

**Monitorizarea calității somnului folosind semnale EEG și arhitecturi
convoluționale 2D cu atenție**

Conducător lucrare:
Conf. dr. ing. Anamaria Rădoi

Student:
Mitrea Bogdan-Gabriel

Cuprins

Rezumat	3
2 Introducere	4
2.1 Metode în literatură	5
3 Metoda propusă	7
3.1 Preprocesarea semnalelor EEG	7
3.2 Structura clasei Dataset și generarea spectrogramelor	8
3.3 Modelul propus	10
4 Experimente	11
4.1 Arhitectura CNN inițială	11
4.2 Îmbunătățiri	11
4.3 Rezultat final	13
5 Concluzii	14
Bibliografie	15

Rezumat

Adnotarea manuală a fazelor somnului pe baza înregistrărilor EEG este un proces consumator de timp și predispus la subiectivitate între tehnicieni de somn diferiți. Această lucrare propune o metodă automată de clasificare a stadiilor somnului utilizând o arhitectură de tip CNN bidimensională, îmbunătățită cu un mecanism de atenție spațială, având ca input spectrograme generate din semnale EEG preprocesate și segmentate. Modelul a fost antrenat și evaluat pe un set de date real, cu segmentare per pacient și ponderare a claselor dezechilibrate. S-a obținut o acuratețe de 83.87%, demonstrând capacitatea rețelei de a distinge între cele cinci stadii ale somnului. Rezultatele sugerează că modelul propus poate reprezenta un pas important către sisteme de analiză automată a somnului în aplicații clinice.

Cuvinte cheie: EEG, somn, spectrogramă, rețele neuronale convoluționale, mecanism de atenție, clasificare automată

Capitolul 2

Introducere

Somnul este un proces fiziologic vital, esențial pentru sănătatea și bunăstarea organismului. O calitate a somnului scăzută, în consecință, este corelată cu un risc mai ridicat de boli cardiace, și cu o productivitate redusă - persoanele care au scoruri scăzute ale calității somnului raportează oboseală pe parcursul zilei și dificultăți de concentrare.

Din punct de vedere *neurofiziologic*, somnul este structurat în mai multe stadii care reflectă diferite niveluri de activitate cerebrală. De regulă, un ciclu complet al somnului este împărțit în cinci faze diferite:

- **Faza W (Wake)** - reprezintă starea de veghe, în care pacientul este conștient și activ, fiind caracterizată de cea mai mare activitate cerebrală.
- **Somnul NREM (Non-Rapid Eye Movement)** - subdivizat la rândul său în trei stadii:
 - **Stadiul N1** - marchează tranziția de la stadiul de veghe la somn; se poate reveni ușor la starea de veghe.
 - **Stadiul N2** - stadiu mai profund, ce face trecerea spre *deep sleep*, activitatea la frecvențe mai ridicate diminuându-se.
 - **Stadiul N3** - stadiul *deep sleep*, extrem de important în regenerarea organismului, în care pacientul prezintă cea mai scăzută activitate cerebrală.
- **Somnul REM (Rapid Eye Movement)** - asemănătoare stării de veghe, este caracterizat de mișcări rapide ale ochilor. În această fază au loc majoritatea viselor.

Fazele somnului au un impact mare în determinarea calității sale. Metrici precum procentul fazei W de-a lungul somnului, numărul de treceri din stadiul N1 în starea de veghe, procentul somnului adânc, contribuie în calcularea calității somnului. Pentru analiza acestor faze, un instrument frecvent utilizat este electroencefalograma (EEG).

Electroencefalograma (EEG) reprezintă un test ce monitorizează activitatea electrică a creierului. Această procedură presupune montarea unei căști EEG cu diferite sisteme de amplasare a electrozilor (sistemul 10-20, 10-10 sau 10-5, cu un număr diferit de electrozi). În cadrul acestui studiu, am folosit sistemul 10-20, în configurația standard de 21 de electrozi, amplasați uniform pe zonele cerebrale.

