

IA et IRM pour mieux comprendre le cerveau après un AVC

Retour d'expérience

Francesca Galassi

Empenn - Inria / IRISA

Journée ARDoISE 2026 - IA, données et santé

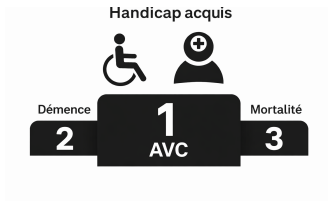
Contexte clinique

Après un AVC : un enjeu majeur de santé publique

- ⇒ Survient à tout âge, mais avec une incidence en forte hausse liée au **vieillessement de la population**



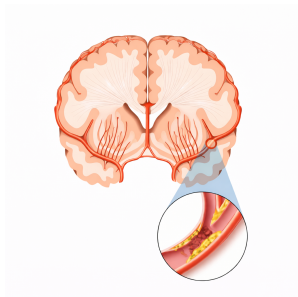
- ⇒ Séquelles motrices et cognitives, **impact durable sur la qualité de vie** et le parcours de soins



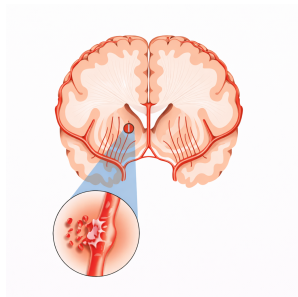
Un AVC : deux mécanismes physiopathologiques

- ⇒ Interruption brutale de la circulation sanguine cérébrale
- ⇒ Deux mécanismes distincts, conduisant à des **déficits neurologiques parfois comparables**

AVC ischémique
(forme majoritaire, ~80–85%)

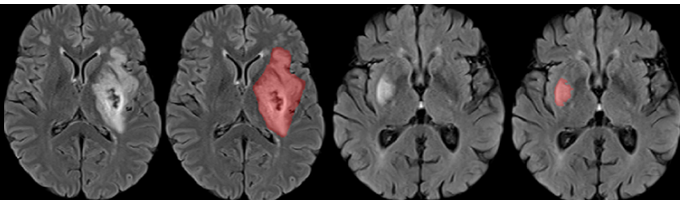


AVC hémorragique
(~15–20%)



IRM post-AVC : la lésion comme point d'entrée

- ⇒ Mesurer le **volume** et la **localisation** de la lésion
- ⇒ Forte variabilité inter-patients et contours parfois ambigus



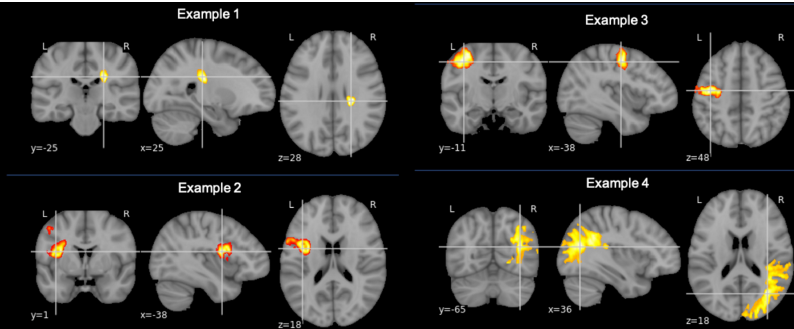
IRM (séquence FLAIR) post-AVC : coupes axiales avec masque de lésion superposé.

- ⇒ Besoin: des mesures **fiables** et **reproductibles** pour comparer des patients

Projet 1 - StrokeSeg

StrokeSeg : segmentation automatique des lésions post-AVC

- ⇒ Approche d'**IA supervisée** entraînée sur des annotations expertes
- ⇒ **Sortie** : masque de lésion → volume et localisation
- ⇒ Adapté aux **données cliniques réelles** (multi-sites, multi-protocoles)

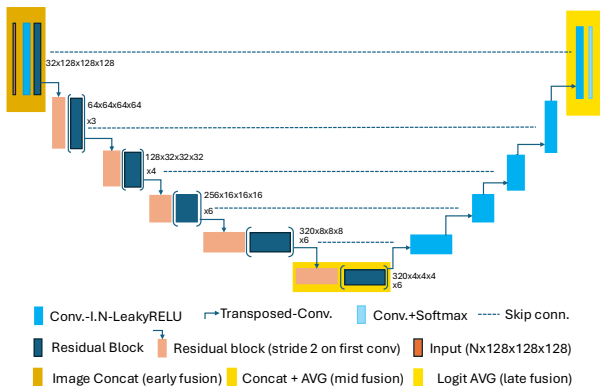


Exemples de masques de lésion post-AVC illustrant la variabilité inter-patients.

