



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan di bidang *speech recognition* telah memunculkan beragam teknologi untuk mempermudah beberapa aspek dalam kehidupan manusia sehari-hari. Dengan kapabilitas untuk mengenal suara manusia, komputer dapat memberi bantuan kepada penyandang disabilitas hingga meningkatkan kualitas hidup masyarakat umum. Aplikasi transkripsi otomatis dan *captioning* menyediakan aksesibilitas media audio dan audio visual untuk individu dengan kesulitan pendengaran. Selain transkripsi dan *captioning*, *speech recognition* juga digunakan untuk mendeteksi perintah suara dalam perangkat-perangkat seperti smartphone dan home assistant.

Untuk pengidap gangguan bicara, sistem *speech recognition* memiliki potensi sebagai alat pembantu komunikasi. Gangguan bicara, contohnya *dysarthria*, pada umumnya disertai gangguan motorik atau neurologis seperti *stroke* [1] sehingga metode komunikasi alternatif sulit digunakan. Walaupun *speech recognition* sudah umum digunakan dalam berbagai aplikasi, sistem *speech recognition* masih belum memiliki kinerja yang optimal dalam melakukan pengenalan dalam konteks *disordered speech*. Karena sebagian besar dari model *speech recognition* dilatih dengan data pembicara tanpa disabilitas, sistem *speech recognition* berkesulitan untuk melakukan transkripsi secara akurat. Adanya perbedaan artikulasi antara pengidap gangguan bicara dan pembicara tanpa disabilitas mengakibatkan model cenderung melakukan kesalahan dalam mengidentifikasi ucapan yang bersifat *disordered speech* [2].

Untuk membangun sistem *speech recognition* yang dapat digunakan oleh pengidap gangguan bicara, model *speech recognition* harus dilatih dengan data *disordered speech*. Sejak terbitnya *The Whitaker database* [3], terdapat beberapa

upaya untuk menanggulangi keterbatasan data *disordered speech* dengan cara pengumpulan data. Beberapa *database* ternama yang mengandung data *disordered speech* seperti TORGO [4] dan *Project Euphonia* [5] memiliki tujuan untuk mempelajari fitur-fitur yang membedakan pembicara dengan gangguan bicara dengan pembicara normal sehingga sistem *speech recognition* untuk pengidap gangguan bicara dapat dikembangkan.

Walaupun ketersediaan data *disordered speech* mulai bertambah, jumlah data yang ada masih dianggap terbatas [6] dan belum cukup untuk digunakan dalam perancangan sistem *speech recognition*. Secara kumulatif, jumlah data *disordered speech* yang terkumpul tidak sebanding dengan jumlah data yang biasa digunakan dalam pelatihan model *speech recognition* [7]. Pada waktu penelitian dilakukan, mayoritas dari data *disordered speech* hanya tersedia dalam Bahasa Inggris dan tidak tersedia dalam Bahasa Indonesia sehingga tidak dapat digunakan untuk melatih model *speech recognition* untuk gangguan bicara dalam Bahasa Indonesia. Pengumpulan data *disordered speech* merupakan proses yang sulit dilakukan karena prosedur pengumpulan data cenderung menantang secara fisik pengidap gangguan bicara sehingga disarankan adanya pendekatan alternatif untuk mendapatkan data *disordered speech* yang dibutuhkan untuk membangun sistem *speech recognition* yang dapat digunakan pengidap gangguan bicara. Oleh karena itu, dibutuhkan pengetahuan mengenai data *disordered speech* yang minimum digunakan pada tahap pelatihan agar model *speech recognition* dapat mengenali ucapan pengidap gangguan bicara dengan akurat.

Model *speech recognition* pada umumnya dibuat dengan pendekatan statistik, yaitu menggunakan model berbasis *hidden markov model* (HMM). Kesederhanaan model membuat HMM mudah untuk diimplementasi pada sistem *speech recognition*, namun model berbasis *neural network*, terutama *convolutional neural network* (CNN) mulai digunakan sebagai pengganti HMM. Model berbasis *neural network* memiliki fleksibilitas dan kemampuan untuk generalisasi data

sehingga mampu untuk mengenal pola bahasa yang baru. Selibhnya, dengan menggunakan proses konvolusi, model CNN dapat menangani variasi fitur *speech* yang dapat terjadi akibat perubahan frekuensi maupun kecepatan pada sinyal. Dengan demikian, model *speech recognition* berbasis CNN semakin banyak digunakan.

1.2. Rumusan Masalah

Masalah yang dapat dirumuskan berdasarkan latar belakang penelitian adalah sebagai berikut:

- 1.2.1. Apakah model *speech recognition* untuk gangguan bicara dapat dilatih menggunakan *training data* yang terdiri dari campuran data *normal speech* dan *disordered speech*?
- 1.2.2. Berapa rasio data *disordered speech* yang harus digunakan dalam *training* pada model *speech recognition* untuk gangguan bicara?

1.3. Tujuan Penelitian

Tujuan yang dapat dicapai dari penelitian adalah sebagai berikut:

- 1.3.1. Mengevaluasi kinerja model *speech recognition* yang dilatih dengan data campuran
- 1.3.2. Mendapatkan rasio data terbaik dalam membangun model *speech recognition* untuk gangguan bicara dengan jumlah data yang terbatas

1.4. Batasan Penelitian

Penelitian yang dilakukan memiliki batasan-batasan sebagai berikut:

- 1.4.1. Metode pemrosesan data dengan *speed perturbation* tidak dapat merepresentasikan setiap karakteristik gangguan bicara secara spesifik
- 1.4.2. *Dataset* yang digunakan untuk melatih model memiliki jumlah data yang terbatas dalam konteks sistem *automatic speech recognition*

1.5. Manfaat Penelitian

Hasil dari penelitian yang dilakukan diharapkan dapat memberikan manfaat sebagai berikut:

1.5.1. Dapat digunakan sebagai dasar dalam perancangan sistem *speech recognition* untuk pengidap gangguan bicara

1.5.2. Dapat digunakan sebagai referensi penelitian lebih lanjut dalam bidang *speech recognition*



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA