

RAPORT

Cerinta 1.7 Tema 1:

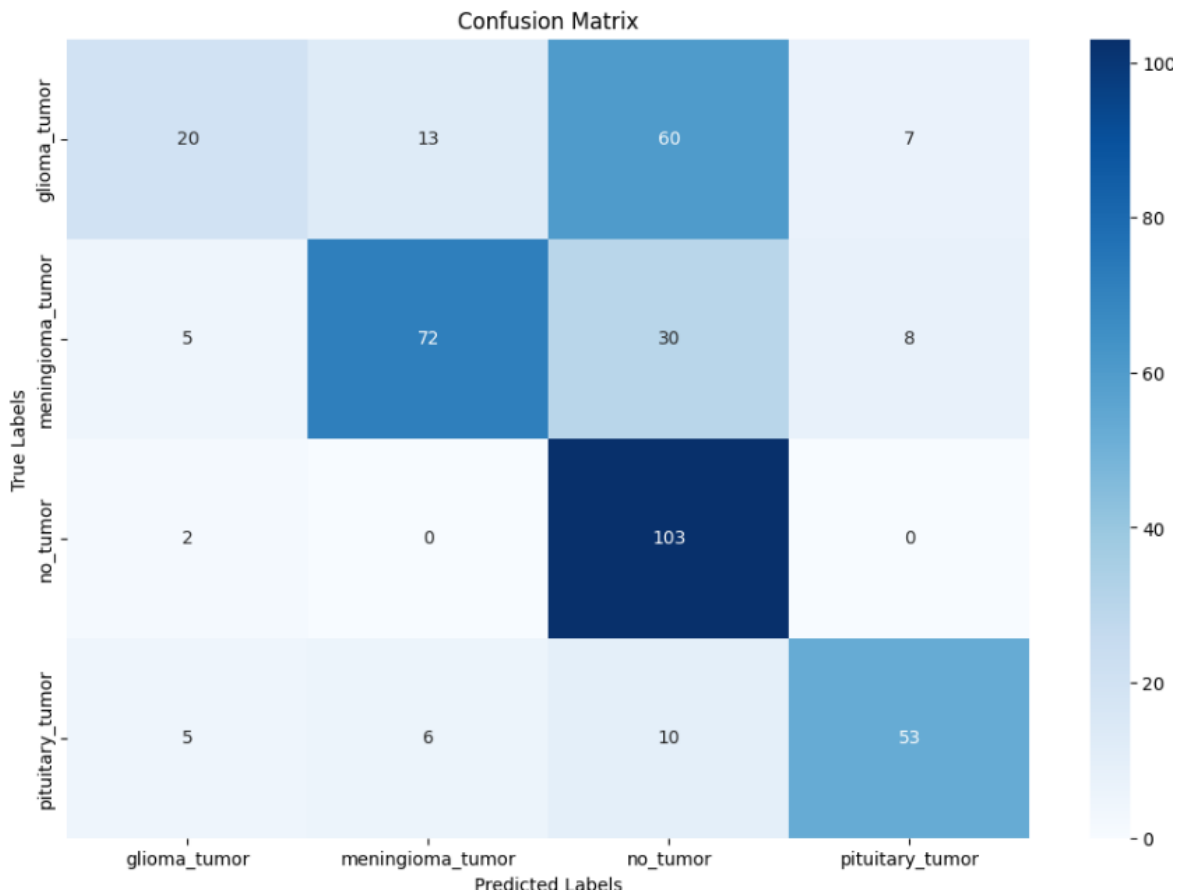
1. Incarcarea datelor: Am folosit functia de la ex anterior pt a imparti datele in validare si antrenare
2. Definirea modelului: Modelul de convolutie folosit este unul simplu, preluat din laboratorul de Notiuni Fundamentale de ml. Avem:
 1. **Două straturi convoluționale:** Acestea extrag caracteristici importante din imagini. Primul strat are 16 filtre, iar al doilea are 32 de filtre. Ambele folosesc o dimensiune a kernelului de 3x3 și padding de 1 pentru a păstra dimensiunile imaginii.
 2. **Două straturi de pooling maxim:** Acestea reduc dimensiunile imaginii și extrag caracteristicile cele mai importante.
 3. **Două straturi complet conectate:** Acestea clasifică imaginea în una dintre cele 4 categorii de tumori. Primul strat are 128 de neuroni, iar al doilea are 4 neuroni (corespunzător celor 4 clase).
 4. **Funcții de activare ReLU:** Acestea introduc neliniaritate în model.
3. Definirea functiei de pierdere si a optimizatorului: Functii preluate din laboratorul de Notiuni Fundamentale de ml. Am folosit funcția de eroare **CrossEntropyLoss**, deoarece aceasta calculează diferența dintre distribuția de probabilitate prezisă de model și distribuția reală a claselor. Optimizatorul folosit este **Adam**, care este un algoritm de optimizare bazat pe gradient.
4. Antrenarea modelului: Modelul de antrenare a fost preluat din laboratorul de Notiuni Fundamentale de ml (functia def Train_Eval). Pt rata de invatare am folosit un lr scheduler **ReduceLROnPlateau**, pentru a reduce rata de învățare atunci când pierderea de pe setul de validare nu se mai îmbunătățește.
5. Evaluarea modelului: tot preluata din laborator,iar **numarul de epoci este 10**

Classification Report:

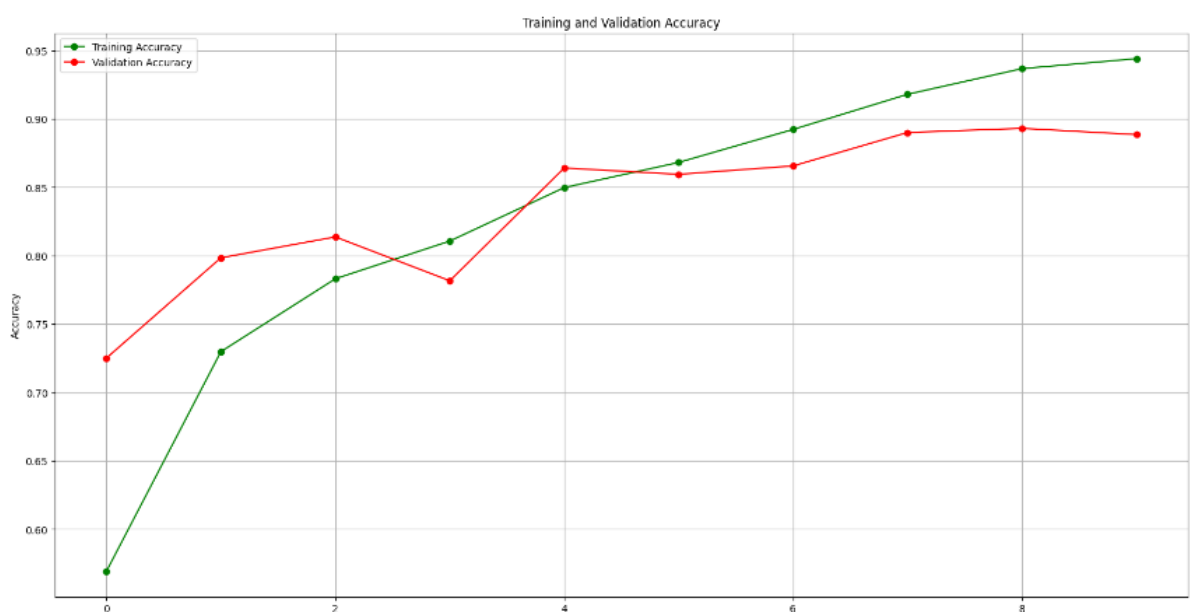
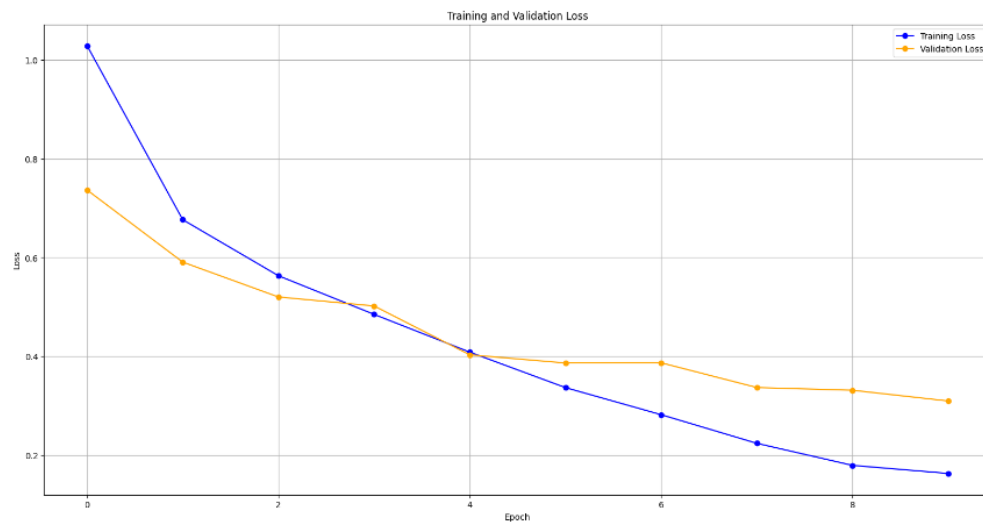
Classification Report:				
	precision	recall	f1-score	support
glioma_tumor	0.62	0.20	0.30	100
meningioma_tumor	0.79	0.63	0.70	115
no_tumor	0.51	0.98	0.67	105
pituitary_tumor	0.78	0.72	0.75	74
accuracy			0.63	394
macro avg	0.68	0.63	0.60	394
weighted avg	0.67	0.63	0.60	394

