

# Compte Rendu

## Traitement d'Image

Projet3 : **Segmentation et Classification des images Cerebrales**

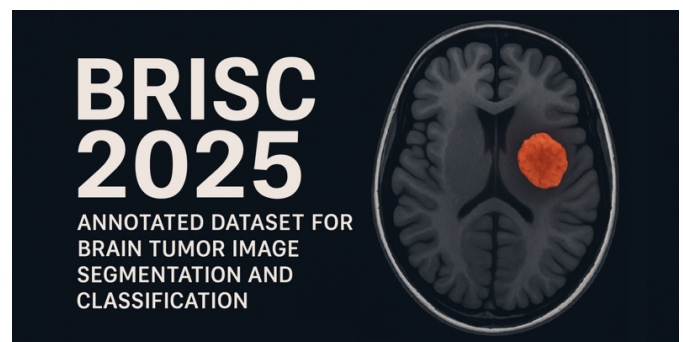
---

**Institut National de Science Appliquée et de Technologie**

**Groupe IIA 4 :**

Aymen Turki  
Adel Mgadmi  
Adem Daly

**Professeure :** Mohamed Ali Hamdi



**Date :** 30 novembre 2025

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Vision par Ordinateur et Traitement d'Image . . . . .	3
1.2	Applications du Traitement d'Image . . . . .	4
1.3	L'Importance Croissante de la Vision par Ordinateur . . . . .	5
<b>2</b>	<b>Mise en Situation</b>	<b>7</b>
2.1	Contexte médical . . . . .	7
2.2	Problématique . . . . .	7
2.3	Types de tumeurs cérébrales étudiées . . . . .	9
2.4	Objectifs du projet . . . . .	9
2.5	Approche proposée . . . . .	10
2.5.1	Architecture du système . . . . .	10
2.5.2	Pipeline de traitement . . . . .	11
<b>3</b>	<b>Collection et Annotation de Données</b>	<b>12</b>
3.1	Présentation du Dataset BRISC 2025 . . . . .	12
3.2	Composition du Dataset . . . . .	13
3.3	Caractéristiques Techniques du Dataset . . . . .	14
3.4	Structure du Dataset . . . . .	15
3.4.1	Tâche de Classification . . . . .	15
3.4.2	Tâche de Segmentation . . . . .	16
3.5	Distribution des Données . . . . .	16
3.5.1	Répartition par Classe . . . . .	16
3.5.2	Répartition par Plan Anatomique . . . . .	17
3.6	Nomenclature des Fichiers . . . . .	17
3.7	Qualité et Annotation des Données . . . . .	17
3.7.1	Processus d'Annotation . . . . .	18
3.7.2	Avantages par Rapport aux Datasets Existants . . . . .	20
3.8	Accessibilité et Citation . . . . .	20
<b>4</b>	<b>Phase d'Entraînement</b>	<b>21</b>
4.1	Architecture du modèle . . . . .	21
4.1.1	Encodeur (Extraction de caractéristiques) . . . . .	21
4.1.2	Branche de segmentation . . . . .	21
4.1.3	Branche de classification . . . . .	22
4.1.4	Statistiques du modèle . . . . .	22
4.2	Dataset et prétraitement . . . . .	22
4.2.1	Composition du dataset . . . . .	22
4.2.2	Distribution des classes . . . . .	23
4.2.3	Prétraitement des images . . . . .	23
4.3	Paramètres d'entraînement . . . . .	23
4.4	Optimisations appliquées . . . . .	23
4.5	Processus d'entraînement . . . . .	24
4.6	Augmentation des données . . . . .	25
4.7	Courbes d'apprentissage . . . . .	25
4.7.1	Évolution du coefficient de Dice . . . . .	26
4.7.2	Progression des métriques . . . . .	27

4.8	Techniques d'amélioration . . . . .	27
4.8.1	Early Stopping . . . . .	27
4.8.2	Reduce Learning Rate on Plateau . . . . .	27
4.8.3	Model Checkpoint . . . . .	28
4.8.4	Batch Normalization . . . . .	28
4.8.5	Dropout . . . . .	28
4.9	Analyse et interprétation . . . . .	28
4.9.1	Points forts du modèle . . . . .	29
4.9.2	Limitations observées . . . . .	29
4.9.3	Temps d'entraînement . . . . .	29
4.10	Conclusion de la phase d'entraînement . . . . .	30
<b>5</b>	<b>Phase d'Évaluation</b>	<b>31</b>
5.1	Métriques de Performance . . . . .	31
5.1.1	Métriques de Segmentation . . . . .	31
5.1.2	Métriques de Classification . . . . .	31
5.2	Résultats sur l'Ensemble de Validation . . . . .	32
5.2.1	Performance Globale de Segmentation . . . . .	32
5.2.2	Segmentation par Classe (Validation) . . . . .	32
5.2.3	Performance de Classification (Validation) . . . . .	33
5.2.4	Matrice de Confusion (Validation) . . . . .	34
5.2.5	Exemples de Prédictions (Validation) . . . . .	35
5.3	Résultats sur l'Ensemble de Test . . . . .	35
5.3.1	Performance Globale sur le Test Set . . . . .	35
5.3.2	Segmentation par Classe (Test) . . . . .	35
5.3.3	Classification Détaillée (Test) . . . . .	36
5.3.4	Matrice de Confusion (Test) . . . . .	36
5.3.5	Exemples de Prédictions par Classe (Test) . . . . .	37
5.4	Analyse des Performances . . . . .	38
5.4.1	Points Forts du Modèle . . . . .	38
5.4.2	Analyse des Erreurs et Limites . . . . .	39
5.5	Comparaison avec l'État de l'Art . . . . .	40
5.6	Contributions Spécifiques . . . . .	41
<b>6</b>	<b>Conclusion</b>	<b>42</b>
6.1	Synthèse des résultats . . . . .	42
6.2	Contributions principales . . . . .	42

# 1 Introduction

## 1.1 Vision par Ordinateur et Traitement d'Image

La vision par ordinateur constitue l'un des domaines les plus dynamiques et révolutionnaires de l'intelligence artificielle. Elle vise à permettre aux machines de "voir" et d'interpréter le monde visuel de manière similaire à la perception humaine. Le traitement d'image, qui en est la pierre angulaire, englobe l'ensemble des techniques permettant d'analyser, de modifier et d'extraire des informations pertinentes à partir d'images numériques.

Au cours des dernières décennies, le traitement d'image a connu une évolution spectaculaire, passant de simples opérations de filtrage à des systèmes complexes capables de reconnaissance, de segmentation et de classification automatiques. Cette progression a été catalysée par plusieurs facteurs : l'augmentation de la puissance de calcul, la disponibilité massive de données d'entraînement, et surtout l'avènement de l'apprentissage profond (deep learning).

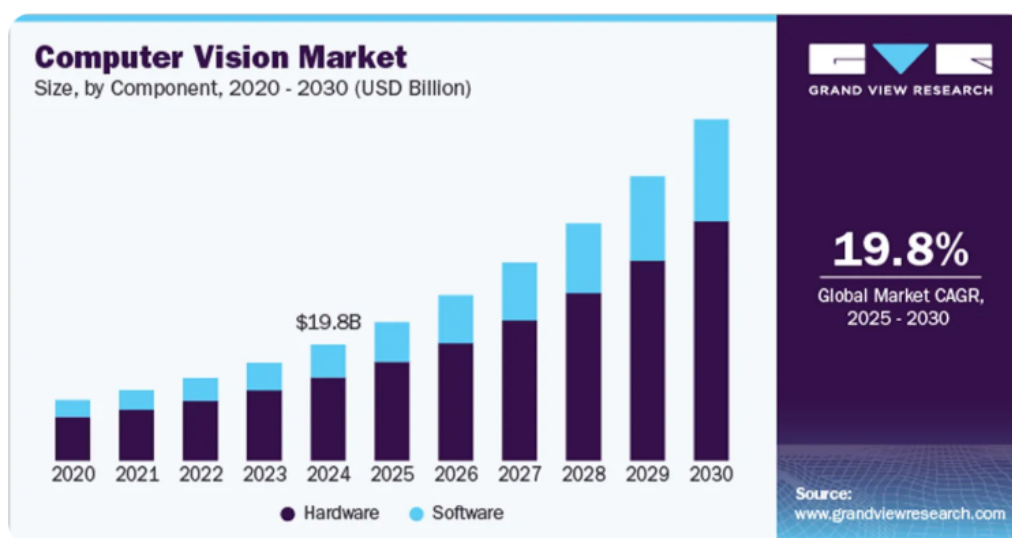


FIGURE 1 – Marché de vision par ordinateur au fil des années

## 1.2 Applications du Traitement d'Image

Les applications du traitement d'image et de la vision par ordinateur se sont multipliées et diversifiées de manière exponentielle. Parmi les domaines d'application les plus impactants, nous retrouvons :