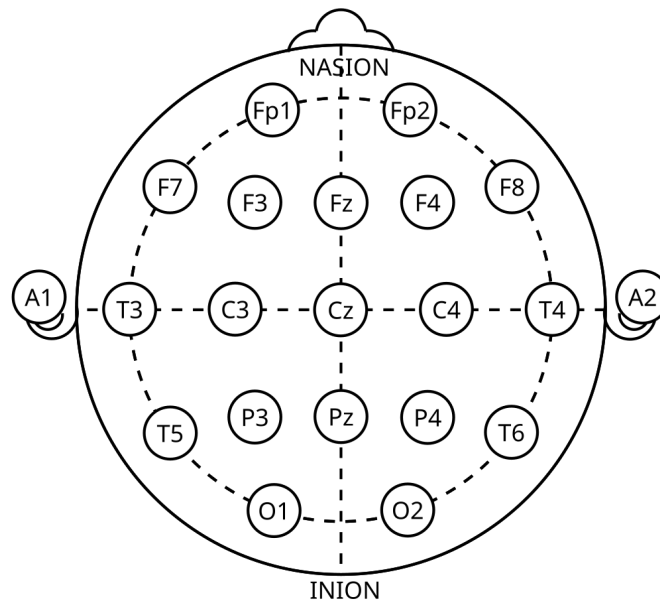


Figura 2.1: Sistemul internațional 10–20 pentru plasarea electrozilor EEG [1]

În mod tradițional, analiza somnului este realizată **manual** de către un tehnician. Acesta se uită pe înregistrările EEG de-a lungul somnului subiecților, și adnotează, pe ferestre a câte 30 de secunde, stadiul în care se află aceștia. Această procedură are dezavantaje clare, precum:

- **consumă foarte mult timp** - poate dura câteva ore;
- **intervine eroarea umană** - fiind subiectivă, poate varia între scoreri;
- necesită un **personal specializat**.

Din aceste motive, este nevoie de o metodă automată de stabilire a fazelor somnului. Odată cu dezvoltarea inteligenței artificiale, în ramurile medicinei au fost dezvoltate metode automate bazate pe Machine Learning și Deep Learning, utilizând modele CNN, RNN sau combinații ale acestora.

În cadrul acestei lucrări, propunem o arhitectură CNN 2D îmbunătățită cu mecanisme de atenție, utilizând spectrograme generate din semnale EEG. Spectrogramele reprezintă o metodă eficientă de vizualizare și de antrenare deoarece surprind informația esențială atât în domeniul timp, cât și în domeniul frecvență.

2.1 Metode în literatură

În ultimii ani, a fost abordată din ce în ce mai des problema clasificării stadiilor somnului folosind semnale EEG și metode de învățare automată. Setul de date **ANPHY-Sleep** [2] se numără printre cele mai recente contribuții în acest sens, oferind cercetătorilor acces la înregistrări EEG de înaltă densitate (înregistrări pe 83 de canale EEG). Acesta permite dezvoltarea și testarea de algoritmi pentru adnotarea automată a somnului în medii controlate.

O altă direcție întâlnită în literatura de specialitate este reprezentarea semnalului EEG sub formă de spectrogramă, obținută prin aplicarea Transformatei Fourier de tip STFT. În [3], a fost pro-

pusă o metodă de clasificare a episoadelor **epileptice** ale pacienților utilizând un model CNN antrenat pe spectrograme, cu rezultate de peste 98% acuratete. Deși aplicată într-un context clinic diferit, abordarea evidențiază avantajele transformării semnalului în reprezentări bidimensionale.

În [4], este propus un model *CNN + Bi-LSTM* pentru clasificarea stadiilor somnului la intrarea în rețeaua neurală. Modelul este evaluat pe mai multe seturi de date consacrate și reușește să îmbunătățească scorurile obținute în special pentru clasa N1, ce prezintă dificultăți în a fi corect clasificată.

De asemenea, conform [5], utilizând același set de date s-au investigat caracteristicile frecvențelor în tranziția de la somn la starea de veghe. Se observă o creștere marcată a frecvențelor înalte în EEG, creștere variabilă în timp și variabil distribuită în regiunile cerebrale. Rezultatele obținute confirmă utilitatea utilizării reprezentării spectrale în contextul semnalelor EEG.

În cadrul lucrării [6], autorii investighează biomarkeri EEG oscilatorii și non-oscilatorii în somnul NREM, utilizați pentru a identifica zone epileptogene la pacienți cu epilepsie focală. În [7], este analizată interacțiunea bidirecțională dintre epilepsie și somn, de la nivel molecular la rețele neuronale, subliniindu-se rolul oscilațiilor lente în reglarea excitabilității cerebrale, evidențiind relevanța EEG ca instrument de investigare a acestor procese.

Capitolul 3

Metoda propusă

Acest capitol detaliază pașii urmați în vederea implementării metodei propuse, cuprinzând pre-procesarea semnalelor EEG, generarea spectrogramelor folosite în antrenare, și construirea și antrenarea modelului CNN cu mecanism de atenție spațială.

3.1 Preprocesarea semnalelor EEG

Setul de date ales **OSF - ANPHY-Sleep** [8] conține înregistrări de tip *high-density EEG*, cu o configurație de 83 de canale de electrozi, de la 29 de pacienți adulți, sănătoși, cu vârste cuprinse între 18 și 45 de ani, ce oferă o posibilitate de analiză profundă a fiziologiei somnului la o rezoluție ridicată.

Înregistrările din această bază de date sunt în format *.edf*, și reprezintă semnalele EEG *nealterate* ale fiecărui pacient pe durata întregii nopți, însoțit de un **artifact detection matrix**, menit să semnaleze existența artefactelor de-a lungul înregistrării pe unele canale, și un **fișier de adnotare** întocmit de către un tehnician de somn certificat, pe durate a câte 30 de secunde, exact în modul în care s-ar realiza manual o astfel de analiză.

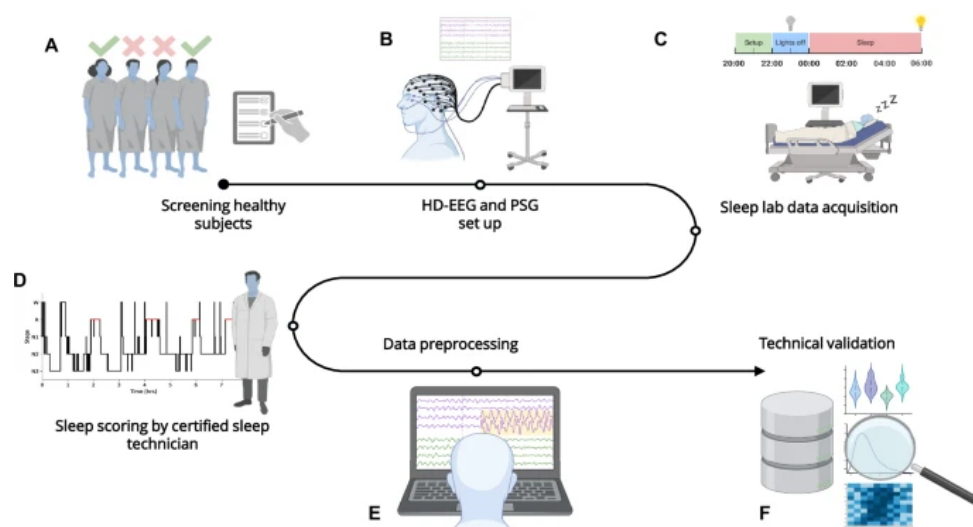


Figura 3.1: Procesul de pregătire a setului de date [2]

Primul pas a fost segmentarea înregistrărilor în fișiere a câte 30 de secunde, pas esențial pentru antrenarea ulterioară și corelarea acestora cu fișierul de adnotare corespunzător fiecărui pacient. Apoi, a fost nevoie de *eliminarea epocilor invalide* (ex. canale corupte, interferențe cu mediul exterior), folosind **artifact detection matrix-ul**. După această selecție au rezultat ~ 40 de epoci / stadiu de somn, deci 200 de epoci per pacient. Per total, setul de date procesat, gata de prelucrare, a avut 5800 de epoci ale căror clase au fost uniform distribuite. În final, segmentele procesate, alături de etichetele corespunzătoare, au fost stocate într-un fișier **.csv** final, pentru încărcare rapidă.

3.2 Structura clasei Dataset și generarea spectrogramelor

Mai departe, am creat clasa EEGDataset standard, necesară pentru antrenare. Aceasta primește fișierul .csv, pipeline-ul de transformări și canalele selectate, pentru a fi posibilă atât o antrenare folosind sistemul 10-20 de amplasare al electrozilor (23 de canale), cât și un sistem cu 83 de canale (mai costisitor din punct de vedere al resurselor și al timpului).