Se poate observa pe raportul de clasificare o acuratețe de aproximativ 0.63, ce este destul de bună bazată pe nr de imagini date. Precision-ul arată procentajul cu care a prezis algoritmul fiecare clasă, recall-ul măsoară proporția de exemple corect clasificate pentru o anumită clasă din totalul exemplurilor reale din acea clasă.

Confusion Matrix:



Aici se poate observa mai bine cât de bine a prezis algoritmul fiecare tip de tumoră, pe diagonală fiind cele prezise corect, cea mai bine prezisă a fost cea de `no_tumor`, apoi cea de `meningioma_tumor`, apoi `pituitary` și la final `glioma`.



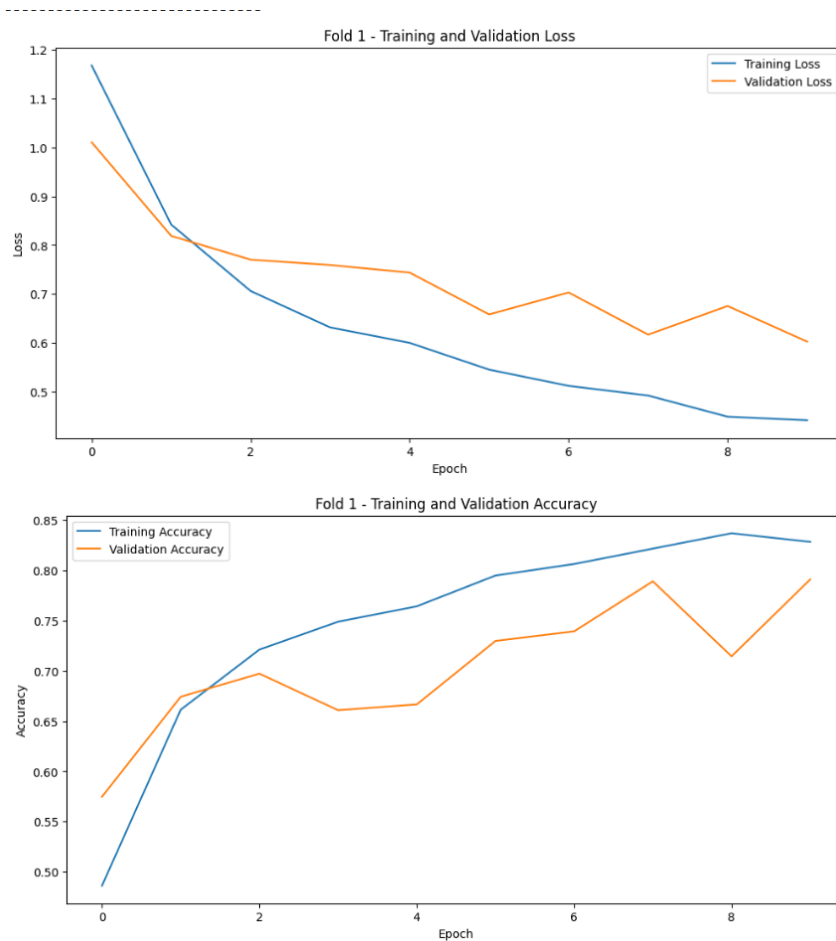
Se poate observa faptul ca odata cu trecerea la o epoca mai avansata acuratetea e din ce in ce mai mare, iar loss-ul din ce in ce mai mic, si pe validation si pe training, ceea ce inseamna ca nu se face overfitting.

CERINTA 1:

Am implementat metoda de impartire a datelor care sa distribuie echitabil exemplele intre cele k fold-uri, pastrand proportiile claselor din setul de date original, utilizand k-fold cross-validation, utilizand biblioteca **StratifiedKFold**.

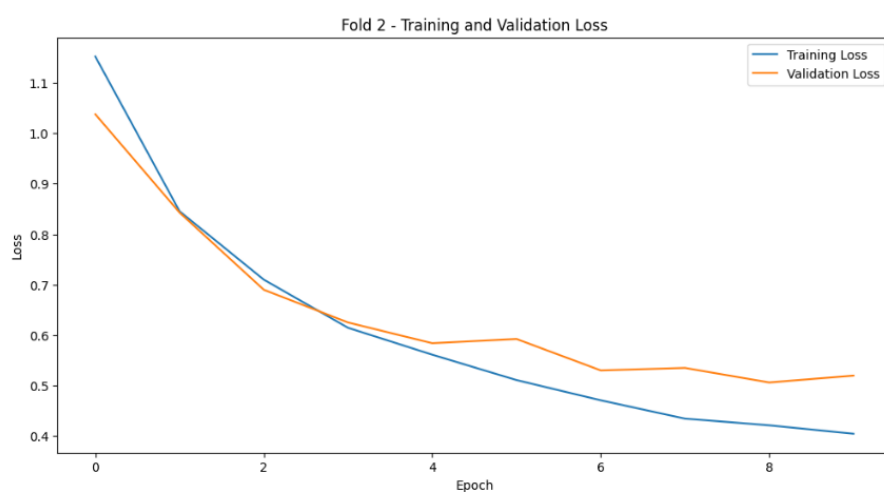
Am impartit in 5 fold-uri, iar pt fiecare fold am folosit acelasi model de antrenare implementat la 1.7 (functia Train_Eval)

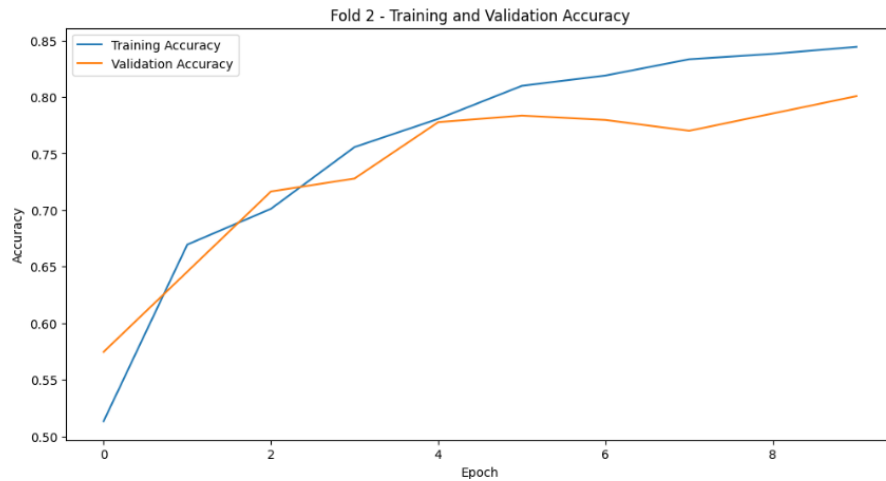
Fold-ul 1:



Pentru primul fold, graficele arata ok, nu se face overfitting deoarece pt loss training-ul si validarea scad, iar la acc training-ul si validarea cresc.

