Pengembangan Sistem Klasifikasi Ekspresi Wajah Untuk Monitoring Kondisi Work Life Balance Menggunakan Teknologi Image Processing Dan Deep Learning (CNN)

1st Pamela Alysia
Faculty of Electrical Engineering Telkom University
Bandung, Indonesia
pamelamel@student.telkomuniversity.ac.id

2nd Nabiel Prama Wijaya K.
Faculty of Electrical Engineering Telkom University
Bandung, Indonesia
nabielprm@student.telkomuniversity.ac.id

3rd Inung Wijayanto
Lecturer in Telecommunication Engineering, Telkom University
Bandung, Indonesia

4th R. Yunendah Nur Fu'adah
Lecturer in Telecommunication Engineering, Telkom University
Bandung, Indonesia

***Abstrak* —Perkembangan teknologi digital mengubah pola kerja dan meningkatkan tuntutan perusahaan dalam menjaga keseimbangan antara produktivitas dan kondisi emosional pegawai. Pemantauan *work life balance* selama ini masih bergantung pada metode manual seperti observasi dan kuesioner yang bersifat subjektif dan kurang akurat. Keterbatasan tersebut dapat menghambat deteksi dini permasalahan kesejahteraan mental karyawan. Penelitian ini bertujuan mengembangkan sistem otomatis untuk mengklasifikasikan ekspresi wajah sebagai indikator kondisi emosional pegawai. Sistem dirancang menggunakan metode *deep learning* dengan arsitektur MobileNetV2 dan pendekatan model *valence–arousal*. Dataset FER-2013 digunakan dalam proses pelatihan dan pengujian model. Evaluasi dilakukan terhadap beberapa parameter, meliputi jumlah epoch, ukuran batch, optimizer, dan *learning rate*. Hasil pengujian menunjukkan bahwa konfigurasi terbaik diperoleh pada 30 epoch, ukuran batch 16, optimizer Adam, dan *learning rate* 0,001, dengan akurasi pengujian sebesar 89,40% dan nilai *loss* 0,0520. Hasil tersebut menunjukkan bahwa model memiliki kinerja yang stabil dan efisien. Sistem yang dikembangkan berpotensi menjadi alat pendukung awal dalam pemantauan kondisi emosional pegawai secara objektif dan berkelanjutan.**

Kata kunci — *deep learning*, ekspresi wajah, fer-2013, mobilenetv2, *work life balance*

I. PENDAHULUAN

Perkembangan teknologi digital yang pesat telah mengubah pola kerja masyarakat modern dan semakin mengaburkan batas antara kehidupan profesional dan kehidupan pribadi. Kondisi ini menyebabkan isu *Work Life Balance* (WLB) menjadi perhatian penting, khususnya karena ketidakseimbangan antara tuntutan pekerjaan dan kehidupan personal dapat berdampak langsung terhadap kondisi emosional serta kesehatan mental pekerja. Berbagai studi menunjukkan bahwa stres kerja, kelelahan emosional, dan *burnout* merupakan konsekuensi umum dari buruknya *Work Life Balance* dan berpotensi menurunkan produktivitas serta kualitas hidup pekerja Namun demikian, penerapan CNN konvensional sering menghadapi kendala berupa kebutuhan komputasi yang tinggi dan ketergantungan pada dataset berukuran besar, sehingga kurang optimal untuk diimplementasikan pada perangkat dengan sumber daya terbatas. Untuk mengatasi permasalahan tersebut, pendekatan *transfer learning* dengan arsitektur ringan seperti *MobileNetV2* menjadi solusi yang efektif. *MobileNetV2* dirancang untuk memberikan keseimbangan antara akurasi dan efisiensi komputasi, sehingga cocok diterapkan pada sistem monitoring berbasis aplikasi bergerak

Penelitian ini bertujuan mengembangkan sistem klasifikasi ekspresi wajah berbasis *deep learning* sebagai alat bantu pemantauan kondisi emosional untuk mendukung evaluasi *Work Life Balance* secara objektif, efisien, dan berkelanjutan

II. KAJIAN TEORI

A. *Work Life Balance* (WLB)

Work Life Balance (WLB) merupakan kondisi keseimbangan antara tuntutan pekerjaan dan kehidupan pribadi yang memungkinkan individu mempertahankan kesejahteraan psikologis serta produktivitas kerja. Ketidakseimbangan WLB diketahui berkorelasi dengan meningkatnya stres kerja, kelelahan emosional (*burnout*), dan penurunan kualitas hidup pekerja. Kondisi emosional menjadi indikator penting dalam menilai *Work Life Balance* karena emosi negatif yang berlangsung secara berkelanjutan dapat mempengaruhi kinerja dan kesehatan mental individu [1].

