



## Implementasi Speech Recognition Pada Aplikasi E-Prescribing Menggunakan Algoritme Convolutional Neural Network

Nur Azis<sup>1,\*</sup>, Herwanto<sup>2</sup>, Fathurrahman Ramadhani<sup>2</sup>

<sup>1</sup> Teknik, Sistem Informasi, Universitas Krisnadwipayana, Jakarta, Indonesia

<sup>2</sup> Teknik, Teknik Informatika, Universitas Krisnadwipayana, Jakarta, Indonesia

Email: <sup>1,\*</sup> [nuraziz@unkris.ac.id](mailto:nuraziz@unkris.ac.id), <sup>2</sup> [Herwanto@unkris.ac.id](mailto:Herwanto@unkris.ac.id), <sup>3</sup> [fathurrahmanramadhani@gmail.com](mailto:fathurrahmanramadhani@gmail.com)

Email Penulis Korespondensi: [nuraziz@unkris.ac.id](mailto:nuraziz@unkris.ac.id)

**Abstrak**—Proses peresepan obat secara manual oleh dokter dapat menimbulkan beberapa permasalahan, diantaranya dokter tidak mengetahui obat apa saja yang tersedia dan memerlukan waktu untuk mengetahui apa saja obat yang tersedia di farmasi. Pengenalan ucapan (speech recognition) pada saat ini sudah banyak digunakan dalam berbagai hal, yang dapat membantu dalam memudahkan pekerjaan. Penerapan pengenalan ucapan dapat dilakukan pada aplikasi e-prescribing dengan metode jaringan syaraf tiruan menggunakan algoritme Convolutional Neural Network (CNN), merupakan metode dasar pada Deep Learning. Penelitian ini bertujuan untuk memudahkan dokter dalam mengisi data obat di aplikasi e-prescribing dengan menggunakan pengenalan ucapan. Data yang digunakan dalam penelitian ini didapatkan dari dataset opensource yang disediakan oleh Google dan melakukan pengumpulan dataset mandiri. Dari hasil percobaan yang telah dilakukan akurasi yang dicapai dengan 40 epoch dan 40 pengentesan langsung dengan kata yang berbeda sebesar 90%. Dimana kata yang berhasil dikenali dengan 36 kata dari 40 kata.

**Kata Kunci:** Peresepan Obat; Aplikasi e-prescribing; Speech Recognition; Convolutional Neural Network; Dataset

**Abstract**—The process of manually prescribing drugs by doctors can cause several problems, including doctors not knowing what drugs are available and it takes time to find out what drugs are available in the pharmacy. Speech recognition is now widely used in various ways, which can help facilitate work. The application of speech recognition can be done in the e-prescribing application with the neural network method using the Convolutional Neural Network (CNN) algorithm, which is the basic method of deep learning. This study aims to facilitate physicians in filling out drug data in e-prescribing applications using speech recognition. The data used in this study were obtained from the open source dataset provided by Google and collected independent datasets. From the results of experiments that have been carried out, the accuracy achieved with 40 epochs and 40 direct impressions with different words is 90%. Where words are successfully recognized 36 words out of 40 words.

**Keywords:** Prescription Drugs; Electronic Prescription Application; Speech Recognition; Convolutional Neural Network; Dataset

