

# Analyse critique du projet: un signal dénaturé !

## 1. Analyse EEG d'un patient

### 1. Analyse EEG original

#### CONTEXTE : QU'EST-CE QU'UN EEG ?

L'électroencéphalogramme (EEG) mesure l'activité électrique du cerveau via des électrodes placées sur le cuir chevelu. Imaginez votre cerveau comme une ville avec 86 milliards de neurones qui communiquent par signaux électriques. L'EEG capture ces "conversations" neurales en temps réel, comme un micro qui enregistrerait le brouhaha d'une foule.

**Pourquoi c'est révolutionnaire pour la psychiatrie ?** Les troubles comme la dépression, ADHD, ou TOC modifient ces patterns électriques. C'est comme si chaque maladie mentale avait sa propre "signature électrique" dans le cerveau.

---

#### CE QUE CONTIENT NOTRE FICHER TDBRAIN

Données massives et précises :

- **240 secondes d'enregistrement** = 4 minutes de cerveau en action
- **60,003 mesures par canal** = une photo du cerveau toutes les 4 millisecondes
- **26 canaux EEG + 7 canaux auxiliaires** = 33 "micros" différents sur la tête
- **3 Mo d'information brute** par patient = équivalent d'un livre de 500 pages

Résolution temporelle exceptionnelle :

Avec une mesure toutes les **4 millisecondes**, on capture :

- Les battements de cœur (1000ms entre battements)
  - Les clignements d'yeux (150-300ms)
  - Les micro-expressions faciales (40-200ms)
  - **L'activité neuronale pure (1-10ms)** ← C'EST ÇA QU'ON VEUT !
- 

#### TOPOGRAPHIE CÉRÉBRALE : UN CASQUE HIGH-TECH

Le système **10-20** place les électrodes selon un standard international, comme un casque de réalité virtuelle ultra-précis :

### Zones frontales (Fp1, Fp2, F3, F4, F7, F8) :

- **Rôle** : Fonctions exécutives, prise de décision, personnalité
- **Pathologies** : Dépression, troubles bipolaires, ADHD
- **Analogie** : Le "CEO" du cerveau qui prend les décisions

### Zones centrales (C3, Cz, C4) :

- **Rôle** : Contrôle moteur, mouvement volontaire
- **Pathologies** : Parkinson, troubles moteurs
- **Analogie** : Le "chef d'orchestre" qui coordonne les mouvements

### Zones pariétales (P3, Pz, P4, P7, P8) :

- **Rôle** : Intégration sensorielle, attention spatiale
- **Pathologies** : Troubles de l'attention, dyslexie
- **Analogie** : Le "GPS" du cerveau qui nous situe dans l'espace

### Zones occipitales (O1, Oz, O2) :

- **Rôle** : Traitement visuel
- **Pathologies** : Troubles visuels, hallucinations
- **Analogie** : Le "processeur graphique" du cerveau

### Zones temporales (T7, T8) :

- **Rôle** : Audition, mémoire, émotions
- **Pathologies** : Épilepsie, troubles de la mémoire
- **Analogie** : Le "disque dur" émotionnel du cerveau



## LES 5 FRÉQUENCES DU CERVEAU : UNE SYMPHONIE ÉLECTRIQUE

Le cerveau produit simultanément 5 "notes musicales" électriques différentes :



### Ondes Delta (0.5-4 Hz) : Le sommeil profond

- **Fréquence** : 0.5-4 battements par seconde
- **État** : Sommeil profond, réparation cérébrale
- **Pathologie** : Troubles du sommeil, traumatismes crâniens
- **Analogie** : Les graves d'une basse qui résonnent lentement

## ● Ondes Theta (4-8 Hz) : La créativité

- **Fréquence** : 4-8 battements par seconde
- **État** : Créativité, méditation, rêverie
- **Pathologie** : ADHD (trop de theta = inattention)
- **Analogie** : Le rythme d'une ballade tranquille

## ● Ondes Alpha (8-13 Hz) : La relaxation

- **Fréquence** : 8-13 battements par seconde
- **État** : Relaxation éveillée, yeux fermés
- **Pathologie** : Anxiété, dépression (alpha déséquilibré)
- **Analogie** : Le tempo d'une chanson pop classique

## ● Ondes Beta (13-30 Hz) : La concentration

- **Fréquence** : 13-30 battements par seconde
- **État** : Concentration, résolution de problèmes
- **Pathologie** : Anxiété (trop de beta = stress)
- **Analogie** : Le rythme rapide d'une musique électronique

