

BAB 5. PENUTUP

5.1. Kesimpulan

Berdasarkan penelitian ini, dapat disimpulkan bahwa aplikasi yang dirancang dan dibangun berhasil menjawab kebutuhan pengguna dengan menghadirkan solusi pendeteksi *Deepfake* yang dapat memprediksi secara cepat, aplikasi berukuran ringan, dan dapat membagikan hasil prediksi ke media sosial eksternal dengan tingkat penerimaan berada pada kategori puas.

1. Perancangan dan pembangunan aplikasi dirancang agar model yang digunakan dapat berjalan secara *on-device* dengan parameter yang sedikit, sehingga hanya membutuhkan penyimpanan yang sedikit dan dapat berjalan secara cepat. Waktu yang dibutuhkan untuk melakukan inferensi sebesar 0,9 detik atau dibawah 1 detik, sehingga tetap menjaga kenyamanan pengguna. Pembangunan aplikasi menggunakan perintah ‘flutter build apk --split-per-abi’ membentuk tiga aplikasi untuk tiga jenis arsitektur yang berbeda, serta membuat ukuran aplikasi menjadi lebih kecil. Ukuran terbesar dari aplikasi yang dibangun berada pada angka 45.88 MB, yang masih ditoleransi oleh pengguna. Penggunaan *plugin share_plus* dari *Framework Flutter* membuat aplikasi memiliki kemampuan untuk membagikan hasil prediksi ke media sosial.
2. Tingkat kualitas sistem, informasi, dan antarmuka pengguna dinilai berdasarkan pengujian bersama calon pengguna menggunakan metode *Post-Study System Usability Questionnaire* (PSSUQ). Dari hasil pengujian, diketahui pengguna merasa puas pada semua skala. Skala kualitas sistem (SysQual) mendapatkan rata-rata nilai sebesar 1,88, kualitas informasi (InfoQual) dengan rata-rata nilai sebesar 1,86, serta kualitas antarmuka (IntQual) dari aplikasi yang dibangun dengan rata-rata nilai sebesar 2,01. Selain itu, penilaian secara keseluruhan (*overall*) dari aplikasi juga membuktikan bahwa pengguna merasa puas dengan aplikasi yang dirancang dan dibangun dengan rata-rata nilai sebesar 1,90. Hal ini membuktikan bahwa perancangan dan pembangunan aplikasi pada penelitian ini secara keseluruhan telah berhasil memenuhi kebutuhan serta memberikan tingkat kepuasan yang optimal bagi pengguna.

5.2. Saran

Berdasarkan proses pengembangan aplikasi dan evaluasi yang telah dilakukan, terdapat beberapa rekomendasi untuk menyempurnakan sistem pada penelitian berikutnya.

1. Tingkatkan Performa Deteksi pada Kelas *Deepfake*
Hasil evaluasi model terpilih menunjukkan bahwa *recall* pada kelas *Deepfake* berada pada angka 75%, yang berarti terdapat sekitar 1 dari 4 konten *Deepfake* yang berpotensi tidak terdeteksi. Untuk penelitian berikutnya, disarankan untuk mengeksplorasi strategi peningkatan *recall* kelas *Deepfake*, misalnya melalui

penyesuaian *threshold* klasifikasi, penggunaan *loss function* yang lebih sensitif terhadap ketidakseimbangan kelas, atau perluasan variasi *dataset*.

2. Perbanyak Variasi *Dataset* dengan Berbagai Resolusi dan Wajah Indonesia
Resolusi yang ada pada *dataset* Celeb-DF versi 2 cenderung sama dikarenakan diproduksi oleh tempat yang sama. Maka dari itu diperlukan peningkatan variasi. Selain itu, pada skema 3 yang dilakukan percobaan ini, terlihat bahwa performa model menurun akibat redundansi *frame*. Maka dari itu, untuk penelitian berikutnya dapat menggunakan variasi *dataset* yang lebih beragam untuk wajah-wajah Indonesia, baik pada data pelatihan atau pengujian.
3. Tambah Dukungan Deteksi dari Konten Video
Pada penelitian ini model masih hanya menerima inputan berupa gambar saja, dan memiliki potensi untuk menghasilkan prediksi yang berbeda tiap *frame*-nya. Penelitian kedepannya dapat dilakukan dengan menambahkan dukungan deteksi lewat video dengan model yang berjalan secara *on-device*.
4. Integrasikan Model dengan *Explainable AI* (XAI) dan *Large Language Model* (LLM)
Potensi solusi kedepannya dapat mengintegrasikan model dengan XAI, dan hasil dari XAI diolah oleh LLM untuk memberikan penjelasan ke pengguna mengenai alasan mengapa gambar atau konten tersebut merupakan *Deepfake*. Hal ini dapat meningkatkan pengalaman pengguna sekaligus memberikan edukasi.