### 1. PENDAHULUAN

Era teknologi yang berkembang sangat cepat ini, banyak teknologi baru yang bermunculan untuk memudahkan dan membantu dalam melakukan suatu pekerjaan. Salah satu kemajuan teknologi informasi merambah pada bidang kesehatan seperti kedokteran[1], Pada era digital, masyarakat pun semakin menyadari bahwa teknologi komunikasi merupakan salah satu alat yang penting dalam mengatasi cepatnya penyebaran arus informasi[2], begitu pula komunikasi dengan seorang dokter juga tak kalah penting dalam memantau kondisi Kesehatan Kita. Hal itu dilakukan demi meminimalkan *medication error*. Adapun pengertian medication error sendiri adalah kejadian yang merugikan pasien yang diakibatkan pemakaian obat selama dalam penanganan tenaga kesehatan[3] maka dalam hal ini sangat dibutuhkan sebuah teknologi yang membantu didalam dunia Kesehatan. Salah satu teknologi yang dibutuhkan dalam dunia Kesehatan adalah *speech recognition* secara otomatis. Hal ini diperlukan untuk membantu kinerja dalam proses *Prescription Drugs*. Mengacu pada penelitian yang pernah dilakukan sebelumnya didalam pengembangan teknologi yang berkaitan dalam bidang Kesehatan. oleh Riza Adrianti Supono [4] dalam penelitiannya membahas tentang peluang dan hambatan penerapan teknologi internet untuk pengobatan jarak jauh, kemudian penelitaian yang dilakukan oleh Karyo Budi Utomo dan Kawan-kawan[5] dalam penelitiannya membahas tentang bagaimana mendeteksi penyakit paru-paru dengan metode *Forward Chaining*. Berdasarkan beberapa penelitian yang sudah di lakukan oleh beberapa peneliti diatas maka penelitian ini akan mencoba meneliti penerapan teknologi yang berbeda.

Adapun focus penelitian dalam paper ini adalah meneliti tentang bagaimana agar peresepan obat dapat dilakukan secara otomatis dengan berbekal ucapan dari seorang dokter pemeriksa. Hal itu dilatarbelakangi oleh sebuah resep yang ditulis oleh seorang dokter terkadang tidak terbaca oleh perawat. Sehingga terkadang memerlukan kerja atau proses yang lebih lama. Adapun penelitian yang ada saat ini seperti dilakukan oleh Adhitya Yoga Pratama Idwal dan kawan-kawan[6] yang meneliti tentang pengenalan suara dalam Bahasa Indonesia untuk aksan Melayu Pontianak dan Sunda Garut, kemudian penelitian oleh Jenny Putri Hapsari[7] bagaimana sebuah suara menjadi kunci dalam pengaksesan Sistem Informasi Akademik, selanjutnya oleh Muhammad Bobbi Kurniawan Nasution[8] yang meneliti tentang pengenalan suara teks bebas dengan menggunakan metode Algoritma SVM dan yang terakhir oleh Handri Al Fani dkk[9] yang merancang sebuah alat monitoring pendekteksi suara diruangan bayi Berbasis Arduino.



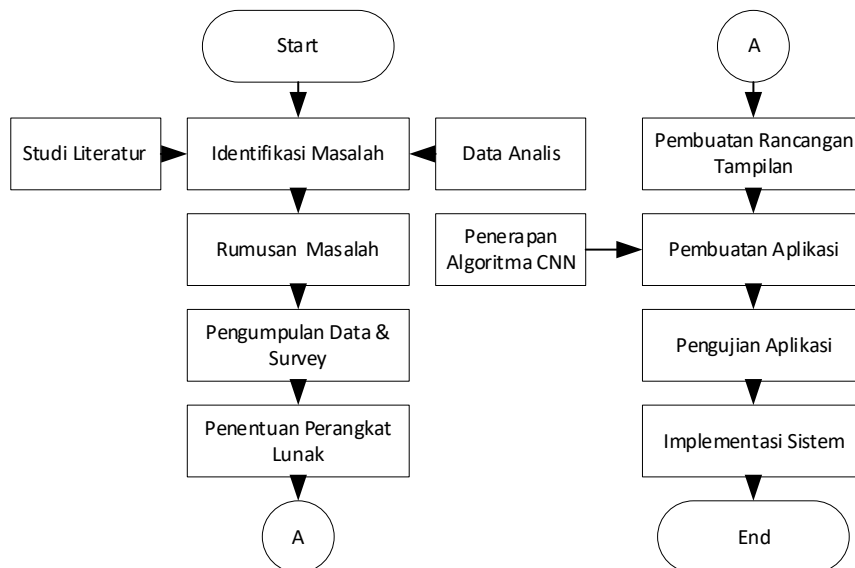
Didalam penelitian ini Sistem pereseapan elektronik (*e-prescribing*)[10] didesain untuk mempermudah dalam melakukan pelayanan dimulai dari tahap penulisan resep, hal tersebut dilakukan untuk agar tidak terjadi kesalahan pada saat proses pereseapan obat atau penulisan resep (*prescribing*)[11], tahap kedua adalah pembacaan resep (*transcribing*)[12] pada tahap ini diperlukan sebuah metode yang berfungsi untuk mengenali suara Adapun metode yang digunakan adalah *Convolutional Neural Networks* (CNN), seperti yang di sampaikan oleh I Wayan Suartika E. P dkk [13] sebuah CNN dapat melakukan pengenalan citra dengan akurasi yang menyaingi manusia pada dataset tertentu, tahap ketiga menyiapkan, membuat, atau meracik obat, dan mendistribusikan obat kepada penderita[14] pada proses ini disebut dengan (*Dispensing*), tahap keempat proses penggunaan obat pada tahap ini Penggunaan obat harus sesuai dengan penyakit, oleh karena itu diagnosis yang ditegakkan harus tepat, patofisiologi penyakit, keterkaitan farmakologi obat dengan patofisiologi penyakit dan dosis yang diberikan dan waktu pemberian yang tepat[15] dan yang terakhir adalah mengenali dan mengevaluasi perkembangan yang terjadi akibat tindakan atau disebut dengan monitoring[16].