Source: Aamodt et al., *Front. Aging Neurosci.*, 2021.

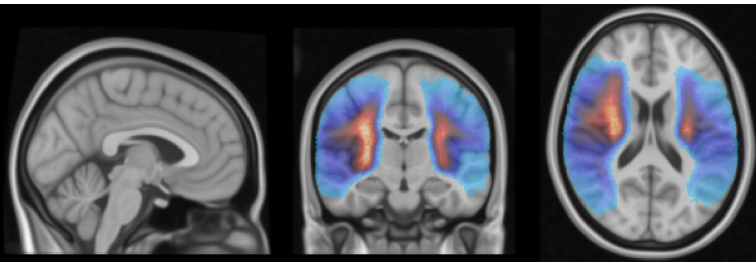
StrokeSeg : modèle de segmentation

- ⇒ Segmentation supervisée **automatique** des lésions
- ⇒ Architecture standard en imagerie médicale : **nnU-Net**



Architecture du modèle de segmentation (nnU-Net / ResEnc). Source: Di Matteo et al., Pattern Recognition Letters, 2025.

StrokeSeg : jeu de données de référence

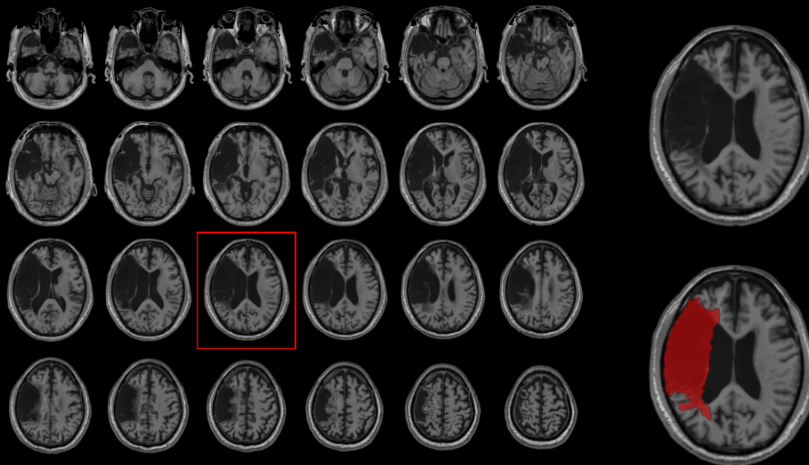


Carte de fréquence des lésions. Source : <https://neurovault.org/images/706022/>.

ATLAS v2.0 (IRM T1, public)

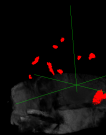
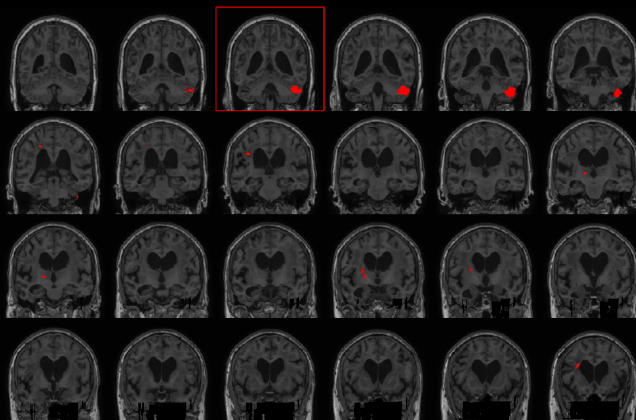
- 655 sujets d'entraînement, annotations expertes manuelles
- Images normalisées (MNI), résolution 1 mm isotrope
- Volumes de lésions très hétérogènes (0.01 à 500 mL)

Exemple: lésion unique, volumineuse (sub-r004s004)



Coupes axiales (T1) + masque GT (annotation experte)

Exemple : multi-lésions (sub-r009s126)



Vue 3D, 9 lésions

Coupes coronales (T1) + masque GT (annotation experte)

Cohorte clinique interne : performance globale

Point de départ : modèle entraîné sur ATLAS v2.0 (CV 5-fold) : Dice = 0.64, F1 = 0.67 (performances comparables à l'état de l'art).

