

Évaluation de la dilatation ventriculaire post-hémorragique chez les prématurés sur IRM cérébrale assistée par intelligence artificielle - LIP6 & IMAG2 Necker

Autrice : Garance Lucas

Encadrement : Isabelle Bloch & Sarah Stricker

Avec l'aide de : Giammarco La Barbera, Enzo Bonnot & Thomas Isla

Contexte médical

Contexte médical

Hydrocéphalie

- Accumulation anormale de liquide céphalo-rachidien dans les ventricules => **dilatation**
- Possible hémorragie intra-ventriculaire
- Segmentation manuelle chronophage
- Volume ventriculaire important dans la prise de décision médicale

Segmentation automatique des ventricules et calcul des volumes

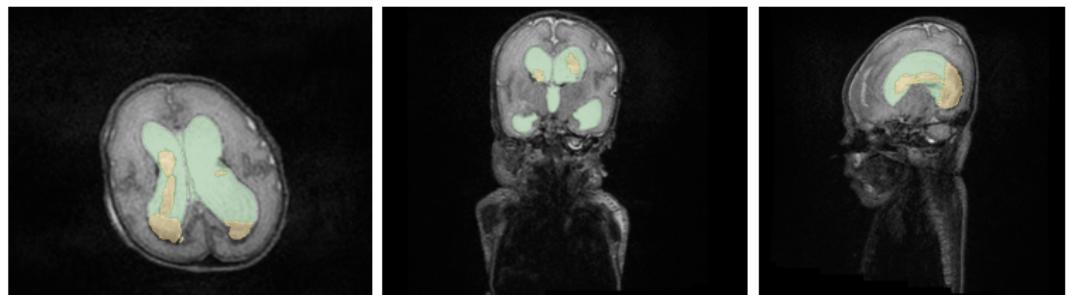


Figure 1 : IRM d'un sujet atteint d'hydrocéphalie avec hémorragie intra-ventriculaire
Vert : ventricules ; Jaune : hémorragie

Caractéristiques des images

Caractéristiques des images

Caractéristiques

- Différentes séquences d'acquisition des images IRM (T1/T2)
- Développement rapide des structures du cerveau chez les bébés => grandes variations dans les images
- Peu de données (encore moins pour sujets avec hémorragie)
- Images avec artefacts

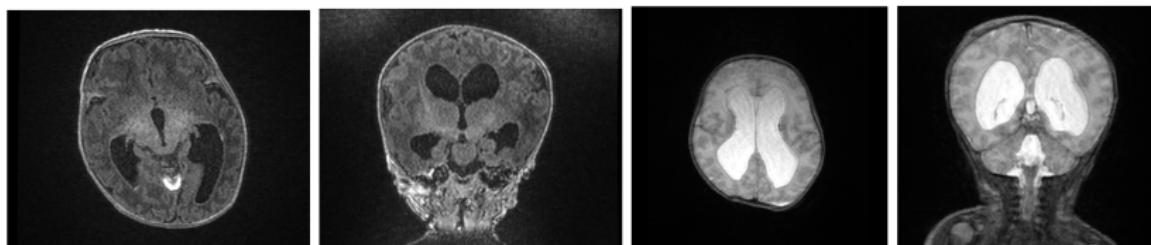


Figure 2 : Gauche : T1 IRM | Droite : T2 IRM - plans axial et coronal

Jeux de données

Jeux de données

Données de Necker

- Total: **40** sujets malades et 99 IRM
- 39 T1 et 60 T2 | 47 pre-op et 52 post-op
- **pre et post IRM très différentes => 60 sujets 'relatifs'**
- Âge gestationnel moyen à la naissance: 31,5 semaines ($\pm 5\text{sem}$) avec min 24sem et max 41sem.
- Dimensions images hétérogènes
- Segmentation manuelle : ventricules, hémorragie => **2 tissus**

Necker dataset	Pre-surgery only	Post-surgery only	Pre & post-surgery	Total
T1 only	12	3	0	15
T2 only	5	19	6 (12)	36
T1 & T2	6 (12)	6 (12)	6 (24)	48
Total	29	34	36	99