## ● Ondes Gamma (30-100 Hz) : L'illumination

- **Fréquence** : 30-100 battements par seconde
- **État** : Moments "eureka", conscience unifiée
- **Pathologie** : Schizophrénie, troubles cognitifs
- **Analogie** : Les aigus perçants d'une mélodie complexe



## RICHESSSE INFORMATIONNELLE : UN TRÉSOR DE DONNÉES

### Entropie de Shannon : La "surprise" du cerveau

- **Valeur mesurée** : 5.02 bits par canal
- **Signification** : Chaque mesure EEG contient 5 bits d'information unique
- **Analogie** : C'est comme si chaque mot d'un livre contenait 5 "surprises" différentes
- **Implication** : Plus l'entropie est élevée, plus le signal est complexe et riche

### Corrélations inter-canaux : La connectivité naturelle

- **Valeur mesurée** : 0.361 de corrélation moyenne
- **Signification** : Les régions cérébrales sont modérément connectées
- **Analogie** : Comme un orchestre où chaque musicien écoute les autres (36% de synchronisation)
- **Implication clinique** : Les pathologies modifient ces connexions naturelles

## Volume de données : Une mine d'or informationnelle

- **3 Mo par patient** = 24 millions de bits d'information
- **60,000 mesures par seconde** sur 26 canaux = 1.5 million de points de données
- **Comparaison** : C'est 100x plus de données qu'une radiographie !

---

## ARTEFACTS : LES "PARASITES" DU SIGNAL CÉRÉBRAL

### Électrooculogramme (EOG) - VPVA, VNVB, HPHL, HNHR :

- **Origine** : Mouvements des yeux et clignements
- **Amplitude** : 10x plus forte que le signal EEG
- **Problème** : Masque l'activité cérébrale frontale
- **Analogie** : Comme un klaxon qui couvre une conversation

### Électrocardiogramme (ECG) - Erbs :

- **Origine** : Battements cardiaques (1 par seconde)
- **Propagation** : Se diffuse dans tout le corps, y compris la tête
- **Problème** : Rythme régulier qui "pollue" toutes les fréquences
- **Analogie** : Comme un métronome constant qui perturbe une mélodie

### Électromyogramme (EMG) - OrbOcc, Mass :

- **Origine** : Contractions musculaires (mâchoires, front, cou)
- **Fréquence** : Couvre toutes les bandes EEG
- **Problème** : Impossible à distinguer de l'activité neuronale
- **Analogie** : Comme du bruit de fond qui couvre toutes les conversations

---

## POTENTIEL DIAGNOSTIQUE : POURQUOI C'EST RÉVOLUTIONNAIRE

### Signatures temporelles fines :

Chaque trouble psychiatrique a ses patterns temporels uniques :

- **Dépression** : Asymétrie alpha frontal (F3 vs F4)
- **ADHD** : Excès de theta + déficit beta
- **TOC** : Hyperconnectivité fronto-striatal
- **Résolution nécessaire** : Millisecondes (pas secondes !)

## Dynamiques trans-fréquentielles :

Les vraies pathologies se cachent dans les interactions entre fréquences :

- **Couplage phase-amplitude** : Quand gamma "surfe" sur theta
- **Connectivité dynamique** : Comment alpha frontal influence beta pariétal
- **Transitions critiques** : Moments où le cerveau "bascule" d'état

## Patterns spatiaux complexes :

L'information cruciale est dans la **géographie cérébrale** :

- **Gradients antéro-postérieurs** : Front vs arrière du cerveau
- **Asymétries hémisphériques** : Gauche vs droite
- **Réseaux fonctionnels** : Comment les régions collaborent

## CONCLUSION : UN SIGNAL D'UNE RICHESSE EXTRAORDINAIRE

Ce fichier EEG TDBrain contient **240 secondes de cerveau humain en action** avec une précision millimétrique. C'est un **documentaire en ultra-haute définition** de l'activité neuronale.