Dice = recouvrement ; F1 lésion = détection ; ASD = erreur frontière (mm).

Agrégation : moyenne voxel-par-voxel des prédictions de modèles entraînés en validation croisée.

Modèle	Dice ↑	F1 lésion ↑	ASD (mm) ↓
Modèle de base	0.68	0.73	6.95
Modèle de base + agrégation	0.70 ^{***}	0.79 ^{***}	3.47
Adaptation sur données cliniques (T1)	0.71 [*]	0.76 [*]	7.19
Adaptation (T1) + agrég	0.72 ^{***}	0.78 ^{**}	3.37
T1+FLAIR fusion précoce + agrég	0.74 ^{***}	0.79 [*]	3.14
T1+FLAIR fusion interm + agrég	0.69	0.77 [*]	7.68
T1+FLAIR fusion tardive + agrég	0.75^{***}	0.80^{**}	2.94

Wilcoxon apparié (bilatéral), vs **modèle de base** : ^{*} $p < 0.05$, ^{**} $p < 0.01$, ^{***} $p < 0.001$.

- Cohorte interne (IRM T1 et FLAIR), 45 sujets, multi-scanners
- Volumes de lésions hétérogènes (0.04 à 300 mL)

→ **FLAIR + fusion tardive + agrégation** ⇒ meilleure performance moyenne.

Cohorte clinique interne : Dice selon la taille de lésion

Modèle	Petites(≤ 15 mL, $n=14$)	Moyennes(15-70 mL, $n=18$)	Grandes(> 70 mL, $n=13$)
Modèle de base + agrég	0.60	0.71	0.79
Adaptation (T1) + agrég	0.58	0.76 ^{**}	0.82 ^{**}
T1+FLAIR fusion précoce + agrég	0.59	0.78 ^{***}	0.82 ^{***}
T1+FLAIR fusion interm + agrég	0.48	0.77 ^{**}	0.82
T1+FLAIR fusion tardive + agrég	0.65 ^{**}	0.77 ^{***}	0.83 [*]

Wilcoxon apparié (bilatéral), vs **modèle de base + agrégation** (référence). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

- **Gain principal sur les petites lésions** : late fusion plus robuste aux contrastes faibles et au désalignement.
- **Moyennes / grandes** : amélioration plus modérée, toutes les fusions se rapprochent.

- ⇒ Outil **open-source** / **open access** pour la **recherche** post-AVC
- ⇒ Documentation complète pour **reproduire** le pipeline (données → masque)
- ⇒ Pensé pour la **reproductibilité** et la **réutilisation** sur d'autres cohortes



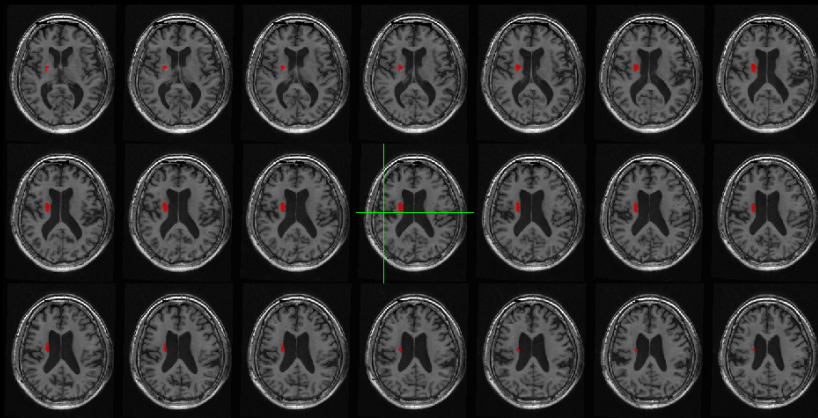
StrokeSeg (research only) -
<https://strokeseg-doc.readthedocs.io/>

Usage recherche uniquement (pas un dispositif médical).