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Tahapan penelitian pembuatan Aplikasi E-Prescribing merupakan tahapan-tahapan yang akan dilakukan dalam menyelesaikan penelitian. Dengan dilakukannya tahapan-tahapan dengan terstruktur dan sistematis agar penelitian ini dapat berjalan dengan baik.

Berikut merupakan tahapan-tahapan dalam membangun Aplikasi E-Prescribing pada penelitian ini, yang dapat digambarkan melalui diagram alur penelitian seperti pada gambar 1.



**Gambar 1.** Diagram Alur Penelitian

Penjelasan gambar 1 Diagram Alur Penelitian :

a. Identifikasi Masalah

Pada tahapan ini, dilakukan pengidentifikasian masalah yang ada, sehingga dapat ditemukannya solusi alternatif untuk memecahkan permasalahan. Pada pengidentifikasian ini dilakukan dengan datang langsung ke rumah sakit, dan melakukan perumusan masalah, selanjutnya menentukan tujuan penelitian yang mana tujuan penelitian ini akan digunakan sebagai parameter untuk memecahkan permasalahan yang ada..

b. Rumusan Masalah

Pada tahapan ini, dilakukan perumusan masalah berdasarkan masalah yang sudah di identifikasi. Yaitu merupakan penjabaran dari identifikasi masalah seperti pertanyaan lengkap dan terfperinci tentang ruang lingkup masalah yang akan diteliti berdasarkan identifikasi masalah tersebut. Dan akan menjadi penentu apa bahasan yang akan dilakukan dalam sebuah penelitian.

c. Pengumpulan Data & Survey

Pada tahapan ini, dilakukan pengumpulan data yang akan digunakan sebagai bahan dan patokan dalam pembuatan aplikasi. Data disini juga termasuk pengumpulan data mentah atau data set yang digunakan untuk proses pengenalan suara, dengan melakukan perekaman pengucapan yang kemudian disimpan dalam suatu database.

d. Penentuan Perangkat Lunak

Pada tahap ini peneliti melakukan analisa dan mencari aplikasi yang dapat dijadikan sebagai bahan informasi untuk penelitian.



e. Pembuatan Rancangan Tampilan

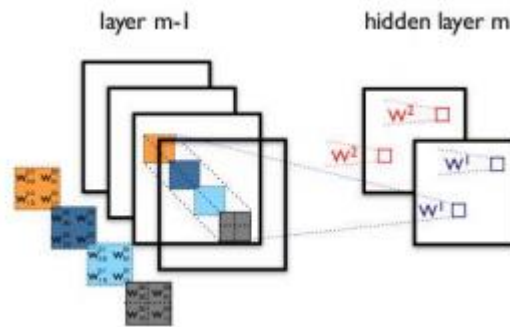
Pada tahapan ini, membuat suatu permodelan atau blueprint aplikasi yang akan dibuat, sehingga dalam proses pembuatannya akan menjadi efektif dan efisien. Dan juga pada hasil akhir pembuatan aplikasi diharapkan sesuai dengan apa yang telah direncanakan,

