

BAB I PENDAHULUAN

1.1 Latar Belakang

Video call atau panggilan video merupakan salah satu hasil pengembangan lebih lanjut dalam bidang komunikasi dengan memanfaatkan hasil tangkapan kamera yang berupa video atau gambar yang dapat bergerak serta terdapat juga suara yang akan ditransmisikan secara *real time* melalui jaringan internet (Partiwi, 2017). Salah satu manfaat dari panggilan video yaitu memudahkan interaksi manusia dalam melakukan komunikasi jarak jauh secara tatap muka. Dalam perkembangannya, teknologi panggilan video sudah diterapkan ke beberapa aplikasi komunikasi salah satunya yaitu whatsapp, Line dan sebagainya.

Selain itu panggilan video juga menjadi alat penting dalam hal penggunaannya bagi beberapa orang, termasuk salah satunya bagi para psikolog yang menggunakannya untuk konsultasi jarak jauh dengan pasien. Pemanfaat panggilan video sebagai tempat melakukan konsultasi jarak jauh atau konsultasi *online* memiliki potensi yang cukup besar dengan 93% para remaja tertarik melakukan konsultasi online (Mirawati, 2015). Melihat potensi yang begitu besar ini, perlu adanya inovasi teknologi untuk mendukung proses konsultasi, khususnya dalam hal analisis emosi. Menurut Steinberg (2003) remaja merupakan masa seseorang mengalami gejolak emosional dan sosial yang tinggi (Mirawati, 2015). Oleh karena itu, ketepatan dalam mengenali dan merespons emosi remaja selama konsultasi sangat krusial. Jika penanganan emosi tidak tepat, maka solusi yang diberikan pun berpotensi tidak efektif dan malah memperburuk situasi. Salah satu teknologi yang memiliki potensi yang dapat digunakan pada panggilan video yaitu *facial emotion recognition*. Dengan teknologi ini, konselor diharapkan menjadi lebih memahami kondisi emosional klien secara *real-time* dan memberikan respon yang lebih tepat. Hal ini sangat penting mengingat 70% komunikasi nonverbal terjadi melalui ekspresi wajah (Collier & Collier, 2014).

Secara umum pengenalan emosi pada wajah atau yang disebut juga *Facial Expression Recognition* (FER) sudah banyak diimplementasikan secara luas dalam berbagai sektor seperti interaksi komputer manusia (*Human Computer*

Interaction), identifikasi biometrik, analisis emosi manusia, dan berbagai bidang lainnya (Mahmud & Mamun, 2017). Maka dari itu secara teori, FER dapat diimplementasikan pada saat *video call* dengan menampilkan emosi secara *real time* pada psikolog atau pendengar.

Pada fitur pengenalan emosi, untuk pengembangannya akan didasarkan pada pembelajaran mimik wajah manusia yang dilatih menggunakan algoritma *Deep Generative Adversarial Networks* (DGANs) serta menerapkan metode *Two-branch Disentangled* dalam menjalankan proses klasifikasinya. Penggunaan algoritma serta metode tersebut didasarkan pada penelitian yang telah dilakukan pada jurnal yang ditulis oleh Siyue Xie pada tahun 2021, yang menjelaskan manfaat dari pendalaman gambar detail pada algoritma GANs untuk membantu proses model klasifikasi emosi wajah (Xie dkk, 2021). Dengan mengkombinasikan arsitektur model klasifikasi yang tepat dan penggabungan atau *transfer learning* dengan model GANs secara langsung dapat membantu proses klasifikasi menjadi lebih efektif dengan akurasi yang didapatkan sebesar 97.53% untuk *dataset* CK+, 97,20% untuk *dataset* TFEID dan 99.32% untuk *dataset* RaFD (Xie dkk, 2021).

Salah satu kekurangan dari algoritma GANs yaitu dibutuhkan waktu dan data yang banyak dalam melakukan pembelajarannya. Banyak sekali *dataset* FER yang tersedia di internet, akan tetapi kebanyakan data tersedia merupakan wajah orang barat, afrika dan asia timur akan tetapi sangatlah sedikit *dataset* wajah orang indonesia atau asia selatan. Sementara itu, *dataset* yang tersedia mempunyai kemiripan wajah dengan kebanyakan orang indonesia sangatlah rendah. Justru banyak sekali perbedaannya seperti perbedaaan warna rambut, pola mata, warna kulit dan sebagainya. Pada jurnal (Xie dkk, 2021) terdapat pembahasan yang menjelaskan suatu konsep yang cukup menarik yaitu konsep penggabungan antara 2 gambar dengan memisahkan antara ekspresi wajah dengan atribut wajah lainnya. Maka dari itu, pada penelitian ini akan dilakukan pengembangan lebih lanjut yang bertujuan untuk menjawab sebuah kemungkinan potensi bahwasannya jika penggabungan tersebut dapat dilakukan dalam hal mengkombinasi data wajah yang akan dikumpulkan dengan wajah dari *dataset* kaggle akankah berpengaruh

terhadap akurasi deteksi emosi wajah khususnya pada penggunaan oleh orang indonesia.