Projet 2 - Détection d'anomalies non supervisée

IRM post-AVC : au-delà de la lésion focale

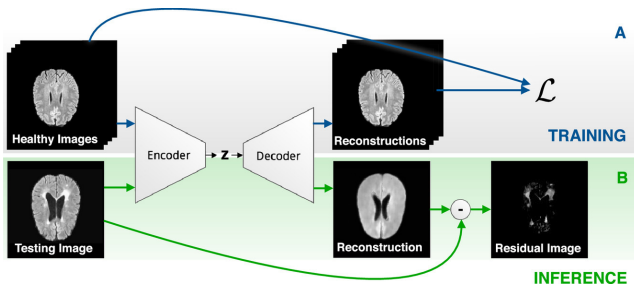
- L'IRM post-AVC montre des **altérations multiples**, pas uniquement la lésion
- Altérations secondaires : **atrophie, dilatation ventriculaire, changements corticaux**
- Plusieurs de ces marqueurs sont associés à des **scores cliniques et cognitifs**



Exemple ATLAS : dilatation ventriculaire et élargissement des sillons, visibles mais non annotés.

Apprendre la normalité

- ⇒ Idée clé : apprendre la **variabilité normale** sur des IRM saines
- ⇒ À l'inférence : produire une **reconstruction saine** de l'image patient
- ⇒ Carte d'anomalie = image réelle – reconstruction



Principe général de la détection d'anomalies par reconstruction (Baur et al., 2020).

État de l'art UAD - en imagerie cérébrale

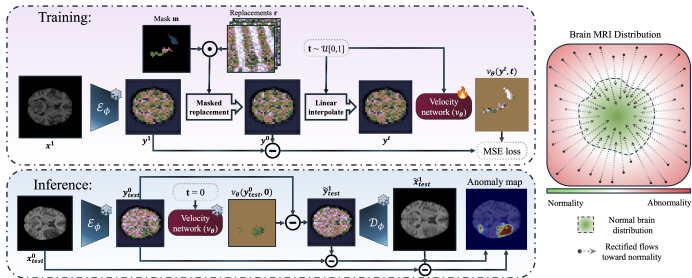
Pathologie	Famille	Dice	AUPRC	Réfs clés
Tumeurs BraTS (4 séq.) ; in-house	AE	0.390– 0.773	0.300– 0.833	Lu'24 ; Kascenas'23
	VAE	0.301– 0.650	0.279– 0.819	Lambert'21 ; Lüth'23
	GAN	0.630– 0.770	–	Nguyen'21
	Diffusion / Flow	0.371–0.851	0.417–0.590	Beizaae'26 ; Iqbal'23
SEP MSSEG ; MSLUB (1–4 séq.)	AE	0.173– 0.605	0.203	Baur'18 ; Ghorbel'23
	VAE	0.133– 0.378	0.272	Pinaya'22
	GAN	0.482	–	Dey'21
	Diffusion / Flow	0.055– 0.247	0.067– 0.106	Behrendt'23 ; Pinaya'22
WMH WMH 2017 (3 séq.)	VAE	0.133– 0.463	0.320	Lambert'21 ; Pinaya'22
	Diffusion / Flow	0.132– 0.569	–	Fontanella'24
AVC (lésion) ATLAS ; ISLES (1–2 séq.)	AE	–	0.705	Qu'26
	VAE	0.331	0.256– 0.549	Bengs'21 ; Lüth'23
	Diffusion / Flow	0.148–0.416	0.145	Beizaae'26 ; Bercea'23

→ La majorité des benchmarks UAD mesurent la performance sur des **anomalies focales**

→ **Objectif** : **évaluer un modèle existant** et tester l'impact des **données d'entraînement** sur ce qu'il détecte.