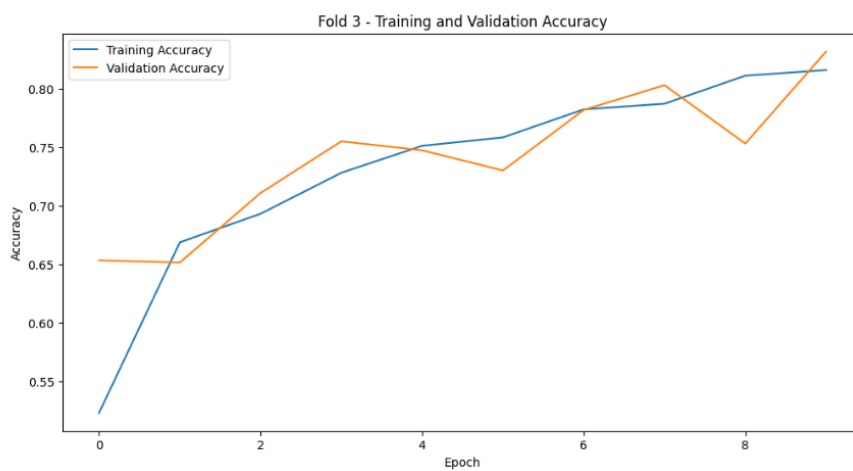
Fold-ul 2:





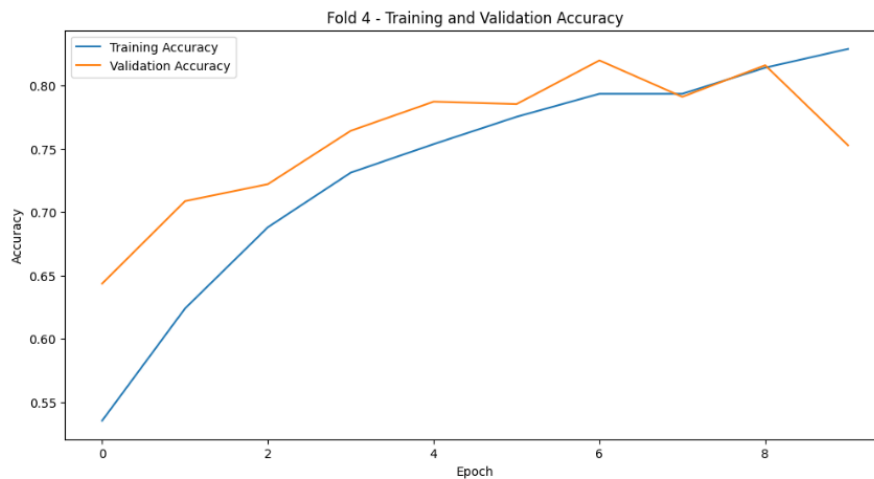
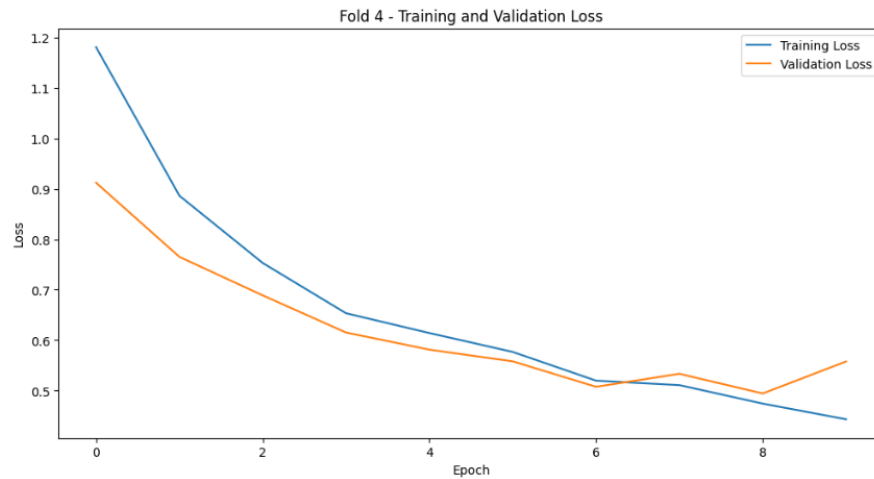
Pentru al 2lea fold, graficele arata ok, nu se face overfitting deoarece pt loss training-ul si validarea scad, iar la acc training-ul si validarea cresc.

Fold-ul 3:



Pentru al 3lea fold, graficele arata ok, nu se face overfitting deoarece pt loss training-ul si validarea scad, iar la acc training-ul si validarea cresc.

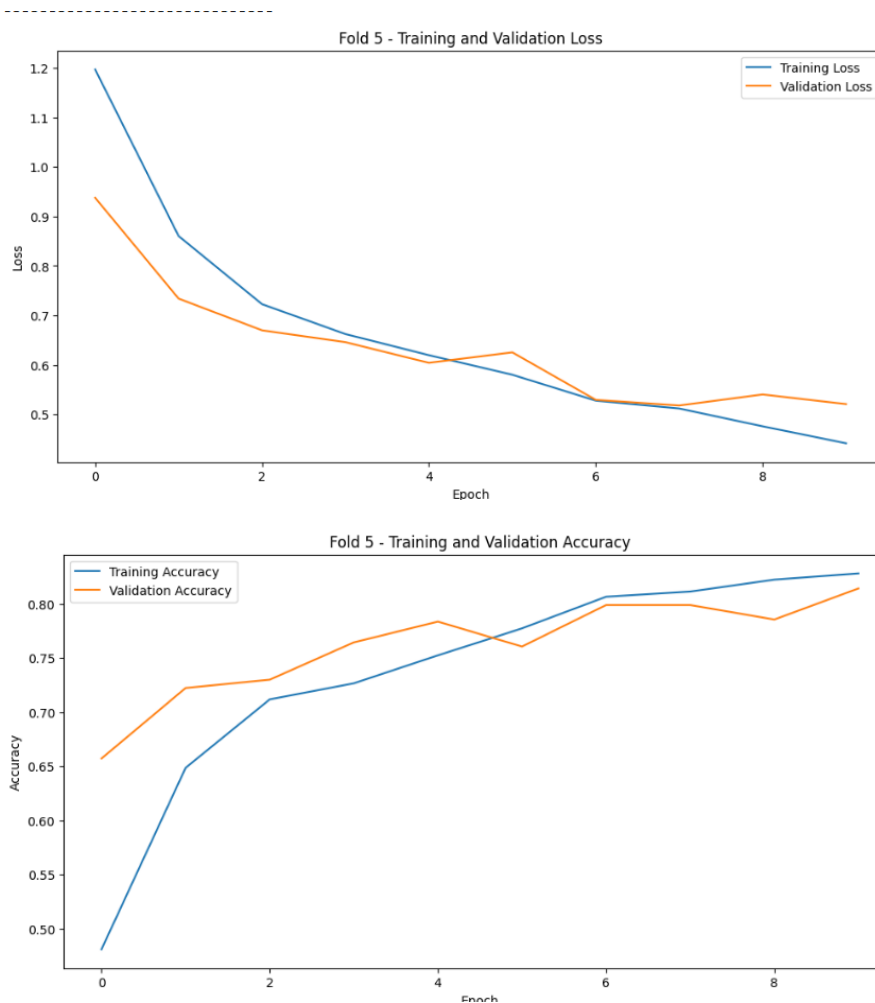
Fold-ul 4:



Fold 5/5

Pentru al 4lea fold, graficele arata ok, nu se face overfitting deoarece pt loss training-ul si validarea scad, iar la acc training-ul si validarea cresc.

Fold-ul 5:



Pentru al 5lea fold, graficele arata ok, nu se face overfitting deoarece pt loss training-ul si validarea scad, iar la acc training-ul si validarea cresc.

- Dupa finalizarea tuturor fold-urilor:

```
Precision - Mean: 0.5419, Std: 0.0236
Recall - Mean: 0.4426, Std: 0.0402
F1_score - Mean: 0.4016, Std: 0.0449
Accuracy - Mean: 0.4426, Std: 0.0402
```

Precision: modelul a fost in medie 54% precis la identificarea corecta in toate foldurile.

Deviatia standard este mica -> variabilitate intre fold-uri mica, model consistent.

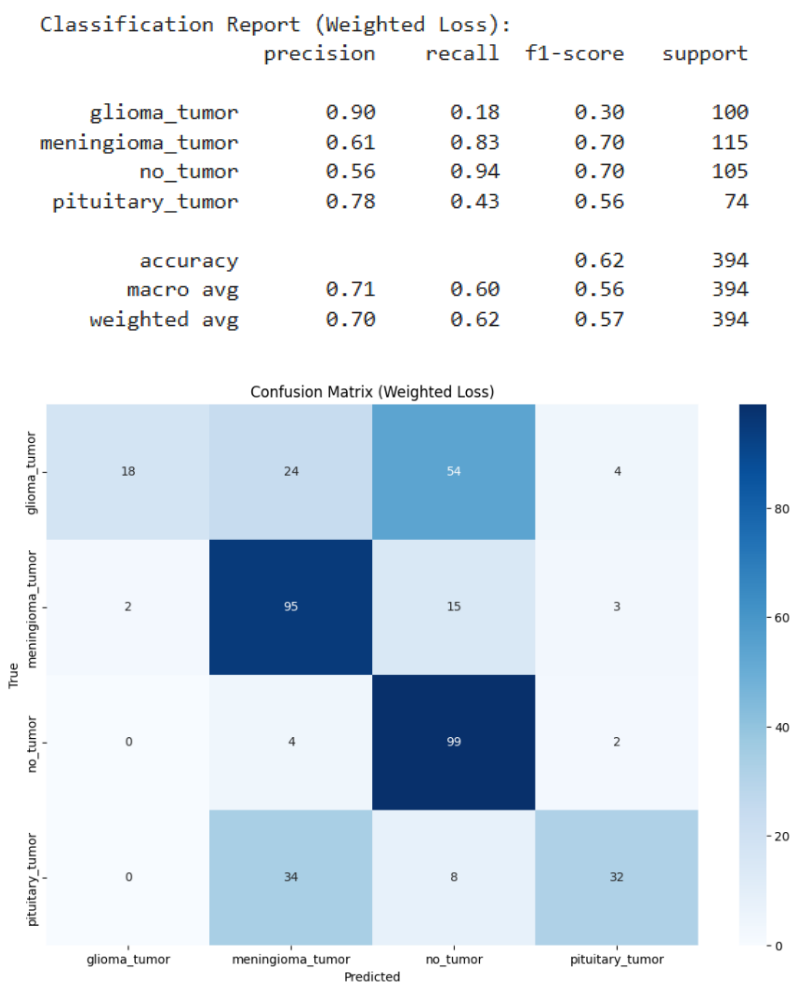
Recall: modelul a identificat corect in medie 44.26% dintre exemplele.

F1-score: performanta modelul mai scazuta

Accuracy: modelul are in medie un accuracy de 44% pe toate fold-urile.

Cerinta 2:

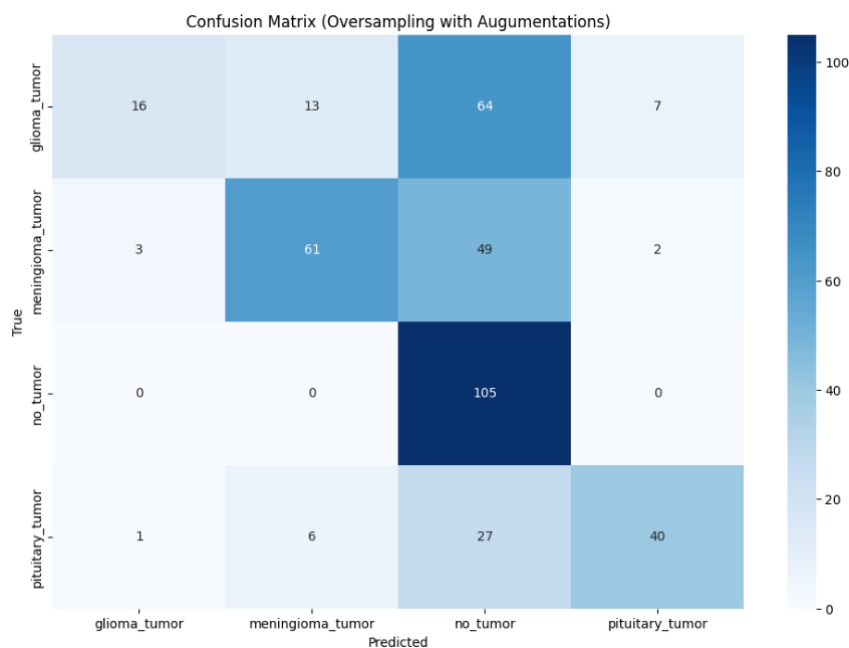
Evaluarea Functiei de Pierdere cu Ponderi:



Cu functia de weighted loss are o performanta egala cu cea normala (da rasta cred ca e din cauza modelul care si el are deja implementat o functie de loss, deci presupun ca nu fac o diferenta asa de mare)

Evaluare Oversampling cu augmentari

Classification Report (Oversampling with Augmentations):				
	precision	recall	f1-score	support
glioma_tumor	0.80	0.16	0.27	100
meningioma_tumor	0.76	0.53	0.63	115
no_tumor	0.43	1.00	0.60	105
pituitary_tumor	0.82	0.54	0.65	74
accuracy			0.56	394
macro avg	0.70	0.56	0.54	394
weighted avg	0.69	0.56	0.53	394



Modelul este mai prost decat cel fara oversamplingcu augmentari, iar acest lucru poate avea mai multe motive:

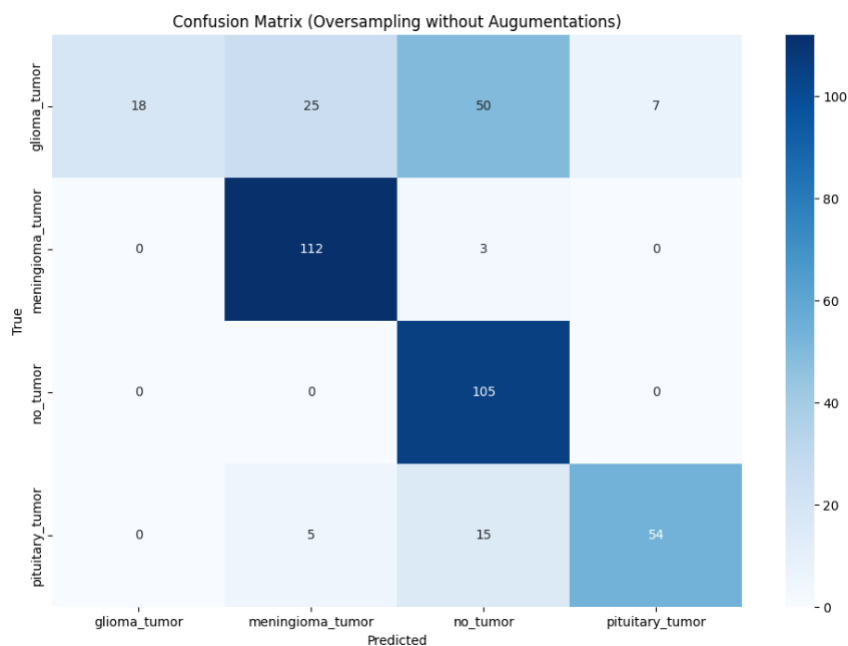
1. Augmentări nepotrivite/prea multe/adaugare prea mult de zgomot
2. Am facut o supra antrenare deoarece am augmentat datele, modelul ar putea supra-învăța pe aceste date augmentate specifice, ceea ce duce la o performanță mai slabă pe date noi și nevăzute.

