



INTISARI

DETEKSI KEKERASAN MULTIMODAL BERBASIS DEEP LEARNING PADA GAME ONLINE BER-RATING RAMAH ANAK

Oleh

Jasson Prestiliano

22/500409/SPA/00854

Sebagian *game online* memiliki rating yang ramah anak (*child friendly*), yaitu di mana pengembang *game* sudah mengatur agar konten di dalam *game* sesuai dengan apa yang diperbolehkan untuk dimainkan oleh anak. Namun demikian, kekerasan dalam *game* tersebut masih sering ditemui, misalnya secara visual pada adegan kekerasan yang dimungkinkan karena adanya fitur *user-generated content* dan secara verbal pada fitur chat pada *game* tersebut. Hal tersebut seringkali terlepas dari pengamatan orang tua saat anak-anak bermain *game online*.

Penelitian ini mengusulkan suatu model pendeteksi kekerasan secara multimodal, yaitu pada sisi visual dan verbal pada *game* berbasis *deep learning*. Penelitian ini menggunakan empat kelas untuk merepresentasikan kekerasan, yaitu kekerasan visual, kekerasan verbal, kekerasan visual dan verbal atau tidak ditemukan kekerasan. Masukan dari model ini berupa video hasil *screen capture* dari *game online* pada saat dimainkan oleh anak. Selanjutnya dilakukan ekstraksi fitur pada sisi visual dan sisi verbal. Berikutnya, data yang sudah diekstraksi kemudian ditraining dengan menggunakan 3DCNN, BiLSTM dan Attention Mechanics untuk mendeteksi kekerasan dari sisi visual serta BERT dan BiLSTM untuk mendeteksi kekerasan dari sisi verbal.

Hasil dari modalitas visual mencapai rata-rata akurasi 99.03% setelah dilatih dan divalidasi pada 3 dataset yaitu *Hockey Dataset*, *Violent Movie Dataset* dan *Online Game Violence Dataset*. Sedangkan hasil dari modalitas verbal mencapai akurasi 94.30% setelah dilatih dan divalidasi menggunakan *Indonesian Chat Dataset*. Pendekatan multimodal dengan *hybrid late fusion* memiliki hasil akurasi dari modalitas visual dan verbal sebesar 96.67%.

Kata kunci: *Deteksi kekerasan, Multimodal, Deep Learning, Game Online*



ABSTRACT

MULTIMODAL DEEP LEARNING VIOLENCE DETECTION IN CHILD-FRIENDLY RATED ONLINE GAMES

By:

Jasson Prestiliano

22/500409/SPA/00854

Online games rated as child-friendly are expected to be developed with children's safety in mind. As a result, parents tend to trust these games and ease their supervision during gameplay. Unfortunately, user-generated features in these games could expose underage players to verbal and visual violence, which can be detrimental to children's psychological well-being.

This research proposes a multimodal violence detection model on deep learning-based games' visual and verbal sides. Violence is defined into four classes, namely visual violence, verbal violence, visual and verbal violence, or no violence. The input from this model is a video of a screen capture of an online game when played by children. Next, feature extraction was carried out on the visual and verbal sides. Next, the extracted data was trained using 3DCNN, BiLSTM, and Attention Mechanics to detect visual violence and BERT and BiLSTM to detect verbal violence.

The visual modality's results reached an average accuracy of 99.03% after being trained and validated on three datasets: the Hockey Dataset, the Violent Movie Dataset, and the Online Game Violence Dataset. Meanwhile, the verbal modality's results reached an accuracy of 94.30% after being trained and validated using the Indonesian Chat Dataset. The multimodal approach with *hybrid late fusion* had visual and verbal modality accuracy of 96.67%.

Keywords: *Violence Detection, Multimodal, Deep Learning, Game Online*