

BAB I

PENDAHULUAN

1.1. Latar Belakang

Aktivitas yang men-distraksi saat mengemudi merupakan salah satu faktor utama penyebab kecelakaan lalu lintas di berbagai negara. Aktivitas seperti menggunakan ponsel, berbicara dengan penumpang, minum, atau merias wajah dapat mengalihkan perhatian pengemudi sehingga meningkatkan risiko kecelakaan, cedera, bahkan kematian. Organisasi Kesehatan Dunia (WHO) melaporkan bahwa sekitar 1,35 juta orang meninggal setiap tahun akibat kecelakaan lalu lintas. Di Indonesia, data Korlantas Polri tahun 2022 mencatat sebanyak 94.600 kasus kecelakaan, dan perhatian pengemudi yang terdistraksi merupakan salah satu penyebab yang paling dominan [1][2]. Data ini menunjukkan bahwa aktivitas yang men-distraksi fokus saat berkendara masih menjadi masalah serius yang memerlukan penanganan segera dalam konteks keselamatan jalan raya.

Perkembangan teknologi *computer vision* dalam beberapa tahun terakhir telah mendorong munculnya berbagai pendekatan berbasis *deep learning* untuk mengidentifikasi aktivitas pengemudi yang terdistraksi menggunakan citra, hal ini dilakukan dalam upaya untuk mengurangi risiko kecelakaan lalu lintas yang diakibatkan oleh manusia yang terdistraksi dalam mengemudi [3]. Tantangan utama pada *domain* ini meliputi variasi aktivitas distraksi yang dapat mempersulit perancangan model *universal* yang dapat diterapkan pada semua skenario, serta kebutuhan model yang efisien untuk dapat digunakan pada skenario *real-time* karena identifikasi pengemudi yang terdistraksi secara tepat waktu sangatlah penting [4]. Model *deep learning* dengan ukuran besar umumnya memberikan akurasi tinggi, tetapi membutuhkan sumber daya komputasi yang besar sehingga kurang sesuai untuk aplikasi bergerak atau perangkat dengan kemampuan terbatas [5]. Oleh karena itu, dibutuhkan model yang ringan namun tetap memiliki kemampuan untuk mengekstraksi fitur yang akurat dan *robust*.

Dalam beberapa tahun terakhir, terdapat tren penelitian yang signifikan pada pengembangan model *deep learning* ringan untuk tugas klasifikasi aktivitas

distraksi pengemudi. Penelitian terbaru berfokus pada arsitektur *Convolutional Neural Network* (CNN) efisien seperti *EfficientNetB0* yang dapat digunakan dan menghasilkan akurasi yang baik walau dengan sumber daya komputasi yang terbatas [6]. Namun penelitian tersebut dan penelitian terdahulu lainnya [7] dalam konteks klasifikasi aktivitas distraksi pengemudi ini, semua hanya terbatas menggunakan dataset sekunder yang kondisi pencahayaannya relatif sama semua. Seluruh penelitian terdahulu tersebut menggunakan dataset sekunder *open-source* yang bernama *State Farm Distracted Driver dataset*, yang berisi 10 label aktivitas distraksi pengemudi, yaitu c0 (*safe driving*), c1 (*Texting-right*), c2 (*Calling on the phone-right*), c3 (*Texting-left*), c4 (*Calling on the phone-left*), c5 (*Operating the Radio*), c6 (*Drinking*), c7 (*Reaching behind*), c8 (*Hair and makeup*), dan c9 (*Talking to the passenger*).