Pe lângă normalizarea semnalului per canal, am aplicat o serie de augmentări relevante în prelucrarea semnalelor EEG, pentru a crește diversitatea datelor de antrenament, și pentru a îmbunătăți generalizarea modelului la variații ne semnificative ale semnalului:

- **RandomNoise** — constă în adăugarea de zgomot gaussian alb la semnal, cu scopul de a simula artefacte minore sau variații de fond. Matematic, semnalul devine:

$$\tilde{x}[n] = x[n] + \epsilon[n], \quad \epsilon[n] \sim \mathcal{N}(0, \sigma^2) \quad (3.1)$$

- **RandomMask** - presupune anularea unor secvențe scurte din semnal prin aplicarea unei măști binare. Această augmentare forțează modelul să devină mai robust la lipsa temporară a informației:

$$\tilde{x}[n] = x[n] \cdot m[n], \quad m[n] \in \{0, 1\} \quad (3.2)$$

- **RandomSignFlip** - simulează o inversare de semn (similară cu o inversare de polaritate, posibilă în EEG real din cauza poziționării electrozilor). Se aplică astfel:

$$\tilde{x}[n] = s \cdot x[n], \quad s \in \{-1, 1\} \quad (3.3)$$

cu o anumită probabilitate ca $s = -1$.

- **RandomTimeReversal** - presupune inversarea completă în timp a unei secvențe de semnal, ceea ce poate ajuta rețeaua să învețe trăsături independente de direcția temporală:

$$\tilde{x}[n] = x[N - 1 - n] \quad (3.4)$$

unde N este lungimea totală a ferestrei analizate.

Apoi, tot aici are loc și generarea spectrogramelor, folosind modulul **scipy.signal** pentru procesare de semnal și analiză spectrală. Spectrogramele sunt generate per canal, cu o frecvență de eșantionare de 1000Hz și cu o rezoluție de 257×257 , cu ajutorul STFT (Short-time Fourier transform). Transformata Fourier de scurtă durată (STFT) este utilizată pentru a analiza conținutul

spectral al unui semnal EEG într-un mod dependent de timp. În cazul discret, pentru un semnal $x[n]$, STFT se calculează astfel:

$$X(t, f) = \sum_{n=0}^{N-1} x[n] \cdot w[n - t] \cdot e^{-j2\pi f n} \quad (3.5)$$

unde t indică poziția ferestrei în timp, iar f este frecvența exprimată în Hz.

Reprezentarea densității spectrale de putere (PSD) a semnalului, forma cea mai frecvent utilizată în vizualizarea spectrogramelor, este de fapt pătratul amplitudinii STFT:

$$\text{spectrogram}\{x(t)\}(\tau, \omega) \equiv |X(\tau, \omega)|^2 \quad (3.6)$$

unde τ și ω corespund notărilor t și f din ecuația (3.5), reprezentând, respectiv, poziția temporală a ferestrei și frecvența analizată.

Spectrogramele canalelor de interes sunt apoi concatenate, forma finală a inputului trimis către antrenare fiind $[n_channels, 257, 257]$. Exemple de spectrograme ale semnalelor EEG corespunzătoare diverselor etape de somn sunt prezentate în Figura 3.2.