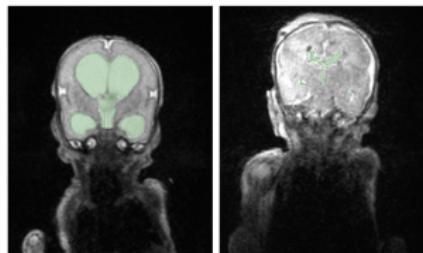


Table 1 : distribution des données de Necker | Figure 3: IRM pre-op et post-op

Jeux de données

Données de Zurich (*Payette et al., 2021*)

- **80** sujets avec 1 IRM **T2** (sains ou dilatation superficielle)
- Données foetales - de 20 à 33 semaines
- Dimensions images: $256 \times 256 \times 256$
- Segmentation : **7 tissus**

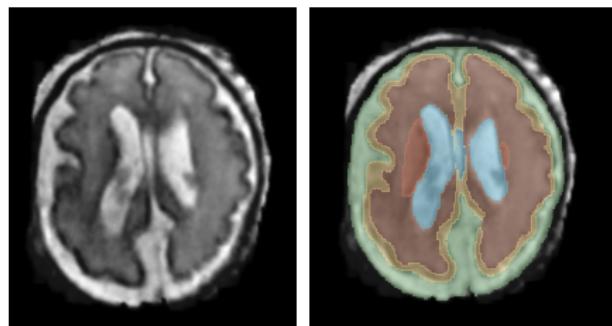


Figure 4 : Exemple Zurich: IRM et segmentation - plan axial
Segmentations : Liquide céphalo-rachidien externe, Matière grise, Matière blanche, Ventricules,
Cervelet, Matière grise profonde, Tronc cérébral

Jeux de données

Données de Vienne (*Payette et al., 2024*)

- **40** sujets avec 1 IRM **T2** (sains ou dilatation superficielle)
- Données foetales - de 20 à 36 semaines
- Dimensions images: $256 \times 256 \times 256$
- Segmentation : mêmes que Zurich dataset => **7 tissus**
- IRM avec organes de la mère

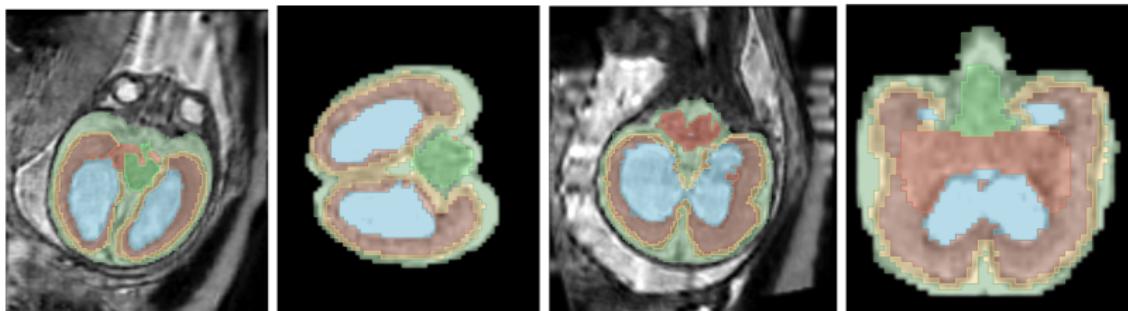


Figure 5 : Exemple Vienne : IRM T2 et segmentation avant et après pré-traitement - plans axial et coronal

État de l'art

État de l'art

SynthSeg (*Billot et al., 2023*)

- Entraîné que sur des **données synthétiques**
- **Agnostique aux contrastes**
- Échoue à segmenter IRM cerveau de bébé

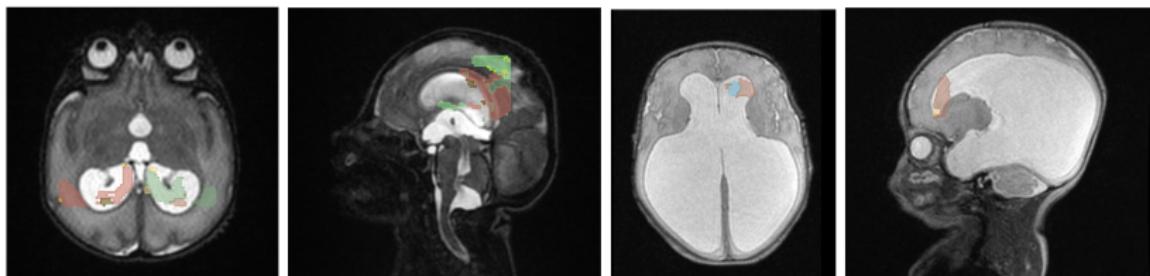


Figure 6 : Test inférence sur sujets Necker de SynthSeg pré-entraîné sur 37 structures

