

BAB I

PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, manfaat dan tujuan, ruang lingkup, dan sistematika penulisan pada skripsi yang berjudul Pengenalan Pola Gerak Bibir dalam Pengucapan Suku Kata Bahasa Indonesia Menggunakan Arsitektur EfficientNetV2-BiLSTM

1.1. Latar Belakang

Pengguna aktif aplikasi *online meeting* meningkat drastis di tahun 2020 (Pratama dkk., 2020). Peningkatan drastis pengguna aktif aplikasi *online meeting* disebabkan oleh faktor utama yaitu pandemi Covid-19. Walaupun saat ini pandemi telah berlalu, Camilleri & Camilleri (2022) menyatakan bahwa penggunaan komunikasi secara daring akan menjadi hal yang normal setelah masa Covid-19.

Pengguna komunikasi daring seperti konferensi video terdiri dari berbagai kalangan, tetapi aksesibilitas untuk beberapa pengguna yang mempunyai kebutuhan khusus terutama penyandang tunarungu ataupun orang-orang yang mempunyai gangguan pendengaran seperti presbikusis masih memiliki berbagai tantangan. Anderson (2021) mengemukakan bahwa terdapat berbagai tantangan penggunaan konferensi video untuk orang yang mempunyai gangguan pendengaran seperti penggunaan bahasa isyarat yang sulit untuk ditafsirkan pada layar, tidak seluruh penyandang tunarungu mampu membaca gerak bibir, dan permasalahan komunikasi antara orang biasa dan penyandang tunarungu.

Jumlah penyandang tunarungu di Indonesia mencapai 250 ribu orang (Badan Pusat Statistik Indonesia, 2022). Jumlah penyandang tunarungu tersebut cukup signifikan dibandingkan dengan jumlah seluruh penduduk Indonesia. Indonesia memiliki jumlah penduduk sebesar 281 juta jiwa menurut data sensus penduduk 2022-2024 (Badan Pusat Statistik Indonesia, 2024). Bahasa Indonesia merupakan bahasa nasional yang digunakan sebagai alat komunikasi dalam kehidupan sehari-hari masyarakatnya Indonesia (Maghfiroh, 2022). Penyandang tunarungu di Indonesia juga memakai bahasa Indonesia pada kehidupan keseharian mereka.

Penyandang tunarungu yang dapat membaca gerak bibir berjumlah sedikit dikarenakan sulitnya penguasaan komunikasi dengan membaca gerak bibir. Altieri dkk. (2011) menyatakan rata-rata kemampuan dari 84 orang untuk membaca gerak bibir hanya mencapai akurasi sebesar 12,4%. Kesulitan manusia untuk membaca gerak bibir dapat ditangani dengan menggunakan mesin.

Hilder dkk. (2009) menunjukkan pembacaan gerak bibir oleh mesin memiliki akurasi lebih tinggi dari manusia dengan pembacaan gerak bibir untuk fonem mencapai 80,27% pada mesin dan 31,6% pada manusia. Penggunaan pembacaan gerak bibir menggunakan mesin akan memudahkan komunikasi antara orang biasa dan orang yang memiliki kekurangan dalam pendengarannya terutama pada media daring dengan mengenali apa yang dibicarakan seseorang hanya dengan gerakan bibirnya.

Penelitian mengenai pengenalan pola gerak bibir sebelumnya telah diteliti oleh sejumlah peneliti. Kahfi dkk. (2020) meneliti pola gerak bibir untuk fonem vokal Bahasa Indonesia dan menghasilkan akurasi sebesar 84% menggunakan 3D CNN. Fonem vokal Bahasa Indonesia yang hanya terdiri dari a, i, u, e, o belum cukup untuk merepresentasikan Bahasa Indonesia dan masih dapat dikembangkan lagi.

Pradiptatmaka dkk. (2021) meneliti pola gerak bibir untuk kata Bahasa Inggris menggunakan CNN-RNN dan menghasilkan Akurasi 82,8%. Akan tetapi RNN memiliki kekurangan yaitu *vanishing gradient*. Permasalahan *vanishing gradient* tersebut dapat diatasi dengan penggunaan LSTM (Althelaya dkk. 2018). Selain itu terdapat arsitektur berbasis CNN yang telah dikembangkan dan dapat menghasilkan fitur citra lebih cepat dan efisien seperti EfficientNet dan EfficientNetV2.

