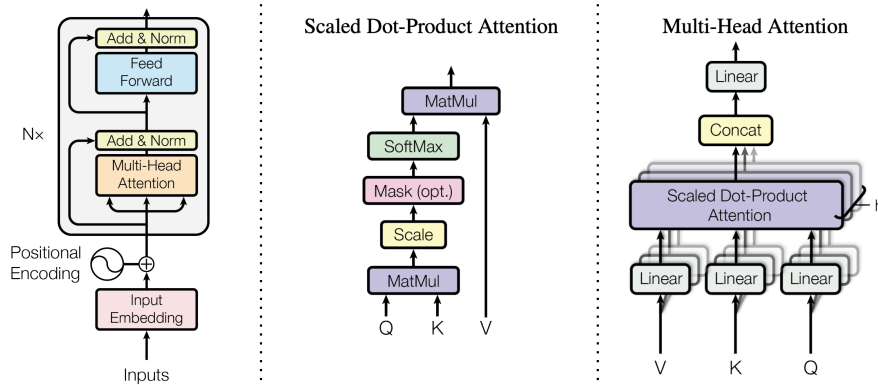


Figura 2.2 (stânga) Arhitectura Transformer [45]. (centru) Atenție bazată pe produs scalar. (dreapta) Atenție multiplă.

grupate în matrice, de exemplu, e.g., $T \in \mathbb{R}^{n \times d}$. În continuare, secvența de descriptori T este procesată de blocul Transformer după cum urmează:

$$Z = \text{Norm}(\text{MultiHeadAttention}(T) + T) \quad (2.4)$$

$$Y = \text{Norm}(\text{FeedForward}(Z) + Z), \quad (2.5)$$

unde $Z \in \mathbb{R}^{n \times d}$, $Y \in \mathbb{R}^{n \times d}$, funcția Norm reprezintă un strat de normalizare [1] iar modulul Feed Forward este de obicei un perceptron cu mai multe straturi. Blocul este aplicat de $N \in \mathbb{N}$ ori, producând caracteristici bazate pe atenție pentru secvența de intrare. **Atenția bazată pe produs scalar.** Elementele de bază ale atenției bazate pe produs scalar sunt evidențiate în centrul Figurii 2.2. Intrarea constă din interogări $Q \in \mathbb{R}^{n \times d_k}$, chei $K \in \mathbb{R}^{n \times d_k}$ și valori $V \in \mathbb{R}^{n \times d}$. Calculăm produsele scalare ale interogării cu toate cheile, împărțim rezultatul la $\sqrt{d_k}$ și aplicăm o funcție softmax pentru a obține ponderile valorilor (matricea de atenție). Blocul este descris formal ca:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.6)$$

Atenția multiplă. Atenția multiplă în arhitectura transformer permite modelului să se concentreze simultan pe diferite părți ale intrării, îmbunătățindu-și capacitatea de a captura relații complexe. Aceasta aplică multiple proiecții liniare asupra intrărilor Q , K , V , calculează atenții independente și concatenează ieșirile pentru rezultatul final. Acest lucru îmbunătățește extragerea caracteristicilor și îmbunătățește performanța în sarcinile care necesită dependențe din zone îndepărtate din imagine și înțelegerea contextuală. Blocul este descris vizual în partea dreaptă a Figurii 2.2.

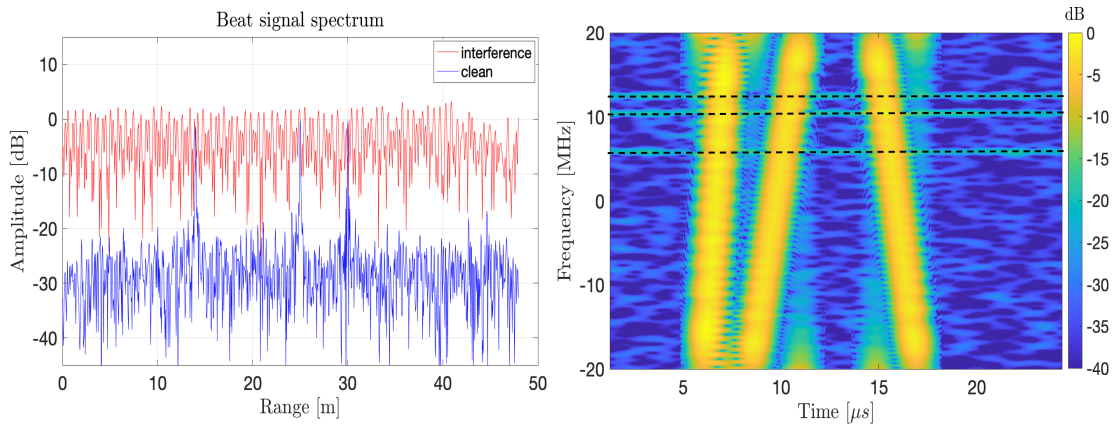


Figura 3.1 (left) The magnitude of the range profile of an FMCW radar sensor is depicted. (right) The magnitude of the STFT corresponding to the above range profile is presented. The targets are represented by thin horizontal lines, while the interference sources manifest as thicker, more prominent diagonal lines.

Capitolul 3

Reducerea Interferențelor Radar Auto

3.1 Introducere

Reducerea interferenței radarului auto este o preocupare majoră în industria auto. Senzorii radar cei mai frecvenți folosiți în industria auto sunt cei cu undă continuă modulată în frecvență (FMCW) / secvențe chirp (CS), care transmit secvențe de semnale chirp liniare. Cu toate acestea, adoptarea tot mai mare a senzorilor radar [17] crește probabilitatea de interferență între senzori proveniți de la diferite vehicule, generând semnale corupte și inutilizabile. În Figura 3.1, prezentăm un profil al unui semnal radar cu interferență (roșu) și fără interferență (albastru), evidențiind faptul că anumite ținte devin absorbite de nivelul de zgomot generat de multiple surse de interferență. Pentru a identifica cu succes aceste ținte, este imperativ să implementăm soluții eficiente pentru a atenua interferența radarului. Pentru a aborda această problemă, cercetătorii au propus diverse tehnici, de la abordări convenționale [18, 44, 14, 48, 2, 19, 27, 26] până la metode de învățare profundă [35, 37, 36, 40, 9, 25, 39].

Extindem cercetările anterioare [35] prin proiectarea unei rețele complet convoluționale (FCN) [22] capabilă să recupereze atât faza, cât și magnitudinea spectrului de bătaie ale semnalului radar și să gestioneze mai multe surse de interferență necoerente. Rețeaua noastră primește ca date de intrare părțile reale și imaginare, precum și magnitudinea Transformatei Fourier pe Ferestre (STFT) a semnalului de bataie cu interferențe. Rețeaua oferă ca rezultate partea reală, partea imaginară și magnitudinea profilului în distanță fără interferențe. În timp ce majoritatea abordărilor de învățare profundă au studiat atenuarea interferențelor radar cu o singură sursă de interferență [35, 9, 26], noi ne propunem să abordăm sarcina de reducere a interferențelor radar (RFI) sub multiple surse de interferență. Pentru a atinge acest obiectiv, creăm un set de date care imită îndeaproape scenariile reale din domeniul auto, implicând multiple surse de interferență. Setul nostru de antrenament include până la trei surse de interferență, în timp ce pentru setul de testare, considerăm scenarii cu până la șase surse de interferență. Mai mult, publicăm două seturi de date pentru eforturile viitoare în domeniul atenuării interferențelor radar.

3.2 Metode Propuse

3.2.1 Modelul de Semnal Radar

În sistemele radar FMCW, semnalul transmis $s_{TX}(t)$ este o secvență de chirpuri. În prezența interferențelor reciproce, antena receptoare colectează un amestec din două semnale, semnalul reflectat și semnalul de interferență. Conform 2.1.1, semnalul analitic de bătaie $s_b(t)$ în prezența interferențelor este exprimat astfel:

$$s_b(t) = \left\{ \sum_{i=0}^{N_t-1} \underline{A}_i \cdot \exp(j2\pi k \tau_i t) + \sum_{l=0}^{N_{int}-1} \underline{A}_{RFI,l} \cdot \exp [j\pi(k - k_{RFI,l})(t - t_{RFI,l})^2] \cdot p\left(\frac{t - t_{RFI,l}}{T_{AAF,l}}\right) \right\} \cdot p\left(\frac{t - \frac{T}{2}}{T}\right), \quad (3.1)$$

unde $\underline{A}_{RFI,l}$ este amplitudinea complexă a semnalului de interferență l și $p(t)$ este funcția fereastră. Ca rezultat, $s_b(t)$ cuprinde o combinație de exponențiale complexe care reprezintă țintele și semnalele interferente sub forma unor chirpuri în banda de bază.

3.2.2 Preprocesarea Datelor

Așa cum este ilustrat în Figura 3.1, calculul STFT-ului discret este esențial pentru a separa țintele de sursele de interferență. Ecuația de mai jos descrie transformarea unui semnal din domeniul timp într-o reprezentare asemănătoare unei imagini utilizând STFT discret, folosind o fereastră *hamming*. În plus, pentru a aduce datele de intrare aproximativ în intervalul $[-1, 1]$, scalăm STFT-ul prin împărțire la $\alpha = 40$, o valoare obținută statistic pe întregul set de antrenament. Pentru a obține profiluri în distanță fără interferență din STFT-ul semnalului de bătaie (prin antrenarea rețelei noastre), efectuăm o Transformare

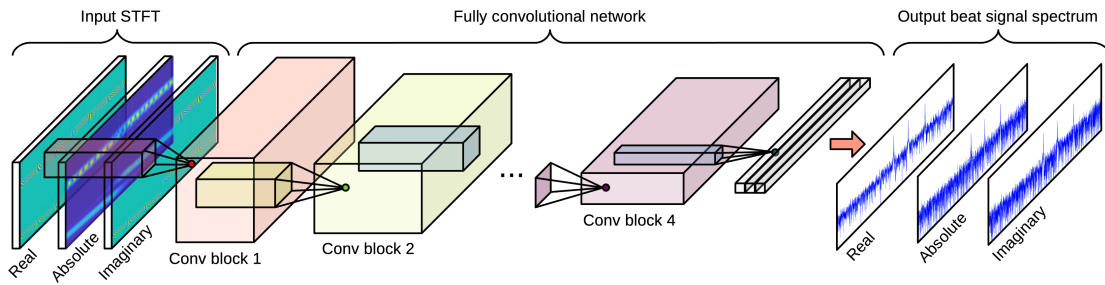


Figura 3.2 Structura modelului nostru FCN implică procesarea intrării STFT printr-o secvență de patru blocuri convoluționale (care cuprind straturi convoluționale și de extragere). Această prelucrare continuă până când dimensiunea verticală este redusă la 1, obținându-se spectrul semnalului de bătaie.