Dar, at the same time, avem o acuratete maxima pe no_tumor, chiar daca celelalte merg mai prost.

Evaluare Oversampling fara augmentari:

Classification Report (Oversampling without Augmentations):

	precision	recall	f1-score	support
glioma_tumor	1.00	0.18	0.31	100
meningioma_tumor	0.79	0.97	0.87	115
no_tumor	0.61	1.00	0.76	105
pituitary_tumor	0.89	0.73	0.80	74
accuracy			0.73	394
macro avg	0.82	0.72	0.68	394
weighted avg	0.81	0.73	0.68	394

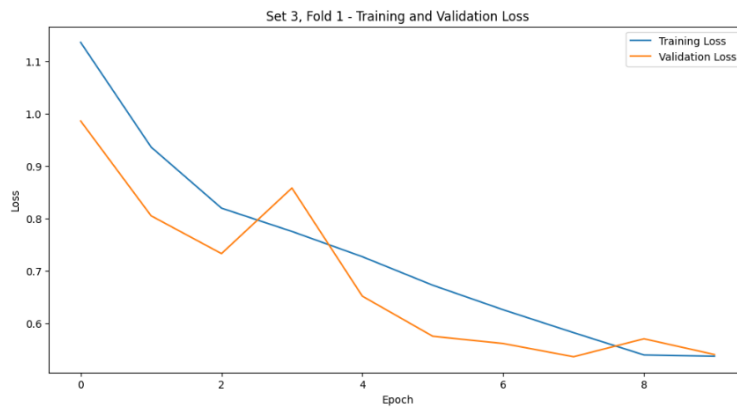


Obs ca oversampling-ul fara augmentari a dus la o imbunatatire a performantei generale a modelului, cu o acuratete crescuta de la 0.66 la 0.73. De asemenea, macro-average F1-score a crescut de la 0.62 la 0.68.

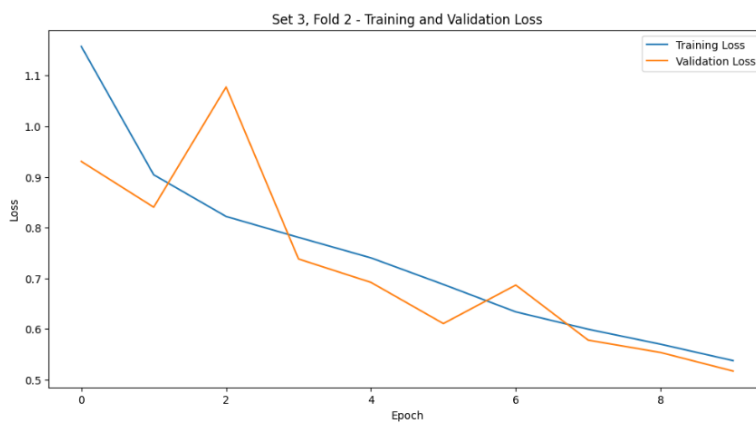
Oversampling-ul fara augmentari a fost un avantaj pe acest model deoarece a ajutat la dezechilibrul claselor. In cazul nostru, oversampling-ul fara augmentari a avut un impact pozitiv asupra performantei modelului, in special pentru clasa glioma_tumor.

Cerinta 3:

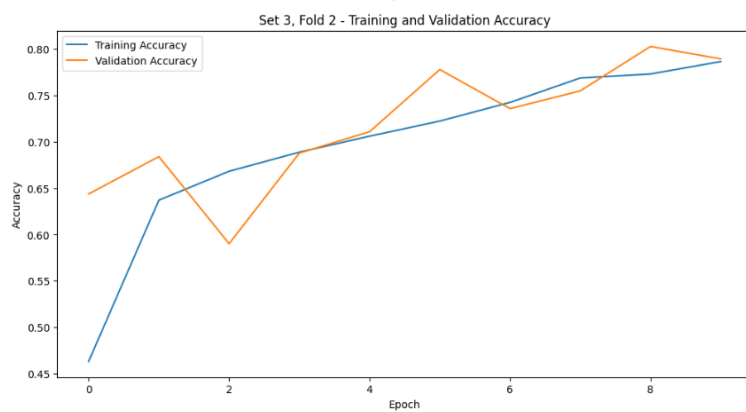
Pt Setul de augmentari 1:

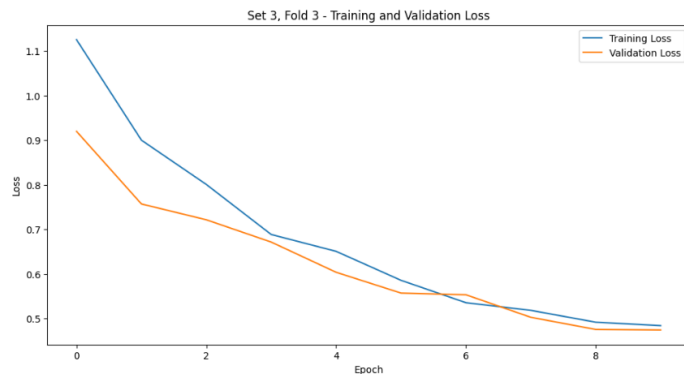


Fold 1:

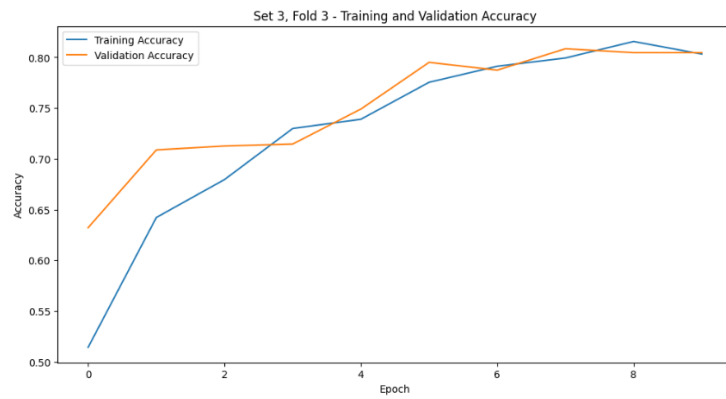


Fold 2:





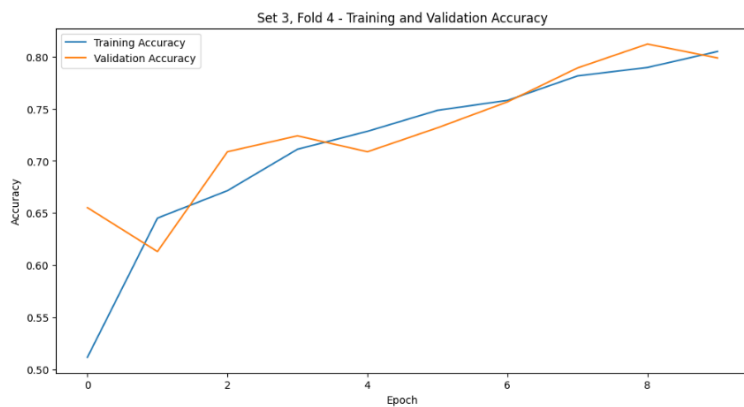
Fold 3:



Epoch



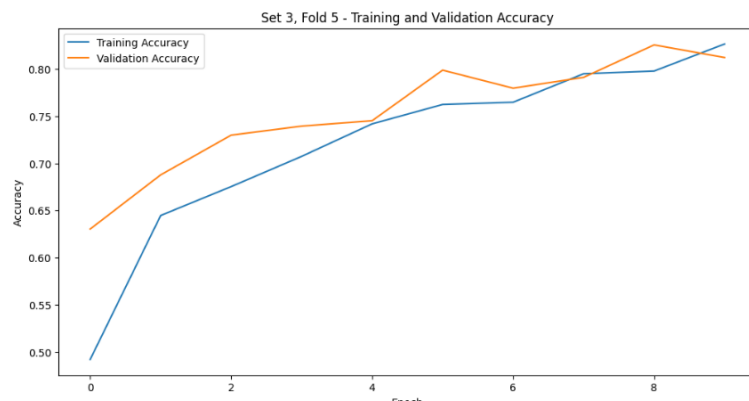
Fold 4:



Epoch

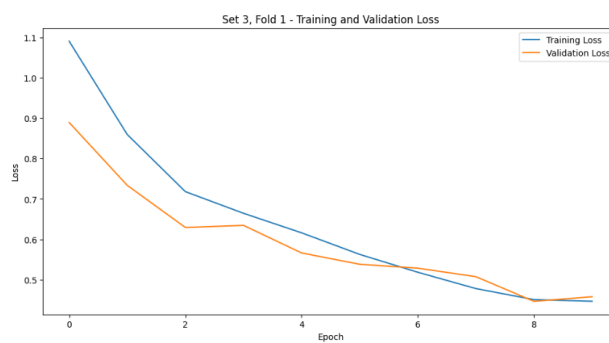


Fold 5:

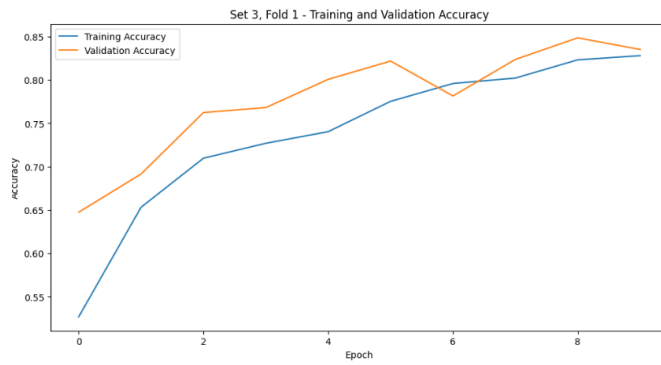


Rezultate pentru setul de augmentări 1: {'precision': (0.5495760278698308, 0.037405053402348766), 'recall': (0.4294416243654823, 0.03251902005694815), 'f1_score': (0.3823222741003377, 0.037045654935508106), 'accuracy': (0.4294416243654823, 0.03251902005694815)}

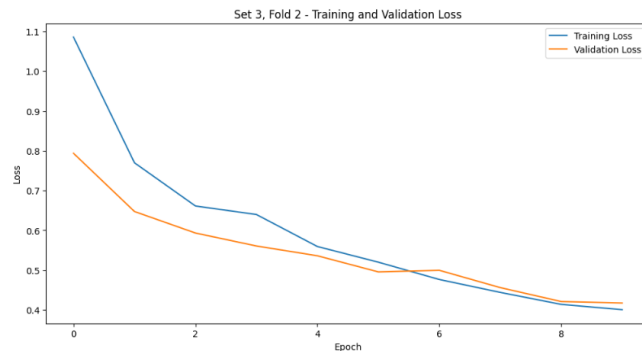
Pentru setul de augmentari 2:



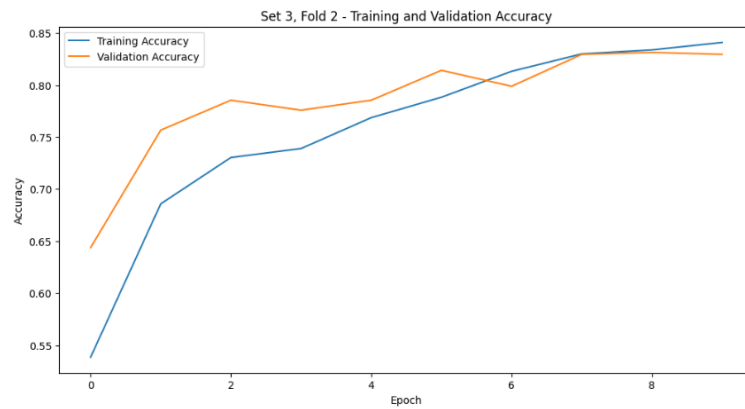
Fold1:



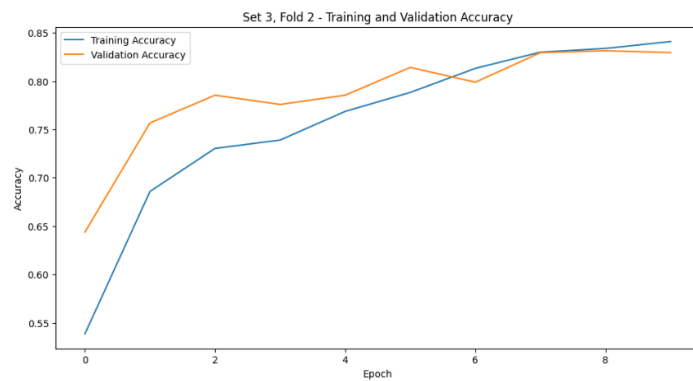
0.14 0.16



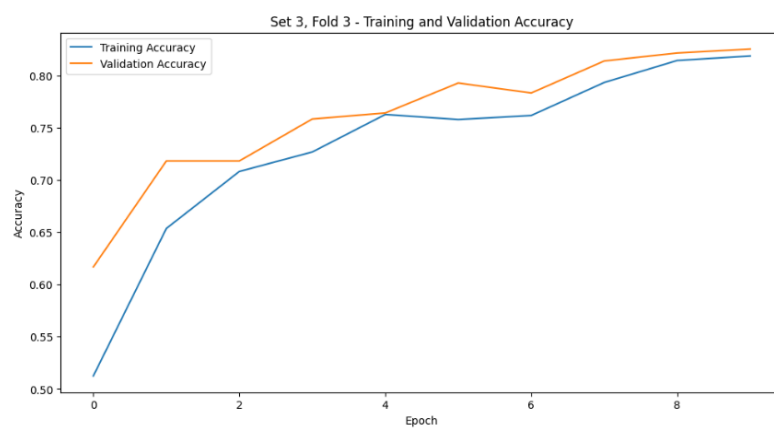
Fold2:



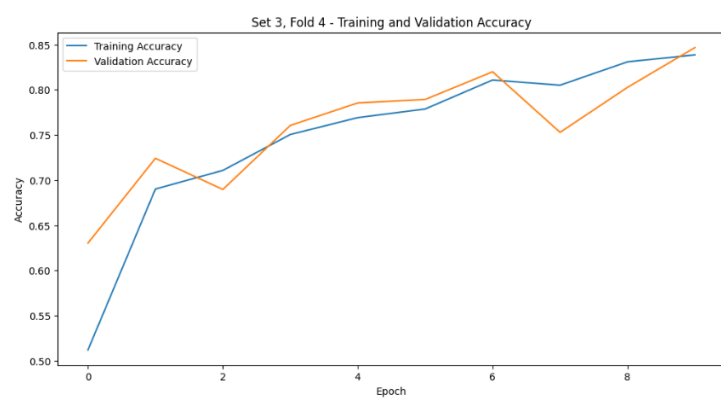
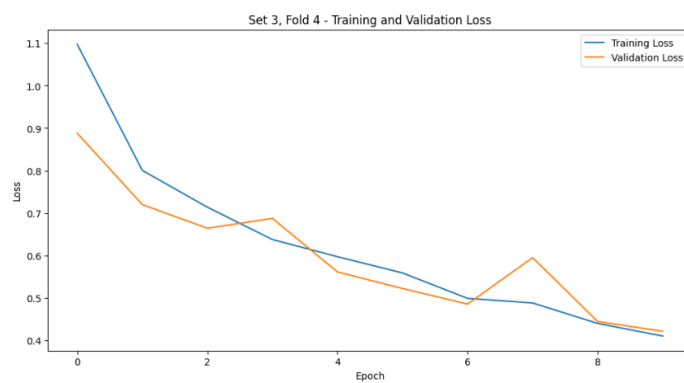
0.14 0.16



Fold3:

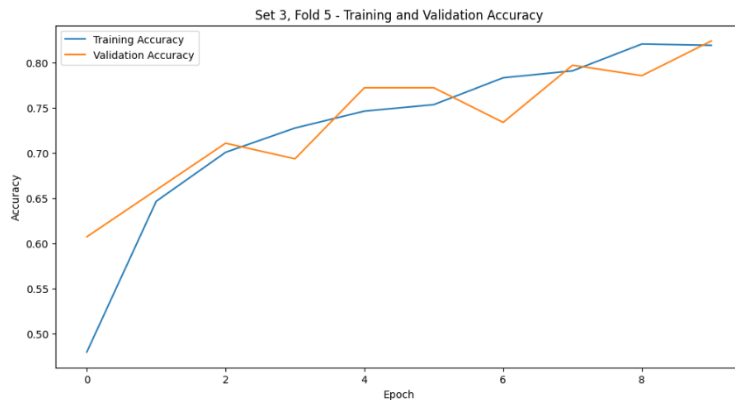


Fold4:



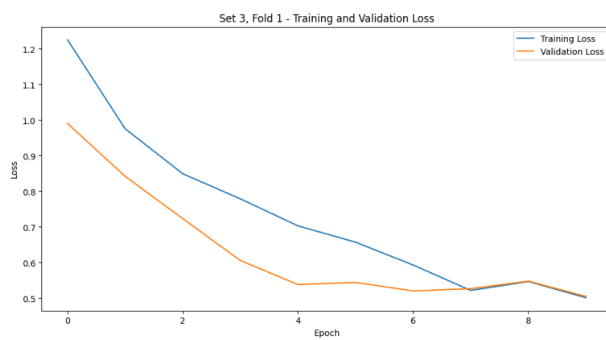


Fold5:

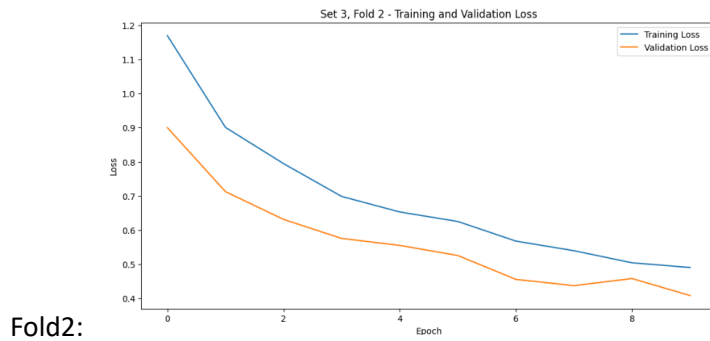


Rezultate pentru setul de augmentări 2: {'precision': (0.5883695027626714, 0.018973319513834054), 'recall': (0.48223350253807107, 0.049968425187105425), 'f1_score': (0.4418569573219931, 0.05184122828676386), 'accuracy': (0.48223350253807107, 0.049968425187105425)}

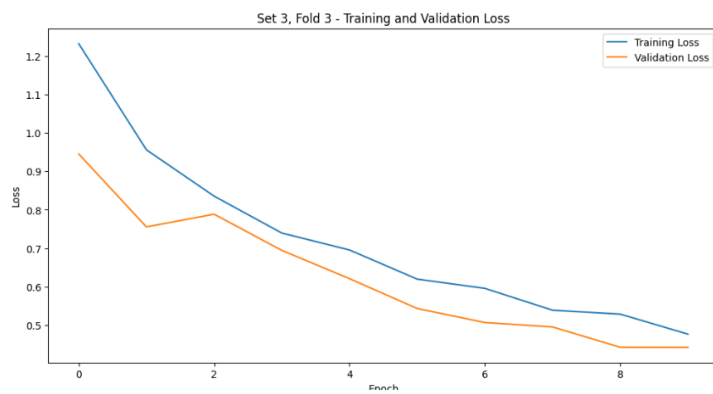
Setul de augmentari 3:



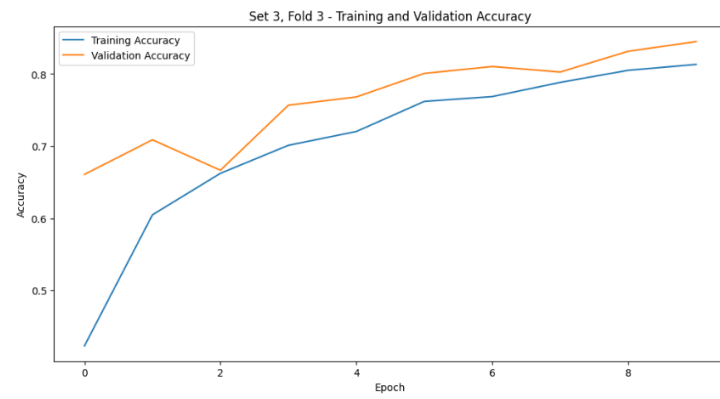
Fold1:



Fold2:



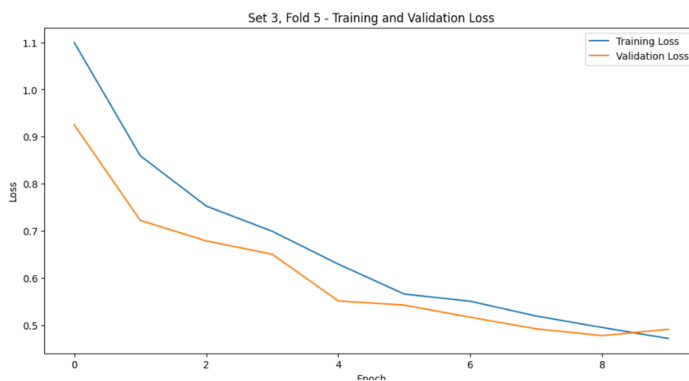
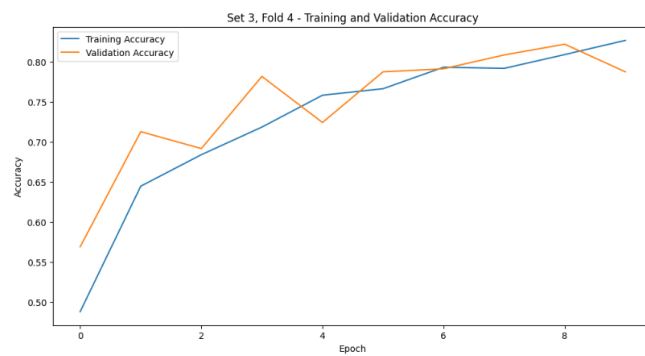
Fold3:



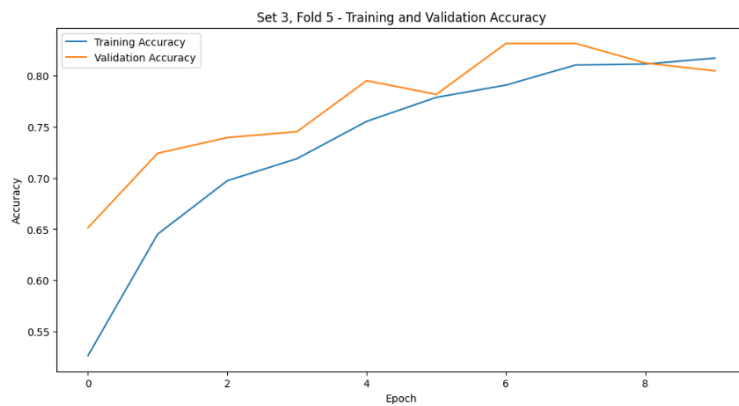
Epoch 9 / 9



Fold4:



Fold5:



Rezultate pentru setul de augmentări 3: {'precision': (0.6591754349569641, 0.03138846955237063), 'recall': (0.5568527918781726, 0.014231315501397149), 'f1_score': (0.5159341324746166, 0.011991847487046091), 'accuracy': (0.5568527918781726, 0.014231315501397149)}

Rezultate comparate pentru seturile de augmentări:

	Set 1	Set 2	Set 3
precision	0.5496 ± 0.0374	0.5884 ± 0.0190	0.6592 ± 0.0314
recall	0.4294 ± 0.0325	0.4822 ± 0.0500	0.5569 ± 0.0142
f1_score	0.3823 ± 0.0370	0.4419 ± 0.0518	0.5159 ± 0.0120
accuracy	0.4294 ± 0.0325	0.4822 ± 0.0500	0.5569 ± 0.0142

Bazat pe fiecare set, se vede ca setul 3 de transformari a obtinut cele mai bune rezultate pentru toate metricele:

1. Precision: Setul 3 are cea mai mare valoare medie pentru precision (0.65), ceea ce indică o proporție mai mare de predicții corecte pentru fiecare clasă.
2. Recall: Setul 3 are, de asemenea, cea mai mare valoare medie pentru recall (0.55), ceea ce indică o proporție mai mare de exemple corect clasificate pentru fiecare clasă.
3. F1-score: Setul 3 are cel mai mare F1-score mediu (0.51), ceea ce reprezintă o medie armonică între precision și recall, oferind o măsură echilibrată a performanței.
4. Accuracy: Setul 3 are cea mai mare acuratețe medie (0.55), ceea ce indică o proporție mai mare de predicții corecte în general.