f. Pembuatan Aplikasi

Pada tahap ini peneliti membuat aplikasi yang berfungsi untuk informasi dan penerapan metode algoritma *Convolutional Neural Networks* (CNN) agar menghasilkan perekaman suara dengan baik. Adapun proses kerja pada CNN, data yang dipropagasikan pada jaringan adalah data dua dimensi, sehingga operasi linear dan parameter bobot pada CNN berbeda. Pada CNN operasi linear menggunakan operasi konvolusi, sedangkan bobot tidak lagi satu dimensi saja, namun berbentuk empat dimensi yang merupakan kumpulan kernel konvolusi seperti pada Gambar.2. Dimensi bobot pada CNN adalah[13]:

$$neuron\ input \times neuron\ output \times tinggi \times lebar$$

Karena sifat proses konvolusi, maka CNN hanya dapat digunakan pada data yang memiliki struktur dua dimensi seperti citra dan suara.



**Gambar 2.** Proses Konvolusi pada CNN

g. Pengujian Aplikasi

Pada Tahapan ini, merupakan tahapan terakhir dalam penelitian ini, dengan mencari tau apakah penelitian ini berhasil atau belum berhasil.

h. Implementasi Sistem

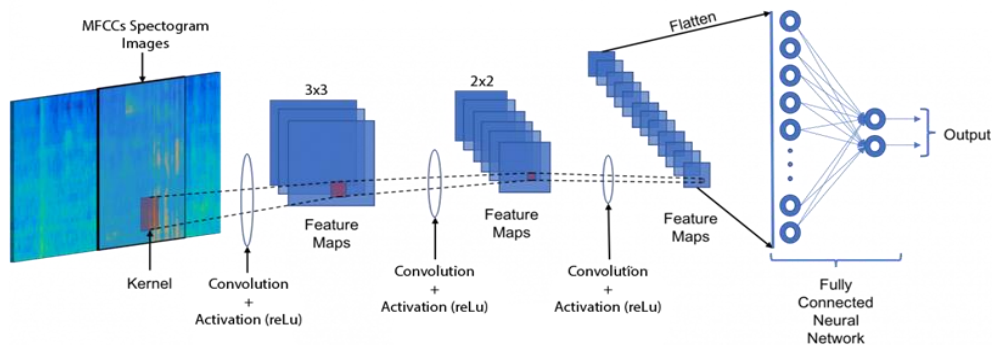
Pada tahap ini aplikasi diterapkan pada rumah sakit.

### 3. HASIL DAN PEMBAHASAN

Implementasi Aplikasi e-prescribing adalah tahap penerapan sekaligus pengujian bagi sistem baru serta merupakan tahap dimana aplikasi siap dioperasikan pada keadaan yang sebenarnya, efektifitas sistem baru akan diketahui secara pasti, juga untuk semua kelebihan dan kekurangan sistem dan aplikasi program.

**a. Convolutional Neural Network**

Untuk melakukan training data menggunakan CNN Langkah pertama yaitu dengan data hasil ekstraksi menggunakan MFCCs, dimana dengan mengubah suara menjadi visual spektrogram per-frekuensi suara. Sehingga dihasilkan data visual yang dapat diklasifikasi oleh CNN. Dalam melakukan klasifikasi diperlukan 4 tahapan yang harus dilakukan, yaitu convolution, pooling, flattening dan full connection



**Gambar 3.** Arsitektur Conlutional Neural Network

i. Convolution

*Convolution Layer* terdiri dari neuron yang tersusun sedemikian rupa sehingga membentuk sebuah filter dengan panjang dan tinggi dalam pixel. Dengan contoh layer berukuran 6x6x3 yang artinya layer memiliki panjang 6 pixel, tinggi 6 pixel, dan tebal atau jumlah 3 channel dari model gambar yaitu RGB. Ketiga filter akan digeser, setiap pergeseran akan dilakukan operasi antara input dan nilai dari filter tersebut, sehingga

menghasilkan activation map atau feature map. Ada beberapa parameter yang harus ditentukan dalam melakukan *convolution* yaitu *stride*, *padding*, dan *aktivasi*.