U-Net (*Valabregue et al., 2023*)

- U-Net: **modèle de référence** en segmentation images médicales
- Compare modèles SynthSeg et U-Net
- Entraîné sur des bébés (sains)

Méthodes

SynthSeg (Billot et al., 2023)

SynthSeg : vue d'ensemble

- Combinaison **modèle génératif + modèle segmentation**
- Génère données synthétiques à partir de segmentations
- Assume l'existence de segmentations

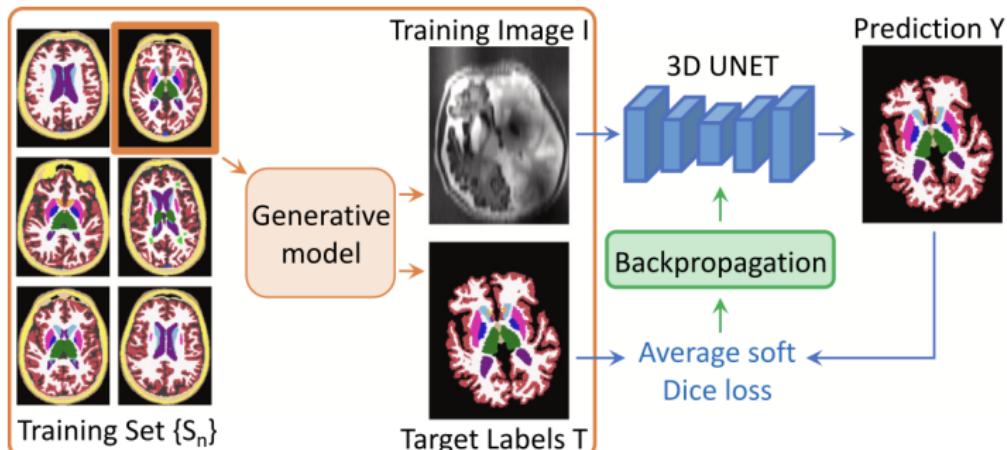


Figure 7 : Modèle SynthSeg : générateur + 3D U-Net

SynthSeg - Modèle génératif (*Billot et al., 2023*)

Modèle génératif

- Segmentation s - K tissus
- **Déformation** de s avec transformation spatiale aléatoire
- **Génération** donnée synthétique à partir GMM conditionné sur la segmentation déformée, avec contraste aléatoire
- **Simulation d'artefacts**
- Simulation de la **variabilité de la résolution**

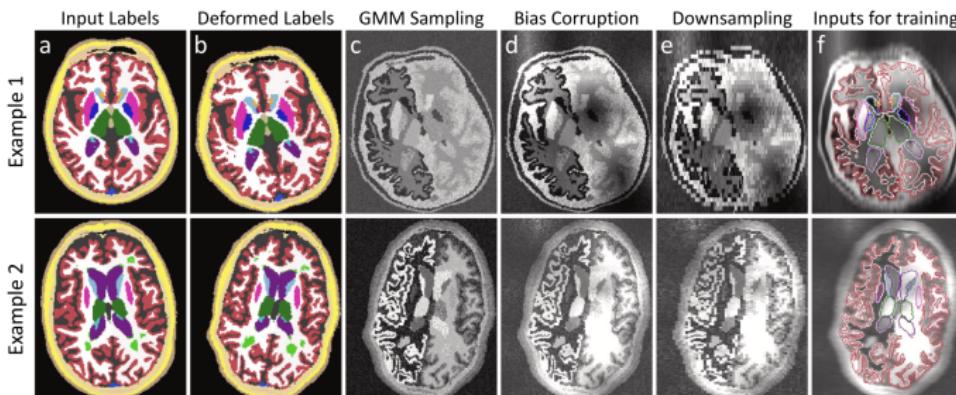


Figure 8 : Étapes intermédiaires du modèle génératif de SynthSeg

U-Net (*Ronneberger et al., 2015*)