Dengan penggunaan *dataset* tambahan dari Kaggle dan penambahan tahap penggabungan antara 2 data wajah untuk memaksimalkan hasil akurasi yang didapatkan, dibandingkan menggunakan algoritma tradisional GANs penelitian ini mengusulkan algoritma lebih lanjut dari GANs yaitu algoritma *Deep Generative Adversarial Network* (DGAN). Hal yang menjadikan DGAN sebagai algoritma utama penelitian ini yaitu karena pada model GANs tradisional menggunakan berbagai fungsi yang berhubungan dengan kepadatan probabilitas dalam memperkirakan suatu distribusi yang membuat tidak dapat bekerja dengan optimal jika digunakan pada distribusi kompleks (Alotaibi, 2020). Sebaliknya dengan menggunakan DGAN dapat memungkinkan mempelajari probabilitas distribusi yang kompleks secara efektif, yang membuat hasil yang didapatkan jauh lebih baik dibandingkan dengan GANs tradisional.

1.2 Rumusan Masalah

Dari latar belakang tersebut maka dapat disimpulkan beberapa poin inti yang akan disusun dalam rumusan masalah berikut ini:

1. Bagaimana penerapan pendeteksi emosi menggunakan algoritma DGANs pada fitur panggilan video di sistem konsultasi ?
2. Bagaimana tingkat akurasi model yang telah dikembangkan pada sistem konsultasi ?
3. Apakah mengkombinasi data dari sistem dengan data wajah orang dapat berpengaruh terhadap hasil akurasi deteksi emosi wajah. ?

1.3 Tujuan

Berdasarkan rumusan masalah pada pembahasan sebelumnya, di rancanglah lebih lanjut terkait tujuan penelitian ini. Maka dari itu, terdapat beberapa tujuan pada penelitian kali ini yang diantaranya yaitu:

1. Mengembangkan model yang dapat melakukan klasifikasi suatu emosi berdasarkan gambar mimik wajah yang sekaligus dapat diterapkan pada sistem konsultasi.

2. Menguji keakuratan model deteksi emosi serta memastikan pada saat diimplementasikan ke dalam sistem konsultasi berjalan dengan baik.
3. Menguji lebih lanjut konsep penggabungan 2 gambar antara gambar wajah yang akan dikumpulkan atau wajah umum masyarakat Indonesia dengan gambar wajah yang bervariasi yang didapatkan dari internet.

Dengan tujuan tersebut, diharapkan pada saat melakukan penelitian bisa lebih fokus dan tidak keluar jalur dari akar permasalahan yang telah dirumuskan sebelumnya.

1.4 Manfaat

Berdasarkan tujuan yang telah dirancang sebelumnya, diharapkan penelitian kali ini memiliki beberapa manfaat diantaranya yaitu:

1. Memudahkan atau memberikan fleksibilitas pada pengguna baik itu pasien maupun psikolog dalam melakukan sesi konsultasi.
2. Diharapkan sistem mampu memberikan gambaran emosi secara akurat sehingga meningkatkan efektivitas konsultasi serta membantu dalam memahami dan merespons kondisi emosional khususnya pada usia remaja.
3. Diharapkan dapat mengurangi beban kinerja dari algoritma GAN dalam melakukan pembelajaran gambar dan klasifikasi.

Dengan beberapa manfaat tersebut, diharapkan juga penelitian ini memiliki dampak baik khususnya dalam masyarakat luas.

1.5 Batasan Masalah

Agar masalah tidak terlalu luas dari tujuan dan manfaat pada penelitian kali ini maka diperlukan beberapa batasan masalah diantaranya yaitu:

1. Sistem ini membatasi layar minim dengan lebar sebesar 400 pixel, batasan ini ditetapkan untuk memastikan responsif dan tampilan yang optimal.
2. Fokus pembahasan utama pada penelitian ini yaitu penerapan pendeteksi emosi pada panggilan video.
3. Pada penelitian ini, untuk menerapkan video conference terdapat kemungkinan untuk menggunakan API pihak ketiga yaitu dari agora.io, [Twilio](https://twilio.com), [Metered.ca](https://metered.ca) dan sebagainya.

4. Untuk pengambilan data pada penelitian ini dilakukan secara mandiri dan sukarela. Dengan Format pengambilan data berupa video dengan durasi minimal 10 detik dan maksimal 30 detik.
5. Alat pengambilan data minimal menggunakan perangkat telepon pintar dan tidak ada batasan berapa megapiksel untuk ukuran kamera. Serta resolusi minimum hasil perekaman yaitu 1280 piksel untuk panjangnya dan 720 piksel untuk lebarnya.
6. Untuk perangkat yang digunakan untuk pelatihan menggunakan Google Collab Pro dengan 200 komputasi unit yang menggunakan jenis *runtime* T4 GPU dan didapatkan total 108 jam komputasi.
7. Sementara untuk pengujian menggunakan komputer pribadi dan Google Collab Gratis. Untuk Spesifikasi komputer pribadi yaitu Intel Core i5 6600 dan total 20 GB DDR4 ram yang digunakan.

Halaman ini sengaja dikosongkan