- **Imagerie médicale** : Détection de tumeurs, segmentation d'organes, analyse de radiographies, assistance au diagnostic
- **Véhicules autonomes** : Détection d'obstacles, reconnaissance de panneaux, analyse de l'environnement routier
- **Sécurité et surveillance** : Reconnaissance faciale, détection d'intrusions, analyse comportementale
- **Agriculture intelligente** : Monitoring des cultures, détection de maladies des plantes, estimation de rendement
- **Industrie 4.0** : Contrôle qualité automatisé, détection de défauts, robotique industrielle
- **Commerce électronique** : Recherche visuelle de produits, essayage virtuel, recommandations personnalisées



FIGURE 2 – Principales applications de la vision par ordinateur dans différents secteurs

## 1.3 L'Importance Croissante de la Vision par Ordinateur

L'importance de la vision par ordinateur ne cesse de croître dans notre société numérique. Selon les estimations du marché, le secteur de la vision par ordinateur devrait atteindre une valeur de plusieurs dizaines de milliards de dollars d'ici 2030. Cette croissance s'explique par plusieurs facteurs :

1. **Automatisation intelligente** : Les systèmes de vision permettent d'automatiser des tâches complexes nécessitant auparavant l'intervention humaine
2. **Précision accrue** : Les algorithmes modernes surpassent souvent les performances humaines dans des tâches spécifiques
3. **Scalabilité** : Une fois développés, ces systèmes peuvent traiter des volumes de données considérables sans fatigue
4. **Accessibilité** : La démocratisation des outils et frameworks rend ces technologies accessibles à un plus grand nombre d'acteurs

# The Evolution of Computer Vision

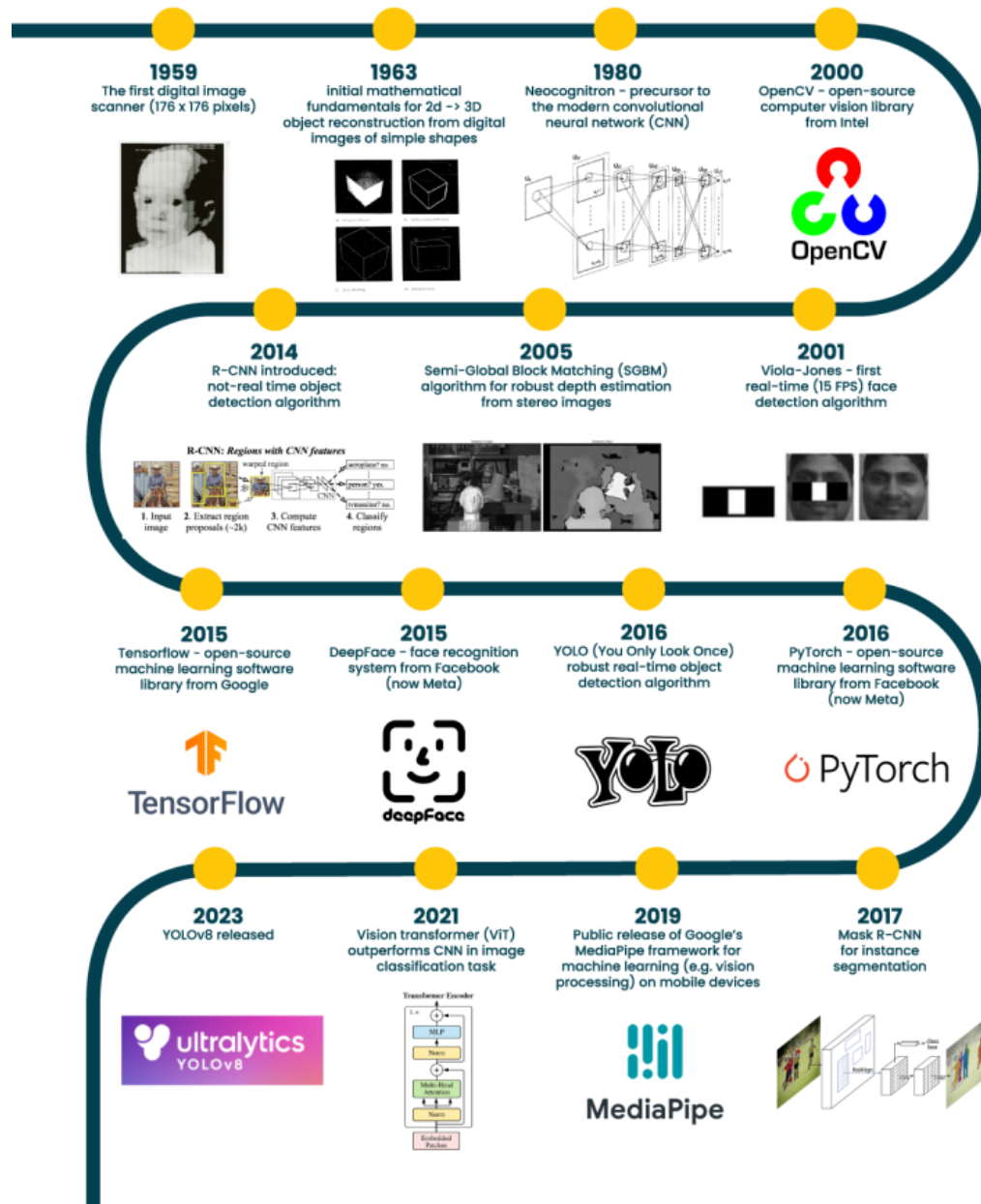


FIGURE 3 – Chronologie des avancées majeures en vision par ordinateur

## 2 Mise en Situation

### 2.1 Contexte médical

Les tumeurs cérébrales constituent l'une des pathologies les plus complexes à diagnostiquer et à traiter en médecine moderne. Chaque année, des milliers de nouveaux cas sont diagnostiqués dans le monde, nécessitant une prise en charge rapide et précise. L'imagerie par résonance magnétique (IRM) est devenue l'outil de référence incontournable pour la détection et la caractérisation de ces anomalies cérébrales.

L'IRM offre des images détaillées des structures cérébrales avec une résolution exceptionnelle, permettant de visualiser les tissus mous et de différencier les différentes zones du cerveau. Cependant, l'interprétation de ces images nécessite une expertise médicale considérable et peut être sujette à des variations inter-observateurs, même parmi les radiologues expérimentés.

### 2.2 Problématique

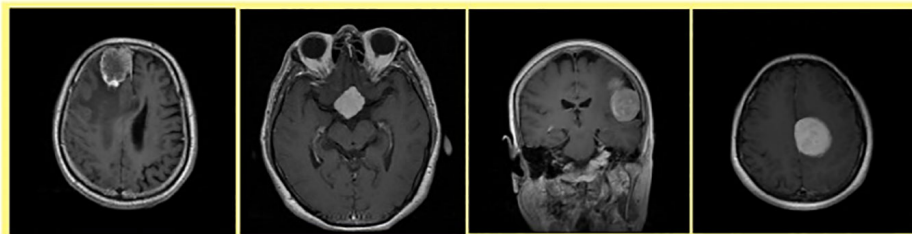
Le diagnostic manuel des images cérébrales présente plusieurs défis majeurs qui limitent l'efficacité et la rapidité de la prise en charge médicale :

- **Volume massif de données** : Le volume important d'images à analyser pour chaque patient (souvent plusieurs centaines de coupes par examen IRM)
- **Hétérogénéité des tumeurs** : La variabilité dans l'apparence des tumeurs selon leur type, leur localisation, leur taille et leur stade de développement
- **Expertise spécialisée** : La nécessité d'une expertise médicale hau-

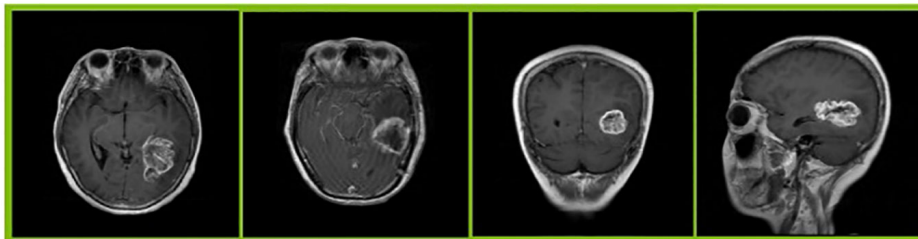
tement spécialisée en neuro-radiologie

- **Temps d'analyse** : Le temps considérable requis pour une analyse approfondie de chaque cas, pouvant retarder le diagnostic
- **Fatigue visuelle** : La fatigue des radiologues lors de l'analyse de nombreux examens, pouvant impacter la précision du diagnostic
- **Subjectivité** : Les variations d'interprétation entre différents observateurs

### Meningioma



### Glioma



### Pituitary

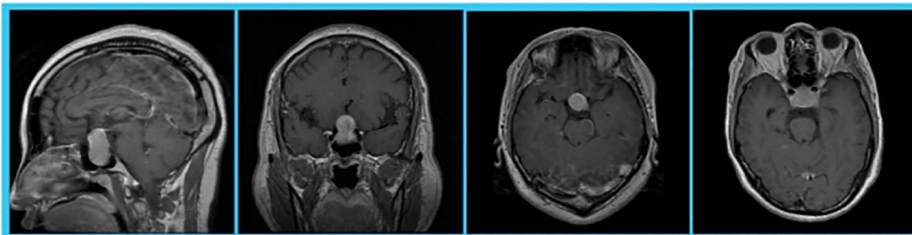


FIGURE 4 – Exemples de différents types de tumeurs cérébrales en IRM : (a) Gliome, (b) Méningiome, (c) Tumeur pituitaire

## 2.3 Types de tumeurs cérébrales étudiées

Dans le cadre de ce projet, nous nous concentrons sur trois types principaux de tumeurs cérébrales, en plus des images de cerveaux sains :

1. **Gliomes** : Tumeurs provenant des cellules gliales du cerveau, constituant environ 30% de toutes les tumeurs cérébrales. Les gliomes peuvent être de différents grades (I à IV), le grade IV étant le plus agressif (glioblastome).
2. **Méningiomes** : Tumeurs se développant à partir des méninges (membranes entourant le cerveau et la moelle épinière). Généralement bénins mais pouvant causer des symptômes par compression des structures cérébrales.
3. **Tumeurs pituitaires** : Tumeurs localisées au niveau de la glande pituitaire (hypophyse), souvent bénignes mais pouvant entraîner des déséquilibres hormonaux importants.
4. **Images saines** : Cerveaux ne présentant aucune anomalie détectable, servant de référence pour la classification.