2D et 3D U-Net

- 5 niveaux de convolutions et sous-échantillonnage
- **Combinaison Dice et Cross-Entropy loss**
- Pour modèle 3D : taille patches = $128 \times 128 \times 128$
- Pour modèle 2D : taille images = 256×256

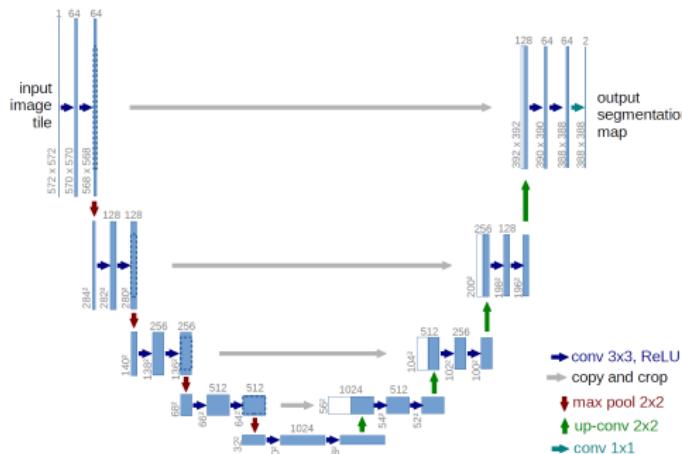


Figure 9 : U-Net architecture pour les modèles 2D et 3D

Pré-traitement des données

Pré-traitement des données

Pré-traitements

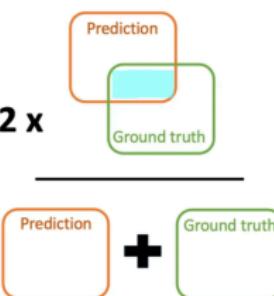
- **Masque** appliqué pour cacher les organes de la mère sur les données Vienne
- **Réduction du bruit** : filtre médian pour préserver les contours
- **Cropping** : réduire background focus structures sur les 3 [modèles 3D] ou 2 premières [modèle 2D] dimensions
- **Normalisation** : des intensités entre 0 et 1
- **Padding** : image cubique
- **Rééchantillonnage** : image $256 \times 256 \times 256$ [modèles 3D] ou $256 \times 256 \times .$ [modèle 2D]
- Pour modèle 2D : **extraction coupes 2D** (256×256) - coupe axiale - entraînement : 1/3

Résultats

Mesures d'évaluation

Mesures d'évaluation

- Score de Dice moyen sur les structures : $\frac{2 \times |P \cap R|}{|P| + |R|}$
P : segmentation prédite et R : segmentation manuelle (référence)
- Score de Dice sur les ventricules
- Précision et rappel
- Volume ventricules (cm^3)
- Volume erreur : $volume\ segmentation\ manuelle - volume\ prédit\ (cm^3)$

$$\text{Dice} = \frac{2 \times \text{Area of overlap}}{\text{Total area}} = \frac{2 \times \text{Area of overlap}}{\text{Prediction} + \text{Ground truth}}$$


Résultats préliminaires

	Scores	2D U-Net	3D U-Net
Entrainé sur FeTA - 7 tissus	Mean V_Dice_score	0.543	0.757
	Standard deviation	0.233	0.18
	Precision	0.547	0.734
	Rappel	0.487	0.667
Entrainé sur FeTA - 1 tissu (ventricules)	Mean V_Dice_score	0.545	0.572
	Standard deviation	0.292	0.267
	Precision	0.642	0.773
	Rappel	0.521	0.52

Figure 10 : Comparaison performances 3D et 2D
U-Net testés sur FeTA avec 7 et 1 tissus

=> Performances 3D U-Net >> performances 2D U-Net

Expériences

Expérience (E1)

- Comparaison 3D U-Net et Synthseg sur 7 tissus
- Entrainement sur FeTA : 100 IRM T2 (70 Zurich et 30 Vienne)

Expérience (E2)