Fourier Rapidă (FFT) pe etichetele noastre din domeniul timp pentru a obține profilurile de distanță dorite.

3.2.3 Rețeaua Neuronală Propusă

Scopul nostru este să dezvoltăm un model de rețea neurală capabil să abordeze eficient RFI-ul prin translatarea precisă a unei intrări STFT zgomotoase în FFT dorit a semnalului de bătaie pentru orice semnal dat, luând în considerare atât magnitudinea, cât faza. Pentru a îndeplini acest obiectiv, am introdus o arhitectură FCN, reprezentată vizual în Figura 3.2. În timp ce lucrările anterioare, cum ar fi [27], au explorat soluții ce fac trecerea de la STFT la FFT, aceste abordări nu folosesc tehnici de învățare profundă.

Noutatea arhitecturii noastre neurale constă în structura intrării și ieșirii, fiecare cuprinzând o reprezentare compusă din trei canale distincte. Primul și al treilea canal al intrării corespund părților reale și imaginare ale STFT, în timp ce canalul din mijloc reprezintă magnitudinea STFT. Magnitudinea STFT captează cele mai semnificative informații vizuale și poate fi văzută ca o hartă de atenție [47, 46], care, în loc să fie calculată de rețea, este furnizată ca un canal de intrare. Ieșirea rețelei urmează un design similar în ceea ce privește canalele, cu distincția principală constând în dimensiunea sa spațială, așa cum este descris mai jos. Deși rețeaua noastră nu calculează explicit faza, aceasta poate fi calculată din părțile reale și imaginare.

Modelul nostru neural este structurat pentru a gestiona un tensor de intrare cu dimensiunile $154 \times 2048 \times 3$ și pentru a produce un tensor de ieșire cu dimensiunile $1 \times 2048 \times 3$. Rețeaua reduce sistematic dimensiunea de-a lungul axei verticale (154), reprezentând numărul de intervale de timp pentru calculul STFT, până când ajunge la dimensiunea de 1. În același timp, menține dimensiunea constantă de-a lungul celeilalte axe, specific numărul de puncte FFT (N_x).

Arhitectura noastră, ilustrată în Figura 3.2, constă în 10 straturi convoluționale organizate în 4 blocuri. Fiecare dintre primele două blocuri include 3 straturi convoluționale urmate de max-pooling, în timp ce al treilea și al patrulea bloc au câte 2 straturi convoluționale fiecare, ultimul fiind lipsind de pooling. LeakyReLU [23] urmează fiecare strat convoluțional, cu excepția ultimelor două. Numărul de filtre începe de la

Tabel 3.1 Valorile minime și maxime pentru fiecare parametru utilizate pentru generarea semnalelor din baza noastră de date.

Parameter	Minimum	Maximum	Step
Interference sources	1	3	1
SNR [dB]	5	40	5
SIR [dB]	-5	40	-
Relative interference signal slope	0	1.5	-
Number of targets	1	4	-
Target amplitude	0.01	1	-
Target distance [m]	2	95	-
Target phase [rad]	$-\pi$	π	-

32 și crește cu 32 în blocurile ulterioare, ajungând la 128 în blocul final. Dimensiunea nucleului scade de la 13×13 la 5×5 de-a lungul rețelei, excepție fiind ultimul strat convoluțional cu un nucleu de 1×1 . Filtrele convoluționale au un pas de 1×1 cu bordare circulară, iar filtrele de extragere sunt de 2×1 , înjumătățind dimensiunea verticală. Bordarea cu zero asigură hărți de activare de dimensiuni egale pentru max-pooling.

3.2.4 Constrangerea Ponderilor

Pentru reducerea interferențelor radar am dezvoltat o metodă de constrangere a ponderilor [11, 21], care are două faze. Prima fază începe cu o parte standard de antrenament. Ulterior, se execută o parte de antrenament constrânsă, special concepută pentru a tăia zgomotul intern al rețelei și pentru a îmbunătăți raportul semnal-zgomot al acesteia. Se remarcă faptul că, arhitectura rețelei rămâne neschimbată în timpul testării. În a doua etapă, metoda de constrangere reduce zgomotul intern al rețelei prin reducerea la zero a ponderilor mai mici decât un anumit prag. Pragul este determinat prin sortarea ponderilor în conformitate cu valoarea și reducerea la zero a celor mai mici r procente din ponderi.

3.3 Seturi de Date

Am extins setul de date ARIM [35] pentru a include diverse scenarii auto ce imita lumea reală, cu multiple surse de interferență. În publicația noastră recentă [36], am introdus o bază de date nouă și extinsă numită ARIM-v2, care cuprinde 144,000 de semnale generate sintetic ce replică fidel scenariile auto din lumea reală. Generarea fiecărui eșantion implică selecția aleatorie a valorilor dintr-un set de parametri realistici detaliați în Tabelul 3.1.

Datele generate sintetic oferă control complet asupra procesului de antrenare, furnizând informații cuprinzătoare pentru soluții superioare de reducere a interferențelor radar. Semnalele fara interferența servesc ca etichete în timpul antrenamentului modelului de învățare automată, facilitând evaluarea obiectivă a performanței prin compararea ieșirii prezise cu ieșirea așteptată.

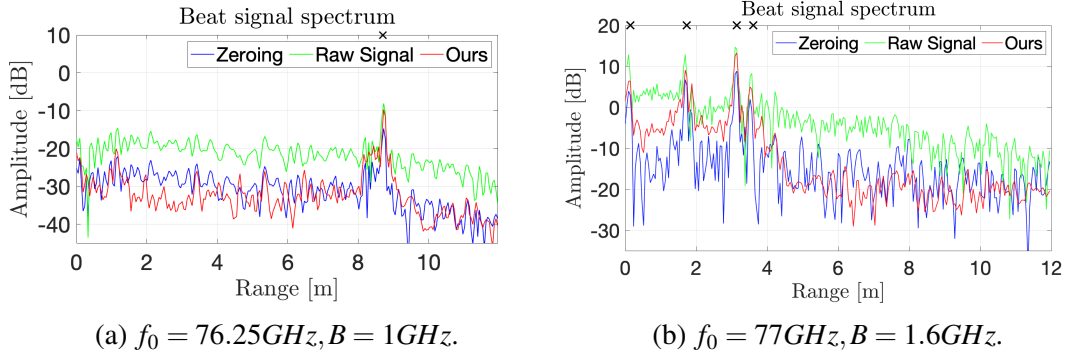


Figura 3.3 Rezultatele calitative ale modelului nostru FCN+pruning. Semnalele reale sunt obținute folosind următoarele configurări: (a) cu un senzor radar de la FAU și (b) cu transceiverul radar NXP TEF810X de 77 GHz. Aici, parametrul f_0 denotă frecvența centrală a interferenței, iar B semnifică lățimea de bandă a chirpului.

Tabel 3.2 Rezultatele pe seturile de validare și test ARIM-v2, care conțin până la trei surse de interferență. Amplitudinea este măsurată în decibeli (dB), în timp ce faza este măsurată în grade. Simbolul \uparrow indică faptul că valorile mai mari sunt preferabile, în timp ce \downarrow semnifică faptul că valorile mai mici sunt mai bune.

Method	Validation set				Test set			
	$\overline{\Delta\text{SNR}} \uparrow$	AUC \uparrow	MAE \downarrow Ampl.	MAE \downarrow Phase	$\overline{\Delta\text{SNR}} \uparrow$	AUC \uparrow	MAE \downarrow Ampl.	MAE \downarrow Phase
Oracle (true labels)	16.87	0.971	0	0	17.15	0.970	0	0
Zeroing	8.64	0.930	2.11	12.63	8.94	0.929	2.13	12.55
CNN [40]	12.64	0.952	2.14	8.25	12.94	0.953	2.17	8.13
FCN (ours)	15.02	0.961	1.39	7.32	15.28	0.959	1.40	7.44
FCN + pruning (ours)	15.09	0.963	1.25	6.41	15.36	0.961	1.27	6.58

3.4 Rezultate

3.4.1 Metrici de Evaluare

Metodele de evaluare includ AUC, eroarea absolută medie a amplitudinii (MAE pentru amplitudine), eroarea absolută medie a fazei (MAE pentru fază) și îmbunătățirea medie a raportului semnal-zgomot ($\overline{\Delta\text{SNR}}$). Îmbunătățirea medie a SNR-ului este calculată pentru ținta cu cea mai mare amplitudine, reprezentând diferența dintre SNR înainte și după atenuarea interferenței în profilul de distanță.

3.4.2 Alegerea Hiper-Parametrilor

În toate experimentele, am utilizat mini-seturi de 16 eșantioane, folosind optimizatorul Adam [15] cu o rată de învățare de $5 \cdot 10^{-5}$ și o reducere a ponderilor de 10^{-5} . Raportul de reducere a zgomotului (r) este validat pentru setul de valori $\{0.15, 0.3, 0.45\}$, cu performanța optimă observată la $r = 0.3$. Rețeaua este optimizată folosind eroarea absolută medie pentru toate cele trei canale considerate.

3.4.3 Rezultate pe ARIM-v2

În setul de date ARIM-v2, am comparat FCN-ul nostru sub regimuri convenționale și de constrangere a ponderilor cu etichetele, baseline-ul de zeroing și CNN-ul de la Rock *et al.*[40]. Rezultatele din Tabelul 3.2 arată că abordarea noastră a obținut performanțe superioare în toate metricile, apropiate de nivelurile oracolului (etichetelor). În scenarii dificile cu multiple surse de interferență, FCN-ul nostru obține erori semnificativ mai mici comparativ cu zeroing. Metoda noastră depășește metodele competitive, fiind o metodă robustă chiar și pentru multiple surse de interferență.