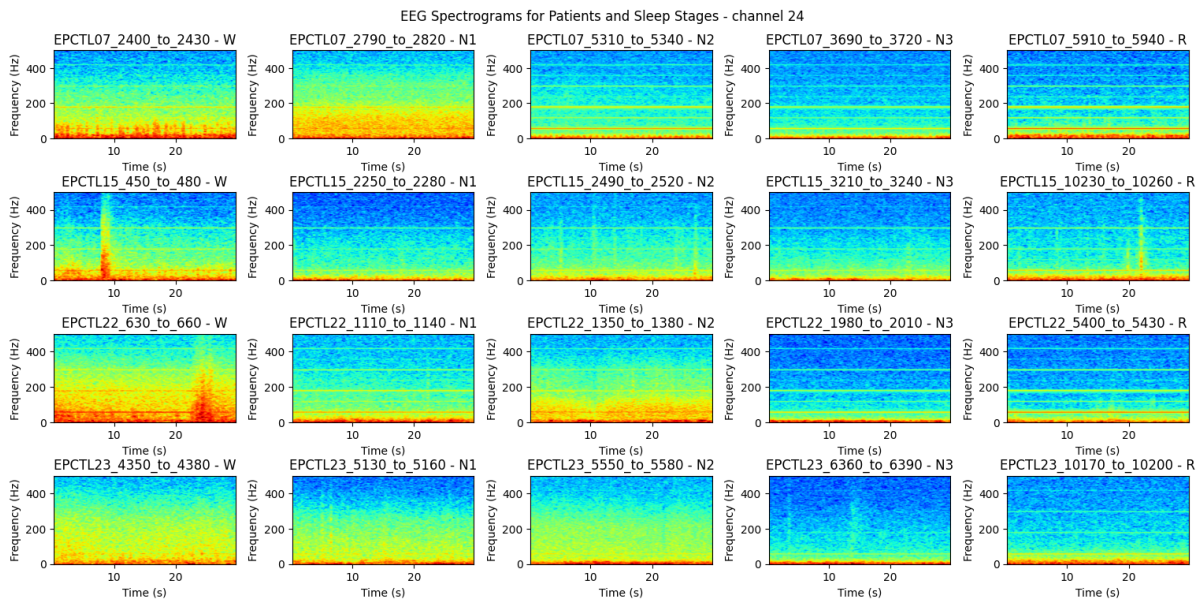


Figura 3.2: Exemplu de prelucrare a semnalului EEG într-o spectrogramă

3.3 Modelul propus

Arhitectura rețelei constă în 3 blocuri convoluționale, fiecare având: Conv2D → BatchNorm → ReLU → MaxPool. Între cele 3 blocuri, am adăugat **skip-connections** de tip ResNet, cu convoluție 1x1, pentru a permite modelului să învețe mai ușor și pentru a nu pierde informație. Structura generală a rețelei convoluționale implementate este prezentată în Figura 3.3.

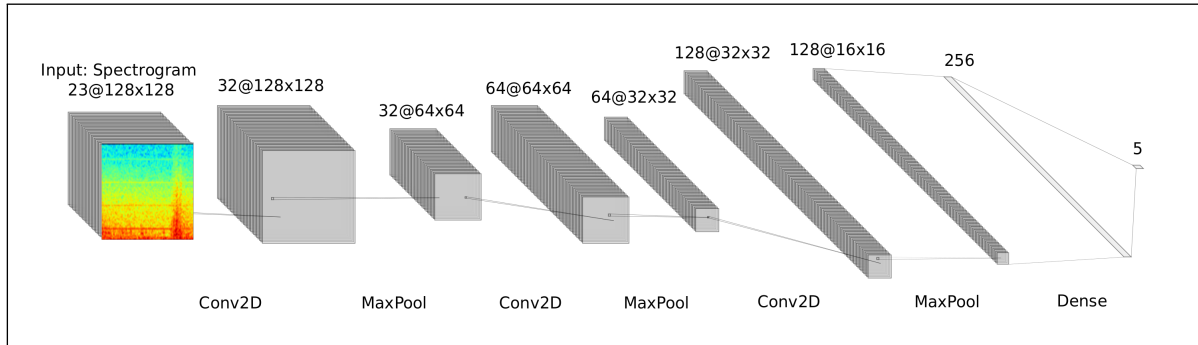


Figura 3.3: Arhitectura rețelei neuronale CNN aplicată pe spectrogramă EEG.

Stratul convoluțional 2D extrage caracteristici din spectrogramă, precum modele temporale sau de frecvență, fiecare astfel de filtru învățând un tip diferit de tipar.

Rolul **batch normalization** existent în fiecare strat este de a stabili distribuția activărilor în timpul antrenării, reducând sensibilitatea la inițializarea greșită a ponderilor și, totodată, riscul de overfitting.

ReLU (Rectified Linear Unit) este o funcție de activare ce accelerează și simplifică antrenarea prin introducerea neliniarității în model. În finalul blocurilor convoluționale se află un strat **MaxPool** ce reduce dimensiunea caracteristicilor spațiale ale reprezentărilor intermediare, prin selectarea valorii maxime din diferite regiuni locale.

A fost adăugat și un mecanism de **atenție spațială**, pentru ca modelul să învețe să acorde mai multă atenție zonelor relevante din spectrograme (ex. frecvențele joase sunt esențiale în fiziologia somnului). Prin acest mecanism, rețeaua este ghidată să evidențieze regiunile informative din reprezentarea spațială a datelor și să diminueze influența celor mai puțin relevante. Acest lucru se traduce prin aplicarea unei hărți de atenție A peste activările intermediare x , obținând:

$$x_{\text{out}} = x \odot A \quad (3.7)$$

unde simbolul $\odot$ reprezintă înmulțirea element cu element între activări și harta de atenție.