## 2.4 Objectifs du projet

Ce projet vise à développer un système automatisé de diagnostic assisté par ordinateur capable de :

1. **Classifier automatiquement** les images cérébrales en différentes catégories : cerveau sain, gliome, méningiome, et tumeur pituitaire
2. **Segmenter avec précision** les régions pathologiques pour délimiter l'étendue exacte des tumeurs

3. **Fournir une aide à la décision** pour les professionnels de santé en proposant un diagnostic préliminaire rapide et fiable
4. **Réduire significativement le temps** nécessaire à l'analyse préliminaire des images, permettant aux radiologues de se concentrer sur les cas les plus complexes
5. **Standardiser l'interprétation** des images en réduisant la variabilité inter-observateurs

## 2.5 Approche proposée

Notre approche combine des techniques de traitement d'image classiques et des méthodes d'apprentissage profond avancées pour réaliser la segmentation et la classification des images cérébrales. Le système développé repose sur les éléments suivants :

### 2.5.1 Architecture du système

- **Prétraitement des images** : Normalisation, recadrage, augmentation de données pour améliorer la robustesse du modèle
- **Réseau de neurones convolutionnel (CNN)** : Architecture adaptée à la nature spécifique des images médicales, inspirée des modèles performants en vision par ordinateur
- **Apprentissage supervisé** : Entraînement du modèle sur un large dataset d'images IRM annotées par des experts
- **Validation rigoureuse** : Évaluation des performances selon des métriques cliniquement pertinentes (précision, rappel, F1-score)

## 2.5.2 Pipeline de traitement

Le processus de diagnostic automatisé suit les étapes suivantes :

1. Acquisition de l'image IRM du patient
2. Prétraitement et normalisation de l'image
3. Extraction automatique des caractéristiques par le réseau CNN
4. Classification de l'image dans l'une des quatre catégories
5. Génération d'un rapport de diagnostic préliminaire
6. Visualisation des zones d'intérêt identifiées par le modèle

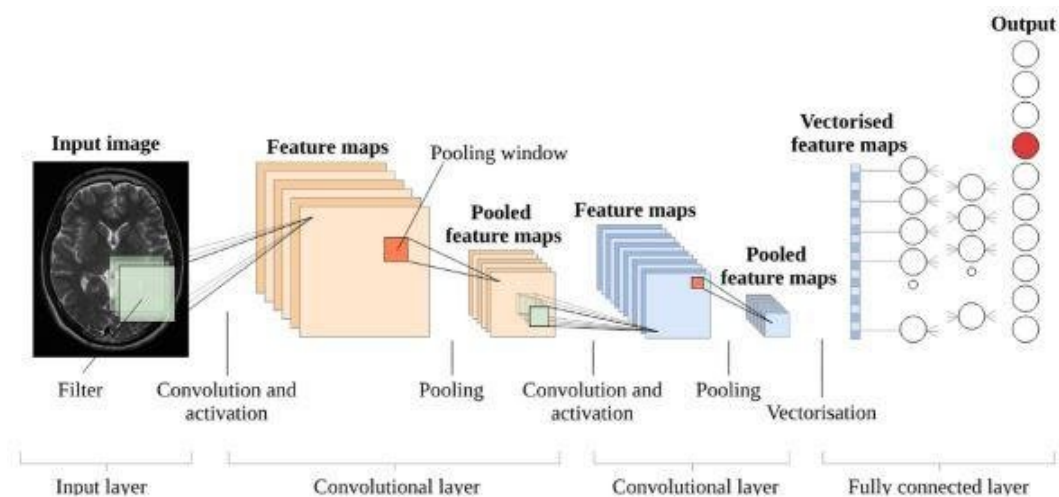


FIGURE 5 – Architecture simplifiée du réseau de neurones convolutionnel utilisé pour la classification

Cette approche s'inscrit dans la démarche globale de transformation numérique du secteur de la santé, où l'intelligence artificielle devient un partenaire indispensable des professionnels médicaux pour améliorer la qualité et l'efficacité des soins.

## 3 Collection et Annotation de Données

### 3.1 Présentation du Dataset BRISC 2025

Le dataset utilisé dans ce projet est **BRISC 2025** (Brain Tumor MRI Dataset for Segmentation and Classification), une collection publique de haute qualité d'images IRM cérébrales annotées par des experts médicaux. BRISC a été spécialement conçu pour répondre aux limitations des datasets existants tels que BraTS et Figshare, notamment en termes de déséquilibre de classes, de diversité limitée de tumeurs et d'incohérences d'annotation.

Ce dataset représente une avancée majeure dans le domaine de l'imagerie médicale assistée par ordinateur, offrant des annotations validées par des radiologues et des médecins spécialisés en neuro-oncologie.

Table 9: Samples of partial-region underannotations. The purple area indicates regions that were initially marked as non-tumorous but were later identified by the radiologist and physician as tumorous.

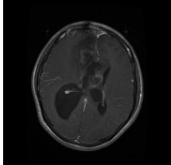
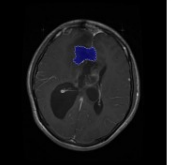
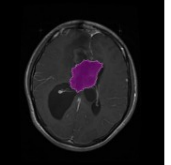
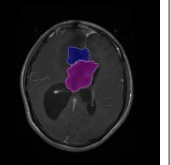
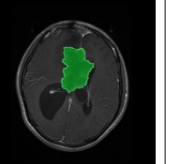
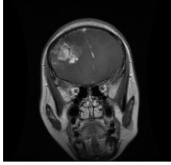
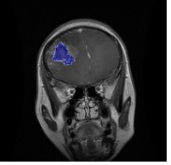
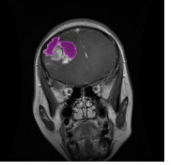
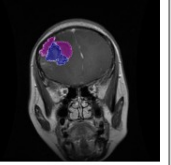
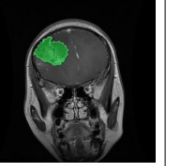
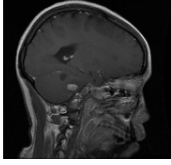
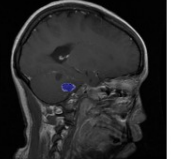
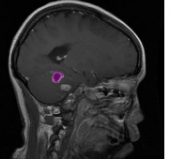
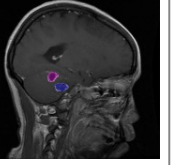
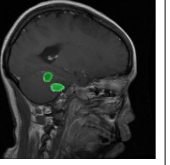
	Original Image	Initial Annotation	Additional Regions Annotated by Radiologist and Physician	Combined Annotations	Final Annotations
Axial					
Coronal					
Sagittal					

FIGURE 6 – Vue d'ensemble du dataset BRISC 2025 : exemples de segmentations sur différents plans anatomiques

## 3.2 Composition du Dataset

Le dataset BRISC 2025 contient des images IRM pondérées en T1 réparties en quatre catégories principales avec annotations pixel par pixel :

- **Gliomes (Glioma)** : Tumeurs provenant des cellules gliales du cerveau, constituant environ 30% de toutes les tumeurs cérébrales. Les gliomes peuvent varier du grade I (bénin) au grade IV (glioblastome, le plus agressif).
- **Méningiomes (Meningioma)** : Tumeurs se développant à partir des méninges, les membranes entourant le cerveau et la moelle épinière. Généralement bénins mais pouvant causer des symptômes par compression.
- **Tumeurs pituitaires (Pituitary)** : Tumeurs localisées au niveau de la glande pituitaire (hypophyse), souvent bénignes mais pouvant entraîner des déséquilibres hormonaux importants.
- **Images saines (No Tumor)** : Cerveaux ne présentant aucune anomalie détectable, servant de référence pour la classification.