3.4.4 Generalizare pe Date Reale

Considerația principală atunci când antrenăm o rețea neurală pe semnale generate sintetic este capacitatea sa de a generaliza la date din lumea reală. În consecință, evaluăm capacitatea de generalizare a FCN-ului nostru testându-l pe semnale reale obținute de la doi senzori radar distincti. În Figura 3.3, prezentăm rezultate calitative pentru două semnale reale care contin interferențe, comparând metoda noastră cu abordarea de zeroing. Primul grafic, ilustrat în Figura 3.3a, este generat folosind date reale de la FAU [10]. La inspectarea rezultatelor, devine evident că rețeaua noastră reușește să ofere estimări mai precise ale amplitudinilor țintei, atenuând eficient interferențele și reducând nivelul de zgomot de fundal. Al doilea grafic, reprezentat în Figura 3.3b, conține date furnizate de compania NXP, capturate folosind transceiverul radar NXP TEF810X de 77 GHz în timpul unor experimente exterioare pe un drum cu două benzi. În aceste experimente, radarul victimă a fost fixat pe bara de protecție a unei mașini, în timp ce radarul interferent a fost plasat pe un trepied într-o locație fixă în afara drumului. Abordarea noastră atenuază cu succes interferențele, oferind rezultate superioare în ceea ce privește amplitudinea țintei comparativ cu algoritmul de zeroing.

3.5 Discutii

În această lucrare, prezentăm o contribuție în domeniul prelucrării semnalelor radar pentru autovehicule prin introducerea unei noi rețele complet convoluționale capabilă să estimeze cu acuratețe atât magnitudinea cât și faza semnalelor radar, în special în prezența mai multor surse de interferență. Mai mult decât atât, am creat o baza de date cuprinzătoare care include semnale radar simulate în scenarii realiste și complexe. De asemenea, modelul nostru este capabil să generalizeze pentru doi senzori radar distincti fără nicio ajustare.

Capitolul 4

Analiza Trasaturilor Oceanice

4.1 Introducere

În aceasta teză, ne bazăm pe metodele propuse în [33, 34] pentru a dezvolta soluția de cautare a imaginilor oceanice în mod nesupervizat. Mai exact, am integrat algoritmul de descompunere pe subaperturi (SD) cu învățarea profundă nesupervizată (UDL). Prin utilizarea UDL pentru motorul de cautare a imaginilor, eliminăm necesitatea datelor etichetate. În plus, utilizând algoritmul de preprocesare bazat pe SD, îmbunătățim semnificativ acuratețea cautării, apropiindu-ne de nivelurile de acuratețe obținute de abordările de învățare supervizată. De asemenea, evaluăm capacitatea modelului nostru de a se adapta la algoritmi de teledetecție bazati pe modele fizice, cum ar fi imagini cu centroide Doppler. Utilizând tehnicile bazate pe modele fizice descrise, am dezvoltat un algoritm eficient de interogare pe baza de imagine, conceput pentru a asista experții în identificarea fenomenelor similare de pe suprafața oceanului. Fiecare imagine SAR este caracterizată printr-un vector de trasaturi calculat folosind o rețea neurală profundă antrenată în prealabil, antrenată într-un mod nesupervizat. În plus, am extins aplicabilitatea interogării prin imagine la o abordare mai complexă numită interogare prin parametri fizici. În mod specific, am estimat imaginile Doppler pe subaperturile imaginilor SAR SLC și le-am utilizat pentru sistemul de interogare.

4.2 Metode Propuse

4.2.1 Descompunerea în Subaperturi

Având în vedere natura extrem de nestationară a suprafeței oceanului, observarea acesteia din unghiuri diferite poate oferi informații suplimentare despre zona iluminată. Prin urmare, adoptăm o strategie de descompunere a imaginii SAR în subaperturi, fiecare corespunzând unei imagini formate utilizând doar o porțiune din unghiul de azimut total. Această descompunere ne permite să simulăm diferite unghiuri de observație ale aceleiași scene, adunând astfel informații mai cuprinzătoare. Algoritmul SD este ilustrat în prima parte a Figurii 2.1.

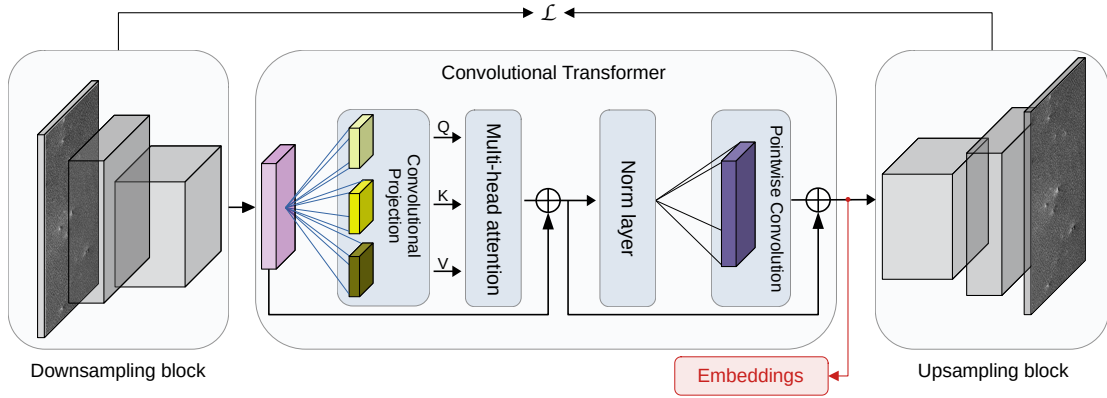


Figura 4.1 Arhitectura propusă de auto-codificator numită CyTran. Săgeata roșie indică stratul considerat pentru extragerea trăsăturilor.

4.2.2 Estimarea Centroidelor Doppler

În această teză am propus o tehnică de preprocesare bazată pe modele fizice pentru suprafața oceanului, numită estimarea centroidelor Doppler (DCE). Fiecare imagine SAR sau subapertură generată de SD, ar putea fi ulterior procesată pentru a obține centroidele Doppler. Această etapă de procesare transformă datele SAR în informații cu caracter fizic, legate de mișcarea zonei iluminate. Procesarea este descrisă formal în Secțiunea 2.1.4. O ilustrație a acestui proces este reprezentată în zona portocalie din Figura 2.1. Când DCE este integrat în fluxul de lucru, fiecare imagine Doppler este estimată pe o subapertură și ulterior decimată.

4.2.3 Rețele Neuronale Nesupervizate

În studiul nostru, am utilizat arhitectura generativă CyTran, ce este compusă dintr-un bloc de reducere a dimensiunilor convoluțional, un bloc transformer convoluțional și un bloc de mărire a dimensiunilor, așa cum este ilustrat în Figura 4.1. Inspirat de blocul propus în lucrarea noastră [38], blocul transformer convoluțional este folosit ca punct intern de procesare pentru arhitectură. Fie $X \in \mathbb{R}^{c \times m \times n}$ tensorul de intrare pentru blocul transformer, unde c reprezintă numărul de canale, iar m și n denotă lățimea și înălțimea. Dimensiunile spațiale ale descriptorilor vizuali sunt determinate de câmpul receptiv al filtrelor din stratul de proiecție convoluțional, așa cum este ilustrat în Figura 4.1. Proiecția convoluțională cuprinde trei blocuri de proiecție aproape identice, fiecare cu parametri separați. Fie W_Q , W_K și W_V parametrii învățabili ai celor trei straturi de proiecție. Descriptorii pentru interogare (Q), cheie (K) și valoare (V) sunt calculate după cum urmează:

$$\begin{aligned}
 Q &= \text{conv_projection}(X, W_Q), \\
 K &= \text{conv_projection}(X, W_K), \\
 V &= \text{conv_projection}(X, W_V),
 \end{aligned}
 \tag{4.1}$$

Tabel 4.1 Rezultatele de acuratețe pentru un model ResNet18 testat pe setul de test TenGeoP-SARwv. Prin ‘‘Subaperture (1)’’ intelegem ca intrarea este doar prima subapertura. Rezultatele semnificativ mai bune (pentru un nivel 0.01), conform unui test McNemar aplicat pe perechi, sunt marcate cu †.

Vignette	98.0
Subaperture (1)	94.0
Subapertures	98.9 [†]
DCE Vignette	78.6
DCE Subapertures (1)	75.3
DCE Subapertures	93.3 [†]

unde $Q \in \mathbb{R}^{n_q \times d_q}$, $K \in \mathbb{R}^{n_k \times d_k}$ si $V \in \mathbb{R}^{n_v \times d_v}$. Pentru operaiunea ulterioara care implica nmuliri de matrici, avem nevoie ca $d_q = d_k$, i $n_k = n_v$. Datorita numarului egal de filtre n convoluia punct la punct n toate cele trei blocuri, $d_q = d_k = d_v$. Interogrile, cheile i valorile generate sunt trimise catre un strat de atenie paralel, cu scopul de a captura interaciunea dintre toti descriptorii prin codarea fiecarei entitai n termenii informaiilor contextuale globale.

Folosim arhitectura CyTran ntr-un mod nesupervizat, avnd ca obiectiv aproximarea funciei identitate prin auto-codarea intrrii. Am considerat eroarea medie patratice ca funcie obiectiv i am extras descriptorii aa cum este ilustrat n Figura 4.1.

4.2.4 Cutarea Bazata pe Coninut a Imaginilor

Avnd n vedere o baza de date foarte mare cu imagini SAR oceanice, introducem un algoritm nesupervizat conceput pentru a identifica imaginile SAR similare. Subliniem ca algoritmul nostru este versatil, neavnd nevoie de etichete i nefiind constrns la niciun tip particular de date de intrare. Pentru a demonstra versatilitatea abordrii noastre, lum n considerare doua tipuri distincte de date ca intrare: subaperturile SAR i centroizii Doppler estimai din subaperturi. n consecina, efectum cutarea imaginilor folosind att datele SAR de amplitudine cat i reprezentrile bazate pe modele fizice. Exploatnd acest ultim tip de date, algoritmul poate construi un motor de cutare mai complex capabil sa identifice fenomene bazate pe caracteristici fizice specifice, cum ar fi curenii oceanici cu viteze particulare.