B. *Convolutional Neural Network (CNN)*

Convolutional Neural Network (CNN) merupakan arsitektur jaringan syaraf tiruan yang dirancang untuk pengolahan data visual. CNN bekerja melalui lapisan konvolusi dan *pooling* untuk mengekstraksi fitur visual dari tingkat sederhana hingga kompleks. Dalam klasifikasi ekspresi wajah, CNN mampu mengenali pola penting seperti kontur wajah, tekstur, dan perubahan ekspresi tanpa memerlukan perancangan fitur secara manual. Penelitian sebelumnya menunjukkan bahwa CNN mampu mencapai performa tinggi pada dataset standar FER seperti FER-2013 dan JAFFE [2].

C. *MobileNetV2* dan *FER-2013*

MobileNetV2 merupakan arsitektur CNN ringan yang dirancang untuk efisiensi komputasi dengan

menggunakan *depthwise separable convolution* dan *inverted residual block*. Arsitektur ini mampu memberikan keseimbangan antara akurasi dan kecepatan pemrosesan, sehingga sesuai untuk aplikasi berbasis perangkat bergerak dan sistem monitoring real-time. Penggunaan MobileNetV2 dalam sistem FER terbukti mampu mempertahankan performa klasifikasi dengan kebutuhan komputasi yang lebih rendah dibandingkan CNN berukuran besar [3].

Facial Expression Recognition (FER) merupakan bidang dalam *computer vision* yang bertujuan mengidentifikasi kondisi emosional manusia melalui analisis citra wajah. Pendekatan awal FER menggunakan metode konvensional seperti *Haar Cascade*, *Principal Component Analysis* (PCA), dan *Local Binary Pattern* (LBP) yang dikombinasikan dengan algoritma klasifikasi seperti *Support Vector Machine* (SVM). Meskipun metode tersebut ringan secara komputasi, performanya cenderung menurun pada kondisi pencahayaan yang bervariasi dan pose wajah yang kompleks [4].

D. *Firebase* dan *Flutter*

Firebase digunakan sebagai *database real-time* dan autentikasi pengguna. Aplikasi *mobile* dikembangkan menggunakan *Flutter* karena mendukung lintas platform (Android/iOS) dan mudah dikembangkan. *Firebase* RTDB memungkinkan pembaruan data sensor secara langsung, dan *Flutter* mempermudah tampilan antarmuka pengguna yang inovatif dan interaktif.

III. METODE

Penelitian ini menggunakan pendekatan eksperimental dengan metode *deep learning* untuk mengembangkan sistem klasifikasi ekspresi wajah sebagai indikator kondisi emosional yang berkaitan dengan *Work Life Balance*. Desain penelitian meliputi tahapan pengumpulan dataset, pra-pemrosesan citra, pelatihan model *Convolutional Neural Network* (CNN) berbasis *transfer learning*, pengujian performa model, serta pemetaan hasil klasifikasi ke dalam model Valence–Arousal. Alur penelitian dirancang untuk menghasilkan sistem yang akurat, efisien secara komputasi, dan aplikatif pada perangkat dengan spesifikasi menengah.



GAMBAR 1
(Diagram Alir Pengerjaan Sistem Klasifikasi Wajah)

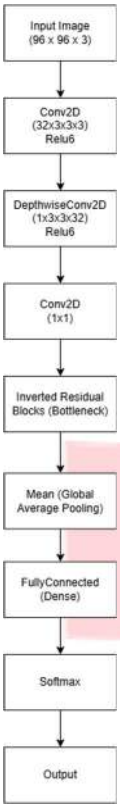
Gambar 1 menunjukkan Diagram alir tahapan pengerjaan sistem klasifikasi ekspresi wajah berbasis MobileNetV2, mulai dari persiapan dataset hingga analisis hasil pengujian sistem.