- Comparaison 3D U-Net et Synthseg sur 1 tissu
- Entrainement sur FeTA : 100 IRM T2 (70 Zurich et 30 Vienne)

Expérience (E3)

- Comparaison 3D U-Net et Synthseg sur 1 tissu
- Entrainement sur FeTA + Necker : 148 IRM T2 (70 de Zurich, 30 de Vienne et 48 de Necker)

Test sur 44 IRM (32 T2 et 12 T1) : 10 Zurich, 10 Vienne et 24 Necker

Résultat entrainement sur FeTA avec 7 tissus (E1)

		Structures Dice score	Precision score	Recall score	Ventricles Dice score	Mean ventricles volumes	Mean ventricles error	
SynthSeg 1 pre-training	Total mean	0.336	0.477	0.545	0.394	103.621	-182.715	
	Total standard deviation	0.17	0.165	0.151	0.199			
3D U-Net 2 dataloaders	Total mean	0.556	0.54	0.717	0.628	189.132	-97.2	
	Total standard deviation	0.246	0.176	0.138	0.245			
Zurich dataset Vienna dataset Necker dataset Necker T1 Necker T2 Pre-op T2 Necker Post-op T2 Necker								
SynthSeg 1 pre-training	Mean V_Dice_score	0.683	0.204	0.293	0.231	0.356	0.304	0.409
	Std V_Dice_score	0.257	0.124	0.215	0.214	0.206	0.258	0.143
3D U-Net 2 dataloaders	Mean V_Dice_score	0.844	0.67	0.37	0.092	0.648	0.605	0.691
	Std V_Dice_score	0.105	0.255	0.376	0.25	0.246	0.305	0.188

Figure 11 : Résultats meilleurs modèles 3D U-Net et SynthSeg entraînés sur 7 tissus sur FeTA - volumes et erreurs donnés en cm^3

3D U-Net vs SynthSeg

- Performances 3D U-Net >> performances SynthSeg
- SynthSeg robuste aux différentes séquences / 3D U-Net n'arrive pas à segmenter T1
- Meilleurs résultats sur les données de Zurich

Résultat entraînement sur FeTA avec 1 tissu (E2)

		Structures Dice score	Precision score	Recall score	Ventricles Dice score	Mean ventricles volumes	Mean ventricles error
SynthSeg 1 pre-training	Total mean	0.213	0.143	0.615	0.213	1089.346	803.01
	Total standard deviation	0.175	0.129	0.322	0.175		
3D U-Net 2 dataloaders	Total mean	0.572	0.773	0.52	0.572	154.177	-132.159
	Total standard deviation	0.267	0.236	0.276	0.267		
	Zurich dataset	Vienna dataset	Necker dataset	Necker T1	Necker T2	Pre-op T2 Necker	Post-op T2 Necker
SynthSeg 1 pre-training	Mean V_Dice_score	0.209	0.305	0.125	0.019	0.232	0.357
	Std V_Dice_score	0.114	0.233	0.178	0.037	0.199	0.161
3D U-Net 2 dataloaders	Mean V_Dice_score	0.819	0.564	0.333	0.094	0.572	0.529
	Std V_Dice_score	0.18	0.276	0.345	0.251	0.247	0.314

Figure 12 : Résultats des meilleurs modèles 3D U-Net et SynthSeg entraînés sur 1 tissu sur FeTA - volumes et erreurs donnés en cm^3

3D U-Net vs SynthSeg

- Performances 3D U-Net >> performances SynthSeg
- SynthSeg non robuste aux différentes séquences
- SyntheSeg sur-segmente ventricules quand entraîné sur 1 tissu (et sous-segmente quand entraîné sur 7 tissus)

Résultat entraînement sur FeTA + Necker avec 1 tissu (E3)

		Structures Dice score	Precision score	Recall score	Ventricles Dice score	Mean ventricles volumes	Mean ventricles error
SynthSeg 1 pre-training	Total mean	0.221	0.144	0.789	0.221	1663.398	1377.063
	Total standard deviation	0.161	0.119	0.192	0.161		
3D U-Net 2 dataloaders	Total mean	0.605	0.82	0.556	0.605	158.513	-127.822
	Total standard deviation	0.294	0.23	0.305	0.294		