Penelitian terbaru sebelumnya [6] dalam konteks klasifikasi aktivitas distraksi pengemudi ini menggunakan arsitektur *EfficientNetB0* dengan menambahkan teknik *attention mechanisms Channel Attention* (CA), dan terbukti mampu meningkatkan sensitivitas model terhadap fitur penting tanpa menambah parameter yang signifikan dengan nilai akurasi sebesar 97% jika tanpa menambahkan CA dan 99% dengan menambahkan CA. Pada penelitian ini juga sudah membandingkan *EfficientNetB0+CA* dengan arsitektur CNN lainnya seperti *VGG16*, *ResNet50*, *Xception*, dan *InceptionV3*, dengan tambahan CA untuk masing-masing arsitektur dan tanpa tambahan CA. Hasil penelitian ini tetap menunjukkan bahwa model yang diusulkan, yaitu *EfficientNetB0+CA* memiliki nilai akurasi tertinggi sehingga lebih disarankan untuk klasifikasi aktivitas distraksi pengemudi. Penelitian terdahulu lainnya [7] juga melakukan klasifikasi aktivitas distraksi pengemudi, tetapi pada penelitian ini membandingkan arsitektur CNN dasar, *VGG16*, *ResNet50*, dan *MobileNetV2*. Hasil penelitian ini menunjukkan bahwa *MobileNetV2* memiliki nilai akurasi tertinggi sebesar 98%.

Dari penelitian terdahulu untuk klasifikasi aktivitas distraksi pengemudi tersebut, dapat diketahui bahwa penelitian terdahulu tersebut hanya terbatas menggunakan dataset sekunder yang kondisi pencahayaannya relatif sama semua, yaitu dataset *State Farm Distracted Driver dataset*. Hal tersebut menjadi masalah

karena secara realistis tidak semua orang mengemudi hanya pada satu kondisi pencahayaan saja (seperti pagi atau siang hari saja), tetapi orang mengemudi pada kondisi pencahayaan yang beragam, karena orang mengemudi di waktu yang berbeda-beda. Variasi pencahayaan nyata seperti kondisi pagi dengan bayangan tajam, siang yang cerah, sore dengan arah sinar miring, dan malam dengan kontras rendah dapat mengubah distribusi visual citra secara drastis dan berdampak pada kemampuan arsitektur CNN untuk mengenali pola yang sama, sehingga memunculkan tantangan tersendiri dalam penerapan model klasifikasi aktivitas pengemudi pada kondisi dunia nyata [8][9]. Hal ini menyebabkan *gap* yang perlu diteliti karena arsitektur CNN seperti *EfficientNetB0* kinerjanya dapat terpengaruhi dengan kondisi pencahayaan yang beragam, sedangkan penelitian sebelumnya hanya terbatas menggunakan dataset yang kondisi pencahayaannya relatif seragam.

EfficientNetB0 menawarkan efisiensi melalui desain *compound scaling* yang mengoptimalkan akurasi dan ekstraksi fitur multiskala menggunakan *inverted residual blocks* dengan ekspansi linier yang memungkinkan pengenalan pola mulai dari tekstur mikro hingga karakteristik makro [10]. Berdasarkan penelitian terbaru sebelumnya [6] untuk klasifikasi aktivitas distraksi pengemudi juga ditunjukkan bahwa *EfficientNetB0* dengan tambahan CA menghasilkan akurasi tertinggi. Keunggulan *EfficientNetB0* ini juga telah dibuktikan oleh beberapa penelitian lain, seperti penelitian [11] yang menunjukkan bahwa tingkat akurasi *EfficientNetB0* (85,12%) hampir menyamai *ResNet-50* (85,79%) dengan parameter yang kira-kira empat kali lebih sedikit untuk klasifikasi cacar monyet berdasarkan gambar lesi kulit. Kemudian penelitian terbaru [12] juga menunjukkan bahwa *EfficientNetB0* secara umum lebih baik dibandingkan dengan arsitektur CNN ringan lainnya seperti *MobileNetV2* yang juga sering digunakan, dengan nilai akurasi sebesar 80% untuk *EfficientNetB0* dan 75% untuk *MobileNetV2* dalam klasifikasi buah berbasis citra digital. Hal tersebut juga didukung oleh penelitian [13] yang membandingkan arsitektur *EfficientNetB0* dengan *MobileNetV2* untuk klasifikasi koloni bakteri, yang menunjukkan bahwa *EfficientNetB0* memiliki nilai akurasi yang lebih tinggi (94% untuk *EfficientNetB0* dan 86% untuk *MobileNetV2*) dan lebih stabil.