Khadijah dkk. (2021) meneliti klasifikasi muka bermasker dan menghasilkan akurasi sebesar 99% menggunakan EfficientNet serta lebih efisien dibandingkan dengan model lainnya yang berbasis CNN seperti ResNet-50. Hal tersebut dikarenakan kecilnya parameter *network* pada EfficientNet. Namun EfficientNet mempunyai kesulitan saat awal mulai *training* yang disebabkan oleh arsitekturnya. Oleh karena itu telah dikembangkan EfficientNetV2 untuk mengatasi masalah tersebut.

Kurniawan & Suyanto (2020) meneliti pola gerak bibir untuk suku kata Bahasa Indonesia dan menghasilkan akurasi sebesar 100% menggunakan StCNN-BiGRU. Penggunaan *Bidirectional* pada GRU dapat menangkap konteks dasar lebih baik dengan menggunakan *forward direction input* dan *backward direction input* (Siami-Namini dkk., 2019).

Penelitian pada skripsi ini akan membahas pengenalan pola gerak bibir pada suku kata Bahasa Indonesia menggunakan arsitektur EfficientNetV2-BiLSTM. Penelitian ini diharapkan membantu perkembangan pengenalan pola gerak bibir pada Bahasa Indonesia serta dapat digunakan untuk pengenalan ucapan pada video.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, dapat dirumuskan sebuah permasalahan yaitu bagaimana menghasilkan model dengan arsitektur EfficientNetV2-BiLSTM yang dapat mengenali gerak bibir pada pengucapan suku kata Bahasa Indonesia dan dapat mengklasifikasikannya sesuai dengan suku kata yang diucapkan.

1.3. Tujuan dan Manfaat

Tujuan dari penelitian pada skripsi ini adalah menghasilkan model dengan arsitektur EfficientNetV2-BiLSTM yang dapat mengenali pola gerak bibir pada pengucapan suku kata Bahasa Indonesia dan mengevaluasi kinerja model dalam mengklasifikasikan suku kata.

Manfaat dari penelitian pada skripsi ini adalah model yang dihasilkan dapat dijadikan sebagai acuan untuk pembuatan otomatisasi pembacaan gerak bibir yang nantinya dapat diterapkan pada *online meeting* ataupun bahan ajaran pembacaan gerak bibir.

1.4. Ruang Lingkup

Ruang lingkup penelitian skripsi ini dibatasi oleh aspek-aspek sebagai berikut:

1. Pengumpulan *dataset* menyesuaikan metodologi yang digunakan oleh Kurniawan dan Suyanto (2020). *Dataset* berupa video dengan *frame rate* sejumlah 30 fps yang terdiri dari pengucapan 5 kalimat bahasa Indonesia dengan total 29 kelas yang terdiri dari 28 suku kata dan 1 representasi diam dan diucapkan oleh 5 pria dan 5 wanita berumur 20-60 tahun.
2. Implementasi dilakukan menggunakan bahasa pemrograman *python*.

1.5. Sistematika Penulisan

Sistematika penulisan yang digunakan dalam skripsi ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I	Bab ini menyajikan latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan dalam pengenalan pola gerak bibir pada pengucapan suku kata Bahasa Indonesia.
BAB II	LANDASAN TEORI Bab ini menyajikan tentang tinjauan pustaka (penelitian terkait) dan dasar teori yang digunakan untuk pengenalan pola gerak bibir pada pengucapan suku kata Bahasa Indonesia.
BAB III	METODOLOGI PENELITIAN Bab ini menyajikan tentang metodologi penelitian yang dilakukan pada pengenalan pola gerak bibir di pengucapan suku kata Bahasa Indonesia. Metodologi tersebut berupa pengumpulan data, <i>preprocessing</i> data, ekstraksi fitur, pelatihan model dan evaluasi.
BAB IV	HASIL DAN PEMBAHASAN Bab ini menyajikan tentang hasil penelitian dari penerapan metode arsitektur EfficientNetV2-BiLSTM terhadap pengenalan pola gerak bibir pada pengucapan suku kata Bahasa Indonesia.
BAB V	PENUTUP Penutup berisi kesimpulan dan saran dari penelitian pengenalan pola gerak bibir pada pengucapan suku kata Bahasa Indonesia dan saran-saran untuk pengembangan penelitian selanjutnya.