A. *Dataset*

Dataset yang digunakan dalam penelitian ini adalah *dataset* publik FER-2013, yang berisi citra wajah beresolusi 48×48 piksel dalam format *grayscale* dengan tujuh kelas ekspresi dasar, yaitu *happy, neutral, sad, angry, disgust, fear,* dan *surprise*. *Dataset* FER-2013 dipilih karena merupakan *dataset* standar yang banyak digunakan dalam penelitian *Facial Expression Recognition* (FER) dan telah terbukti efektif dalam evaluasi performa model CNN. Penggunaan dataset publik juga bertujuan menjaga reprodusibilitas dan validitas penelitian.

B. Arsitektur Model

Model klasifikasi ekspresi wajah dibangun menggunakan arsitektur *Convolutional Neural Network* dengan pendekatan *transfer learning* berbasis MobileNetV2. MobileNetV2 dipilih karena memiliki jumlah parameter yang relatif kecil dan efisien secara komputasi, sehingga sesuai untuk implementasi pada sistem monitoring berbasis aplikasi bergerak.Pada skema *transfer learning*. Lapisan awal model dipertahankan, sementara lapisan akhir disesuaikan untuk melakukan klasifikasi tujuh kelas ekspresi wajah. Pendekatan ini memungkinkan model mencapai performa optimal meskipun menggunakan *dataset* terbatas.



GAMBAR 2
(Arsitektur Model MobileNetV2)

Gambar 2 menunjukkan arsitektur model MobileNetV2 yang digunakan untuk klasifikasi ekspresi wajah, terdiri dari lapisan konvolusi awal, *inverted residual blocks*, *global average pooling*, dan *softmax* sebagai lapisan keluaran.

A. Pelatihan Model

Proses pelatihan model dilakukan dengan membagi dataset menjadi data pelatihan dan data pengujian. Model dilatih menggunakan optimizer *Adam* dengan *learning rate* sebesar 0,001, *batch size* 16, dan jumlah *epoch* sebanyak 30. Pemilihan parameter ini didasarkan pada hasil evaluasi eksperimental yang menunjukkan keseimbangan terbaik antara akurasi dan stabilitas model. Selama proses pelatihan, fungsi *loss* dan nilai akurasi dipantau untuk memastikan konvergensi model dan meminimalkan risiko *overfitting.*

B. Pra-Pemrosesan Citra

Pra-pemrosesan citra dilakukan untuk meningkatkan kualitas data masukan serta memastikan keseragaman input model. Tahapan pra-pemrosesan meliputi deteksi wajah, *cropping* area wajah, *resize* citra ke ukuran 48×48 piksel, normalisasi nilai piksel, serta konversi warna dari RGB ke *grayscale*. Selain itu, dilakukan *data augmentation* berupa rotasi ringan dan *horizontal flipping* untuk meningkatkan variasi data latih dan mengurangi risiko *overfitting*. Tahapan ini berperan penting dalam meningkatkan stabilitas dan akurasi model klasifikasi ekspresi wajah.

C. Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan metrik kuantitatif yang umum digunakan dalam sistem FER, yaitu *accuracy* dan *loss*. Model dikatakan memiliki performa yang baik apabila mampu mencapai tingkat akurasi di atas 80% dengan nilai *loss* yang rendah dan stabil. Pengujian dilakukan pada data uji yang tidak dilibatkan dalam proses pelatihan untuk memastikan kemampuan generalisasi model terhadap data baru.

D. Hasil dan Pembahasan

Proses pelatihan model klasifikasi ekspresi wajah dilakukan menggunakan arsitektur *Convolutional Neural Network* berbasis *MobileNetV2* dengan pendekatan *transfer learning*. *Dataset* yang digunakan adalah FER-2013 dengan citra wajah beresolusi 48×48 piksel dalam format *grayscale*. Model dilatih menggunakan *optimizer* Adam dengan *learning rate* 0,001, *batch size* 16, dan jumlah *epoch* sebanyak 30.