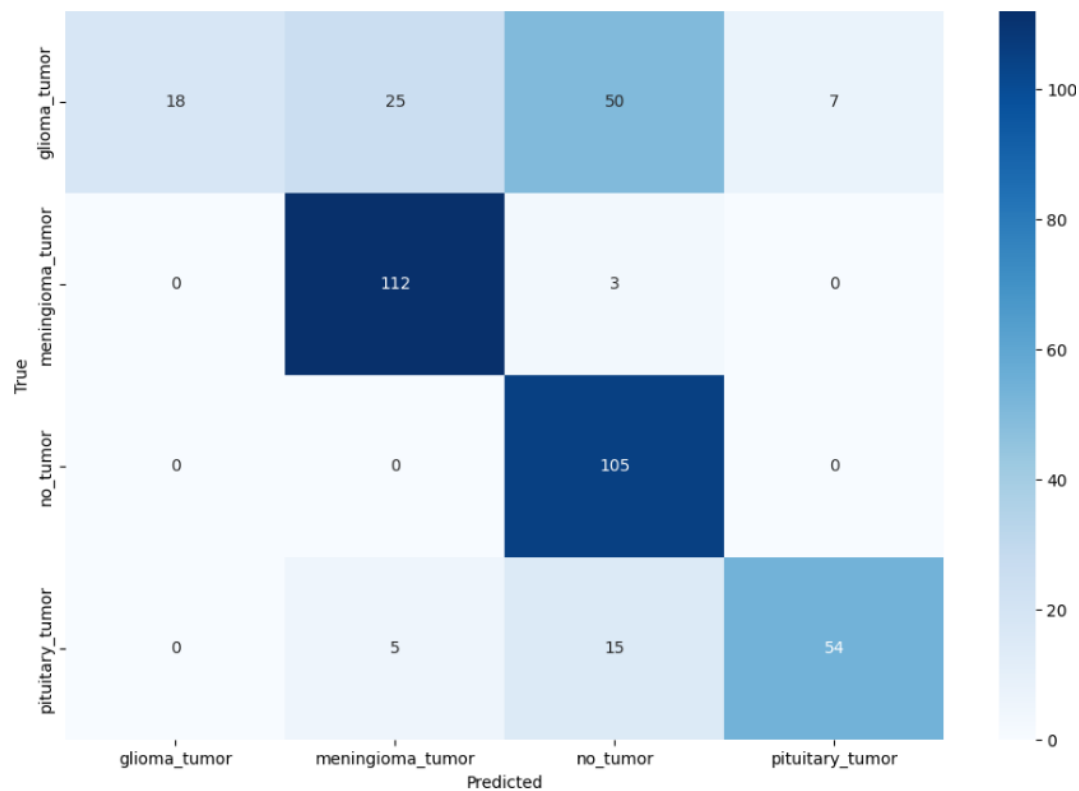
Setul 3 de transformari a reusit sa aduca cele mai bune performante datorita faptului ca a diversificat datele, a acoperit mai mult a variatiile si datorita normalizarii si evitarii supra-antrenarii.

Cerinta 4:

Early Stopping:

Classification Report:

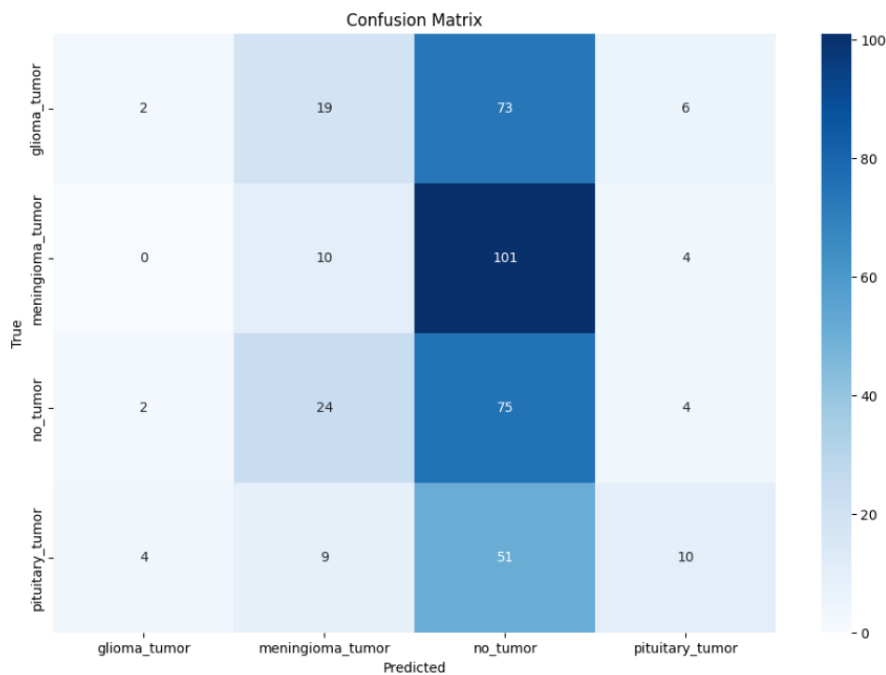
	precision	recall	f1-score	support
glioma_tumor	1.00	0.18	0.31	100
meningioma_tumor	0.79	0.97	0.87	115
no_tumor	0.61	1.00	0.76	105
pituitary_tumor	0.89	0.73	0.80	74
accuracy			0.73	394
macro avg	0.82	0.72	0.68	394
weighted avg	0.81	0.73	0.68	394



STEP LR:

Classification Report:

	precision	recall	f1-score	support
glioma_tumor	0.25	0.02	0.04	100
meningioma_tumor	0.16	0.09	0.11	115
no_tumor	0.25	0.71	0.37	105
pituitary_tumor	0.42	0.14	0.20	74
accuracy			0.25	394
macro avg	0.27	0.24	0.18	394
weighted avg	0.26	0.25	0.18	394



Comparare Rezultate:

Comparison of Results:

	Technique	Train Loss	Validation Loss	Train Accuracy \
0	Early Stopping	0.143594	0.40885	0.947797
1	StepLR	1.048735	0.93656	0.591954

	Validation Accuracy	Time
0	0.881226	111.772117
1	0.655172	89.597452

Obs o diferenta semnificativa in performanța celor două tehnici. Early Stopping a obținut o acuratețe mult mai mare (0.73) și un scor F1 macro mai bun (0.68) în comparație cu Step LR (0.25 și 0.18, respectiv). De asemenea, Early Stopping a performat mai bine pe toate cele patru clase de tumori, în special pentru glioma_tumor și meningioma_tumor.

Early Stopping: Este o tehnică simplă și eficientă pentru a preveni supra-antrenare (overfitting), oprind antrenamentul modelului atunci când performanța pe setul de validare începe să scadă.

În acest context de clasificare a tumorilor cerebrale, unde datele de antrenament pot fi limitate, Early Stopping este o alegere bună pentru a evita supra-antrenare și pentru a obține o performanță mai bună pe date noi și nevăzute.

Step LR: Poate ajuta modelul să scape de minimele locale și să converge către un minim global mai bun, prin reducerea ratei de învățare în timpul antrenării.

În acest caz particular, Step LR nu a reușit să îmbunătățească performanța modelului, posibil din cauza unei ajustări nepotrivite a parametrilor sau a interacțiunii cu alți hiperparametri. Este posibil ca o ajustare fină a parametrilor Step LR să poată duce la o performanță mai bună, dar în acest caz, Early Stopping a fost o alegere mai eficientă.