Penambahan *attention mechanisms* dapat meningkatkan performa arsitektur CNN dengan memfokuskan area yang seharusnya diidentifikasi [14]. Kombinasi arsitektur CNN dengan teknik *attention mechanisms* telah menunjukkan hasil yang menjanjikan dalam meningkatkan performa model dengan memfokuskan bagian yang paling relevan dari data yang diberikan [15]. Penelitian sebelumnya [16] menerapkan teknik *attention mechanisms* pada arsitektur *EfficientNet* untuk meningkatkan klasifikasi citra melalui *feature enhancement* yang menyeimbangkan informasi fitur sehingga hasil penelitian menunjukkan nilai akurasi tertinggi sebesar 97% untuk klasifikasi makanan. Penambahan *attention mechanism* seperti *Channel Attention (CA)* pada *EfficientNetB0* untuk klasifikasi aktivitas distraksi pengemudi juga telah terbukti dapat meningkatkan kemampuan model hingga mencapai akurasi sebesar 99% dengan kecepatan inferensi mencapai puluhan *frame* per detik [6]. Pengembangan *attention mechanisms* juga ditunjukkan membantu memperbaiki citra dalam variasi intensitas cahaya rendah, sehingga dapat membantu model klasifikasi menangani masalah variasi tingkat intensitas pencahayaan [17][18].

Selain *attention mechanisms*, teknik augmentasi data juga telah terbukti menjadi strategi yang efektif dalam meningkatkan kinerja arsitektur CNN terhadap variasi kondisi visual, termasuk variasi pencahayaan. Augmentasi data bekerja dengan memperluas distribusi data pelatihan secara artifisial melalui transformasi seperti *brightness adjustment*, *contrast adjustment*, *rotating*, *color jitter*, serta penambahan *blur* dan *noise*, sehingga model dapat belajar mengenali pola pada perubahan kondisi visual secara konsisten [19]. Penelitian sebelumnya [20] menunjukkan bahwa penggunaan teknik augmentasi data pada arsitektur *EfficientNetB0* dapat menghasilkan akurasi sebesar 98% untuk klasifikasi leukimia, hal ini menunjukkan bahwa augmentasi data dapat bekerja dengan baik pada arsitektur *EfficientNetB0*. Kemudian teknik seperti *Contrast Limited Adaptive Histogram Equalization (CLAHE)* yang diintegrasikan sebagai bagian dari *pipeline* augmentasi juga terbukti mampu meningkatkan visibilitas fitur pada citra dengan pencahayaan yang beragam, sehingga dapat memperbaiki kinerja model klasifikasi

secara keseluruhan [21]. Oleh karena itu penggunaan teknik augmentasi data juga digunakan untuk menangani masalah kondisi pencahayaan yang beragam.