Tehnic, algoritmul de cautare asociaza un vector de trasaturi pentru toate imaginile din baza de date. Vectorul de trasaturi este obtinut prin codarea imaginii cu reeaua CyTran. Ulterior, avnd o imagine de interogare, obtinem descriptorul de interogare, calculm distana cosinus ntre interogare i ntreaga baza de date i gasim cele mai similare N exemple.

Tabel 4.2 Rezultatele cautarii pentru datele de intrare bazate pe SD și DCE. Am inclus rezultatele pentru solutia supervizata (S-) și nesupervizata (U-) bazate pe arhitectura propusă CyTran.

	S-Vig	S-Subap	U-Vig	U-Subap	S-Dop Vig	S-Dop Subap	U-Dop Vig	U-Dop Subap
P@5	98.1	98.9	47.4	72.6	76.9	91.3	51.1	66.7
P@50	97.4	97.2	32.6	57.3	69.8	88.0	36.0	52.0

4.3 Rezultate

4.3.1 Setul de Date

Setul de date TenGeoP-SARwv cuprinde peste 37,000 de imagini SAR oceanice reprezentând 10 fenomene geofizice distincte pe suprafața oceanului. Urmând metodologia descrisă în [33, 34], am împărțit datele în mod aleatoriu în seturi de antrenament (70%), validare (15%) și test (15%). În plus, pentru experimentele care implică analize bazate pe Doppler, am urmat procedura ilustrată în Figura 2.1.

4.3.2 Alegerea Hiper-Parametrilor

În experimentul de clasificare, am calibrat hiper-parametrii similar cu [33]. În ceea ce privește modelul CyTran, am urmat valorile propuse în [34], făcând ajustări doar la numărul de canale de intrare și ieșire pentru a se alinia cu tipul de intrare considerat. Am antrenat modelul pentru 100 de epoci, am utilizat optimizatorul Adam și o dimensiune de mini-set de 16. Pentru DCE, am setat $d_1 = d_2 = 32$ pentru filtrul de medie.

4.3.3 Metrici de Evaluare

Am raportat acuratețea pentru sarcina de clasificare și am efectuat teste statistice McNemar pentru a demonstra ca îmbunătățirile sunt statistic semnificative. În ceea ce privește sarcina de cautare, având în vedere concentrarea noastră pe fluxuri de date mari, am raportat scoruri de precizie pentru 5 (P@5) și 50 (P@50) exemple. Fiecare scor a fost mediat pe baza a 100 de interogări. Mai precis, am calculat P@5 și P@50 pentru 100 de imagini de interogare și apoi am mediat rezultatele.

4.3.4 Rezultate Clasificare

Extindem rezultatele prezentate în [33] în Tabelul 4.1, unde prezentăm acuratețea clasificării realizată de modelul ResNet18 pe setul de test TenGeoP-SARwv, luând în considerare diverse tipuri de date de intrare. Observăm că algoritmul SD îmbunătățește acuratețea atât pentru datele SAR de amplitudine cât și pentru DCE.

4.3.5 Rezultate pentru Antrenare Nesupervizată

Am antrenat modelul auto-codificator CyTran [38] pe setul de antrenament TenGeoP-SARwv și am selectat modelul cu cele mai bune performanțe bazat pe valoarea costului de reconstrucție pe setul de evaluare. Este important de menționat că, deși am încercat

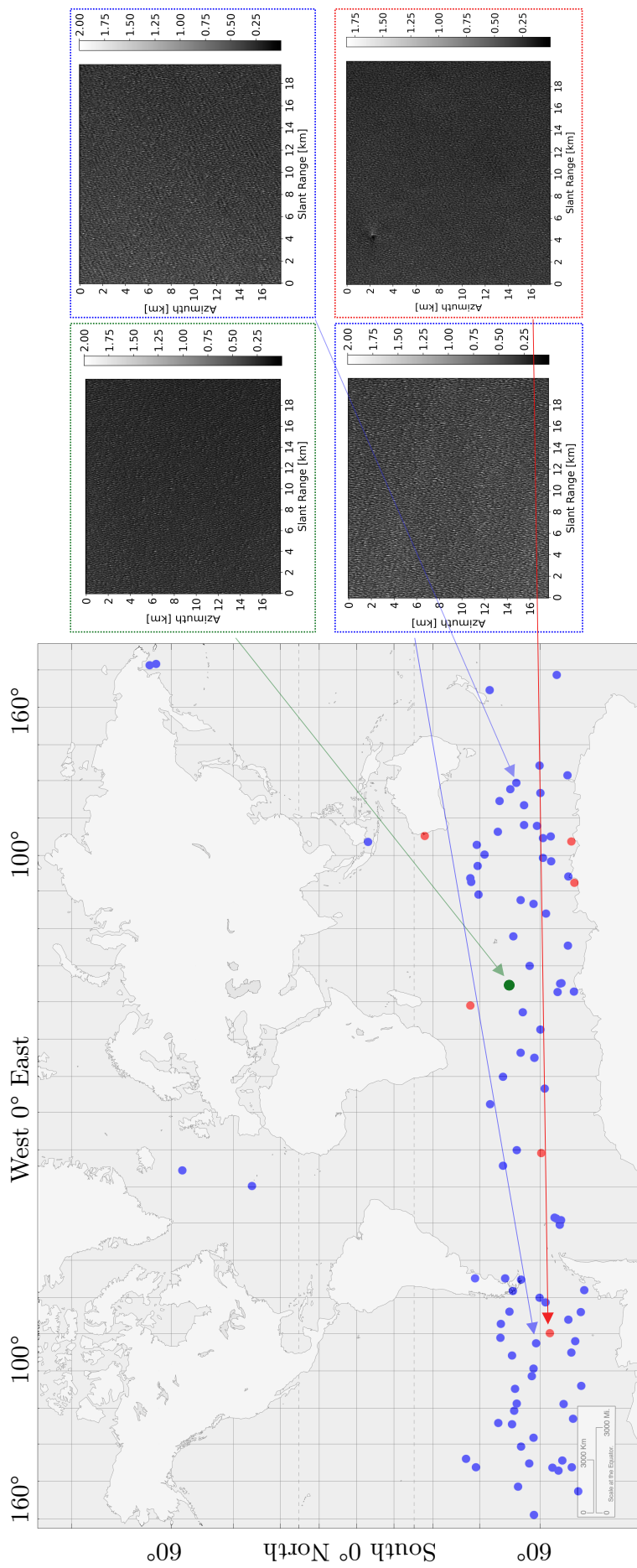


Figura 4.2 Rezultatele pentru sarcina de cautare pe baza descriptorilor din modelul CyTran, antrenat pe toate subaperturile din imaginea SAR, sunt prezentate. Prezentăm cele mai similare $N_{\max} = 100$ imagini împreună cu informațiile lor de localizare. În vizualizare, imaginea de interogare este marcată cu verde, imaginile din aceeași clasă sunt cu albastru, iar imaginile din clase diferite sunt marcate cu roșu. Pe partea dreaptă, imaginile SAR originale sunt afișate: verde și albastru reprezintă Valuri Pure Oceanice, în timp ce roșu reprezintă imaginile cu Iceberguri.

mai multe modele (cum ar fi auto-codificatorul ResNet și U-Net), acestea nu au reușit să convergă către rezultate optime de reconstrucție, astfel au fost excluse din experimentele urmatoare.

4.3.6 Rezultate Tehnici Cautare

În Tabelul 4.2, prezentăm performanța de cautare folosind descriptorii generati de rețeaua CyTran antrenată atât pe imaginile SAR originale cât și pe subaperturi. Performanta este comparata cu descriptorii produși de un model ResNet18 antrenat în mod supervizat. Când comparăm trasaturile supervizate pe imaginile SAR originale (S-Vig) și subaperturi (S-Subap), rezultatele arată diferențe neglijabile, cu variații mai mici de 1% pentru ambele $P@5$ și $P@50$. Cu toate acestea, introducerea algoritmului SD îmbunătățește constant precizia în rezultatele de cautare bazate pe trasaturi nesupervizate. Această reducere a diferenței de performanță de cautare între abordările supervizate și nesupervizate subliniază impactul semnificativ al algoritmului SD.

De asemenea, în Tabelul 4.2, prezentăm performanța de cautare folosind descriptorii generati de rețeaua CyTran antrenată pe DCE pe imaginile SAR originale și subaperturi, comparându-le cu descriptorii produși de un model ResNet18 antrenat în mod supervizat. Așa cum era de așteptat din experimentul de clasificare, performanța de cautare înregistrează o îmbunătățire semnificativă atunci când se utilizează trasaturi nesupervizate bazate pe subaperturi.

Am ilustrat rezultatele de cautare pentru trasaturile nesupervizate antrenate pe subaperturi pentru o imagine de interogare în Figura 4.2 (clasa de valuri oceanice pure). Observăm că imaginile returnate, reprezentând imaginile cele mai similare, sunt dispersate pe întreaga zonă geografică unde fenomenul ar putea să apară. Acest lucru sugerează că modelul nesupervizat nu prezintă suprainvatare în ceea ce privește locația geografică. În plus, am observat similarități structurale între imaginile returnate cu etichete incorecte (evidențiate de punctele roșii în Figura 4.2). Această observație ar putea indica prezența a două fenomene în aceeași locație sau alte similarități intrinseci.

4.4 Discutii

În acest studiu, am extins cercetările anterioare efectuate în [33] prin integrarea algoritmului SD în fluxurile de învățare a caracteristicilor nesupervizate cu rețele transformer. Obiectivul nostru a fost de a folosi SD ca o etapă de preprocesare pentru modelele destinate cautării imaginilor SAR pe suprafața oceanului. Rezultatele obtinute au relevat impactul metodei SD asupra performanței de cautare, în special când se încorporează algoritmi bazati pe modele fizice, cum ar fi DCE. Acesta dezvoltare experimentală permite dezvoltarea unor motoare de căutare mai complexe, capabile să identifice parametri fizici similari, mai degrabă decât structuri de continut. În esență, abordarea noastră s-a concentrat pe o metodologie axata pe date, destinată să sporească eficacitatea algoritmilor de clasificare și cautare atât supervizati cât și nesupervizati.