**La question critique** : XAIguiFormer va-t-il préserver cette richesse ou la détruire dans son preprocessing ?

**Prochaine étape** : Analyser exactement ce que fait XAIguiFormer à ce trésor de données... et proposer une alternative qui respecte la nature temporelle du signal cérébral ! 🧠✨

## 2. Analyse critique du preprocessing des EEGs :

### CONTEXTE : DE L'EEG BRUT AU TOKEN ARTIFICIEL

Nous avons vu la richesse extraordinaire du signal EEG original : **240 secondes de cerveau en action, 26 canaux haute résolution, 5 bandes de fréquences simultanées**. Maintenant, découvrons ce que XAIguiFormer fait subir à ce trésor de données avant de l'analyser...

**Spoiler** : C'est un massacre ! 💧

### PIPELINE DE TRANSFORMATION : 7 ÉTAPES DESTRUCTRICES

#### ÉTAPE 1 : NETTOYAGE BRUTAL DES "MAUVAIS" CANAUX

**Script** : preprocessing.py - Step 1

### Ce qui se passe :

- **Détection automatique** des canaux "bruyants" via PrepPipeline
- **Suppression pure et simple** des électrodes jugées défaillantes
- **Critère** : Variance trop élevée ou corrélation trop faible

**Analogie** : Imaginez qu'on coupe les micros d'un orchestre parce qu'ils captent "trop de bruit"

### Problème critique :

- **Et si le "bruit" était l'information pathologique ?**
- Les troubles psychiatriques se manifestent souvent par des **anomalies électriques**
- On supprime potentiellement les **biomarqueurs** qu'on cherche !

---

## ÉTAPE 2 : FILTRAGE FRÉQUENTIEL DESTRUCTEUR

**Script** : preprocessing.py - Step 2

### Ce qui se passe :

python

```
EEG_bandpass = EEG_raw.filter(l_freq=1, h_freq=45, n_jobs=n_jobs)
```

### Transformation appliquée :

- **Coupe-bas à 1 Hz** : Supprime les très basses fréquences
- **Coupe-haut à 45 Hz** : Supprime les hautes fréquences
- **Filtre FIR** : Modification irréversible du signal

### Ce qu'on perd :

- **Fréquences < 1 Hz** : Variations lentes du cerveau (attention soutenue)
- **Fréquences > 45 Hz** : Activité gamma haute (conscience, traitement complexe)
- **Transitions spectrales** : Moments où le cerveau change de "mode"

**Analogie** : C'est comme écouter un concert en coupant les basses ET les aigus !

---

## ÉTAPE 3 : DÉCOUPAGE TEMPOREL ARBITRAIRE

**Script** : preprocessing.py - Step 3

### Le massacre temporel :

python

```
# Découpage 240s → segments de 30s
```

```
EEG_raw.custom_events = mne.make_fixed_length_events(EEG_raw, duration=40,
overlap=10)
```

```
# Puis recoupe 40s → 30s utiles (supprime 5s début + 5s fin)
```

```
EEG_eyes_close = mne.Epochs(..., tmin=5, tmax=35, ...)
```

### Transformations destructrices :

1. **Signal continu** → **segments artificiels** de 30 secondes
2. **Suppression des transitions** entre segments
3. **Perte de la cohérence temporelle** long-terme

### Ce qu'on détruit :

- **Dynamiques lentes** : Évolution de l'état mental sur minutes
- **Patterns cycliques** : Rythmes naturels du cerveau (>30s)
- **Contexte temporel** : Avant/après chaque segment

**Analogie** : Découper un film en clips de 30s et jeter les transitions !

---

## ÉTAPE 4 : NETTOYAGE ICA - LA LOBOTOMIE NUMÉRIQUE

**Script** : preprocessing.py - Step 4

### Le processus "purificateur" :

```
python
```

```
# Analyse en Composantes Indépendantes
```

```
ica = mne.preprocessing.ICA(n_components=15-20)
```

```
ica.fit(EEG_eyes_close)
```

```
# Classification automatique : "cerveau" vs "artefact"
```

```
ic_info = mne_icalabel.label_components(..., method='iclabel')
```

```
# SUPPRESSION des composantes jugées "non-cérébrales"
```

```
ica.exclude = [composantes classées "muscle", "yeux", "cœur", etc.]
```

```
EEG_ica = ica.apply(EEG_eyes_close)
```

### Le problème fondamental :

- **ICA sépare** les sources d'activité électrique
- **Un algorithme** décide ce qui est "cerveau" vs "artefact"
- **Suppression irréversible** de l'activité jugée "parasité"

**Mais attention !** En psychiatrie :

- **ADHD** : Hyperactivité musculaire = symptôme
- **Anxiété** : Variations cardiaques = biomarqueur
- **TOC** : Micro-mouvements oculaires = indicateur

