

Plan prévisionnel

Dataset retenu

Le dataset choisi pour ce projet est **MAESTRO (MIDI and Audio Edited for Synchronous TRacks and Organization)**. Il contient environ 200 heures d'enregistrements de piano alignés avec leurs transcriptions MIDI correspondantes. Il est largement utilisé pour les tâches de transcription musicale automatique, car il fournit un alignement précis entre l'audio et les événements MIDI, permettant une évaluation rigoureuse des modèles de transcription.

Modèle envisagé

Le modèle récent choisi pour cette étude est **Basic Pitch**, développé par Spotify. Il repose sur un réseau de neurones convolutionnel léger et est conçu pour extraire des événements de pitch polyphoniques à partir d'un signal audio en temps réel. Contrairement aux modèles plus anciens, Basic Pitch offre une meilleure généralisation sur différents types d'instruments et fonctionne avec une faible latence, ce qui le rend attractif pour des applications temps réel.

La baseline retenue pour la comparaison est **Melodia**, une méthode heuristique plus ancienne basée sur le suivi des harmoniques. Melodia est efficace pour extraire la mélodie dominante d'un fichier audio, mais son approche repose sur des techniques de traitement du signal plutôt que sur l'apprentissage profond, ce qui peut limiter ses performances sur des morceaux polyphoniques complexes.

Références bibliographiques

1. **Salamon, J., & Gómez, E. (2012).** "Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics." *IEEE Transactions on Audio, Speech, and Language Processing*. [Article sur Melodia](#) et [ici](#)
2. **Hawthorne, C., et al. (2018).** "Onsets and Frames: Dual-Objective Piano Transcription." *ISMIR 2018*. [Article sur les modèles de transcription MIDI](#)
3. **Bittner, R., Bosch, J. J., Rubinstein, D., Meseguer-Brocal, G., & Ewert, S. (2022).** "A Lightweight Instrument-Agnostic Model for Polyphonic Note Transcription and Multipitch Estimation." *ICASSP 2022*. [Article sur Basic Pitch](#)

Explication de votre démarche de test du nouvel algorithme (votre preuve de concept)

La preuve de concept consiste à comparer la performance de **Basic Pitch** et de **Melodia** sur le dataset **MAESTRO**. La méthode suivie inclut les étapes suivantes :

1. **Prétraitement des données** : Extraction des fichiers audio et MIDI du dataset MAESTRO, segmentation des morceaux en unités plus petites si nécessaire.
2. **Application des modèles** :
 - Basic Pitch sera appliqué aux fichiers audio pour extraire les notes MIDI prédites.
 - Melodia sera utilisé pour extraire les pitches dominants, puis convertis en format MIDI.
3. **Métriques de comparaison** :
 - F1-score** sur l'alignement MIDI entre la prédiction et la vérité terrain.
 - Notes Correct** sur la base des notes totales
4. **Visualisation des résultats** :
 - Utilisation de **Streamlit** pour développer un **dashboard interactif** permettant de comparer les transcriptions MIDI de chaque modèle avec l'original.