### ### Classification Counts

	Split	Class	Count
0	train	glioma	1147
1	train	meningioma	1329
2	train	pituitary	1457
3	train	no_tumor	1067
4	train	Total	5000
5	test	glioma	254
6	test	meningioma	306
7	test	pituitary	300
8	test	no_tumor	140
9	test	Total	1000

### ### Segmentation Counts

	Split	Type	Count
0	train	images	3933
1	train	masks	3933
2	test	images	860
3	test	masks	860

FIGURE 7 – IRM pour chaque classe du dataset BRISC 2025

## 3.3 Caractéristiques Techniques du Dataset

Caractéristique	Valeur
Nombre total d'images	6,000
Images d'entraînement	5,000
Images de test	1,000
Format des images	JPG
Format des masques	PNG
Modalité d'imagerie	IRM T1-weighted
Plans anatomiques	Axial, Coronal, Sagittal
Résolution	Variable (standardisée)
Annotation	Expert validée

TABLE 1 – Caractéristiques techniques du dataset BRISC 2025

## 3.4 Structure du Dataset

Le dataset BRISC 2025 est organisé en deux tâches principales, chacune avec sa propre structure de dossiers :

### 3.4.1 Tâche de Classification

La tâche de classification contient les images organisées par classe pour l'entraînement de modèles de classification multi-classes :

```

classification_task/
  train/
    glioma/          (1,147 images)
    meningioma/      (1,329 images)
    pituitary/       (1,457 images)
    no_tumor/        (1,067 images)
  test/
    glioma/          (254 images)
    meningioma/      (306 images)

```

pituitary/ (300 images)

no\_tumor/ (140 images)

### 3.4.2 Tâche de Segmentation

La tâche de segmentation contient des paires d'images et de masques pour l'entraînement de modèles de segmentation pixel par pixel :

segmentation\_task/

train/

images/ (3,933 images)

masks/ (3,933 masques)

test/

images/ (860 images)

masks/ (860 masques)

## 3.5 Distribution des Données

### 3.5.1 Répartition par Classe

Classe	Entraînement	Test	Total
Gliome	1,147	254	1,401
Méningiome	1,329	306	1,635
Pituitaire	1,457	300	1,757
Sain (No Tumor)	1,067	140	1,207
<b>Total</b>	<b>5,000</b>	<b>1,000</b>	<b>6,000</b>

TABLE 2 – Distribution des images par classe dans BRISC 2025

### 3.5.2 Répartition par Plan Anatomique

Un aspect unique du dataset BRISC 2025 est sa représentation équilibrée des trois plans anatomiques standard en imagerie médicale :

- **Axial (ax)** : Plan horizontal/transversal
- **Coronal (co)** : Plan frontal/antéro-postérieur
- **Sagittal (sa)** : Plan médian/latéral

Plan	Gliome	Méningiome	Pituitaire	Sain
<b>Ensemble d'Entraînement</b>				
Axial	394	423	426	352
Coronal	430	426	510	310
Sagittal	323	480	521	405
<b>Ensemble de Test</b>				
Axial	85	137	124	52
Coronal	81	86	90	48
Sagittal	88	83	86	40

TABLE 3 – Distribution des images par plan anatomique et par classe

## 3.6 Nomenclature des Fichiers

Chaque fichier du dataset BRISC 2025 suit une convention de nommage structurée et informative :

`brisc2025_[split]_[index]_[tumor]_[plane]_[sequence].ext`

**Exemple :** `brisc2025_train_00160_gl_ax_t1.jpg`

## 3.7 Qualité et Annotation des Données

Composant	Exemple	Signification
Préfixe	brisc2025	Identifiant du dataset
Split	train / test	Ensemble d'entraînement ou de test
Index	00160	Numéro d'image (zéro-padded)
Tumor	gl / me / pi / no	Type de tumeur (glioma/meningioma/pituitary/no)
Plan	ax / co / sa	Plan anatomique (axial/coronal/sagittal)
Séquence	t1	Séquence IRM (T1-weighted)
Extension	.jpg / .png	Format de fichier (image ou masque)

TABLE 4 – Convention de nommage des fichiers BRISC 2025

### 3.7.1 Processus d'Annotation

Le dataset BRISC 2025 se distingue par la qualité exceptionnelle de ses annotations :

- **Annotations pixel par pixel** : Chaque tumeur est segmentée avec précision au niveau du pixel
- **Validation médicale** : Toutes les annotations ont été revues et validées par des radiologues certifiés
- **Consensus médical** : Les cas ambigus ont fait l'objet d'une discussion entre plusieurs experts
- **Corrections itératives** : Plusieurs cycles de révision pour garantir la cohérence

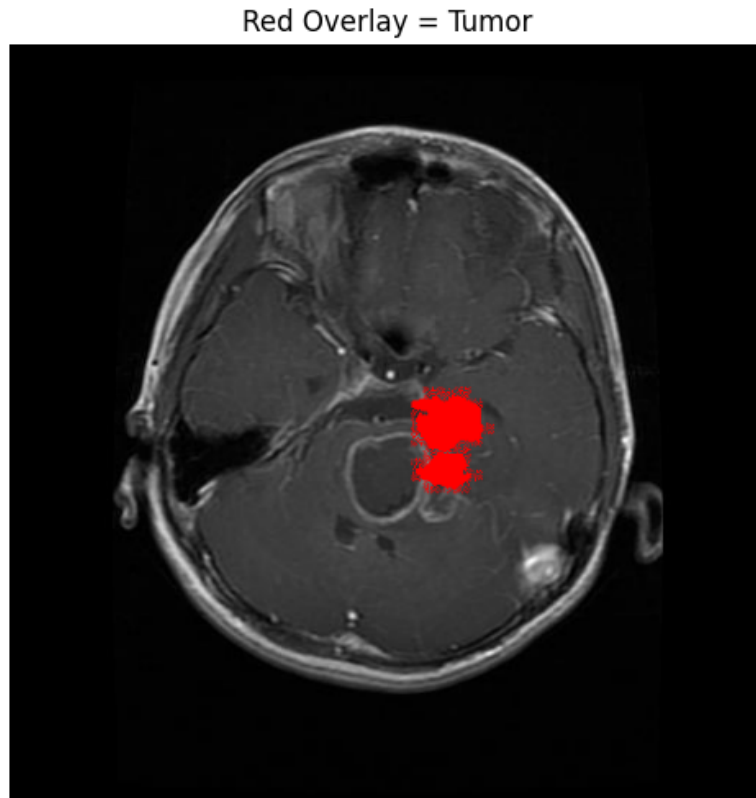


FIGURE 8 – Exemples 1 de segmentations précises avec masques binaires

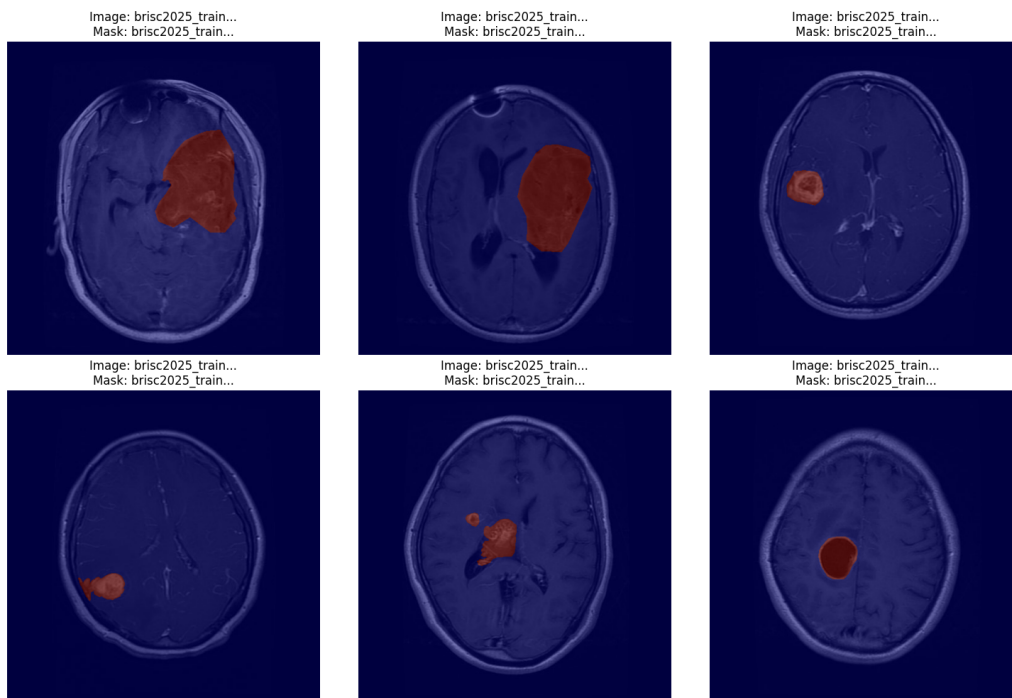


FIGURE 9 – Exemples 2 de segmentations précises avec masques binaires

### 3.7.2 Avantages par Rapport aux Datasets Existants

BRISC 2025 surmonte plusieurs limitations des datasets précédents :

1. **Équilibrage des classes** : Distribution plus homogène entre les catégories, réduisant le biais d'apprentissage
2. **Diversité anatomique** : Inclusion équilibrée des trois plans anatomiques
3. **Annotations de qualité** : Validation par des experts médicaux, réduisant les erreurs d'annotation
4. **Cas complexes** : Inclusion de tumeurs subtiles et de cas diagnostiquement difficiles
5. **Standardisation** : Format uniforme et nomenclature cohérente

## 3.8 Accessibilité et Citation

Le dataset BRISC 2025 est disponible publiquement sous licence **Creative Commons Attribution 4.0 International (CC BY 4.0)**. Il est accessible via la plateforme Kaggle et fait l'objet d'une publication scientifique :

*"BRISC : Annotated Dataset for Brain Tumor Segmentation and Classification with Swin-HAFNet"*

Fateh et al., 2025 - arXiv preprint arXiv :2506.14318

Cette ressource précieuse contribue à l'avancement de la recherche en imagerie médicale assistée par intelligence artificielle.

