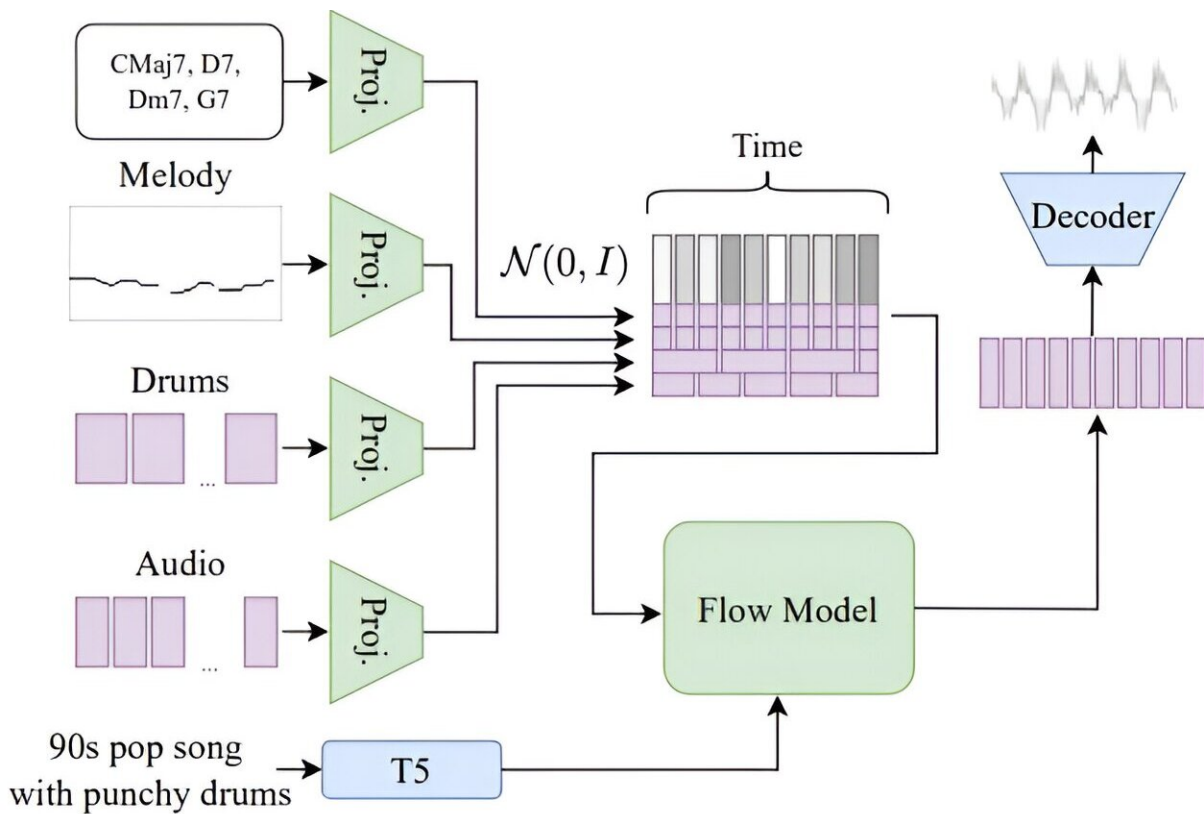# Meta releases four new publicly available AI models for developer use

July 3 2024, by Bob Yirka

Top figure presents the temporal blurring process, showcasing source separation, pooling and broadcasting. Bottom figure presents a high level presentation of JASCO. Conditions are first being projected to low dimensional representation and are concatenated over the channel dimensions. Green blocks have learnable parameters while blue block are frozen. Credit: *arXiv* (2024). DOI: 10.48550/arxiv.2406.10970

A team of AI researchers at Meta's Fundamental AI Research team are making four new AI models publicly available to researchers and developers creating new applications. The team has posted a paper on the *arXiv* preprint server outlining one of the new models, JASCO, and how it might be used.

As interest in AI applications grows, major players in the field are creating AI models that can be used by other entities to add AI capabilities to their own applications. In this new effort, the team at Meta has made available four new models: JASCO, AudioSeal and two versions of Chameleon.

JASCO has been designed to accept different types of audio input and create an improved sound. The model, the team says, allows users to adjust characteristics such as the sound of drums, guitar chords or even melodies to craft a tune. The model can also accept text input and will use it to flavor a tune.

An example would be to ask the model to generate a bluesy tune with a lot of bass and drums. That would then be followed by similar descriptions regarding other instruments. The team at Meta also compared JASCO with other systems designed to do much the same thing and found that JASCO outperformed them across three major metrics.

AudioSeal can be used to add watermarks to speech generated by an AI app, allowing the results to be easily identified as artificially generated. They note it can also be used to watermark segments of AI speech that have been added to real speech and that it will come with a commercial license.

The two Chameleon models both convert text to visual depictions and are being released with limited capabilities. The versions, 7B and 34B,

the team notes, both require the models to gain a sense of understanding of both text and images. Because of that, they can do reverse processing, such as generating captions of pictures.

More information: Or Tal et al, Joint Audio and Symbolic Conditioning for Temporally Controlled Text-to-Music Generation, *arXiv* (2024). DOI: 10.48550/arxiv.2406.10970

Demo page: pages.cs.huji.ac.il/adiyoss-lab/JASCO/

Citation: Meta releases four new publicly available AI models for developer use (2024, July 3) retrieved 8 October 2024 from https://techxplore.com/news/2024-07-meta-ai.html