Modèle utilisé : REFLECT - principe général

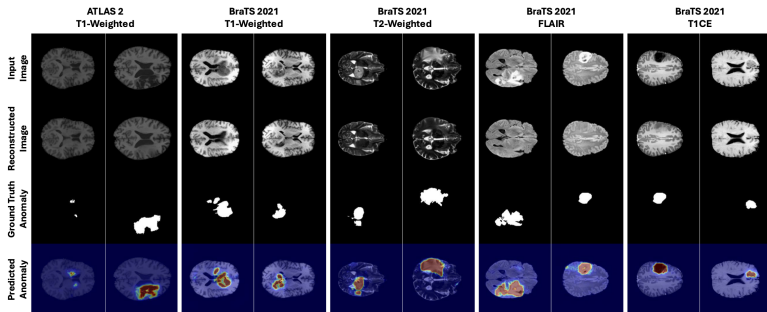
- ⇒ Modèle génératif non supervisé basé sur les **rectified flows**
- ⇒ Apprend un **transport continu** vers une distribution d'IRM saines
- ⇒ Les anomalies correspondent aux **écarts à la normalité apprise**



Apprentissage auto-supervisé à partir d'anomalies synthétiques locales générées dans l'espace latent, permettant l'apprentissage d'un transport direct vers une distribution saine (Beizaee et al., 2025).

Modèle utilisé : REFLECT - exemple de sortie

- ⇒ Reconstruction plus proche de la **normalité anatomique**
- ⇒ Carte d'anomalie : **écart à la normalité apprise**
- ⇒ Mise en évidence de lésions et d'

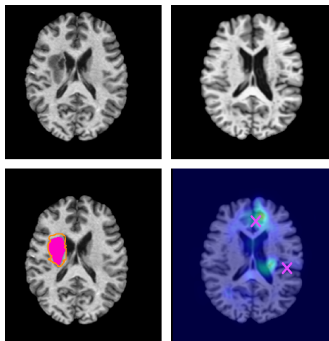


Exemples de reconstructions et cartes d'anomalie produites par REFLECT. De haut en bas : image d'entrée, reconstruction, référence (si disponible), carte d'anomalie prédite.

Détection d'anomalies : impact des données d'entraînement

Évaluation objet-level (FROC) sur la cohorte ATLAS. $FPPI = 0.25-1.5$, seuil $T = 0.036$.

Modèle	Lésions	Anomalies non-lésionnelles	Toutes anomalies
Entraîné sur ATLAS	0.91	0.43	0.72
Entraîné sur IXI (sains)	0.94	0.62	0.84



Segmentation de la lésion (orange : référence, rose : prédiction) et cartes d'anomalie avec annotations expertes (X)

- ⇒ La détection non supervisée permet d'explorer des anomalies **au-delà de la lésion focale**
- ⇒ Apprendre sur des cerveaux **réellement sains** améliore la sensibilité (surtout non-lésionnel)
- ⇒ Les régions détectées sont globalement cohérentes avec l'**interprétation clinique**

Limite et piste méthodologique

- ⇒ Entraînement basé sur des **corruptions synthétiques locales** ⇒ biais vers le **focal**
- ⇒ Donc : très bon pour des anomalies **contrastées**, plus difficile pour des altérations **diffuses** (atrophie, morphologie)
- ⇒ Pistes : corruptions **plus morphologiques** + passage au **3D** (contexte spatial)

Conclusions

Conclusion

- ⇒ L'IRM post-AVC contient des **informations riches**, mais **partiellement annotées**
- ⇒ Les approches **supervisées** structurent la **lésion focale** de façon robuste
- ⇒ Les modèles **génératifs non supervisés** ouvrent une voie **exploratoire au-delà**

Perspectives de recherche

- ⇒ Passer de modèles **2D à 3D** pour intégrer le contexte spatial
- ⇒ Concevoir des **perturbations morphologiques** plus réalistes à l'entraînement
- ⇒ Explorer la **complémentarité multimodale** (IRM + données cliniques)
- ⇒ Utiliser ces représentations pour des **modèles prédictifs interprétables**

Objectif : mieux comprendre et suivre les trajectoires post-AVC

Remerciements et collaborations



- **Clinique** : CHU de Rennes *données, relectures, questions terrain*
- **Industriel** : Siemens Healthineers *pont vers l'usage, contraintes et transfert*

Merci tout particulièrement à l'équipe projet :

Youwan Mahé *conduite de l'étude UAD & StrokeSeg*

Alessandro Di Matteo *stratégie segmentation*

Yann Kerverdo *packaging StrokeSeg*

Florent Leray *ingénierie, packaging & documentation (StrokeSeg)*

Stéphanie Leplaideur *médecine de rééducation, relecture clinique et annotations*

Elisa Fromont *méthodes UAD*

Elise Banner *interprétation des données IRM et expertise imagerie*

