

Fiche explicative – Riffusion

♦ **Nom du modèle :**

Riffusion

♦ **Type :**

Diffusion model appliqué aux spectrogrammes pour la **génération de musique en continu**.

♦ **Développeur :**

Seth Forsgren et Hayk Martiros (projet indépendant)

♦ **Date de sortie :**

Décembre 2022

Objectif

Riffusion vise à **générer de la musique en temps réel** à partir de **prompts textuels**.

Il génère des **spectrogrammes musicaux** via un **modèle de diffusion** adapté à l'audio, puis les reconstruit en fichiers sonores **.wav**.

✓ Résultat : création rapide de morceaux courts et **transition fluide entre styles**.

Fonctionnement simplifié

Étape	Description
Entrée	Prompt texte (ex : "Jazz piano", "Electronic dance")
Génération	Modèle de diffusion génère un spectrogramme basé sur Stable Diffusion
Reconstruction	Transforme le spectrogramme en audio avec la transformée de Fourier inverse (ISTFT)

Techniques utilisées :

- Adaptation de **Stable Diffusion** sur l'espace des **spectrogrammes audio**
 - **Interpolation latente** entre prompts pour créer des transitions musicales continues
 - Reconstruction du signal audio via **ISTFT (Inverse Short-Time Fourier Transform)**
-

Applications concrètes

- Génération de **boucles musicales** pour DJ, beatmakers, producteurs
 - Création de **transitions fluides** entre différents styles musicaux
 - Exploration de **nouvelles idées musicales** à partir de simples mots clés
 - Utilisation pour des **projets artistiques génératifs interactifs**
-


Exemples d'usage

Domaine	Exemple
Production musicale	Générer des loops "house", "jazz", "ambient" instantanément
DJing en live	Créer des transitions fluides entre deux styles musicaux
Recherche audio IA	Étudier comment l'IA gère le morphing entre genres sonores

Détails techniques

Caractéristique	Valeur
Architecture	Diffusion latente sur spectrogrammes
Framework	PyTorch
Reconstruction audio	ISTFT (Inverse Short-Time Fourier Transform)
Dataset d'entraînement	Musique instrumentale issue de bases publiques
Durée de génération	Quelques secondes pour 5-10 secondes d'audio
Objectif	Génération musicale continue, pilotée par prompts texte

Ressources officielles et utiles

-  [Publication / Projet officiel Riffusion](#)
-  [Code source Riffusion sur GitHub](#)

Démonstrations & alternatives pratiques

Démo pratique à tester

-  [Site officiel interactif Riffusion](#)
(Écrire un prompt, générer une musique directement !)

Google Colab utilisables aujourd'hui

-  [Colab officiel Riffusion simple](#)
(Permet de générer ses propres spectrogrammes et sons à partir de textes.)

Tableau des avantages / inconvénients

Avantages

Génération rapide de musique basée sur du texte

Transitions fluides entre styles possibles

Modèle léger (peut tourner en Colab)

Open-source, facile à modifier

Inconvénients

Audio limité en durée (~10 secondes)

Résolution sonore parfois moyenne (artefacts)

Moins performant pour des compositions longues ou complexes

Sons parfois répétitifs sans prompts complexes