Partea finală a rețelei este alcătuită dintr-un strat de Flatten, care transformă activările bidimensionale într-un vector unidimensional, urmat de un strat complet conectat cu 256 de neuroni, ce permite extragerea unei reprezentări dense a caracteristicilor învățate anterior. Pentru a reduce riscul de overfitting, se aplică un mecanism de regularizare de tip dropout cu o rată de 30%. Ultimul strat este unul de ieșire, care aplică funcția Softmax pentru a genera un vector de probabilități corespunzător celor cinci faze ale somnului.

Capitolul 4

Experimente

În această secțiune sunt prezentate experimentele care au condus la rezultatul optim, final, și scorurile obținute în urma antrenării.

În cadrul antrenării, am împărțit pacienții în train/test, păstrând o proporție adecvată - 80% pentru antrenare și 20% pentru testare, și urmărind o abordare user-independent. După transformările aplicate, antrenarea se face cu un batch size de 32, 100 de epoci, utilizând optimizatorul Adam și un learning rate de 0.0001, iar funcția de pierderi aleasă a fost CrossEntropyLoss cu o pondere mai mare pe clasa N1, existând o tendință a modelului de a învăța mai greu această clasă. Clasa N1 este mai greu de diferențiat față de celelalte clase, fapt vizibil și pe spectrogramele din figura 3.2.

Antrenarea a fost realizată pe un GPU NVIDIA T4, disponibil în mediul Google Colab.

4.1 Arhitectura CNN inițială

În prima fază, am pornit de la un model CNN, fără atenție și fără skip-connections, care conținea doar cele 3 blocuri convoluționale Conv2D → ReLU → MaxPool, urmate de Flatten și Softmax. Am obținut o acuratețe de **73.26%**, însă cu un *F1 score* pe clasa **N1** destul de redus - **50%**, și un *recall* mic în dreptul clasei **N2** - **61%**. De asemenea, *precizia* clasei **REM** era doar de **64%**.

Apoi, după stabilirea unei structuri **user-independent** (pacienți diferiți la antrenare față de testare), acuratețea a crescut puțin - **75.05%**, însă problemele au rămas pe aceleași zone.

4.2 Îmbunătățiri

În următoarele experimente, am adăugat, pe rând: batch normalization, skip connections, atenție spectrală, dropout, urmând o serie de ajustări ale ponderilor funcției de pierderi CrossEntropyLoss:

- **Batch normalization** - Acuratețea a continuat să crească, ajungând la un scor de **78.09%**. S-a observat o creștere a *recall*-ului clasei **W** până la **94%**, o foarte ușoară îmbunătățire a

recall-ului clasei **N1**, însă de cele mai vizibile progrese au avut parte clasele **N2** și **REM**, cu o îmbunătățire a *F1 score* de aproximativ **+10%**.

- **Skip connections** - Crește acuratețea generală la **81.02%**, în timp ce este vizibilă o îmbunătățire în mod egal pe toate clasele.
- **Atenție spectrală** - A adus distribuția scorurilor la o formă mai stabilă, cu un *F1 score* pe clasa **N1** de **65%**, deja mult mai bine față de iterațiile anterioare. În ceea ce privește clasa **N2**, scorul final a scăzut ușor cu un procent - **73%**, printr-o scădere a *recall*-ului.
- **Dropout + clase ponderate** - După integrarea mecanismului de *dropout*, au fost adăugate ponderi pe clasele **N1** și **N2**. S-a observat un progres promițător al clasei **N1**, motiv pentru care au fost explorate ponderi diferite pentru acea clasă: **1.4**, **1.8**, **2.2** și **2.4**, și o pondere fixă pentru clasa **N2**, de **1.3**, datorită dificultății moderate de clasificare.

În urma acestui studiu, am ales modelul cu ponderile $[1, 2.2, 1.3, 1, 1]$, ținând cont atât de acuratețea globală optimă, cât și de a alege varianta care are un *F1 score* bun atât pentru clasa **N1**, cât și pentru clasa **N2**.

