

Estimation robuste de la DoA large bande par CNN2D et Critère Normalisé Temps-Fréquence Pondéré

Ce projet implémente une chaîne complète d'estimation de la direction d'arrivée (DoA) d'un signal acoustique à partir de signaux multi-antennes simulés. L'approche repose sur un réseau de neurones convolutionnel profond (CNN2D) enrichi par des blocs résiduels et Squeeze-and-Excitation (SE), et optimisé via un critère de perte pondéré appliqué à une représentation temps-fréquence normalisée.

Structure des fichiers

- **generate_data.py** : Génère des signaux simulés avec bruit et DoA aléatoire.
- **dataset.py** : Applique la STFT à chaque signal d'antenne, normalise les tenseurs.
- **model.py** : Définit le modèle CNN2D avec blocs résiduels et SE.
- **train.py** : Entraîne le modèle et sauvegarde les poids les plus performants.
- **test.py** : Évalue le modèle sur le jeu de test et calcule l'erreur absolue moyenne.

Instructions d'utilisation

1. Génération des données

```
python generate_data.py
```

- Génère 10 000 paires (signal multi-antennes + DoA) dans le dossier `dataset/`.

2. Entraînement du modèle

```
python train.py
```

```
Epoch 094, LR: 6.25e-05, Train Loss: 388.9873, Validation Loss: 394.6313  
Epoch 095, LR: 6.25e-05, Train Loss: 387.5716, Validation Loss: 393.4338  
Epoch 096, LR: 6.25e-05, Train Loss: 399.0940, Validation Loss: 392.9662
```

```
Epoch 097, LR: 6.25e-05, Train Loss: 393.8169, Validation Loss: 403.2364
Epoch 098, LR: 3.13e-05, Train Loss: 391.1696, Validation Loss: 393.6171
Epoch 099, LR: 3.13e-05, Train Loss: 390.2861, Validation Loss: 393.5433
Epoch 100, LR: 3.13e-05, Train Loss: 388.3509, Validation Loss: 393.1766
Training complete. Best Validation Loss: 392.6108 at epoch 82.
```

- Sauvegarde les poids du meilleur modèle dans `cnn2d_wideband_best.pt` and `cnn2d_wideband_last.pt`.

3. Test du modèle

```
python test.py
```

- Affiche la MAE et les erreurs DoA prédites vs vraies.

Dépendances

```
pip install numpy scipy torch pyroomacoustics matplotlib
```

Architecture du Modèle CNN2D

- 4 blocs résiduels convolutionnels (32 → 256 canaux)
- Blocs SE adaptatifs par canal
- Pooling global + 2 couches fully-connected
- Activation LeakyReLU, Dropout
- Perte : Mean Squared Error pondérée

Résultats de l'entraînement

- Meilleure validation atteinte : **392.61** à l'époque 82
- Comportement typique d'une convergence progressive avec stabilisation

Prédictions sur le jeu de test

```
Pred: 30.01° / True: 25.99° → Error: 4.02°
Pred: 30.47° / True: 44.99° → Error: 14.53°
```

Pred: 73.05° / True: 62.40° → Error: 10.65°

...

Mean Absolute Error: 10.96°

- **Erreurs faibles (<5°)** : Fréquentes
- **Erreurs modérées (5-20°)** : Présentes en zone intermédiaire
- **Erreurs fortes (>20°)** : Cas extrêmes (angles aux limites ou bruit élevé)

Analyse visuelle des données (figures)

- **Carte de chaleur moyenne temps-antennes** : [Heatmap_Temps_Antennes.png](#)
- **Écart-type (bruit) par antenne** : [STD_Bruit_Antennes_Exemple_1297.png](#)
- **Signal temporel multi-antennes (DoA = 13.96°)** : [Signal_Antennes_DoA_13.96.png](#)
- **Vue globale des signaux temporels** : [Signal_Temporel_Antennes.png](#)
- **Moyenne temporelle multi-antennes** : [Signal_Temporel_Moyen_Antennes.png](#)
- **Spectre STFT individuel (Antenne 0)** : [Spectre_STFT_Antenne_0.png](#)
- **STFT moyen sur l'antenne 0** : [Spectre_STFT_Moyen_Antenne_0.png](#)
- **Zoom temporel sur l'antenne 0** : [Zoom_Antenne_0.png](#)
- **Zoom temps-fréquence sur l'antenne 0** : [Zoom_TF_Antenne_0.png](#)

Conclusion

Le modèle CNN2D exploitant la représentation STFT normalisée par canal et enrichi par des blocs résiduels et SE, montre une capacité robuste d'estimation du DoA sur des données simulées bruitées. Les performances obtenues indiquent une erreur moyenne raisonnable à ~11°, et une généralisation stable sur l'ensemble de test. Le pipeline est modulaire et facilement extensible à d'autres configurations d'antennes ou de réseaux profonds.