Capitolul 5

Segmentarea Gheții Oceanice din Imagini SAR

5.1 Introducere

Zonele înghetate sunt o componentă crucială a sistemului climatic al Pământului, fiind foarte sensibile la schimbările de temperatură și condițiile atmosferice. Evaluarea precisă a parametrilor gheții oceanice este vitală pentru înțelegerea și predicția variațiilor climatice. Totuși, analiza manuală a tuturor datelor satelitare care acoperă zonele înghetate este impracticabilă. Prin urmare, algoritmi automatizați sunt esențiali pentru a utiliza pe deplin fluxurile continue de date satelitare. În acest studiu, am dezvoltat două arhitecturi, una bazată pe CyTran-ul descris anterior și a doua numită UT-MHTA, o arhitectură inovatoare bazată pe transformer UNet, concepută pentru segmentarea gheții marine folosind imagistica satelitară SAR. UT-MHTA înlocuiește blocul convențional de atenție paralel (MHA) cu o atenție transpusă paralelă (MHTA), permițând captarea interacțiunilor la distanță din aceeași imagine, în timp ce se menține adecvat pentru procesarea imaginilor mari. Performanța modelelor noastre depășește metodele de ultimă oră, oferind rezultate superioare fără a crește semnificativ complexitatea computațională. În mod specific, cea mai bună rețea a noastră, UT-MHTA, atinge o medie a intersecției peste uniune (mIoU) de 68.76% pe setul de date AI4Arctic [4], cu un timp de inferență de 865ms pentru un produs de 400 km².

5.2 Metoda Propusa

5.2.1 Preprocesarea Datelor

Având în vedere dimensiunile mari ale produselor din setul de date AI4Arctic [4], antrenarea rețelelor neurale pe întreaga imagine devine impracticabilă. În schimb, adoptăm o strategie de extragere a ferestrelor de marime fixă, de dimensiuni 512 × 512.

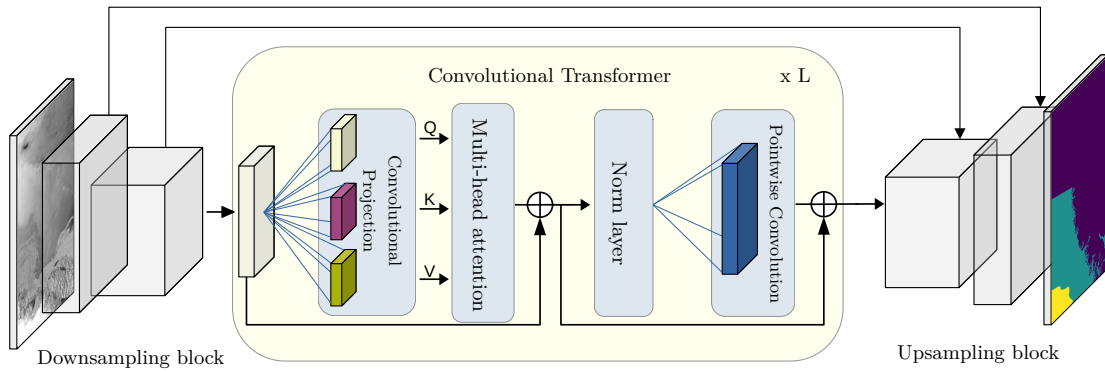


Figura 5.1 Arhitectura de segmentare U-CyTran. Modelul conține un bloc de reducere a dimensiunilor, un bloc transformer și un alt bloc de mărire a dimensiunilor.

Fiecare imagine de antrenament, de dimensiuni $512 \times 512 \times 3$, include atât polarizările HH, HV, cât și unghiul de incidență.

5.2.2 Arhitectura Reziduala Bazata pe CyTran

Arhitectura U-CyTran se bazează pe succesul modelului CyTran [38] de a fi optimizat atât pentru performanță cât și pentru viteză. Deoarece obiectivul nostru este să dezvoltăm un sistem care să poată procesa datele aproape în timp real (în zeci de milisecunde), am considerat potrivită arhitectura CyTran, care accelerează timpul de procesare prin reducerea dimensionalității. Arhitectura este construită din blocuri de reducere și marire a dimensionalității, având ca punct intern de procesare un modul transformer convoluțional. Arhitectura generală este descrisă vizual în Figura 5.1.

5.2.3 Transformer Bazat pe Atenție Transpusă Paralelă

5.2.4 Arhitectura UT-MHTA

Date fiind datele de intrare $I \in \mathbb{R}^{H \times W \times 3}$, UT-MHTA aplică inițial un bloc convoluțional pentru a obține caracteristici abstracte $F \in \mathbb{R}^{H \times W \times C}$, unde $H, W, C \in \mathbb{N}^+$. Atât blocul convoluțional de intrare cât și cel de ieșire sunt construite prin intermediul unui strat convoluțional, având dimensiunea nucleului 3 și bordarea și saltul de 1. Stratul convoluțional este urmat de normalizare (batch-norm) și funcția de activare ReLU. Apoi, F trece printr-o arhitectură asemănătoare UNet cu patru niveluri, unde modulele de procesare sunt blocuri Transformer, așa cum este prezentat în Figura 5.3. Straturile din partea de reducere a arhitecturii sunt compuse exclusiv din blocuri transformer bazate pe MHTA, operația de reducere dimensională fiind efectuată de blocuri de amestecare a pixelilor (evidențiate de săgețile roșii în Figura 5.3). În ceea ce privește partea de marire dimensională a datelor, pe lângă blocul transformer identic, am utilizat și un bloc de

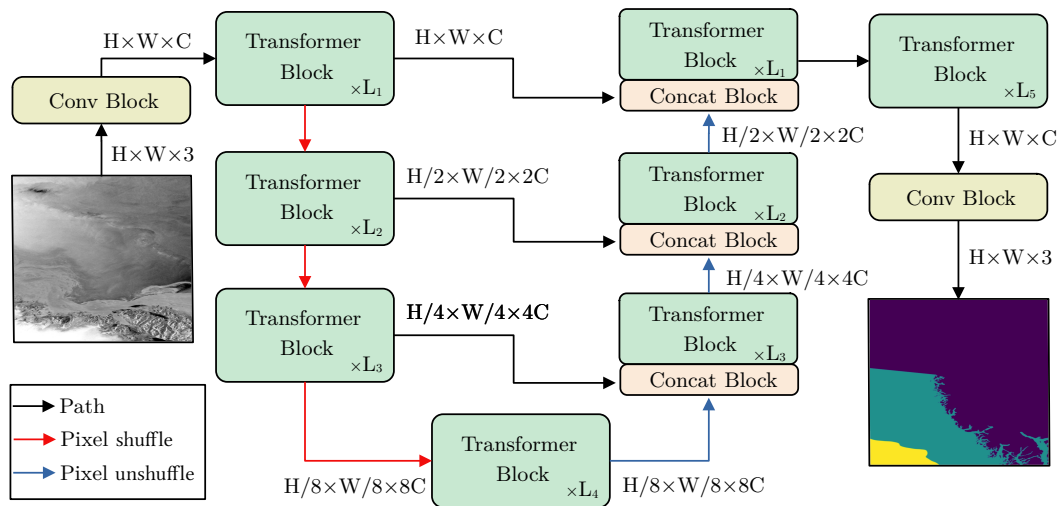


Figura 5.2 Arhitectura de segmentare UT-MHTA cuprinde o secvență de blocuri transformer, înconjurată de două blocuri convoluționale atât la începutul cât și la sfârșitul arhitecturii. Semnificația fiecărei săgeți este ilustrată în colțul din stânga jos.

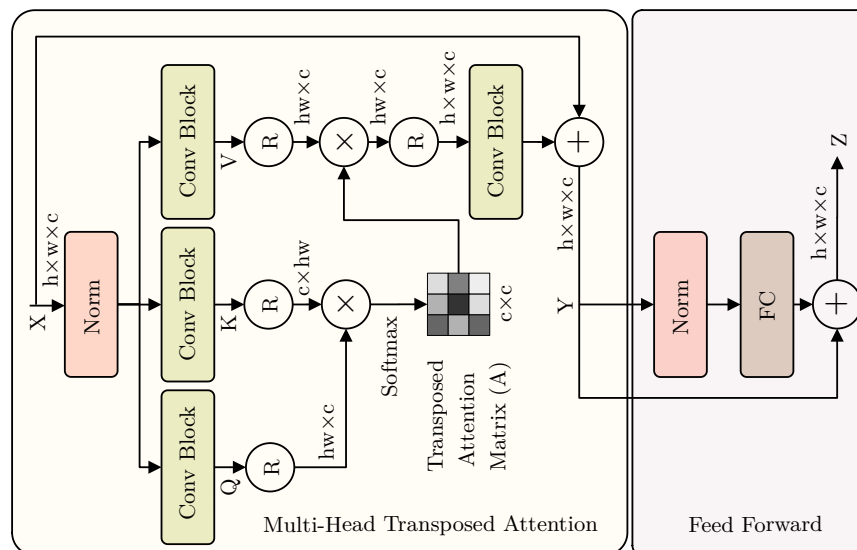


Figura 5.3 Blocul transformer prezent în UT-MHTA. Figura descrie fluxul de procesare pentru blocul transformer.

concatenare, pentru a agrega informațiile provenite de la straturile anterioare. Similar, pentru operația de marire dimensională am utilizat operația de desfășurare a pixelilor. Operațiile de amestecare și desfășurare a pixelilor sunt utilizate pentru a preveni pierderea informațiilor de-a lungul rețelei. În final, ieșirea ultimului bloc Transformer din sistemul UNet este trecută într-un bloc Transformer final, urmat de blocul convoluțional de ieșire, care prezice hărțile de segmentare. Similar cu MHA convențional [8], împărțim numărul de canale în capete și învățăm hărți de atenție separate în paralel.