**Analogie** : Un robot qui décide quels musiciens de l'orchestre peuvent jouer !

---

## ÉTAPE 5 : INTERPOLATION - INVENTION DE DONNÉES

**Script** : preprocessing.py - Step 5

**Reconstruction artificielle :**

```
python
EEG_interp = EEG_ica.interpolate_bads()
```

**Ce qui se passe :**

- Les canaux supprimés à l'étape 1 sont "**reconstruits**"
- **Interpolation spatiale** : moyenne des électrodes voisines
- **Invention pure** de données qui n'ont jamais existé

**Problème philosophique :**

- On **supprime** d'abord l'information réelle
- Puis on **invente** une information artificielle
- Le signal final n'a **jamais existé** dans le cerveau !

**Analogie** : Effacer des mots d'un livre puis deviner ce qu'ils disaient !

---

## ÉTAPE 6 : CONSTRUCTION DES CONNECTOMES - LA GRANDE RÉDUCTION

**Script** : constructFC.py

**Le processus de connectivité :**

```
python
# NOUVEAU découpage : 30s → segments de 3s !
eyes_close_short_epoch = mne.make_fixed_length_epochs(..., duration=3)

# Construction connectome sur chaque 3s
```



```
connectome = mne_connectivity.spectral_connectivity_epochs(
    eyes_close_short_epoch,
    method=['coh', 'wpli'], # Cohérence + wPLI
    fmin=lower_boundary,    # Bandes de fréquences fixes
    fmax=upper_boundary,
    faverage=True           # MOYENNAGE !
)
```

### Les 9 bandes de fréquences fixes :

- Delta (2-4 Hz), Theta (4-8 Hz), Alpha bas/haut (8-10, 10-12 Hz)
- Beta bas/moyen/haut (12-18, 18-21, 21-30 Hz), Gamma (30-45 Hz)
- **Theta/Beta ratio** (artificiel)

### Le paradoxe de la réduction :

3 Mo de signal brut (240s × 26 canaux × 250 Hz)

↓

Segments de 3s → Connectomes moyennés

↓

~500 valeurs de connectivité par segment

↓

99.98% de l'information PERDUE !

### Ce qu'on calcule :

- **Cohérence** : Corrélation entre canaux dans chaque bande
- **wPLI** : Connectivité dirigée (qui influence qui)
- **Matrices 26×26** par bande de fréquence

### Ce qu'on perd :

- **L'activité individuelle** de chaque électrode
- **Les dynamiques temporelles fines** (millisecondes)
- **Les interactions trans-fréquentielles** (alpha qui module gamma)
- **La variabilité temporelle** (l'information est moyennée !)

---

## ÉTAPE 7 : TOKENISATION FINALE - L'ABSTRACTION ULTIME

**Scripts** : EEGBenchmarkDataset.py + connectome\_encoder.py

### Transformation en graphes :

python

*# Connectomes → Graphes PyTorch Geometric*

```
for coh, wpli in zip(cohs, wplis):
```

```
    # Node features = coh rence moyenn e
```

```
    node_feat = coh.reshape((-1, coh.shape[2]))
```

```
    # Edge features = wPLI moyenn 
```

```
    edge_attr = wpli[edge_indices]
```

```
    # Cr ation graphe artificiel
```

```
    data = FreqbandData(x=node_feat, edge_index=..., edge_attr=...)
```

**Puis GNN → Tokens :**

python

*# Passage dans le GNN (connectome\_encoder.py)*

```
freqband_repr = self.gnns(x, data.edge_index, edge_attr)
```

*# Pooling final → 9 tokens de 128 dimensions*

```
freqband_repr = scatter(freqband_repr, data.freqband_order, reduce='mean')
```

```
input_repr = freqband_repr.reshape(-1, 9, 128)
```

**R sultat final :**

- **240 secondes de cerveau → 9 tokens de 128 dimensions**
- **1,560,078  chantillons originaux → 1,152 valeurs** (99.93% de perte !)
- **Richesse temporelle native → Moyennes statistiques statiques**



## BILAN DU MASSACRE : ANALYSE CRITIQUE



### Pertes informationnelles catastrophiques :

#### 1. Perte temporelle :

- **R solution native** : 4 millisecondes
- **R solution finale** : Moyennes sur 3 secondes
- **Ratio de destruction** : 750:1