Dengan keunggulan yang ditunjukkan oleh *EfficientNetB0* dan dengan tambahan *attention mechanisms* seperti *channel attention (CA)* beserta penggunaan teknik augmentasi data, penelitian ini berfokus untuk menerapkan CA dan augmentasi data pada *EfficientNetB0* dalam mengklasifikasikan aktivitas pengemudi terdistraksi berdasarkan citra yang mencakup variasi kondisi pencahayaan nyata di lingkungan Indonesia (pagi, siang, sore, malam). Hal tersebut dilakukan untuk mengisi *gap* dari penelitian sebelumnya yang hanya menggunakan dataset dengan pencahayaan seragam [6][7]. Perbedaan data pada kondisi pencahayaan tersebut menyebabkan penelitian menggunakan CA dan augmentasi data pada *EfficientNetB0* untuk mengklasifikasikan aktivitas pengemudi terdistraksi berdasarkan citra dengan kondisi pencahayaan yang beragam diperlukan, beserta dengan hasil evaluasinya untuk mengetahui lebih lanjut performa model ini pada kondisi data yang lebih realistis. Kemudian implementasi model secara *real-time* juga diperlukan karena dalam konteks klasifikasi aktivitas pengemudi terdistraksi, pengujian dan evaluasi model secara *real-time* dalam kondisi nyata diperlukan untuk mengetahui apakah secara keseluruhan model layak untuk diterapkan dalam kondisi nyata [22]. Oleh karena itu, penelitian ini juga mengimplementasikan model ke dalam *prototype early warning system* berbasis *computer vision* untuk menguji dan mengevaluasi model secara *real-time* dalam kondisi nyata. Penelitian ini diharapkan dapat memberikan pemahaman yang lebih mendalam mengenai pengaruh pencahayaan terhadap model CNN dan juga pengaruh teknik *attention mechanisms* dan augmentasi data pada data citra yang kondisi pencahayaannya beragam, serta memberikan kontribusi ilmiah terhadap pengembangan model *computer vision* modern yang mampu menghadapi kompleksitas variasi lingkungan nyata.

1.2. Rumusan Masalah

Rumusan masalah yang menjadi fokus utama pada penelitian ini adalah sebagai berikut:

1. Bagaimana menerapkan metode *channel attention* dan augmentasi data pada *EfficientNetB0* dalam melakukan klasifikasi aktivitas distraksi pengemudi pada kondisi pencahayaan beragam?
2. Bagaimana evaluasi hasil dari metode *EfficientNetB0* dengan *channel attention* dan augmentasi data dalam melakukan klasifikasi aktivitas distraksi pengemudi pada kondisi pencahayaan beragam?
3. Bagaimana mengimplementasikan hasil dari metode *EfficientNetB0* dengan *channel attention* dan augmentasi data untuk klasifikasi aktivitas distraksi pengemudi secara *real-time*?

1.3. Batasan Masalah

Batasan yang ditetapkan pada masalah yang diamati pada penelitian ini adalah sebagai berikut:

1. Bahasa Program yang digunakan dalam penelitian ini adalah *python*.
2. Data yang digunakan dalam penelitian ini adalah data sekunder dan data primer. Data primer diperoleh dari pengambilan citra pengemudi di Indonesia, Jawa Timur, Surabaya secara pribadi, sementara data sekunder diambil dari *Kaggle.com* yang digunakan oleh penelitian-penelitian sebelumnya, yaitu *State Farm Distracted Driver Detection*.
3. Kelas atau label yang digunakan dalam penelitian ini telah ditentukan, yaitu terdiri dari 10 kelas: c0 (*safe driving*), c1 (*Texting-right*), c2 (*Calling on the phone-right*), c3 (*Texting-left*), c4 (*Calling on the phone-left*), c5 (*Operating the Radio*), c6 (*Drinking*), c7 (*Reaching behind*), c8 (*Hair and makeup*), c9 (*Talking to the passenger*).
4. Data gambar aktivitas pengemudi primer yang digunakan diambil manual menggunakan kamera *webcam*, serta standarisasi pencahayaan dilakukan menggunakan *smarthphone*.
5. Penelitian ini dirancang untuk mengevaluasi kinerja *channel attention* dan augmentasi data pada *EfficientNetB0* dalam klasifikasi aktivitas distraksi pengemudi menggunakan data sekunder dan primer yang kondisi pencahayaannya beragam.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menjawab rumusan masalah serta memberikan arahan yang lebih jelas bagi penelitian ini. Adapun tujuan penelitian ini adalah sebagai berikut:

1. Melakukan proses metode *channel attention* dan augmentasi data pada *EfficientNetB0* untuk mengklasifikasi aktivitas distraksi pengemudi pada kondisi pencahayaan yang beragam.
2. Melakukan proses evaluasi hasil dari metode *EfficientNetB0* dengan *channel attention* dan augmentasi data dalam melakukan klasifikasi aktivitas distraksi pengemudi pada kondisi pencahayaan yang beragam.
3. Mengimplementasikan hasil metode *EfficientNetB0* ke dalam *prototype early warning system* berbasis *computer vision* untuk klasifikasi aktivitas distraksi pengemudi secara *real-time*.