*Stride* merupakan parameter yang menentukan berapa jumlah pergeseran filter, jika nilai *stride* adalah 1 maka *convolution* akan bergeser sebanyak 1 pixel secara horizontal atau vertikal. *Padding* atau zero padding merupakan parameter menentukan jumlah pixel yang berisi 0 yang akan ditambahkan disetiap sisi input dengan tujuan untuk memanipulasi dimensi output dari feature map. Aktivasi merupakan nilai hasil *convolution* dikenakan fungsi aktivasi, fungsi yang biasa digunakan yaitu reLu dikarenakan sifatnya yang lebih berfungsi dengan baik.

ii. *Pooling*

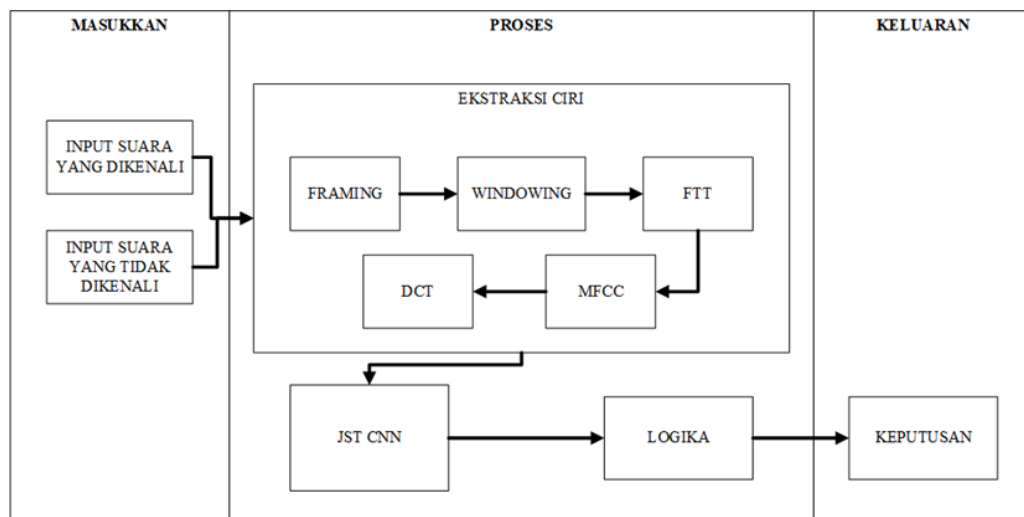
*Pooling Layer* terdiri dari sebuah filter dengan ukuran dan *stride* tertentu yang bergeser pada seluruh area feature map. *Pooling* yang digunakan yaitu Max *Pooling* dan Average *Pooling*, dengan tujuan penggunaan *pooling layer* untuk mengurangi dimensi dari feature map yang dihasilkan dari *convolution*, sehingga mempercepat komputasi karena parameter semakin sedikit dan mengatasi *overfitting*. Bentuk lapisan *pooling* pertama dan kedua yang digunakan filter berukuran 3x3 dan lapisan *pooling* ketiga yang digunakan filter berukuran 2x2 yang diaplikasikan dengan langkah sebanyak 3 kali dan kemudian beroperasi pada setiap irisan input.

iii. *Flattening*

*Flattening* merupakan proses mengubah data *pooling* yang memiliki array 2 dimensi kemudian dikonversi menjadi data satu dimensi single vector. *Flattening* dapat juga diartikan menreshape feature map menjadi sebuah vector agar bisa digunakan sebagai input dari *fully-connected layer*.

iv. *Full Connection*

*Full Connection* merupakan layer dimana semua neuron aktivitas dari lapisan sebelumnya terhubung semua dengan neuron di lapisan selanjutnya seperti halnya jaringan syaraf tiruan bisa. Setiap aktivitas dari lapisan sebelumnya perlu diubah menjadi data satu dimensi sebelum dapat dihubungkan ke semua neuron di lapisan *Fully-Connected*. Hal ini dilakukan bertujuan untuk mengolah data sehingga bisa diklasifikasikan karena *Fully-Connected* memiliki neuron yang secara keseluruhan terhubung



**Gambar 4.** Diagram Proses Pengenalan Ucapan

Keterangan:

**Masukan**

Pada tahap ini pengguna sistem memasukkan data rekaman suara yang tidak dikenali pemiliknya dan data rekaman suara yang sudah diketahui pemiliknya sebagai data pembandingan.

