ABSTRAK

Kemajuan teknologi kecerdasan buatan (AI) dalam hal manipulasi konten visual, khususnya video, menjadi semakin canggih, salah satunya melalui *deepfake*, yaitu metode yang menggunakan jaringan saraf untuk menghasilkan atau mengubah video secara realistis. Meskipun teknologi ini bermanfaat dalam bidang hiburan dan pendidikan, potensi penyalahgunaannya, seperti pembuatan video palsu yang menyesatkan atau merusak reputasi, menjadi ancaman serius. Untuk mengatasi masalah ini, pendekatan berbasis *Generative Adversarial Networks* (GAN) menawarkan solusi yang menjanjikan. GAN terdiri dari generator yang menghasilkan data palsu dan discriminator yang membedakan antara data asli dan palsu. Dalam konteks klasifikasi *deepfake*, GAN dilatih secara kompetitif untuk mengenali manipulasi video secara adaptif, memanfaatkan pola unik dalam *deepfake* yang sulit diidentifikasi oleh manusia. Pendekatan ini tidak hanya meningkatkan akurasi klasifikasi seiring dengan kompleksitas data tetapi juga memungkinkan proses yang lebih cepat dibandingkan metode konvensional. Implementasi GAN dalam klasifikasi video manipulasi memiliki dampak signifikan pada keamanan digital dan sosial, melindungi dari penyebaran informasi palsu, serta meningkatkan kepercayaan publik terhadap media digital. Penelitian ini merupakan langkah penting dalam menjaga integritas informasi di era digital.

Kata kunci : AI, GAN, Video Manipulasi