#### 2. Perte spectrale :

- **Spectre continu** : 0-125 Hz simultan s
- **Bandes artificielles** : 9 bins fixes s par s

- **Interactions perdues** : Couplages cross-fréquentiels

### 3. Perte spatiale :

- **26 électrodes individuelles** → **Corrélations moyennées**
- **Topographie fine** → **Connectivité simplifiée**
- **Activité locale** → **Relations globales**

## Hypothèses discutables :

### 1. L'hypothèse connectome :

*"Les pathologies psychiatriques se manifestent par des altérations de connectivité"*

**MAIS** : Et si l'information cruciale était dans :

- Les **micro-événements temporels** (milliseconde) ?
- L'**activité électrode-spécifique** ?
- Les **variations non-stationnaires** ?

### 2. L'hypothèse de stationnarité :

*"3 secondes d'EEG sont homogènes et moyennables"*

**MAIS** : Le cerveau change d'état en **millisecondes** !

### 3. L'hypothèse des bandes fixes :

*"Delta, Theta, Alpha, Beta, Gamma suffisent"*

**MAIS** : Chaque individu a **ses propres fréquences** !

## Conséquences pour le diagnostic :

**Patterns perdus** :

- **Micro-épilepsie** : Spikes de quelques millisecondes
- **Variabilité ADHD** : Fluctuations sub-secondes d'attention
- **Asymétries dépression** : Différences électrode-spécifiques
- **Transitions TOC** : Basculements rapides d'état mental

---

## LE PARADOXE XAIGUIFORMER

XAIGuiFormer prétend être "explainable" mais :

1. **Détruit 99.93% de l'information** dans le preprocessing

2. **Analyse les 0.07% restants** avec sophistication
3. **"Explique" des moyennes** qui n'existent pas dans le cerveau réel

## ESTIMATION DES PERTES D'INFORMATION (ORDRES DE GRANDEUR)

### Données d'entrée originales :

- **60,000 échantillons × 26 canaux = 1,56 million de points de données**
- **Résolution temporelle** : 4 millisecondes
- **Information brute** : ~3 Mo par patient

### Données finales après preprocessing :

- **9 tokens × 128 dimensions = 1,152 valeurs numériques**
- **Résolution temporelle** : Moyennes sur 3 secondes
- **Information résiduelle** : ~quelques Ko

### Calcul approximatif de la réduction :

Ratio =  $1,560,000 / 1,152 \approx 1,350:1$

Pourcentage conservé  $\approx 1/1,350 \approx 0.07\%$

Pourcentage perdu  $\approx 99.93\%$

**C'est comme analyser un concert en n'écoutant que 3 notes moyennées par minute ! 🎵**

---

## CONCLUSION : UN POTENTIEL GÂCHÉ

Le preprocessing XAlguiFormer transforme un **documentaire 4K du cerveau** en quelques **pixels moyennés**.

**La question fondamentale** : Peut-on vraiment diagnostiquer des **troubles neurologiques complexes** avec des **moyennes statistiques brutales** qui ont perdu toute la **richesse temporelle native** ?

**Prochaine étape** : Proposer une alternative qui préserve la beauté du signal temporel original ! 🧠✨

#### 4. Alternative: EEG TimeSeries Tokenizer

##### ● Foundation Models à considérer :

- **Chronos** (Amazon) : SOTA time series
- **TimesFM** (Google) : Zero-shot forecasting
- **Mantis** comme tu dis
- **PatchTST** : Attention-based

##### ● Architectures spécialisées :

- **TimesNet** : Multi-scale temporal modeling
- **MultiRocket** : Ultra-rapide, très efficace
- **TSMixer** : Simple mais performant

##### ● Avantages clés :

- Préservation signal brut
- Capture dynamiques temporelles
- Pas d'hypothèses sur connectivité

#### L'hypothèse connectome :

Est-ce que les **corrélations entre électrodes** capturent correctement les **biomarqueurs psychiatriques** ?

- **Argumentation PRO** : Les pathologies modifient la connectivité et l'étude seule de celle-ci suffit.
- **Argumentation CONTRE** : Nous ne pouvons pas nous passer de l'information cruciale dans les dynamiques temporelles individuelles des électrodes.