TABEL 1
(Pengujian Variasi *Epoch*)

Arsitektur	*Dataset*	*Epoch*	Akurasi			*Loss*
			Latih	Validasi	Uji	
MobileNetV2	FER-2013	20	97.10%	81.80%	83.40%	0.0812
		30	99.40%	85.30%	88.60%	0.0631
		50	99.60%	83.70%	86.10%	0.0792
		100	99.80%	80.60%	82.20%	0.1063

Berdasarkan hasil pengujian pada Tabel 1, jumlah *epoch* berpengaruh signifikan terhadap performa model. Model dengan 30 epoch menghasilkan akurasi uji tertinggi sebesar 88,60% dan nilai loss terendah sebesar 0,0631. Peningkatan epoch hingga 50 dan 100 tidak memberikan peningkatan performa yang signifikan, bahkan menyebabkan penurunan akurasi uji akibat *overfitting*.

TABEL 2
(Pengujian Variasi *Batch Size*)

Arsitektur	*Dataset*	*Epoch*	*Batch Size*	Akurasi			*Loss*
				Latih	Validasi	Uji	
MobileNetV2	FER-2013	30	16	99.72%	86.45%	89.10%	0.0512
			32	99.40%	85.30%	88.60%	0.0631
			64	98.15%	83.90%	86.40%	0.0824

Berdasarkan hasil pengujian pada tabel 2, dapat disimpulkan bahwa *batch size* berpengaruh signifikan terhadap performa model. *Batch size* yang terlalu besar cenderung menurunkan kemampuan generalisasi model, sedangkan *batch size* yang lebih kecil menghasilkan pembelajaran yang lebih

stabil. Pada penelitian ini, *batch size* 16 memberikan keseimbangan terbaik antara akurasi, stabilitas, dan nilai *loss* yang rendah.

TABEL 3
(Pengujian Variasi *Optimizer*)

Arsitektur	Optimizer	Akurasi			Loss
		Latih	Validasi	Uji	
MobileNetV2	Adam	99.72%	86.45%	89.10%	0.0512
	Nadam	99.85%	86.20%	88.95%	0.0495
	SGD	75.40%	70.10%	72.30%	0.7210

Hasil pengujian *optimizer* pada Tabel 3 menunjukkan bahwa Adam dan Nadam memberikan performa yang lebih baik dibandingkan SGD. Adam menghasilkan akurasi uji tertinggi sebesar 89,10%, sedangkan SGD menunjukkan performa terendah dengan akurasi uji sebesar 72,30% dan nilai *loss* yang tinggi.

TABEL 4
(Pengujian Variasi *Learning Rate*)

Arsitektur	Learning Rate	Akurasi			Loss
		Latih	Validasi	Uji	
MobileNetV2	0.0001	95.20%	84.10%	86.30%	0.1250
	0.001	99.60%	86.10%	89.40%	0.0520
	0.01	45.30%	38.20%	40.10%	1.2400

Berdasarkan hasil pengujian *learning rate* pada Tabel 4, nilai *learning rate* 0.001 menghasilkan akurasi uji tertinggi sebesar 89,40% dan nilai *loss* terendah sebesar 0,0520. *Learning rate* yang terlalu kecil maupun terlalu besar menyebabkan penurunan performa model akibat *underfitting* dan ketidakstabilan pelatihan.

KESIMPULAN

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa sistem klasifikasi ekspresi wajah berbasis *deep learning* menggunakan arsitektur MobileNetV2 berhasil dikembangkan dan menunjukkan performa yang baik dalam mengklasifikasikan ekspresi wajah pada dataset FER-2013. Penerapan pendekatan *transfer learning* terbukti efektif dalam meningkatkan akurasi model sekaligus menjaga efisiensi komputasi. Hasil evaluasi menunjukkan bahwa konfigurasi terbaik diperoleh pada penggunaan 30 *epoch*, *batch size* 16, *optimizer* Adam, dan *learning rate* 0.001, yang menghasilkan akurasi uji sebesar 89,40% dengan nilai *loss* yang rendah dan stabil. Konfigurasi ini memberikan keseimbangan optimal antara kemampuan pembelajaran model dan kemampuan generalisasi terhadap data baru

REFERENSI

[1] International Telecommunication Union, Trends in Telecommunication Reform, ITU Publications, Geneva, 2022
[2] Syahfitri, S. Li et al., "Deep CNN for Facial Expression Recognition," *Pattern Recognition*, 2019
[3] M. Sandler et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *CVPR*, 2018.
[4] Malhotra et al., "Real-Time Face Detection Using Haar Cascade," *International Journal of Computer Vision*, 2021.