**Proses**

Pada proses ini, proses masukan mengolah data suara dengan ekstraksi ciri yang terdiri dari framing, windowing, FFT, Mel-frequency, DCT yang kemudian menghasilkan koefisien MFCC. Masing-masing suara masukan yang sudah mendapat koefisien MFCC dari setiap frame, pengenalan pola menggunakan Jaringan Saraf Tiruan (JST) menggunakan masukan dari koefisien MFCC. Dimana koefisien MFCC dari suara sebagai pembandingan data training dan suara yang tidak dikenal sebagai data uji. Kemudian hasil dari Jaringan Saraf Tiruan dianalisis menggunakan logika keputusan untuk mendapatkan keputusan apakah kedua suara tersebut sama atau tidak.

**Keluaran**

Keluaran dari sistem tersebut yaitu keputusan bahwa suara yang diujikan sama atau tidak dengan suara pembandingan dalam berbentuk text.



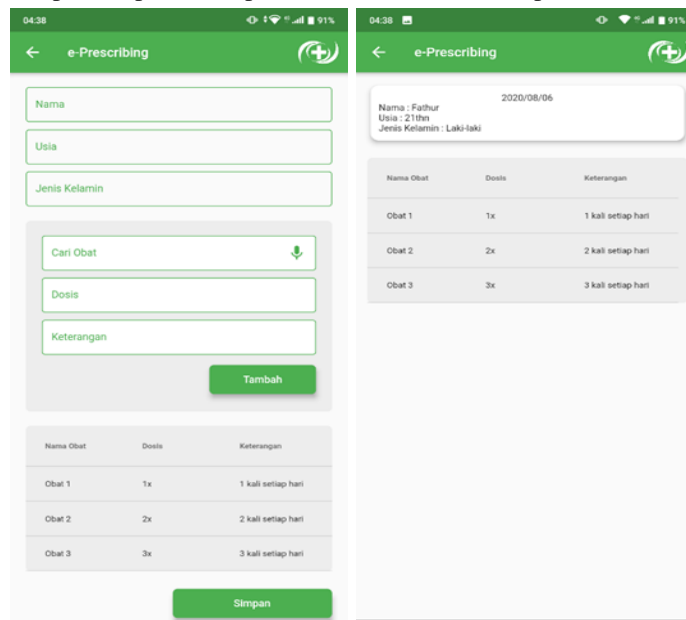
### 3.1 Implementasi Program

Ketika pertama menjalankan aplikasi, maka halaman yang ditampilkan adalah halaman utama. Gambar dari halaman utama dapat dilihat pada gambar di bawah ini.



**Gambar 5.** Tampilan Utama Aplikasi

Dibawah ini merupakan tampilan Input Pereseapan dan Detail Data Peresapan



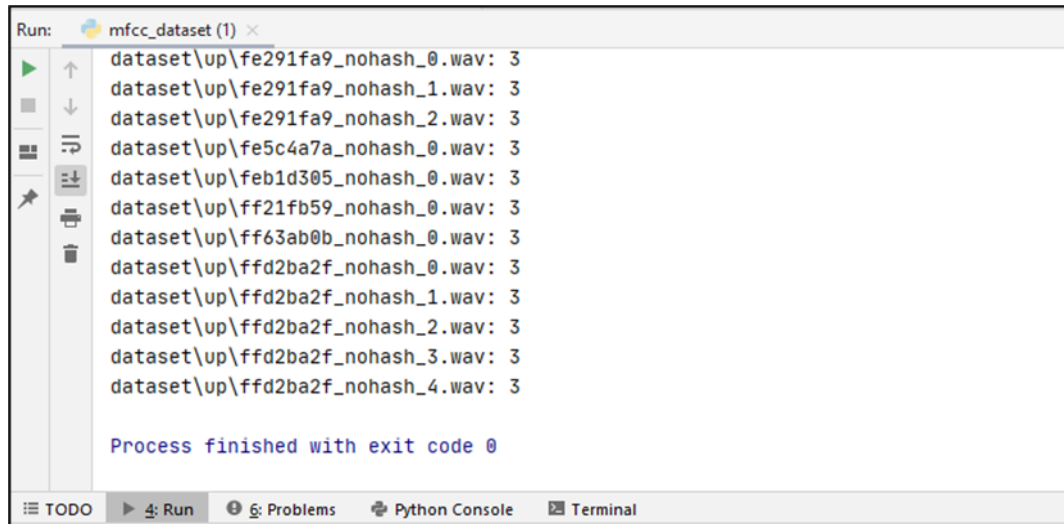
**Gambar 6.** Tampilan Input Pereseapan serta detail hasilnya