1.5. Manfaat Penelitian

Manfaat penelitian yang dapat diperoleh dari hasil penelitian ini adalah sebagai berikut:

1. Bagi bidang keilmuan

Penelitian ini berkontribusi pada pengembangan *computer vision*, khususnya memberikan pemahaman baru mengenai bagaimana model CNN ringan seperti *EfficientNetB0* bereaksi ketika dihadapkan dengan data citra yang kondisi pencahayaannya beragam, seperti pada kasus aktivitas distraksi pengemudi disini. Hal ini memperkaya literatur mengenai *illumination variation* dan *robustness* model *deep learning*. Kemudian penelitian ini juga memberikan *benchmark* baru untuk klasifikasi pengemudi terdistraksi versi Indonesia, karena hingga saat ini penelitian domain Indonesia masih sangat terbatas. Dataset primer yang digunakan dalam penelitian ini dapat menjadi acuan awal bagi peneliti berikutnya. Perbandingan langsung antara arsitektur *EfficientNetB0* yang menggunakan *channel attention* dan augmentasi data dengan yang tidak memberikan wawasan tentang *trade-off* antara akurasi, kompleksitas model, kebutuhan komputasi, serta *robustness* pada aplikasi *real-*

time, beserta dengan pengaruh *attention mechanisms* seperti *channel attention* terhadap kinerja model pada data dengan kondisi pencahayaan beragam.

2. Bagi masyarakat

Penelitian ini dapat memperkuat pendidikan keselamatan berkendara (*road safety awareness*), bahwa aktivitas kecil seperti melihat ponsel dapat meningkatkan risiko kecelakaan, dan hal ini meningkatkan kesadaran akan bahaya aktivitas distraksi pengemudi. Kemudian hasil penelitian ini dapat memberikan dasar awal bagi pengembangan sistem peringatan dini, model klasifikasi yang dihasilkan dapat digunakan sebagai dasar prototipe sistem peringatan seperti pada perangkat IoT mobil yang mendeteksi perilaku berbahaya. Walaupun penelitian ini belum menjadi sistem produksi, hasil penelitian ini dapat mengurangi potensi kecelakaan yang disebabkan oleh *human error*, hasilnya dapat membantu menciptakan teknologi yang berpotensi mencegah kecelakaan, menyelamatkan nyawa, dan mengurangi beban biaya kecelakaan bagi masyarakat.

3. Bagi Pemerintah

Pemerintah (Korlantas Polri, Kemenhub) memiliki program untuk menekan angka kecelakaan akibat kelalaian pengemudi. Penelitian ini memberikan data ilmiah yang dapat mendukung kebijakan tersebut, dan mendukung program keselamatan lalu lintas nasional. Temuan penelitian ini dapat memberikan referensi dasar bagi pemerintah dalam menyusun standar teknologi pendukung keselamatan berbasis kamera, baik untuk kendaraan pribadi maupun kendaraan angkutan umum, dan hal ini memperkuat dasar dalam merancang regulasi keselamatan *smart vehicle*. Dengan membuktikan bahwa model CNN ringan dapat mencapai akurasi tinggi, hal ini dapat mendorong adopsi teknologi keselamatan yang canggih namun tetap ekonomis tanpa membutuhkan perangkat mahal. Secara jangka panjang, teknologi pendeteksi pengemudi terdistraksi dapat membantu menurunkan angka kecelakaan, sehingga menurunkan beban biaya kesehatan, litigasi, dan pemulihan infrastruktur.