Tabelul 4.1: Performanța modelului pentru ponderi diferite ale clasei **N1**

N1_weight	Accuracy	F1 score				
		W	N1	N2	N3	REM
1.4	0.8233	0.8648	0.6318	0.7904	0.9197	0.8933
1.8	0.8326	0.8625	0.6692	0.7932	0.9314	0.9050
2.2	0.8387	0.8609	0.6626	0.8196	0.9328	0.9029
2.4	0.8210	0.8269	0.6362	0.8202	0.9302	0.9006

În **tabelul 4.1**, sunt comparate scorurile celor patru seturi de ponderi. Se observă că pentru clasele **W**, **N3** și **REM**, rezultatele sunt comparabile între cel puțin 3 dintre variante. Diferența este făcută în rândul claselor **N1** și în special **N2**, de unde rezultă modelul cel mai performant.

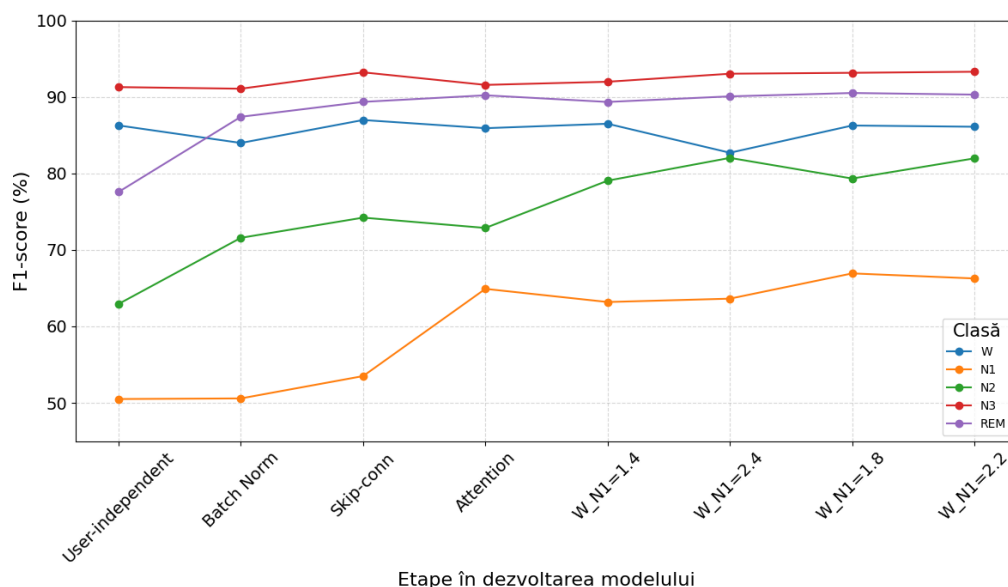


Figura 4.1: Evoluția *F1-score* per clasă în funcție de modificările arhitecturii

În Figura 4.1 se observă distribuția F1 score pe măsură ce s-a adăugat câte un strat convoluțional, pentru a vizualiza componentele cu cel mai mare impact asupra modelului. **Clasa N1** (cea mai dificilă) se îmbunătățește semnificativ după introducerea atenției și a ponderării claselor. De asemenea, **Clasele N3 și REM** au avut performanțe ridicate constant, încă de la primele variante ale modelului.

4.3 Rezultat final

Modelul final rezultat conține cele 3 blocuri convoluționale: Conv2D → BatchNorm → ReLU → MaxPool, la care s-au mai adăugat batch normalization, skip connections, atenție spectrală și pondere mai mare pe clasele **N1 și N2**.

În Tabelul 4.2 este prezentat **raportul complet de clasificare** al modelului propus, ce include valorile de *precision*, *recall* și *F1 score* pentru fiecare clasă, precum și acuratețea globală și mediile macro și ponderate.

Tabelul 4.2: Raport de clasificare al modelului

Clasa	precision	recall	f1-score
W	0.8360	0.8874	0.8609
N1	0.6763	0.6494	0.6626
N2	0.8804	0.7667	0.8196
N3	0.9124	0.9542	0.9328
REM	0.8838	0.9228	0.9029
accuracy			0.8387
macro avg	0.8378	0.8361	0.8358
weighted avg	0.8375	0.8387	0.8370

În Figura 4.2 este ilustrată **matricea de confuzie**, care oferă o perspectivă în detaliu asupra performanței pe fiecare clasă și a erorilor de clasificare.