Blocul Transformer. În Figura 5.3, oferim o prezentare detaliată a blocului transformer utilizat în UT-MHTA. Componenta inițială a acestui bloc este modulul de atenție transpusă paralel, conceput pentru a facilita atenția pe canale în timp ce menține complexitatea liniară a timpului și memoriei în raport cu dimensiunile de intrare $h \times w$. De remarcat este că, toate blocurile convoluționale constau într-un strat convoluțional standard cu dimensiunea nucleului de 3, împreună cu bordare și salt setate la 1. Subliniem că matricea de atenție $A \in \mathbb{R}^{c \times c}$ operează exclusiv pe dimensiunea canalului, asigurând o complexitate redusă. Modulul Feed Forward utilizat aici reflectă este similar cu modulele folosite în arhitecturile transformer obișnuite [8]. Această alegere de proiectare permite o performanță marită și facilitează integrarea fără probleme în cadrul arhitecturii mai largi.

Bloc de Concatenare. Blocul Concat realizează o concatenare obișnuită pe axa canalului, urmată de un strat convoluțional punct la punct. Convoluția punct la punct păstrează același număr de hărți și este destinată să agreghe informații din ramurile de codare și decodare.

5.3 Rezultate

5.3.1 Setul de Date

Setul de date AI4Arctic [4] a fost publicat pentru competiția de gheață marină AI4EO inițiată de Agenția Spațială Europeană (ESA). Acest set de date constă în date SAR cu microunde active Sentinel-1 cuplate cu date de la Radiometrul cu Microunde pasiv de pe senzorul satelitului AMSR2. Produsele GRDM acoperă o regiune de 400 km², având o rezoluție de 90 metri și o spațiere a pixelilor de 40 de metri. Am utilizat aceleași împărțiri ale datelor așa cum sunt descrise în [32].

5.3.2 Alegerea Hiper-Parametrilor

U-CyTran. U-CyTran utilizează optimizatorul Adam cu funcția de pierdere focală [20] pentru a optimiza performanța modelului. Inițiem antrenamentul cu o rată inițială de învățare de 10^{-3} , însoțită de un factor de scadere de 0.1 aplicat după fiecare 10 epoci. Fiecare model este antrenat pentru 50 de epoci, cu mini-seturi ce cuprind 32 de imagini.

UT-MHTA. UT-MHTA este optimizat folosind optimizatorul AdamW în combinație cu funcția de pierdere focală [20]. Antrenamentul începe cu o rată inițială de învățare de 10^{-4} , urmată de un factor de scadere de 0.5 aplicat după fiecare 10 epoci. Modelul este antrenat pentru 50 de epoci, utilizând mini-seturi de 8 imagini. Configurăm numărul de blocuri transformer cu $L_1 = L_2 = L_3 = L_4 = L_5 = 3$, și 5 capete de atenție.

5.3.3 Metrici de Evaluare

În soluția noastră de segmentare semantică care implică trei clase distincte, am identificat media intersecției peste uniune (mIoU) ca fiind metrica cea mai informativă pentru evalu-

Tabel 5.1 Rezultatele segmentării și timpul de inferență pe setul de test AI4Arctic. Am inclus pentru comparație un Transformer clasic și UT fără blocul MHTA.

Method	mIoU (%)	Timp de Inferența (ms)
ResNet AE	53.04	87
UNet [31]	56.43	92
AE	51.23	47
Transformer	64.12	673
CyTran [32]	63.68	120
U-CyTran	63.98	121
UT	62.13	182
UT-MHTA	68.76	865

area performanței modelului. Această metrică se dovedește a fi deosebit de valoroasă în scenarii cu distribuții de clase dezechilibrate. De asemenea, raportăm timpul de inferență pentru procesarea unei scene la rezoluție completă cu dimensiuni spațiale de aproximativ 1100×1100 .

5.3.4 Rezultate Experimentale

U-CyTran. Tabelul 5.1 prezintă o analiză comparativă a performanței U-CyTran față de două metode de bază: un auto-codificator bazat pe ResNet și UNet [41], așa cum este utilizat în [31]. Observăm că U-CyTran depășește ambele metode de bază cu mai mult de 7% în ceea ce privește acuratețea. De remarcă, investigăm impactul diferitelor componente ale rețelei. Când folosim doar blocurile de micșorare și marire a dimensionalității (AE), obținem cea mai mare viteză de inferență. Cu toate acestea, această abordare compromite semnificativ performanța. În schimb, utilizarea doar a blocului transformer duce la o scădere notabilă a vitezei, cu o îmbunătățire marginală a acurateții de aproximativ 0.14%.

UT-MHTA. Tabelul 5.1 prezintă rezultatele comparative ale UT-MHTA față de trei metode de bază: un auto-codificator bazat pe ResNet, UNet [31] și CyTran [32]. UT-MHTA depășește toate metodele de bază cu mai mult de 5% în termen de mIoU. În mod specific, comparativ cu arhitectura pură transformer, UT-MHTA realizează o îmbunătățire notabilă de 4.64%, experimentând o creștere modestă a timpului de inferență de mai puțin de 200ms. Această îmbunătățire subliniază eficacitatea integrării blocului MHTA în arhitectura UT, care crește semnificativ performanța fără a compromite excesiv eficiența computațională.

5.3.5 Generalizare Către Date IW

Având în vedere că setul nostru de date AI4Arctic [4] conține doar date EW GRDM, trebuie să extindem tehnicile noastre de procesare pentru a gestiona datele IW SLC.

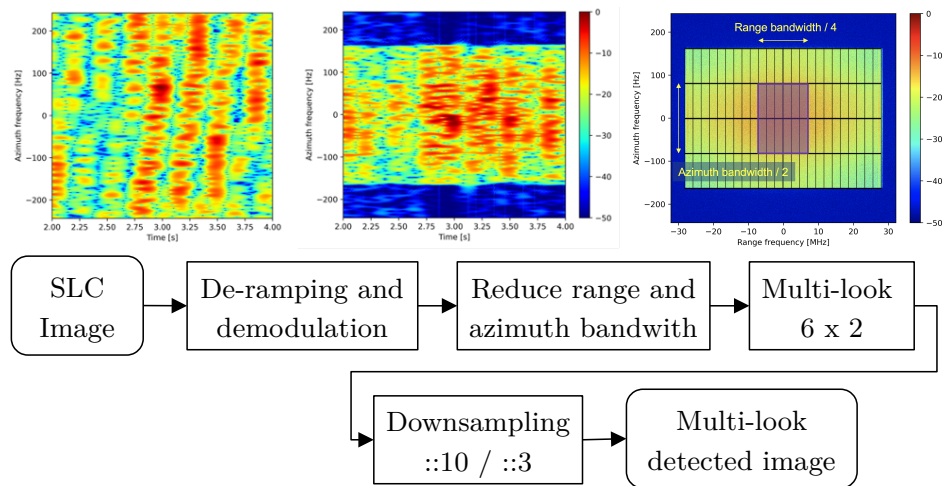


Figura 5.4 Schema de preprocesare pentru adaptarea statisticii IW SLC la EW GRD.

Evităm etichetarea manuală laborioasă efectuată de experți, propunând o tehnică de preprocesare care transformă datele IW SLC într-o imagine asemănătoare cu EW GRD, numită IW-EW GRD. În Figura 5.4, ilustrăm schema de procesare de la imaginea IW SLC la imaginea detectată (IW-EW GRD), aliniindu-i rezoluția și spațierea pixelilor îndeaproape cu datele de antrenament EW GRDM.

Pentru a evalua subiectiv tehnica de preprocesare și modelul UT-MHTA, am inclus figuri care arată rezultatele obținute. Figura 5.5 prezintă modul HV pentru ambele IW GRDH și IW SLC în aceeași zonă geografică. În partea de jos a figurii, am suprapus etichetele prezise de rețeaua UT-MHTA pe produsele HV. Evident, rețeaua demonstrează o generalizare robustă la date nevăzute, cu predicții similare între produsele SLC și GRDH.

5.4 Discuții

În acest capitol am introdus U-CyTran și UT-MHTA, două noi arhitecturi de segmentare a gheții oceanice bazate pe arhitectura transformer UNet, adaptate pentru analiza imaginilor SAR satelitare. În arhitectura U-CyTran, am evoluat soluția prezentată în [32], adăugând conexiuni de tip salt, permițând rețelei să transfere ușor caracteristicile între părțile de intrare și ieșire. În UT-MHTA, inovăm înlocuind mecanismul MHA convențional cu atenția transpusă paralela. Această adaptare permite captarea interacțiunilor pixelilor aflați în regiuni îndepărtate ale imaginii, o capacitate crucială pentru delimitarea precisă a marginilor gheții marine în imagini de mari dimensiuni. Mai mult, am extins aria de utilizare a rețelelor noastre aplicând o schema de preprocesare de la datele IW SLC la date similare EW GRD. Considerând această etapă suplimentară, am observat că predicțiile sunt similare, validând aplicabilitatea rețelei noastre pentru produsele directe de nivel 1 SLC.