### 3.2. Pembahasan

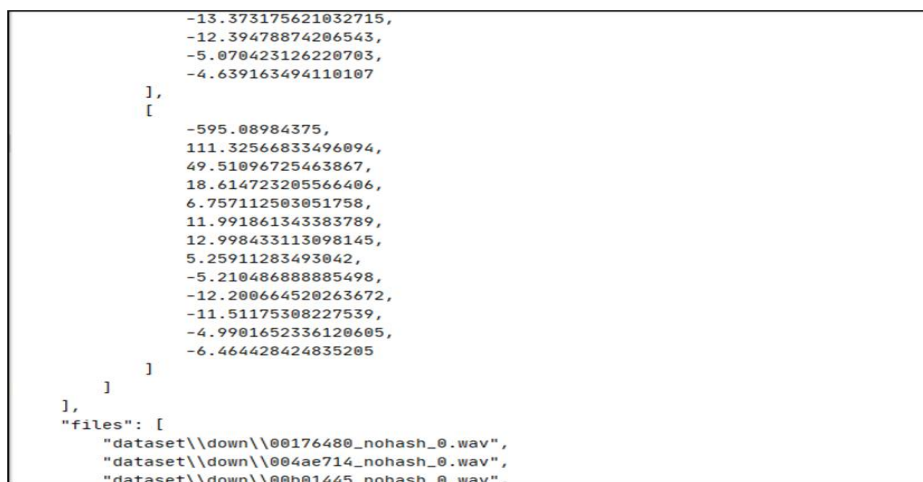
Dalam proses pengenalan suara yang diubah menjadi text, ada beberapa tahapan yang harus dilakukan untuk tercapainya hal tersebut. Pertama mempersiapkan dataset yang berupa file suara berformat .wav yang kemudian dilakukan proses ekstraksi ciri dengan menggunakan FFT dan MFCCs, yang akan mengeluarkan output file .json sebagai data pembandingan yang sudah di ekstraksi. Kedua dilakuan training dataset menggunakan Metode Algoritme CNN dengan melewati 40 epochs dan mendapatkan learning rate dengan accuracy 93%. Terakhir yaitu dengan keyword spotting untuk membandingkan hasil yang ada didataset dan dengan hasil ekstraksi data input yang akan dikenali.

#### a. Implementasi Ekstraksi Ciri Mel-Frequency Cepstral Coefficients

Pada tahapan ini dilakukan Ekstraksi Ciri dengan menggunakan metode Mel-Frequency Cepstral Coefficients dari file suara yang berformat .wav akan di ekstraksi dan di index ke dalam json yang merupakan hasil output dari prosesn ini. File json ini akan digunakan sebagai data yang akan di training.



Gambar 7. Tampilan Hasil Running Proses Ekstraksi



Gambar 8. Tampilan Hasil Data Ekstraksi

**b. Implementasi Algoritme Convolutional Neural Network**

Pada tahapan ini dilakukan Training Dataset yang sudah diekstraksi sebelumnya menggunakan MFCCs. Selanjutnya dilakukan proses training dataset menggunakan Convolutional Neural Network dengan 40 epoch untuk menghasilkan file berformat .h5. Dari penulisan ini dapat diketahui tingkat akurasi dalam proses pengenalan.