#### La question fondamentale :

**Peut-on diagnostiquer la psychiatrie avec des moyennes statistiques ou faut-il préserver les subtilités temporelles natives ?**

# Le choix des TimeSeries : préserver l'intégrité du signal

## Une philosophie de préservation informationnelle

Notre approche alternative s'articule autour d'un principe fondamental : **préserver l'intégrité temporelle native du signal EEG** plutôt que de le dénaturer par un preprocessing destructeur.

### Pourquoi les TimeSeries sont-elles cruciales ?

1. **Résolution temporelle native** : L'EEG offre une résolution de 4 millisecondes (250 Hz), permettant de capturer des événements neurophysiologiques ultra-rapides. Les spikes épileptiques, les micro-réveils, ou les transitions attentionnelles se produisent à cette échelle temporelle fine.
2. **Dynamiques non-stationnaires** : Le cerveau est un système dynamique dont l'état change continuellement. Moyenner sur 3 secondes suppose une stationnarité locale qui n'existe pas dans la réalité neurophysiologique.
3. **Signatures pathologiques temporelles** : Les troubles psychiatriques se manifestent souvent par des altérations dans les dynamiques temporelles : variabilité accrue dans l'ADHD, patterns oscillatoires altérés dans la dépression, ou synchronisation anormale dans l'autisme.
4. **Information multifréquentielle continue** : Plutôt que de séparer artificiellement les bandes de fréquences, une approche TimeSeries préserve le spectre continu et les interactions trans-fréquentielles.
5. **Interdépendances spatio-temporelles** : Un tokenizer TimeSeries peut apprendre à capturer simultanément les patterns temporels individuels et les interdépendances entre électrodes, sans sacrifier l'un pour l'autre.

## Critères de sélection pour l'EEG TimeSeries Tokenizer

Pour identifier le candidat optimal, nous avons défini cinq critères essentiels :

### 1. Préservation temporelle multivarié

Capacité à traiter le signal EEG natif sans moyennage destructeur, tout en capturant non seulement les signaux individuels de chaque électrode mais surtout leurs interdépendances spatio-temporelles complexes.

### 2. Compatibilité architecturale

Intégration fluide avec l'architecture Transformer existante de XAIguiFormer, permettant de conserver l'innovation XAI guided attention tout en remplaçant uniquement le module de tokenisation.

### 3. Performance SOTA

Foundation models ou architectures état-de-l'art en TimeSeries, avec des performances démontrées sur des tâches de classification temporelle complexes.

## 4. Efficacité computationnelle

Complexité raisonnable pour permettre l'entraînement end-to-end sans explosion des coûts computationnels, particulièrement important pour les datasets EEG de grande taille.

## 5. Stabilité d'entraînement

Architecture robuste aux problématiques de vanishing/exploding gradients, crucial car notre EEG TimeSeries Tokenizer sera entraîné conjointement avec le Transformer, nécessitant une propagation stable des gradients à travers l'ensemble de l'architecture.

## Candidats sélectionnés

### Candidat #1 : MANTIS-8M (Foundation Model)

#### Justification du choix :

MANTIS-8M représente le candidat optimal pour notre architecture alternative. Ce foundation model de 8 millions de paramètres, pré-entraîné par apprentissage contrastif sur 2 millions de séries temporelles, répond parfaitement à nos exigences.

#### Avantages clés :

- **Architecture encoder-only** : Intégration native avec notre Transformer existant
- **Préservation temporelle native** : Conçu spécifiquement pour maintenir les patterns temporels fins sans dégradation
- **Gestion multivariée avancée** : Capacité prouvée à capturer les interdépendances complexes entre canaux multiples
- **Scalar embedding module** : Préservation des statistiques fondamentales cruciales pour l'interprétation EEG
- **Stabilité des gradients** : Architecture Transformer éprouvée sans problèmes de vanishing/exploding gradients
- **Performance SOTA** : Meilleur foundation model tant en zero-shot qu'en fine-tuning (84.08% sur 131 datasets)

### Candidat #2 : TimesNet (CNN 2D avec Inception)

#### Justification du choix :

TimesNet propose une approche innovante en transformant les séries temporelles 1D en représentations 2D pour capturer les variations inter-périodes, particulièrement pertinent pour les patterns EEG complexes.