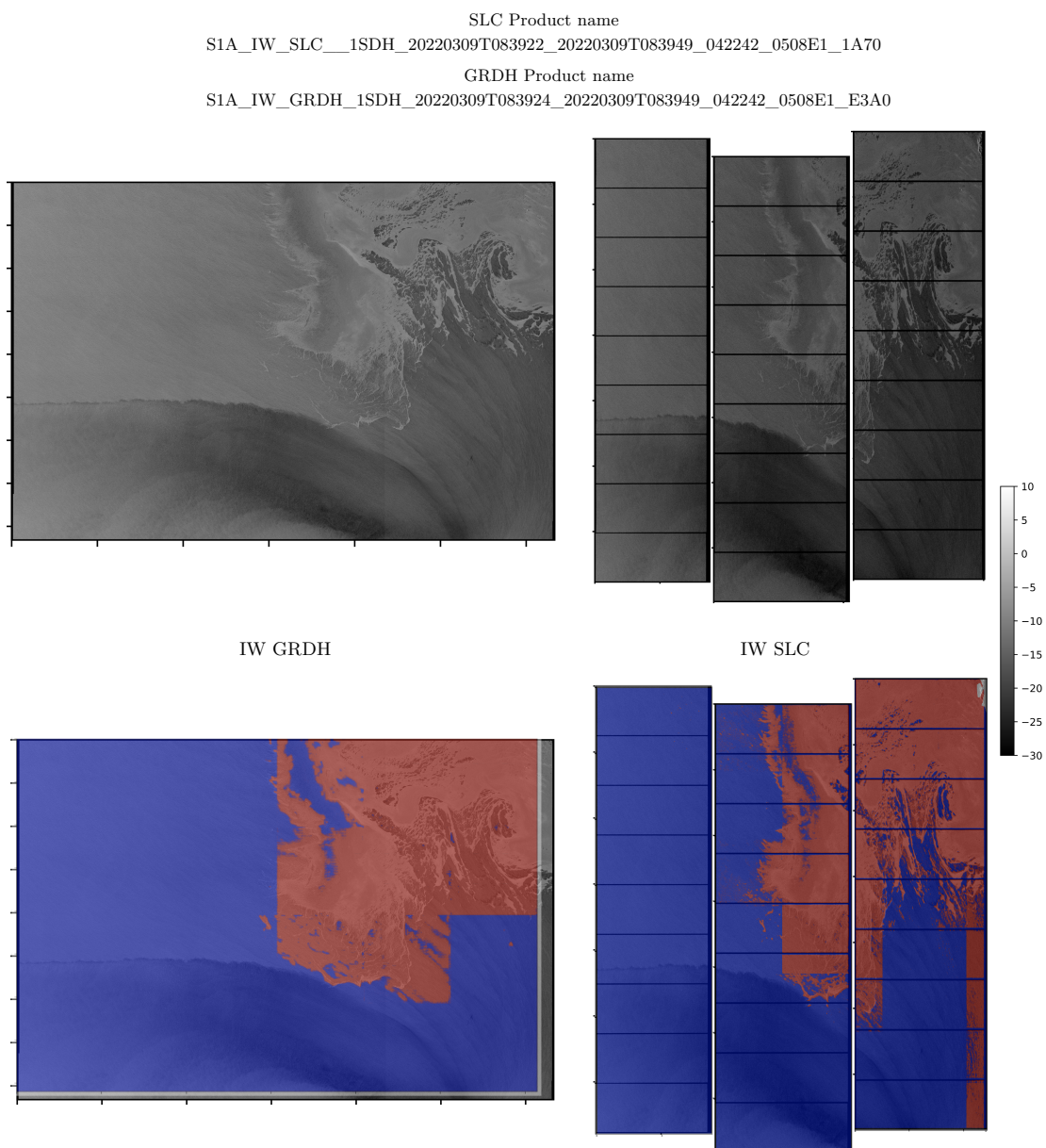


Figura 5.5 În partea stângă sus a figurii ilustrăm imaginea IW GRDH după subeșantionare. În partea dreaptă sus a figurii ilustrăm imaginea IW SLC după schema noastră de preprocesare. În partea de jos a figurii ilustrăm ieșirile corespunzătoare ale rețelei noastre UT-MHTA pentru imaginilor mai sus menționate. Toate produsele sunt în modul HV.

Capitolul 6

Concluzii

În lucrarea noastră ne-am propus să dezvoltăm în continuare domeniul prelucrării semnalelor prin realizarea unei cercetări bazate pe aplicații și aplicând tehnici inovatoare de învățare profundă pentru a depăși lacunele existente. Ne-am concentrat cercetarea în două domenii, atenuarea interferențelor radar auto și observația Pământului. Concluziile sunt detaliate mai jos în conformitate cu domeniul specific.

6.1 Reducerea Interferențelor Radar Auto

Contribuția noastră este după cum urmează:

- Am proiectat o rețea neurală convoluțională completă pentru reducerea interferențelor radar auto în scenarii complexe, cum ar fi sursele multiple de interferențe. FCN-ul proiectat este capabil să estimeze atât magnitudinea, cât și faza spectrului semnalului de batai dorit, fiind compatibil cu blocurile viitoare de prelucrare radar (de exemplu, estimarea vitezei), care necesită faza.
- Am publicat setul de date ARIM-v2 pentru reducerea interferențelor radar auto. Baza de date conține scenarii complexe, precum și surse multiple de interferențe.
- Am desfășurat o serie extinsă de experimente, demonstrând că FCN-ul nostru atinge rezultate de excelență în domeniu. Mai mult, am dovedit empiric capacitatea FCN-ului de generalizare pe date reale de la senzori radar, reducând interferențele pentru doi senzori distincti, fără a fi necesară o ajustare suplimentară.

6.2 Analiza Trăsăturilor Oceanice

Contribuția noastră este după cum urmează:

- Am aplicat decompunerea pe subaperturii pentru a îmbunătăți performanța rețelei atât pentru imaginile SAR, cât și pentru imaginile Doppler.
- Am dezvoltat un transformer hibrid convoluțional pentru cautarea de imagini nesupervizate pe suprafața oceanului.

- Am dezvoltat o abordare bazată pe exploatarea datelor, prin combinarea atât a SD cât și a DCE pentru o căutare nesupervizată mai precisă, precum și a unui motor de căutare bazat pe modele fizice. Abordarea noastră s-a dovedit a fi generică și performantă, permițând modelului nostru să învețe trasaturi complexe, reducând diferența între căutarea de imagini supervizată și nesupervizată.

6.3 Segmentarea Gheții Oceanice din Imagini SAR

Contribuția noastră este după cum urmează:

- Am evoluat arhitectura CyTran pentru a include conexiuni tip salt similare cu UNet pentru rezultate mai precise în segmentarea gheții oceanice. Mai mult, am dezvoltat arhitectura UT-MHTA care înlocuiește modulul tradițional de atenție paralel cu un mecanism de atenție transpus paralel, facilitând capturarea interacțiunilor pixelilor din zone îndepărtate ale imaginii, fiind potrivit pentru imagini de dimensiuni mari.
- Având în vedere nevoia de algoritmi atât preciși, cât și rapizi în domeniu, am optimizat rețelele noastre pentru a obține cea mai bună balanță între precizie și viteză, astfel încât arhitecturile propuse să fie candidați fezabili pentru procesarea în timp aproape real.
- Am dezvoltat un cadru de preprocesare pentru adaptarea produsului IW SLC la imagini asemănătoare cu EW GRD, extinzând aplicabilitatea rețelelor noastre la produsele de nivel 1 SLC.

6.4 Lista de Publicații

Articole de Jurnal

1. Ristea, N.C., Anghel, A. and Ionescu, R.T., 2021. Estimating the Magnitude and Phase of Automotive Radar Signals Under Multiple Interference Sources with Fully Convolutional Networks. *IEEE Access*, 9, pp.153491-153507.
2. Ristea, N.C., Anghel, A., Datcu, M. and Chapron, B., 2023. Guided Unsupervised Learning by Subaperture Decomposition for Ocean SAR Image Retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-11, 2023, Art no. 5207111.
3. Ristea, N.C., Miron, A.I., Savencu, O., Georgescu, M.I., Verga, N., Khan, F.S. and Ionescu, R.T., 2023. CyTran: a Cycle-Consistent Transformer with Multi-Level Consistency for Non-Contrast to Contrast CT Translation. *Neurocomputing*, 538, p.126211.

Articole de Conferință

1. Ristea, N.C., Anghel, A., Datcu, M. and Chapron, B., 2022, July. Guided Deep Learning by Subaperture Decomposition: Ocean Patterns from SAR Imagery. In Proceedings of IEEE International Geoscience and Remote Sensing Symposium (pp. 6825-6828).
2. Chirtu, A., Ristea, N.C. and Radoi, A., 2022, July. Convolutional Transformers for Aerial Image Classification: a General to Specific Learning Curve. In Proceedings of IEEE International Geoscience and Remote Sensing Symposium (pp. 659-662).
3. Ristea, N.C., Anghel, A. and Datcu, M., 2023, July. Sea Ice Segmentation From SAR Data by Convolutional Transformer Networks. In Proceedings of IEEE International Geoscience and Remote Sensing Symposium (pp. 168-171).
4. Ristea, N.C., Anghel, A. and Datcu, M., 2024, July. Multi-Head Transposed Attention Transformer for Sea Ice Segmentation in SAR Imagery. In Proceedings of IEEE International Geoscience and Remote Sensing Symposium.

6.5 Lista de Publicații în Afara Scopului Principal al Tezei

Articole de Jurnal

1. Cutler, R., Saabas, A., Naderi, B., Ristea, N.C., Braun, S. and Branets, S., 2024. ICASSP 2023 Speech Signal Improvement Challenge. IEEE Open Journal of Signal Processing.
2. Cutler, R., Saabas, A., Pärnamaa, T., Purin, M., Indenbom, E., Ristea, N.C., Gužvin, J., Gamper, H., Braun, S. and Aichner, R., 2024. ICASSP 2023 Acoustic Echo Cancellation Challenge. IEEE Open Journal of Signal Processing.
3. Madan, N., Ristea, N.C., Ionescu, R.T., Nasrollahi, K., Khan, F.S., Moeslund, T.B. and Shah, M., 2023. Self-Supervised Masked Convolutional Transformer Block for Anomaly Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 1, pp. 525-542.

Articole de Conferință

1. Ristea, N.C. and Ionescu, R.T., 2023, October. Cascaded Cross-Modal Transformer for Request and Complaint Detection. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 9467-9471).
2. Indenbom, E., Ristea, N.C., Saabas, A., Parnamaa, T., Guzvin, J. and Cutler, R., 2023. DeepVQE: Real Time Deep Voice Quality Enhancement for Joint Acoustic

Echo Cancellation, Noise Suppression and Dereverberation. In Proceedings of INTERSPEECH Conference.

3. Ristea, N.C., Ionescu, R.T. and Khan, F., September, 2022. SepTr: Separable Transformer for Audio Spectrogram Processing. In Proceedings of INTERSPEECH Conference (pp. 4103-4107).