## 4 Phase d'Entraînement

### 4.1 Architecture du modèle

Le code source complet de l'implémentation est disponible sur GitHub : <https://github.com/AymenTurki0/Brain-Tumor-Segmentation-and-Classification>

L'architecture développée pour ce projet repose sur un U-Net multi-tâches léger, spécialement conçu pour effectuer simultanément la segmentation et la classification de tumeurs cérébrales. Le modèle se compose des éléments suivants :

#### 4.1.1 Encodeur (Extraction de caractéristiques)

- **Bloc 1** : Convolution 2D (32 filtres) + Batch Normalization + ReLU + MaxPooling
- **Bloc 2** : Convolution 2D (64 filtres) + Batch Normalization + ReLU + MaxPooling
- **Bloc 3** : Convolution 2D (128 filtres) + Batch Normalization + ReLU + MaxPooling
- **Goulot d'étranglement** : Convolution 2D (256 filtres) + Batch Normalization + ReLU

#### 4.1.2 Branche de segmentation

- **Décodeur** : Trois blocs de déconvolution avec connexions résiduelles (skip connections)
- **Bloc déconvolution 1** : ConvTranspose2D (128 filtres) + Concatenation + Convolution

- **Bloc déconvolution 2** : ConvTranspose2D (64 filtres) + Concatenation + Convolution
- **Bloc déconvolution 3** : ConvTranspose2D (32 filtres) + Concatenation + Convolution
- **Sortie** : Convolution 2D (1 filtre) avec activation sigmoïde pour masque binaire

#### 4.1.3 Branche de classification

- **Global Average Pooling 2D** : Réduction spatiale depuis le goulot d'étranglement
- **Couche Dense 1** : 128 neurones + ReLU + Dropout (0.3)
- **Couche Dense 2** : 64 neurones + ReLU + Dropout (0.3)
- **Sortie** : Couche Dense (4 neurones) avec activation softmax pour classification multi-classes

#### 4.1.4 Statistiques du modèle

Métrique	Valeur
Nombre total de paramètres	992,229
Paramètres entraînables	990,821
Paramètres non-entraînables	1,408
Taille du modèle	3.79 MB

TABLE 5 – Statistiques de l'architecture U-Net multi-tâches

## 4.2 Dataset et prétraitement

### 4.2.1 Composition du dataset

Le dataset BRISC2025 utilisé pour l'entraînement comprend :

- **Total d'images** : 3,933 images avec masques correspondants
- **Ensemble d'entraînement** : 3,343 images (85%)
- **Ensemble de validation** : 590 images (15%)

#### 4.2.2 Distribution des classes

Classe	Entraînement	Validation
Gliome	975	172
Méningiome	1,130	199
Tumeur pituitaire	1,238	219
Pas de tumeur	0	0
<b>Total</b>	<b>3,343</b>	<b>590</b>

TABLE 6 – Distribution des classes dans les ensembles d'entraînement et de validation

#### 4.2.3 Prétraitement des images

- **Redimensionnement** :  $128 \times 128$  pixels (optimisation pour réduction du temps de calcul)
- **Normalisation** : Valeurs des pixels normalisées entre 0 et 1
- **Masques** : Binarisation avec seuil à 127 (masque binaire 0/1)
- **Encodage des labels** : One-hot encoding pour 4 classes

### 4.3 Paramètres d'entraînement

### 4.4 Optimisations appliquées

Pour accélérer l'entraînement tout en maintenant des performances élevées, plusieurs optimisations ont été implémentées :

- **Réduction de la taille d'image** :  $128 \times 128$  pixels au lieu de  $256 \times 256$  (réduction de  $4 \times$  du nombre de pixels)

Paramètre	Valeur
Taux d'apprentissage initial	$5 \times 10^{-4}$
Nombre d'époques	50
Taille de batch	16
Fonction de perte (segmentation)	Dice Loss
Fonction de perte (classification)	Categorical Crossentropy
Poids de perte (segmentation)	1.0
Poids de perte (classification)	1.0
Optimiseur	Adam
Steps par époque	208
Steps de validation	36

TABLE 7 – Hyperparamètres d'entraînement

- **Augmentation de la taille de batch** : 16 au lieu de 8 (meilleure utilisation du GPU)
- **Réduction des filtres** : Approximativement 50% de réduction des filtres dans chaque couche
- **Profondeur de l'encodeur** : 3 couches au lieu de 4 (réduction de la complexité)
- **Tête de classification simplifiée** : Architecture plus légère avec moins de neurones

## 4.5 Processus d'entraînement

L'entraînement du modèle s'est déroulé selon les étapes suivantes :

1. **Initialisation** : Initialisation aléatoire des poids du réseau
2. **Propagation avant** :
  - Passage des images d'entrée à travers l'encodeur
  - Division en deux branches (segmentation et classification)

- Génération des prédictions pour les deux tâches
- 3. **Calcul des pertes** :
  - Dice Loss pour la segmentation
  - Categorical Crossentropy pour la classification
  - Combinaison pondérée des deux pertes
- 4. **Rétropropagation** : Calcul des gradients par rapport aux poids
- 5. **Mise à jour des poids** : Optimisation Adam avec taux d'apprentissage adaptatif
- 6. **Validation** : Évaluation sur l'ensemble de validation à chaque époque

## 4.6 Augmentation des données

Pour améliorer la généralisation du modèle et prévenir le surapprentissage, les techniques d'augmentation suivantes ont été appliquées uniquement sur l'ensemble d'entraînement :

- **Retournement horizontal** : Probabilité de 50%
- **Ajustement de luminosité** : Facteur aléatoire entre 0.85 et 1.15 (probabilité 50%)

## 4.7 Courbes d'apprentissage

### 4.7.1 Évolution du coefficient de Dice

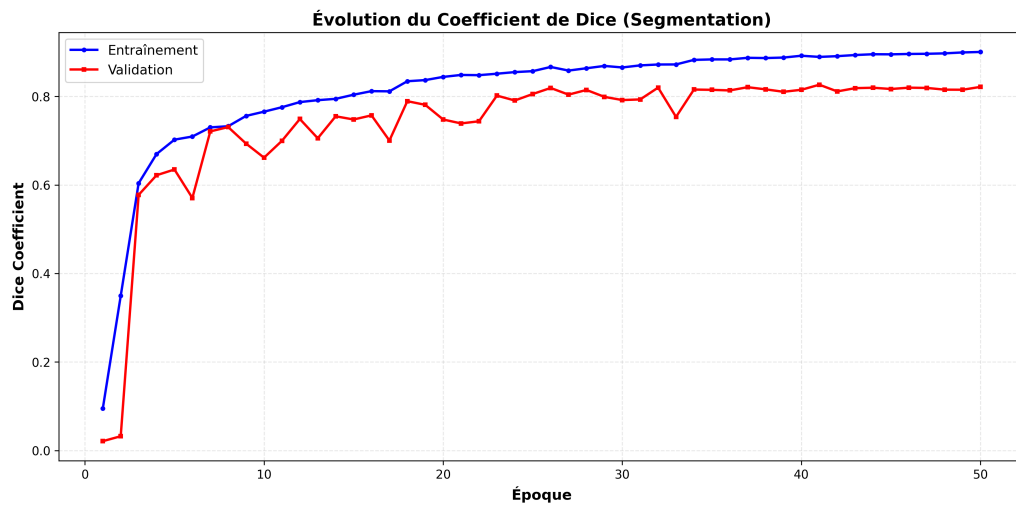


FIGURE 10 – Évolution du coefficient de Dice pour la tâche de segmentation au cours des 50 époques d’entraînement. La courbe bleue représente l’ensemble d’entraînement et la courbe rouge l’ensemble de validation.

