



HAL
open science

Apprentissage profond et génération de musique

Jean-Pierre Briot

► **To cite this version:**

Jean-Pierre Briot. Apprentissage profond et génération de musique. ActuIA, 2020, Dossier Intelligence artificielle et musique, 7, pp.46. hal-03809774

HAL Id: hal-03809774

<https://hal.science/hal-03809774>

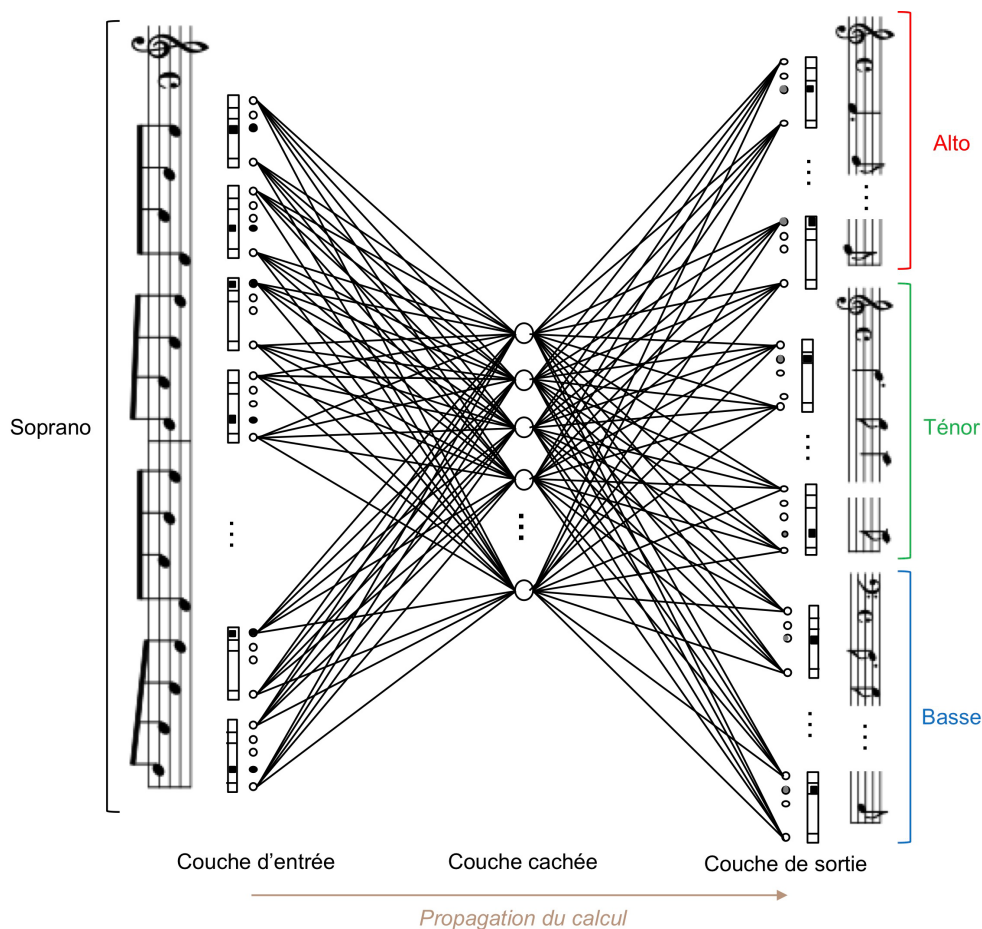
Submitted on 14 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Avec les chaînes de Markov, les réseaux de neurones sont un des deux principaux modèles d'apprentissage machine utilisés pour la musique. Les premiers travaux, préfigurant les travaux actuels [1], datent de la fin des années 80. L'apprentissage profond (*deep learning*), en plus d'avancées majeures en termes de ressources (computationnelles, données et techniques), a récemment introduit des modèles (architectures) spécifiques, dits génératifs, ainsi dédiés, non plus à la classification ou la prédiction, mais à la génération de contenu (musique, mais aussi images ou texte) [2]. Les principaux sont les autoencodeurs variationnels (VAE) et les réseaux antagonistes génératifs (GAN), suivis récemment par le modèle Transformer. Il y a également des modèles dédiés au transfert de style (par exemple pour produire des accompagnements dans un style donné).

Les résultats sont assez impressionnants, malgré le manque de structure produite. Le relatif manque d'originalité, intrinsèque aux approches d'apprentissage malgré des astuces pour sortir de la distribution apprise, n'est pas un problème si l'on considère les techniques actuelles comme un outil d'assistance et non un générateur autonome (sans capacité d'auto-évaluation esthétique). Pour ce faire, des approches de contrôle, fondées sur le conditionnement ou/et sur l'exploration d'un espace latent (dans le cas d'autoencodeurs), sont proposées, ainsi que des approches de génération modulaire et incrémentale (par exemple pour la complétion de certains segments). Une tendance actuelle est la montée en puissance des modèles de génération d'audio (son, par opposition à une représentation symbolique telle qu'une partition ou la norme MIDI).



Exemple d'architecture simple de réseau de neurones *feedforward* (à propagation avant) à une seule couche cachée générant un contrepoint à trois voix (dans le style des Chorales de Bach)

[1] Jean-Pierre Briot. From artificial neural networks to deep learning for music generation – History, concepts and trends. *Neural Computing and Applications (NCAA)*, (33):39–65, janvier 2021. DOI: 10.1007/s00521-020-05399-0.

[2] Jean-Pierre Briot, Apprentissage profond et génération de musique, Hors série Intelligence artificielle, Tangente – L'aventure mathématique, N° 68, pages 30–37, septembre 2019.