

ABSTRAK

Analisis struktur pada musik dilakukan untuk menemukan pengelompokan segmen-segmen yang sama dan membuat suatu struktur. Tantangan utama dalam analisis struktur musik adalah bagaimana menangkap konteks jangka panjang tanpa membebani komputasi secara berlebihan. Model berbasis *deep learning* telah memberikan hasil menjanjikan, namun kompleksitas komputasi dari mekanisme *attention* menjadi hambatan. *Neighborhood Attention* (NA) diperkenalkan sebagai solusi untuk mengurangi kompleksitas waktu dari mekanisme *attention* dengan tetap mempertahankan kemampuan model dalam mengenali pola kontekstual. NA telah terbukti efektif dalam tugas analisis struktur musik pada audio terpisahkan. Akan tetapi, efektivitas tersebut belum dioptimalkan. *Fused Neighborhood Attention* (FNA) dikembangkan untuk mengatasi hal ini dengan menurunkan kompleksitas NA dari semula kuadratik menjadi linier. Penelitian ini mengaplikasikan dan mengkaji efektivitas FNA dalam tugas analisis struktur musik audio terpisahkan dengan memodifikasi model NA dari penelitian sebelumnya. Model FNA dibandingkan dari sisi akurasi segmentasi musik serta efisiensi komputasi. Model FNA memiliki tingkat *F-measure Hit Rate* dengan toleransi 0,5 detik (HR.5F) terbaik sebesar 0,516 dan *Pairwise F-measure* (PwF) sebesar 0,611. Pada ukuran rekaman yang pendek, latensi inferensi FNA lebih kecil dibandingkan NA. Berdasarkan perbandingan tersebut, FNA berhasil meningkatkan kinerja model dengan mengorbankan sedikit pengurangan pada akurasi.

Kata kunci : *Fused Neighborhood Attention*, analisis struktur musik, temu balik informasi musik