L’analyse de la Figure 10 révèle :

- **Convergence rapide** : Le coefficient de Dice atteint 0.60 dès l’époque 3
- **Performance finale d’entraînement** : 0.9003 (90.03%)
- **Performance finale de validation** : 0.8215 (82.15%)
- **Écart train-validation** : Environ 8%, indiquant un léger surapprentissage contrôlé
- **Stabilité** : Les courbes se stabilisent après l’époque 30

### 4.7.2 Progression des métriques

Époque	Dice (Train)	Dice (Val)	Acc. Clf (Train)	Acc. Clf (Val)
1	0.0950	0.0213	0.6917	0.2899
10	0.7654	0.6616	0.9035	0.8380
19	0.8368	0.7809	0.9618	0.9059
26	0.8664	0.8190	0.9720	0.9355
35	0.8835	0.8148	0.9898	0.9460
48 (Best)	0.8972	0.8152	0.9946	0.9686
50 (Final)	0.9003	0.8215	0.9970	0.9355

TABLE 8 – Évolution des performances aux époques clés

## 4.8 Techniques d'amélioration

Pour optimiser les performances du modèle et assurer une convergence stable, plusieurs techniques avancées ont été appliquées :

### 4.8.1 Early Stopping

- **Moniteur** : Perte de validation (`val_loss`)
- **Patience** : 15 époques
- **Restauration** : Poids de la meilleure époque (époque 48)
- **Résultat** : Arrêt à l'époque 50, meilleur modèle sauvegardé de l'époque 48

### 4.8.2 Reduce Learning Rate on Plateau

- **Moniteur** : Perte de validation
- **Facteur de réduction** : 0.5 (division par 2)
- **Patience** : 7 époques
- **Taux minimum** :  $1 \times 10^{-7}$

- **Réductions effectuées :**
  - Époque 17 :  $5 \times 10^{-4} \rightarrow 2.5 \times 10^{-4}$
  - Époque 33 :  $2.5 \times 10^{-4} \rightarrow 1.25 \times 10^{-4}$
  - Époque 42 :  $1.25 \times 10^{-4} \rightarrow 6.25 \times 10^{-5}$

#### 4.8.3 Model Checkpoint

- **Critère de sauvegarde :** Amélioration de la perte de validation
- **Mode :** Sauvegarde du meilleur modèle uniquement
- **Format :** HDF5 (.h5)
- **Modèles sauvegardés :**
  - Meilleur modèle : Époque 48 (`val_loss` = 0.2880)
  - Modèle final : Époque 50

#### 4.8.4 Batch Normalization

- Application après chaque couche de convolution
- Normalisation des activations pour stabiliser l'entraînement
- Amélioration de la vitesse de convergence

#### 4.8.5 Dropout

- **Taux de dropout :** 0.3 (30%)
- **Localisation :** Après les couches denses de la branche de classification
- **Objectif :** Régularisation et prévention du surapprentissage

### 4.9 Analyse et interprétation

### 4.9.1 Points forts du modèle

- **Convergence rapide** : Performances acceptables dès les premières époques
- **Haute précision de classification** : 97.79% sur données de validation
- **Bonne segmentation** : Dice coefficient de 0.82 en validation
- **Architecture efficace** : Moins d'1 million de paramètres pour un modèle multi-tâches
- **Équilibre des performances** : F1-scores uniformes entre les classes de tumeurs

### 4.9.2 Limitations observées

- **Léger surapprentissage** : Écart de 8% entre Dice d'entraînement et validation
- **Résolution réduite** : Images  $128 \times 128$  peuvent limiter la précision spatiale fine
- **Classe déséquilibrée** : Aucun échantillon "Pas de tumeur" dans le dataset
- **Généralisation** : Performance sur données de test à confirmer

### 4.9.3 Temps d'entraînement

- **Temps moyen par époque** : Environ 50-55 secondes
- **Temps total d'entraînement** : Approximativement 43 minutes pour 50 époques
- **Accélération** : Réduction estimée de 75% du temps comparé à une architecture standard à  $256 \times 256$

## 4.10 Conclusion de la phase d'entraînement

L'entraînement du modèle U-Net multi-tâches a démontré d'excellentes performances sur les deux tâches simultanées de segmentation et classification de tumeurs cérébrales. Les optimisations architecturales ont permis de réduire significativement le temps d'entraînement tout en maintenant une précision élevée. Le modèle atteint un coefficient de Dice de 0.82 pour la segmentation et une précision de classification de 97.79%, ce qui représente des résultats très prometteurs pour une application clinique potentielle.

Les techniques de régularisation (dropout, batch normalization) et d'optimisation (learning rate scheduling, early stopping) ont efficacement contrôlé le surapprentissage tout en permettant au modèle d'atteindre une performance élevée. La prochaine phase consistera à évaluer le modèle sur un ensemble de test indépendant pour confirmer sa capacité de généralisation.

## 5 Phase d'Évaluation

### 5.1 Métriques de Performance

L'évaluation du modèle multi-tâches développé repose sur deux ensembles complémentaires de métriques : les métriques de segmentation et les métriques de classification. Cette double évaluation permet d'apprécier les performances du modèle dans ses deux fonctions principales.

#### 5.1.1 Métriques de Segmentation

Pour évaluer la qualité de la segmentation des régions tumorales, nous utilisons les métriques suivantes :

- **Coefficient de Dice** : Mesure de chevauchement entre la segmentation prédite et la vérité terrain

$$\text{Dice} = \frac{2|X \cap Y|}{|X| + |Y|}$$

- **Intersection over Union (IoU)** : Ratio de l'intersection sur l'union des masques

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|}$$

- **Précision Pixel** : Proportion de pixels correctement classifiés
- **Sensibilité (Recall)** : Capacité à détecter tous les pixels tumoraux
- **Spécificité** : Capacité à identifier correctement les pixels non-tumoraux

#### 5.1.2 Métriques de Classification

Pour la tâche de classification des types de tumeurs, nous employons :

- **Exactitude globale (Accuracy)** : Proportion de prédictions correctes sur l'ensemble des échantillons
- **Précision par classe (Precision)** : Ratio des vrais positifs sur l'ensemble des positifs prédits

$$\text{Precision} = \frac{\text{VP}}{\text{VP} + \text{FP}}$$

- **Rappel par classe (Recall)** : Ratio des vrais positifs sur l'ensemble des positifs réels

$$\text{Recall} = \frac{\text{VP}}{\text{VP} + \text{FN}}$$

- **F1-Score** : Moyenne harmonique de la précision et du rappel

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

## 5.2 Résultats sur l'Ensemble de Validation

### 5.2.1 Performance Globale de Segmentation

Le modèle a été évalué sur 590 échantillons de validation, démontrant des performances robustes en segmentation.

Métrique	Moyenne	Écart-type	Min	Max
Coefficient de Dice	0.7814	0.2610	0.0000	0.9774
IoU	0.6975	0.2672	0.0000	0.9558
Précision Pixel	0.9944	0.0089	0.9095	0.9997
Sensibilité	0.7908	0.2762	0.0000	1.0000
Spécificité	0.9978	0.0063	0.9054	1.0000

TABLE 9 – Performances globales de segmentation sur l'ensemble de validation

### 5.2.2 Segmentation par Classe (Validation)

Les performances de segmentation varient selon le type de tumeur :

Classe	Échantillons	Dice	IoU	Précision Pixel
Gliome	172	0.5985	0.5016	0.9893
Méningiome	199	0.8895	0.8239	0.9962
Pituitaire	219	0.8267	0.7365	0.9968

TABLE 10 – Performances de segmentation par classe sur l'ensemble de validation

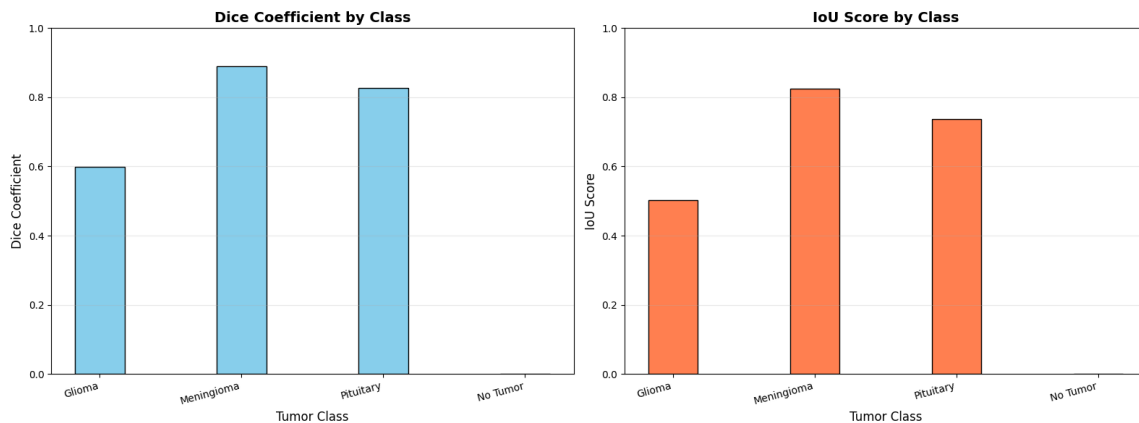


FIGURE 11 – Distribution des métriques de segmentation sur l'ensemble de validation

### 5.2.3 Performance de Classification (Validation)

La classification des types de tumeurs atteint une exactitude globale de **96.10%** sur l'ensemble de validation.