Langkah pertama dalam melakukan data training dataset menggunakan CNN dengan membuat object sequential class, Pada kode diatas, parameter pertama (64) adalah nilai filter nya. parameter kedua adalah dimensi filter nya (3x3), parameter ketiga adalah fungsi aktivasi yang kita gunakan (relu) disini kita menggunakan rectifier function.

Langkah selanjutnya adalah pooling. Pooling bertujuan untuk mereduksi ukuran gambar sebanyak mungkin. Disini juga kita mencoba mereduksi jumlah total node pada layer selanjutnya. Kemudian dengan menambahkan objek classifier kita ke dalam layer pooling. Pooling pertama dan kedua dalam bentuk matrix 3x3 dan pada pooling ketiga dalam bentuk matrix 2x2 sebagai pixel loss minimum dan precise region dimana feature dialokasikan. Selanjutnya adalah proses flattening. Data pooling yang kita miliki berupa array 2 dimensi kemudian dikonversi menjadi data satu dimensi single vector.

Berikutnya adalah proses dense. Dense adalah fungsi untuk menambahkan layer yang fully connected. units menandakan jumlah node yang harus ada di hidden layer, nilainya antara jumlah input node dan output node. Sedangkan fungsi aktivasinya kita gunakan relu. Saatnya kita inialisasi layer output kita, yang hanya berupa satu node karena klasifikasi ini termasuk dalam binary classification. Kode diatas memperlihatkan bahwa kita menggunakan node dengan jumlah dari kata yang diklasifikasikan, dan fungsi aktivasi softmax pada layer terakhir.





- [9] H. Al Fani, S. Sumarno, J. Jalaluddin, D. Hartama, and I. Gunawan, "Perancangan Alat Monitoring Pendeteksi Suara di Ruang Bayi RS Vita Insani Berbasis Arduino Menggunakan Buzzer," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 144, 2020, doi: 10.30865/mib.v4i1.1750.
- [10] F. C. Sabila *et al.*, "Pereseapan Elektronik ( E- Prescribing ) Dalam Menurunkan Kesalahan Penulisan Resep Electronical Prescribing ( E- Prescribing ) to Reduce Prescribing Error," *Med. J. Lampung Univ.*, vol. 7, no. 3, pp. 271–275, 2018, [Online]. Available: <http://juke.kedokteran.unila.ac.id/index.php/majority/article/view/2090>.
- [11] A. Yusuf, V. Fitria, D. Nugraha, and N. Mentari, "Kajian Administrasi dan Farmasetik Resep Pasien Rawat Jalan di RSUD dr. Soekardjo Kota Tasikmalaya Periode 10 Maret – 10 April 2017 Berdasarkan Permenkes Nomor 58 Tahun 2014," *J. Pharm. Sci. Med. Res.*, vol. 2, no. 1, pp. 24–41, 2019, [Online]. Available: <http://e-journal.unipma.ac.id/index.php/pharmed>.
- [12] Amalia, "Rational Drug Prescription Writing," *Juke*, vol. 4, no. 7, pp. 22–30, 2016.
- [13] W. S. Eka Putra, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.
- [14] I. M. A. Widnyana, "KAJIAN TERHADAP DENPASAR SEBAGAI TEMPAT PELAKSANAAN DISPENSING (PENYERAHAN) OBAT OLEH KALANGAN TENAGA MEDIS BERDASARKAN PELEGALAN UNDANG-UNDANG PRAKTEK KEDOKTERAN I," *J. Pendidik. Kesehat. Rekreasi*, vol. 2, pp. 60–68, 2015, doi: 10.1377/hlthaff.2013.0625.
- [15] A. Pratiwi and R. Sinuraya, "Prescribing Analysis for 2–5 Years Old Children in Bandung During Year 2012," *Indones. J. Clin. Pharm.*, vol. 3, no. 1, pp. 18–23, 2014, doi: 10.15416/ijcp.2014.3.1.18.
- [16] A. Herliana and P. M. Rasyid, "Sistem Informasi Monitoring Pengembangan Software Pada Tahap," *J. Inform.*, no. 1, pp. 41–50, 2016.