

Comparing our model, TangoFlux, with other state-of-the-art text-to-audio generation models: TangoFlux achieves better quality audio (measured by CLAP and FD scores) while being approximately 2x faster than the next fastest model, all with fewer trainable parameters!