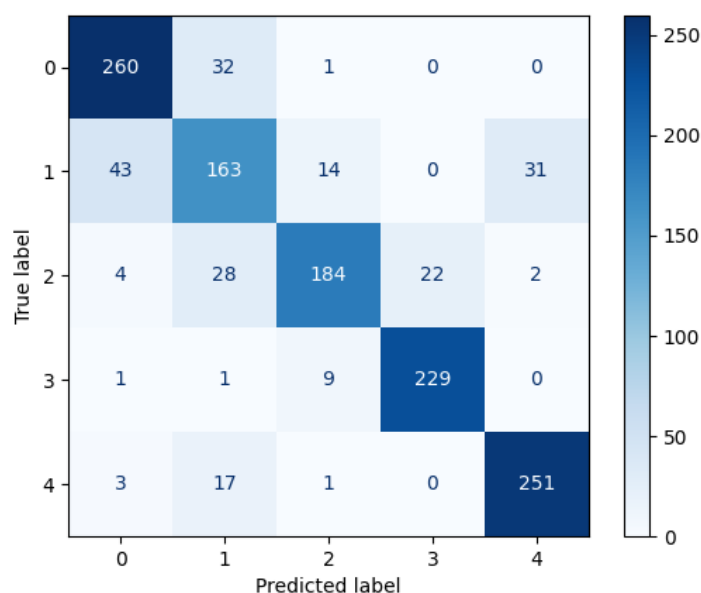


Figura 4.2: Matricea de confuzie a modelului

Capitolul 5

Concluzii

Lucrarea a urmărit dezvoltarea unei metode automate de clasificare a stadiilor somnului, utilizând semnale EEG și arhitecturi convoluționale 2D cu atenție. Modelul propus integrează spectrograme per canal, separare a pacienților și un mecanism de atenție spațială, obținând o acuratețe de 83.87%. Prin analiza detaliată și optimizarea treptată, am evidențiat impactul ponderării dezechilibrului între clase, în special pentru N1, care este, de obicei, cel mai dificil stadiu de detectat. Limitările actuale țin atât de dimensiunea redusă a setului de date, cât și de similaritatea generală între clasa N1 și stadiile vecine, însă metoda propusă oferă o bază solidă pentru extinderea ulterioară a clasificării somnului în contexte clinice reale.

Bibliografie

- [1] Wikimedia Commons, “21 electrodes of International 10–20 system for EEG,” https://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg, 2022.
- [2] X. Wei, T. Avigdor, A. Ho, E. Minato, A. Garcia-Asensi, J. Royer, Y. L. Wang, V. Travnicek, K. Schiller, B. C. Bernhardt, and B. Frauscher, “ANPHY-Sleep: an Open Sleep Database from Healthy Adults Using High-Density Scalp Electroencephalogram,” *Scientific Data*, vol. 11, p. 896, 2024. [Online]. Available: <https://www.nature.com/articles/s41597-024-03722-1>
- [3] S. Bourouis, M. Alharbi, and N. El Hadadi, “An automated classification of EEG signals based on spectrogram and CNN for epilepsy diagnosis,” *2021 International Conference on Intelligent Systems (IS)*, pp. 1–6, 2021.
- [4] N. Ghaboosi, A. Mansouri, M. Sadeghi, and S. Adabi, “A Deep Learning Method Approach for Sleep Stage Classification with EEG Spectrogram,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 10, p. 6322, 2022.
- [5] T. Avigdor, G. Ren, C. Abdallah, F. Dubeau, C. Grova, and B. Frauscher, “The awakening brain is characterized by a widespread and spatiotemporally heterogeneous increase in high frequencies,” *Advanced Science*, p. e2409608, 2025.
- [6] V. Latreille, J. Corbin-Lapointe, L. Peter-Derex, J. Thomas, J. Lina, and B. Frauscher, “Oscillatory and nonoscillatory sleep electroencephalographic biomarkers of the epileptic network,” *Epilepsia*, vol. 65, no. 10, pp. 3038–3051, 2024.
- [7] L. Sheybani, B. Frauscher, C. Bernard, and M. Walker, “Mechanistic insights into the interaction between epilepsy and sleep,” *Nature Reviews Neurology*, vol. 21, no. 4, pp. 177–192, 2025.
- [8] ANPHY Lab, “ANPHY-Sleep Dataset on OSF,” <https://osf.io/r26fh/>, 2025.