

# Fiche explicative – CoverSong

---

♦ **Nom du modèle :**

**CoverSong** (aussi appelé **Voice-to-Voice** ou **Singing Voice Conversion**)

♦ **Type :**

Modèle de **conversion vocale chantée (Singing Voice Conversion Model)**.

♦ **Développeur :**

Plusieurs initiatives :

- Projets open-source comme **So-VITS-SVC** (China community)
- Modèles internes chez **Tencent AI Lab**, **Meta AI** (MusicGen++)
- Plein de forks communautaires (ex : DiffSVC, Mangio-SVC)

♦ **Date de sortie :**

Les premières versions communautaires datent de **2022** (So-VITS-SVC v1-v4)

---

## **Objectif**

CoverSong est destiné à **prendre une voix chantée (ou parlée) et la transformer** pour qu'elle **sonne comme si une autre personne l'avait chantée**, tout en **gardant** :

- la **mélodie**,
- le **rythme**,
- l'**expression vocale**.

✓ Résultat : possibilité de "**rechanter**" n'importe quelle chanson dans **n'importe quelle voix cible**.

---

## **Fonctionnement simplifié**

Étape	Description
Entrée	Fichier audio d'une personne chantant/parlant + voix cible enregistrée
Extraction	Extraction de caractéristiques vocales (pitch, timbre, phonèmes)
Conversion	Ré-échantillonnage et adaptation des caractéristiques dans la voix cible
Reconstruction	Génération du nouvel audio transformé en voix cible

### **Techniques utilisées :**

- **Encoder-Decoder modèle VITS** (Variational Inference Text-to-Speech)
  - **Pitch & phoneme prediction** pour conserver la mélodie
  - **HiFi-GAN** ou vocodeurs GAN pour améliorer la qualité de la sortie
  - **Diffusion models** (nouveaux modèles comme DiffSVC pour plus de réalisme)
- 

## **Applications concrètes**

- **Rechanter** des chansons célèbres avec sa propre voix
  - **Créer des covers IA réalistes** d'artistes (en respectant les droits)
  - **Outils de correction vocale IA** pour les producteurs de musique
  - **Assistant vocal pour les musiciens** (transposer automatiquement un morceau)
-




## Exemples d'usage

Domaine	Exemple
Covers IA	Changer une chanson de Michael Jackson pour qu'elle soit chantée par Eminem (IA)
Création musicale	Adapter des maquettes vocales en différentes voix d'artistes
Accessibilité	Transformer une voix parlée en chanté de manière réaliste

## Détails techniques

Caractéristique	Valeur
Architecture	VITS / Diffusion + vocodeur GAN
Framework	PyTorch
Input	Audio chanté ou parlé + voix cible enregistrée (dataset)
Output	Fichier audio chanté transformé
Dataset d'entraînement	Samples de voix (~5 à 30 minutes de voix cible suffisent)
Objectif	Conversion de voix en conservant mélodie + expression

## Ressources officielles et utiles

-  [Publication So-VITS-SVC \(arXiv\)](#) – Base théorique pour Voice Conversion
  -  [GitHub – So-VITS-SVC v4 \(modèle open-source\)](#)
  -  [GitHub – Mangio-SVC \(simplifié pour Colab\)](#)
-

## Démonstrations & alternatives pratiques

### Google Colab utilisables aujourd'hui

-  [Colab So-VITS-SVC \(Cover de chansons IA\)](#)

*(Permet de générer des covers directement en chargeant un fichier audio + un modèle vocal cible.)*

---

## Tableau des avantages / inconvénients

### Avantages

Génération de covers très réalistes

Outils open-source faciles à utiliser (Colab)

Adaptable à beaucoup de styles (chant, rap, etc.)

Diffusion et HiFi-GAN améliorent beaucoup la qualité

### Inconvénients

Nécessite un peu de dataset voix cible (5-30 min)

Peut créer des problèmes légaux si utilisé sans droits

Qualité dépend fortement de la voix cible et des prétraitements

Nécessite GPU pour un traitement rapide