6.6 Direcții Viitoare

Datorită accentului semnificativ pus pe tehnicile de învățare profundă, arhitecturile se bucura de o îmbunătățire continuă. Având în vedere că soluțiile actuale în domeniul prelucrării semnalelor sunt departe de a fi perfecte, este necesară o dezvoltare pentru a obține performanțe mai bune sau procesari mai rapide, astfel încât modelele să fie implementate fără probleme în aplicații și medii reale. În acest sens, cercetarea noastră viitoare se va concentra pe: (i) integrarea și adaptarea de modele noi în aplicațiile descrise în prezent și (ii) explorarea posibilității de a integra scheme de învățare profundă în aplicații bazate pe prelucrarea semnalelor. În rezumat, cercetarea noastră viitoare se va concentra pe:

- Îmbunătățirea modelele propuse în domeniul de reducere a interferenței radar în ceea ce privește viteza, păstrând în același timp aproximativ aceeași performanță. Această direcție de cercetare este destinată în mod specific pentru a micșora modelele actuale, astfel încât soluțiile să fie viabile pentru a rula pe dispozitive limitate în timp real.
- Îmbunătățirea modelului propus pentru cautarea imaginilor oceanice în ceea ce privește performanța. Deoarece cerințele de viteză ale aplicației sunt mai relaxate, ar trebui dezvoltate îmbunătățiri suplimentare pentru a depăși precizia actuală a cautării.
- Dezvoltarea de noi strategii de învățare nesupervizată, special concepute pentru cautarea imaginilor oceanice bazate pe modele fizice. Explorarea caracteristicilor datelor ar putea duce la regimuri de antrenament nesupervizate mai bune, ceea ce în cele din urmă ar crește performanța cautării.
- Performanța actuală a segmentării gheții oceanice pe produsele IW SLC (prin efectuarea transformării IW-EW GRD) ar putea fi îmbunătățită în continuare prin alinierea și mai precisă a rezoluției produsului obținut la rezoluția EW GRDM. Acest lucru ar putea fi corectat prin adaptarea numărului de sub-looks și efectuarea de suprapuneri între sub-looks.
- Explorarea a noi domenii ale prelucrării semnalelor în care tehnicile de învățare profundă ar putea fi aplicate, de exemplu, estimarea vântului bazată pe date SAR pentru observarea Pământului.

Bibliografie

- [1] Ba, J. L., Kiros, J. R., and Hinton, G. E. (2016). Layer Normalization.
- [2] Bechter, J., Rameez, M., and Waldschmidt, C. (2017). Analytical and experimental investigations on mitigation of interference in a dbf mimo radar. *IEEE Transactions on Microwave Theory and Techniques*, 65(5):1727–1734.
- [3] Brooker, G. M. (2007). Mutual interference of millimeter-wave radar systems. *IEEE Transactions on Electromagnetic Compatibility*, 49(1):170–181.
- [4] Buus Hinkler, J., Wulf, T., Stokholm, A. R., Korosov, A., Saldo, R., and Pedersen, L. T. (2022). AI4Arctic Sea Ice Challenge Dataset. *Technical University of Denmark*.
- [5] Cai, Z., Fan, Q., Feris, R. S., and Vasconcelos, N. (2016). A unified multi-scale deep convolutional neural network for fast object detection. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, pages 354–370. Springer.
- [6] Carter, N. A., Dawson, J., Joyce, J., Ogilvie, A., and Weber, M. (2018). Arctic Corridors and Northern Voices: Governing Marine Transportation in the Canadian Arctic (Pond Inlet, Nunavut Community Report).
- [7] Dhillon, A. and Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, 9(2):85–112.
- [8] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In *Proceedings of ICLR*.
- [9] Fan, W., Zhou, F., Tao, M., Bai, X., Rong, P., Yang, S., and Tian, T. (2019). Interference Mitigation for Synthetic Aperture Radar Based on Deep Residual Network. *Remote Sensing*, 11(14):1654.
- [10] Fuchs, J., Dubey, A., Lübke, M., Weigel, R., and Lurz, F. (2020). Automotive Radar Interference Mitigation using a Convolutional Autoencoder. In *Proceedings of RADAR*.
- [11] Han, S., Pool, J., Tran, J., and Dally, W. (2015). Learning both weights and connections for efficient neural network. *Proceedings of NIPS*, 28:1135–1143.
- [12] He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.

- [13] Jiao, L., Tang, X., Hou, B., and Wang, S. (2015). Sar images retrieval based on semantic classification and region-based similarity measure for earth observation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(8):3876–3891.
- [14] Kim, G., Mun, J., and Lee, J. (2018). A Peer-to-Peer Interference Analysis for Automotive Chirp Sequence Radars. *IEEE Transactions on Vehicular Technology*, 67(9):8110–8117.
- [15] Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of ICLR*.
- [16] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of NIPS*, pages 1097–1105.
- [17] Kunert, M. (2012). The EU project MOSARIM: A general overview of project objectives and conducted work. In *Proceedings of EuRAD*.
- [18] Kunert, M., Bodereau, F., Goppelt, M., Fischer, C., John, A., Wixforth, T., Ossowska, A., Schipper, T., and Pietsch, R. (2010). D1.5 - Study on the state-of-the-art interference mitigation technique, MOre Safety for All by Radar Interference Mitigation (MOSARIM) project. Technical report, Robert Bosch GmbH.
- [19] Laghezza, F., Jansen, F., and Overvest, J. (2019). Enhanced Interference Detection Method in Automotive FMCW Radar Systems. In *Proceedings of IRS*, pages 1–7.
- [20] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal Loss for Dense Object Detection. In *Proceedings of ICCV*, pages 2980–2988.
- [21] Liu, Z., Sun, M., Zhou, T., Huang, G., and Darrell, T. (2019). Rethinking the value of network pruning. *Proceedings of ICLR*.
- [22] Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of CVPR*, pages 3431–3440.
- [23] Maas, A. L., Hannun, A. Y., and Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of WDLASL*.
- [24] Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P. (2016). Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2):645–657.
- [25] Mun, J., Ha, S., and Lee, J. (2020). Automotive Radar Signal Interference Mitigation Using RNN with Self Attention. In *Proceedings of ICASSP*, pages 3802–3806.
- [26] Mun, J., Kim, H., and Lee, J. (2018). A Deep Learning Approach for Automotive Radar Interference Mitigation. In *Proceedings of VTC-Fall*.
- [27] Neemat, S., Krasnov, O., and Yarovoy, A. (2018). An interference mitigation technique for FMCW radar using beat-frequencies interpolation in the STFT domain. *IEEE Transactions on Microwave Theory and Techniques*, 67(3):1207–1220.
- [28] Noh, H., Hong, S., and Han, B. (2015). Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1520–1528.

- [29] Rawat, W. and Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, 29(9):2352–2449.
- [30] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on computer vision and pattern recognition*, pages 779–788.
- [31] Ren, Y., Xu, H., Liu, B., and Li, X. (2020). Sea Ice and Open Water Classification of SAR Images Using a Deep Learning Model. In *Proceedings of IGARSS*, pages 3051–3054. IEEE.
- [32] Ristea, N.-C., Anghel, A., and Datcu, M. (2023a). Sea Ice Segmentation From SAR Data by Convolutional Transformer Networks. In *Proceedings of IGARSS*, pages 168–171. IEEE.
- [33] Ristea, N.-C., Anghel, A., Datcu, M., and Chapron, B. (2022). Guided Deep Learning by Subaperture Decomposition: Ocean Patterns from SAR Imagery. In *Proceedings of IGARSS*, pages 6825–6828. IEEE.
- [34] Ristea, N.-C., Anghel, A., Datcu, M., and Chapron, B. (2023b). Guided Unsupervised Learning by Subaperture Decomposition for Ocean SAR Image Retrieval. *Proceedings of Transactions on Geoscience and Remote Sensing*.
- [35] Ristea, N.-C., Anghel, A., and Ionescu, R. T. (2020). Fully Convolutional Neural Networks for Automotive Radar Interference Mitigation. In *Proceedings of VTC-Fall*.
- [36] Ristea, N.-C., Anghel, A., and Ionescu, R. T. (2021a). Estimating the magnitude and phase of automotive radar signals under multiple interference sources with fully convolutional networks. *IEEE Access*, 9:153491–153507.
- [37] Ristea, N.-C., Anghel, A., Ionescu, R. T., and Eldar, Y. C. (2021b). Automotive radar interference mitigation with unfolded robust pca based on residual overcomplete auto-encoder blocks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3209–3214.
- [38] Ristea, N.-C., Miron, A.-I., Savencu, O., Georgescu, M.-I., Verga, N., Khan, F. S., and Ionescu, R. T. (2023c). Cytran: A cycle-consistent transformer with multi-level consistency for non-contrast to contrast ct translation. *Neurocomputing*, 538:126211.
- [39] Rock, J., Roth, W., Toth, M., Meissner, P., and Pernkopf, F. (2021). Resource-Efficient Deep Neural Networks for Automotive Radar Interference Mitigation. *IEEE Journal of Selected Topics in Signal Processing*, 15(4):927–940.
- [40] Rock, J., Toth, M., Messner, E., Meissner, P., and Pernkopf, F. (2019). Complex Signal Denoising and Interference Mitigation for Automotive Radar Using Convolutional Neural Networks. In *Proceedings of FUSION*.
- [41] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of MICCAI*, pages 234–241. Springer.
- [42] Sharma, N., Jain, V., and Mishra, A. (2018). An analysis of convolutional neural networks for image classification. *Procedia Computer Science*, 132:377–384.
- [43] Stewart, R. H. (1988). Seasat: results of the mission. *Bulletin of the American Meteorological Society*, 69(12):1441–1447.

- [44] Uysal, F. (2019). Synchronous and asynchronous radar interference mitigation. *IEEE Access*, 7:5846–5852.
- [45] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In *Proceedings of NIPS*, pages 5998–6008.
- [46] Wang, W. and Shen, J. (2017). Deep visual attention prediction. *IEEE Transactions on Image Processing*, 27(5):2368–2378.
- [47] Woo, S., Park, J., Lee, J.-Y., and So Kweon, I. (2018). CBAM: Convolutional Block Attention Module. In *Proceedings of ECCV*, pages 3–19.
- [48] Xu, Z. and Shi, Q. (2018). Interference Mitigation for Automotive Radar Using Orthogonal Noise Waveforms. *IEEE Geoscience and Remote Sensing Letters*, 15(1):137–141.