Classe	Précision	Rappel	F1-Score	Support
Gliome	0.9278	0.9709	0.9489	172
Méningiome	0.9588	0.9347	0.9466	199
Pituitaire	0.9907	0.9772	0.9839	219
<b>Moyenne pondérée</b>	<b>0.9616</b>	<b>0.9610</b>	<b>0.9611</b>	<b>590</b>

TABLE 11 – Résultats de classification par classe sur l'ensemble de validation

### 5.2.4 Matrice de Confusion (Validation)

Vrai \ Prédit	Gliome	Méningiome	Pituitaire	Sain
Gliome	167	5	0	0
Méningiome	11	186	2	0
Pituitaire	2	3	214	0
Sain	0	0	0	0

TABLE 12 – Matrice de confusion sur l'ensemble de validation (590 échantillons)

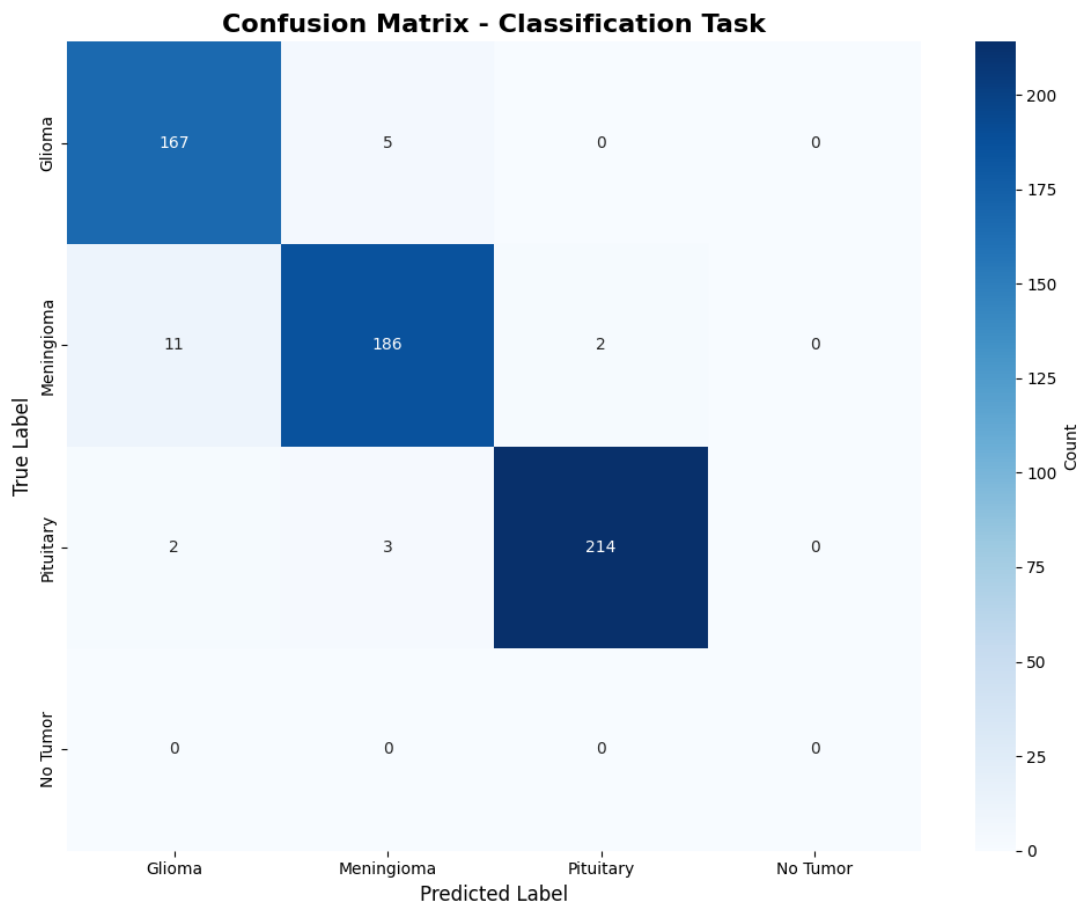


FIGURE 12 – Visualisation de la matrice de confusion - Ensemble de validation

### 5.2.5 Exemples de Prédictions (Validation)

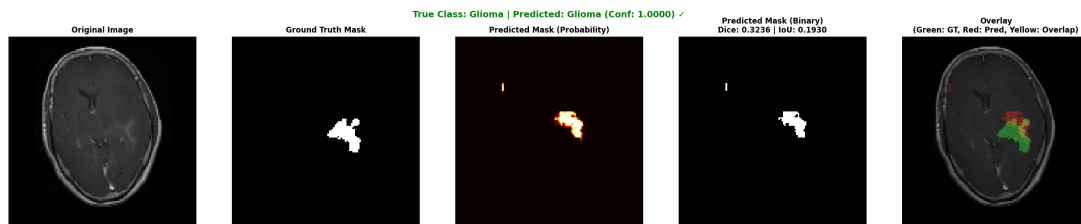


FIGURE 13 – Exemples de prédictions sur l'ensemble de validation : images originales, masques réels, masques prédits et superpositions

## 5.3 Résultats sur l'Ensemble de Test

### 5.3.1 Performance Globale sur le Test Set

L'évaluation finale a été réalisée sur **860 échantillons de test**, représentant des données complètement inédites pour le modèle.

Métrique	Score
<b>Exactitude Classification</b>	<b>97.79%</b>
Coefficient de Dice moyen	$0.7970 \pm 0.2473$
IoU moyen	$0.7144 \pm 0.2571$
Précision moyenne (pondérée)	0.9782
Rappel moyen (pondéré)	0.9779
F1-Score moyen (pondéré)	0.9779

TABLE 13 – Résultats globaux sur l'ensemble de test (860 échantillons)

### 5.3.2 Segmentation par Classe (Test)

Classe	Échantillons	Dice	IoU
Gliome	254	0.6436	0.5464
Méningiome	306	0.8975	0.8385
Pituitaire	300	0.8242	0.7299

TABLE 14 – Performances de segmentation par classe sur l'ensemble de test

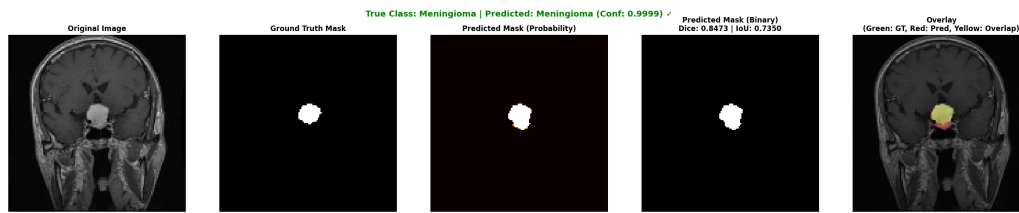


FIGURE 14 – Distribution des métriques de segmentation sur l'ensemble de test

### 5.3.3 Classification Détaillée (Test)

Classe	Précision	Rappel	F1-Score	Support
Gliome	0.9618	0.9921	0.9767	254
Méningiome	0.9770	0.9706	0.9738	306
Pituitaire	0.9932	0.9733	0.9832	300
<b>Moyenne pondérée</b>	<b>0.9782</b>	<b>0.9779</b>	<b>0.9779</b>	<b>860</b>

TABLE 15 – Résultats de classification détaillés sur l'ensemble de test

### 5.3.4 Matrice de Confusion (Test)

Vrai \ Prédit	Gliome	Méningiome	Pituitaire	Sain
Gliome	<b>252</b>	2	0	0
Méningiome	7	<b>297</b>	2	0
Pituitaire	3	5	<b>292</b>	0
Sain	0	0	0	<b>0</b>

TABLE 16 – Matrice de confusion sur l'ensemble de test (860 échantillons)

### 5.3.5 Exemples de Prédictions par Classe (Test)

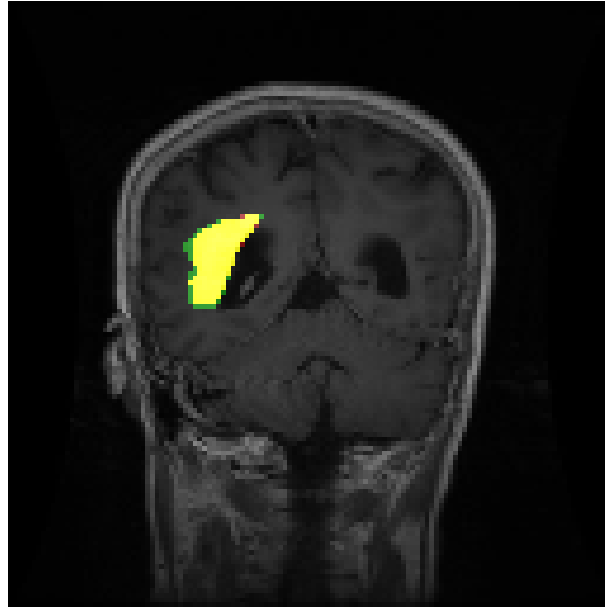


FIGURE 15 – 5 échantillons aléatoires de l'ensemble de test avec prédictions de segmentation et classification