#### Avantages clés :

- **Capture des variations inter-périodes** : Idéal pour les patterns oscillatoires EEG multi-fréquentiels
- **Architecture 2D innovante** : Transformation 1D→2D préservant les interdépendances spatiales entre électrodes

- **Modules Inception multi-échelles** : Capture simultanée de différentes bandes de fréquences EEG
- **Stabilité d'entraînement** : CNN avec connexions résiduelles, architecture robuste
- **Performance démontrée** : 73.6% sur benchmarks TimeSeries, surpassant les transformers classiques

**Limitation** : Nécessite une adaptation architecturale pour l'intégration avec le Transformer existant.

### 🏆 Candidat #3 : MultiROCKET (Random Convolutions)

#### Justification du choix :

MultiROCKET constitue une alternative efficace basée sur des convolutions aléatoires pour l'extraction de features temporelles robustes.

#### Avantages clés :

- **Ultra-efficacité** : Extraction de features très rapide, idéal pour les gros volumes EEG
- **Robustesse temporelle** : Excellente capacité de représentation des patterns temporels
- **Support multivarié** : Gestion native des signaux multi-canaux
- **Gradient-safe** : Pas d'entraînement des convolutions, élimination des problèmes de gradients
- **Performance/coût optimal** : Résultats compétitifs avec une fraction de la complexité computationnelle

**Limitation** : Features fixes non-apprenables, flexibilité réduite par rapport aux approches end-to-end.

## Un problème rédhibitoire : le temps d'entraînement

L'un des défis majeurs de notre approche consistant à remplacer le pipeline destructeur par un tokenizer préservant l'information temporelle native réside dans l'explosion du volume de données à traiter. Là où le pipeline actuel réduit drastiquement les données (de 1,56 million de points à 1,152 valeurs), notre approche conserve la richesse temporelle du signal EEG brut, entraînant une augmentation massive des besoins computationnels.

### Estimation des temps d'entraînement pour les trois candidats

Pour évaluer la faisabilité pratique, nous avons calculé les temps d'entraînement estimés sur GPU A100, considérant le dataset TDBRAIN (910 patients en entraînement, 60,003 échantillons  $\times$  33 canaux par patient).

**MANTIS-8M et TimesNet : prometteurs mais prohibitifs**



MANTIS-8M représente le candidat le plus séduisant conceptuellement. Ce foundation model pré-entraîné sur 2 millions de séries temporelles offre des performances état de l'art (84,08% en fine-tuning) et sa capacité d'apprentissage adaptatif lui permettrait de découvrir des patterns neurophysiologiques complexes. Cependant, avec 1,260,063 points par patient ( $60,003 \times 21$  canaux EEG utiles), 8 millions de paramètres, 100 époques et 15 batches par époque, l'estimation donne :  $(8M \times 1,26M \times 100 \times 15) / \text{throughput\_A100} \approx \mathbf{950 \text{ heures}}$ , soit 40 jours continus.

TimesNet présente une approche innovante transformant les séries 1D en représentations 2D pour capturer les variations inter-périodes, particulièrement adapté aux relations inter-électrodes EEG. Avec 3 millions de paramètres :  $(3M \times 1,26M \times 100 \times 15) / \text{throughput\_A100} \approx \mathbf{350 \text{ heures}}$ , soit 15 jours. Ces temps sont incompatibles avec les contraintes d'un projet de recherche.

## **MultiROCKET : l'efficacité pragmatique**

MultiROCKET se distingue par son approche non-apprenante : ses 500,000 paramètres correspondent au classificateur final, les features étant extraites par convolutions aléatoires fixes. L'estimation donne :  $(0,5M \times 1,26M \times 100 \times 15) / \text{throughput\_A100} \approx \mathbf{60 \text{ heures}}$ . Bien que le plus raisonnable, ce temps reste problématique pour un développement itératif efficace.

## **Optimisations nécessaires : downsampling intelligent et réduction de dataset**

### **Downsampling intelligent basé sur le théorème de Nyquist-Shannon**

Le preprocessing existant applique un filtrage passe-haut à 45 Hz. Selon le théorème de Nyquist-Shannon, pour capturer un signal de fréquence maximale  $f_{\text{max}}$ , il faut échantillonner à  $2 \times f_{\text{max}}$  minimum. Avec  $f_{\text{max}} = 45 \text{ Hz}$ , une fréquence de 90 Hz suffit à préserver toute l'information utile.

Ce downsampling de 250 Hz à 90 Hz (facteur 2,78) préserve toutes les bandes EEG : delta (0,5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), et gamma (jusqu'à 45 Hz). Cette optimisation réduit le temps de 60 à ~22 heures.