	Zurich dataset	Vienna dataset	Necker dataset	Necker T1	Necker T2	Pre-op T2 Necker	Post-op T2 Necker	
SynthSeg 1 pre-training	Mean V_Dice_score	0.248	0.319	0.096	0.052	0.141	0.119	0.082
	Std V_Dice_score	0.166	0.22	0.097	0.053	0.113	0.103	0.096
3D U-Net 2 dataloaders	Mean V_Dice_score	0.819	0.564	0.432	0.088	0.776	0.788	0.764
	Std V_Dice_score	0.18	0.276	0.427	0.248	0.249	0.319	0.186

Figure 13 : Résultats des meilleurs modèles 3D U-Net et SynthSeg entraînés sur 1 tissu sur FeTA + Necker - volumes et erreurs donnés en cm^3

3D U-Net vs SynthSeg

- Performances 3D U-Net >> performances SynthSeg
- SynthSeg non robuste aux différentes séquences
- Gain de performances sur images T2 Necker pour modèle 3D U-Net

Conclusions et perspectives

Conclusions et perspectives

Conclusions

- Performances 3D U-Net >> 2D U-Net et SynthSeg
- Comportement de SynthSeg sur la robustesse dans variation séquences n'est pas vérifié quand entraîné sur 1 tissu
- Meilleures performances des modèles sur les données de Zurich
- En moyenne, 3D U-Net sous-segmente ventricules et SynthSeg sur-segmente ventricules

Perspectives

- Améliorer 3D U-Net (augmentation de données)
- Modifier modèle SynthSeg pour qu'il voit des données réelles lors de l'entraînement et de la validation
- Affiner l'évaluation sur les volumes

Merci pour votre attention !

Et merci à Isabelle, Sarah ainsi que l'équipe de Necker et celle de LFI :)

Annexes

Annexes

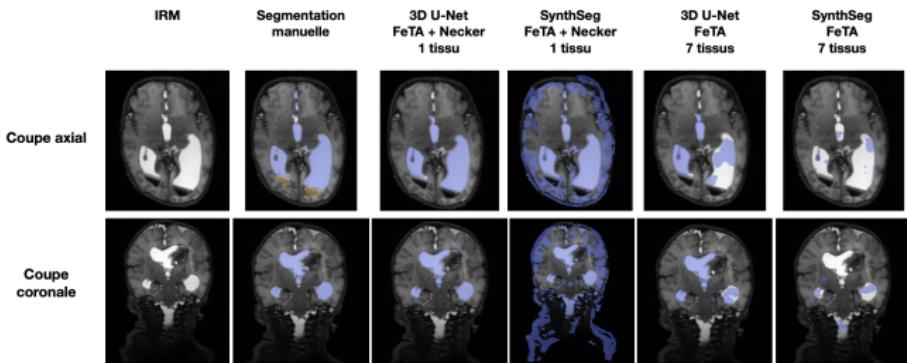


Figure 14 : Exemple sujet Necker pre-op

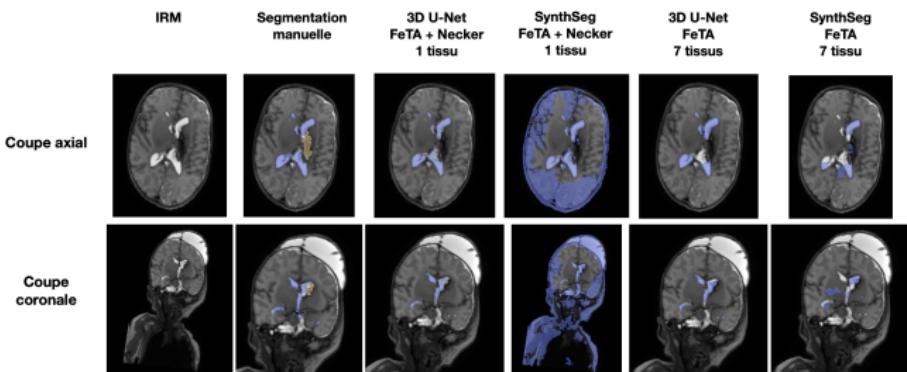


Figure 15 : Exemple sujet Necker post-op