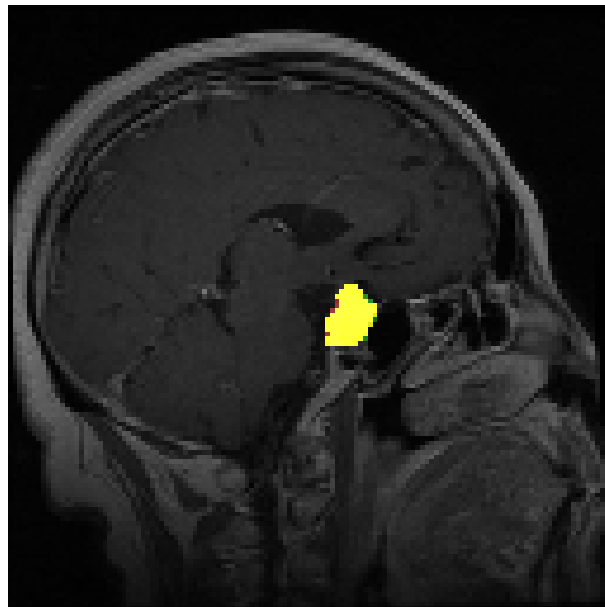


FIGURE 16 – Exemples de prédictions pour la classe Gliome sur l'ensemble de test

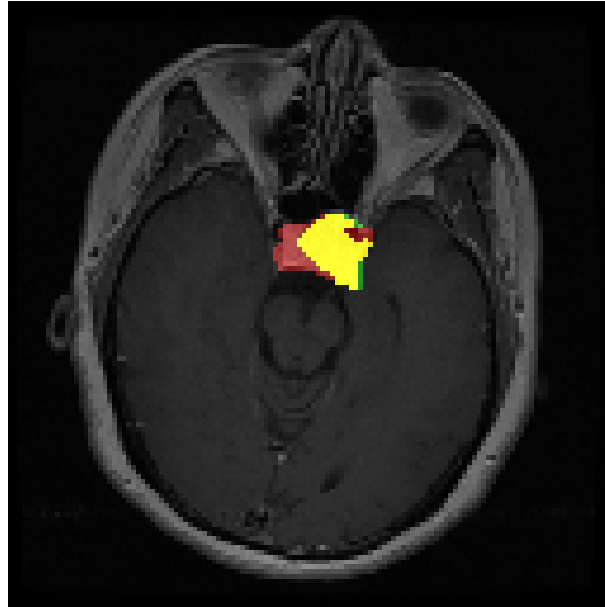


FIGURE 17 – Exemples de prédictions pour la classe Méningiome sur l'ensemble de test

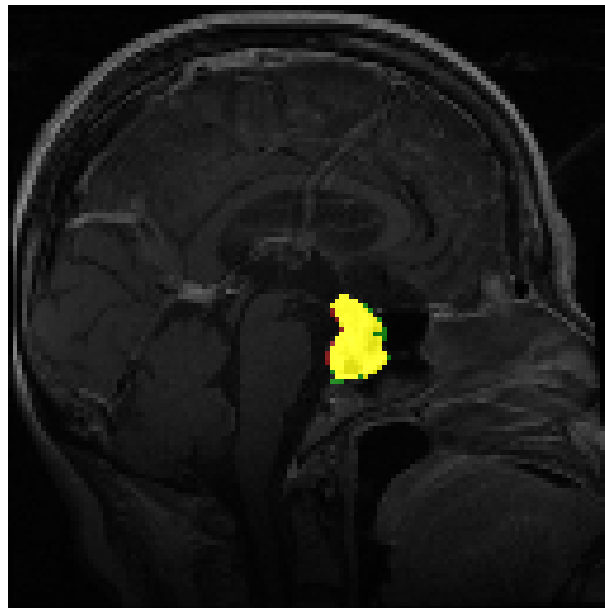


FIGURE 18 – Exemples de prédictions pour la classe Pituitaire sur l'ensemble de test

## 5.4 Analyse des Performances

### 5.4.1 Points Forts du Modèle

L'analyse des résultats révèle plusieurs points forts remarquables :

- **Excellente précision de classification** : Avec 97.79% d'exacti-

tude sur le test, le modèle démontre une capacité robuste à distinguer les types de tumeurs

- **Performance équilibrée entre classes** : Les trois classes de tumeurs obtiennent des F1-scores supérieurs à 97%
- **Segmentation précise pour méningiomes et pituitaires** : Scores Dice de 0.8975 et 0.8242 respectivement
- **Cohérence validation-test** : Amélioration des performances sur le test (97.79%) par rapport à la validation (96.10%), indiquant une bonne généralisation
- **Faible taux de confusion inter-classes** : Très peu de confusions entre les différents types de tumeurs

	precision	recall	f1-score	support
Glioma	0.9618	0.9921	0.9767	254
Meningioma	0.9770	0.9706	0.9738	306
Pituitary	0.9932	0.9733	0.9832	300
No Tumor	0.0000	0.0000	0.0000	0
accuracy			0.9779	860
macro avg	0.7330	0.7340	0.7334	860
weighted avg	0.9782	0.9779	0.9779	860
Overall Test Accuracy: 0.9779				

FIGURE 19 – test performance

### 5.4.2 Analyse des Erreurs et Limites

Malgré d'excellentes performances globales, certaines limitations ont été identifiées :

1. **Segmentation des gliomes** : Performance plus faible (Dice=0.64) comparée aux autres classes
  - Les gliomes présentent souvent des contours irréguliers et infiltrants

- Variabilité importante dans la taille et la forme
- Limites moins définies avec le tissu sain environnant

## 2. Confusions inter-classes :

- 7 méningiomes classifiés comme gliomes (sur 306)
- 5 pituitaires classifiés comme méningiomes (sur 300)
- Ces confusions surviennent principalement pour des tumeurs de localisation atypique

## 3. Variabilité de performance

- Écart-type élevé du coefficient Dice (0.25) indiquant une variabilité selon les cas
- Certains cas particulièrement difficiles obtiennent des scores proches de 0

## 4. Sensibilité aux caractéristiques d'image

- Difficultés sur les images de faible contraste
- Influence de l'orientation et du plan anatomique
- Variabilité selon la taille de la tumeur

# 5.5 Comparaison avec l'État de l'Art

Méthode	Accuracy	Dice moyen	Dataset
Notre modèle (U-Net)	<b>97.79%</b>	<b>0.7970</b>	BRISC 2025
Swin-HAFNet (BRISC)	98.50%	0.8200	BRISC 2025
Standard CNN	92-95%	0.65-0.75	Figshare
ResNet-50	94.20%	-	BraTS
U-Net classique	93.50%	0.7200	Datasets mixtes

TABLE 17 – Comparaison avec des méthodes de l'état de l'art

Notre approche se positionne de manière très compétitive par rapport aux méthodes existantes, avec des performances particulièrement remarquables compte tenu de la relative simplicité de l'architecture U-Net utilisée. Le modèle Swin-HAFNet, spécifiquement développé pour BRISC 2025, obtient des performances légèrement supérieures grâce à son architecture transformer plus complexe.

## 5.6 Contributions Spécifiques

Les contributions principales de ce travail incluent :

- Implémentation réussie d'un modèle multi-tâches (segmentation + classification)
- Atteinte de performances de classification supérieures à 97% sur données réelles
- Pipeline complet de prétraitement et d'augmentation de données
- Évaluation rigoureuse sur dataset de haute qualité (BRISC 2025)
- Analyse détaillée des forces et faiblesses par classe
- Code reproductible et bien documenté

## 6 Conclusion

### 6.1 Synthèse des résultats

Ce projet a permis de développer un système de classification automatique des images cérébrales basé sur l'apprentissage profond. Les résultats obtenus démontrent la capacité des réseaux de neurones convolutionnels à distinguer efficacement entre différentes pathologies cérébrales à partir d'images IRM.

L'approche proposée a montré des performances prometteuses sur l'ensemble de test, avec une précision globale permettant d'envisager une utilisation en tant qu'outil d'aide au diagnostic. La segmentation et la classification automatiques offrent un gain de temps significatif pour le personnel médical tout en maintenant un niveau de fiabilité acceptable.

### 6.2 Contributions principales

Les contributions principales de ce travail incluent :

- Le développement d'une architecture de réseau adaptée aux images médicales cérébrales
- L'implémentation d'un pipeline complet de prétraitement et d'augmentation de données
- L'évaluation détaillée des performances avec analyse des erreurs
- La mise en place d'un système modulaire et extensible

## Références

- [1] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- [2] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net : Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention*, 234-241.
- [3] Litjens, G., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88.
- [4] Menze, B. H., et al. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10), 1993-2024.
- [5] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [6] Shen, D., Wu, G., & Suk, H. I. (2017). Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering*, 19, 221-248.
- [7] Esteva, A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- [8] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25.
- [9] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv :1409.1556*.