### **Réduction stratégique du dataset**

La réduction du dataset d'entraînement de 910 à 50 patients (facteur 18,2), combinée au downsampling, donne :  $60h \div (2,78 \times 18,2) \approx \mathbf{1,2 \text{ heure}}$ , compatible avec nos contraintes expérimentales.

### **Implications méthodologiques**

Cette réduction introduit des biais significatifs : un dataset de 50 patients limite la généralisation et augmente la variance des métriques. De plus, cela avantage structurellement MultiROCKET, plus robuste aux petits datasets que les architectures de deep learning qui nécessitent de grandes quantités de données. Cette asymétrie pourrait fausser la comparaison en notre faveur, biais à considérer dans l'interprétation des résultats comme preuve de concept plutôt que validation définitive.

# MultiROCKET : intégration dans le pipeline et back-propagation

L'intégration de MultiROCKET dans notre architecture soulève des questions techniques importantes concernant la propagation des gradients et l'optimisation end-to-end. En effet, MultiROCKET n'est pas un tokenizer au sens strict mais un extracteur de features utilisant des convolutions aléatoires fixes, ce qui nécessite une réflexion architecturale spécifique.

## Deux approches d'intégration

### Option A : Pipeline séparé en deux étapes

Cette approche consiste à séparer l'extraction de features de l'entraînement du modèle. Dans un premier temps, MultiROCKET transforme les signaux EEG bruts en features fixes qui sont sauvegardées. Dans un second temps, ces features pré-calculées alimentent directement le XAIguiTransformer pour l'entraînement.

Les avantages incluent une séparation claire des responsabilités et l'efficacité computationnelle puisque MultiROCKET n'est calculé qu'une seule fois. De plus, cette approche évite toute complication liée aux gradients. Cependant, elle présente l'inconvénient majeur d'empêcher l'optimisation end-to-end, les features étant figées et non adaptées spécifiquement à notre tâche de classification psychiatrique.

### Option B : Pipeline intégré end-to-end

L'approche intégrée combine MultiROCKET et XAIguiTransformer dans une architecture unique où les signaux EEG bruts traversent successivement l'extracteur de features puis le transformer. Bien que MultiROCKET ne soit pas entraînable, il fait partie intégrante du forward pass.

Cette option offre un pipeline unifié facilitant les comparaisons avec l'architecture actuelle, permet un gradient flow optimal vers le transformer, et simplifie l'implémentation en ne nécessitant qu'un seul modèle. L'inconvénient réside dans une complexité architecturale légèrement supérieure.

## Choix architectural : l'approche end-to-end

Nous optons pour l'option B intégrée pour plusieurs raisons déterminantes. Le pipeline unifié permet une comparaison équitable avec XAIguiFormer actuel en conservant le même format d'entrée (signaux EEG bruts). Le gradient flow optimal permet au transformer de s'adapter aux features spécifiques produites par MultiROCKET, optimisant ainsi l'ensemble de la chaîne de traitement. Enfin, la simplicité d'implémentation avec un seul modèle à entraîner facilite le développement et les expérimentations.

## MultiROCKET : un extracteur de features sans apprentissage

La nature non-apprenante de MultiROCKET pourrait sembler problématique pour la back-propagation, mais cette préoccupation est infondée. MultiROCKET fonctionne comme un extracteur de features utilisant des convolutions aléatoires fixes suivies d'un pooling, produisant des représentations de haute dimension des signaux d'entrée.

### Processus de propagation

Le forward pass suit une séquence logique : les signaux EEG traversent d'abord MultiROCKET qui extrait des features fixes différentiables, puis ces features alimentent le XAIguiTransformer qui produit les prédictions finales. Crucially, bien que MultiROCKET ne modifie pas ses paramètres, ses sorties restent différentiables par rapport aux entrées du transformer.

Le backward pass s'effectue normalement depuis la loss de classification jusqu'aux features MultiROCKET, permettant au transformer de recevoir les gradients nécessaires à son optimisation. La propagation s'arrête naturellement aux features puisque MultiROCKET ne possède pas de paramètres entraînaables, mais cela n'empêche nullement l'optimisation du reste du pipeline.

Cette architecture hybride tire parti de la robustesse de MultiROCKET pour l'extraction de features tout en préservant la capacité d'adaptation du transformer, offrant ainsi un compromis optimal entre efficacité computationnelle et performance prédictive.