

Penerapan Deep Learning Pada Aplikasi Prediksi Penyakit Pneumonia Berbasis Convolutional Neural Networks

Agung Perdananto ^{#1}, Achmad Udin Zailani ^{*2}

Teknik Informatika, Universitas Pamulang

Jl.Surya Kencana No.1, Pamulang. Tangerang Selatan, Indonesia

¹agung.perdananto@gmail.com

²dosen00270@gmail.com

Received on 24-11-2019, revised on dd-mm-yyyy, accepted on 28-11-2019

ABSTRACT

Pneumonia merupakan salah satu penyebab kematian anak-anak tertinggi di dunia. Badan Kesehatan Dunia (WHO) memperkirakan bahwa penyakit ini memicu 15% dari seluruh kematian anak-anak di bawah usia 5 tahun (WHO, 2015). Dalam mengenali foto rontgen terdapat banyak metode yang bisa diimplementasikan. Salah satu dari implementasinya adalah dengan menggunakan Convolutional Neural Network (CNN). Convolutional Neural Network atau yang biasa disebut CNN merupakan bagian dari Deep Learning yang melakukan proses pembelajaran untuk mencari representasi terbaik. Dengan Menggunakan model tersebut yang diimplementasikan dengan aplikasi web didapatkan hasil akurasi sebesar 75% dari uji acak yang dilakukan dengan menggunakan data yang berbeda dari data training.

Keywords: Pneumonia, Deep Learning, Convolutional Neural Networks, Pooling.

I Introduction

Pneumonia merupakan pembunuh utama anak dibawah usia lima tahun (Balita) di dunia, lebih banyak dibandingkan dengan penyakit lain seperti AIDS, Malaria dan Campak. Namun, belum banyak perhatian terhadap penyakit ini. Di dunia, dari 9 juta kematian Balita lebih dari 2 juta Balita meninggal setiap tahun akibat pneumonia atau sama dengan 4 Balita meninggal setiap menitnya. Dari lima kematian Balita, satu diantaranya disebabkan pneumonia. Di Indonesia menunjukkan; prevalensi nasional ISPA: 25,5% (16 provinsi di atas angka nasional), angka kesakitan (morbiditas) pneumonia pada Bayi: 2.2 %, Balita: 3%, angka kematian (mortalitas) pada bayi 23,8%, dan Balita 15,5% (Riset Kesehatan Dasar, 2007) . Pneumonia merupakan salah satu penyebab kematian anak-anak tertinggi di dunia. Badan Kesehatan Dunia (WHO) memperkirakan bahwa penyakit ini memicu 15% dari seluruh kematian anak-anak di bawah usia 5 tahun (WHO, 2015). Pada tahun 2015, terdapat lebih dari 900.000 anak-anak yang meninggal akibat pneumonia. Di Indonesia sendiri,

pneumonia diperkirakan telah merenggut sekitar 25.000 jiwa balita pada tahun 2013. Pneumonia sendiri bisa jadi infeksi paru-paru biasa tapi kadang kala bisa mematikan. Pneumonia bisa terjadi karena berbagai organisme, pneumonia karena bakteri atau virus. Upaya pemerintah dalam menekan angka kematian akibat pneumonia diantaranya melalui penemuan kasus pneumonia Balita sedini mungkin di pelayanan kesehatan dasar, penatalaksanaan kasus dan rujukan. Adanya keterpaduan dengan lintas program melalui pendekatan MTBS di Puskesmas serta penyediaan obat dan peralatan untuk Puskesmas Perawatan dan di daerah terpencil.

Menurut Menkes, pengendalian penyakit ISPA memiliki kendala diantaranya cakupan penemuan masih sangat rendah akibat tingginya mutasi tenaga kesehatan. Selain itu pengendalian pneumonia bukan program prioritas karena di beberapa daerah anggaran untuk pneumonia jumlahnya tidak memadai bahkan tidak ada sama sekali. Untuk pneumonia karena bakteri kita punya pilihan pengobatan yang lebih banyak. Jadi dapat dikatakan bahwa lebih baik terkena pneumoia karena bakteri. Sedangkan pneumonia karena virus, kita semua pernah mendengar mengenai flu. Flu adalah proses infeksi sangat umum yang menyebabkan pneumonia. Dan kita tidak punya banyak pilihan untuk mengobatinya (Dr. Matthew Schreiber, spesialis penyakit paru-paru di RS Medstar Washington). Dengan melihat permasalahan yang ditemukan di atas, penulis tertarik untuk melaksanakan penelitian yang berjudul “Penerapan deep learning pada Aplikasi prediksi penyakit Pneumonia berbasis Convolutional Neural networks”. Dengan harapan penelitian ini memberikan solusi pemecahan masalah dan juga membantu mempercepat diagnosis.

II Literature Review

Penelitian yang berkaitan dengan penelitian yang penulis lakukan diantaranya yaitu:

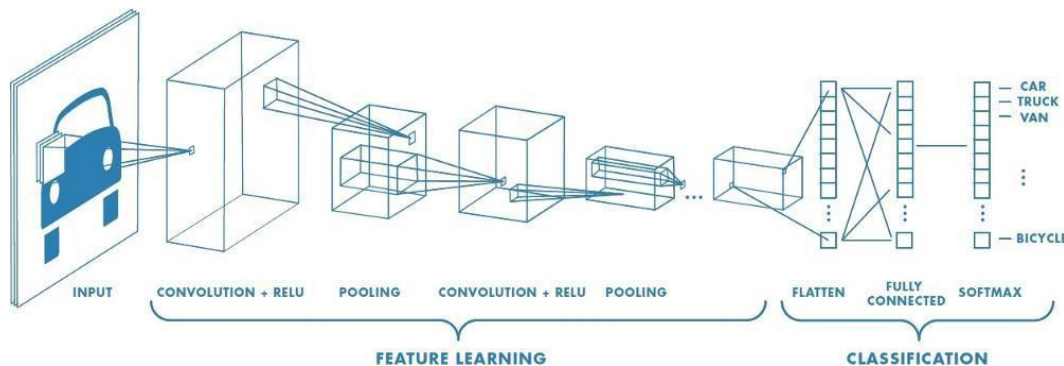
Penelitian yang dilakukan oleh Halprin Abhirawa, Jondri,M.Si., Anditya Arifianto, S.T.,M.T. (2017) dengan judul “Pengenalan Wajah Menggunakan Convolutional Neural Network”, penelitian yang berisikan dalam mengenali wajah terdapat banyak metode yang bisa diimplementasikan. Salah satu dari implementasinya adalah dengan menggunakan Convolutional Neural Network (CNN). Convolutional Neural Network atau yang biasa disebut CNN merupakan bagian dari Deep Learning yang melakukan proses pembelajaran untuk mencari representasi terbaik. Dataset yang digunakan adalah The Extended Yale Face Database B, yang berupa dataset foto wajah. Dengan menggunakan proses dropout diperoleh hasil terbaik dengan tingkat akurasi pengenalan setinggi 89.73%. Sedangkan apabila dilakukan pengujian terhadap data testing akan diperoleh hasil akurasi pengenalan setinggi 75.79%. 7. Dalam penelitian yang dilakukan oleh Naufal Dzaky Anwari, Anditya Arifianto,S.T.,M.T., Jondri,S.Si, M.Si (2018) dengan judul “Multilabel Image Annotation Menggunakan Metode Convolutional Neural Network”, penelitian yang berisikan Dengan berkembangnya sosial media terutama yang memiliki fitur untuk mengunggah foto dan gambar menyebabkan banyaknya gambar yang diunggah pada sosial media. Gambar tersebut dapat digunakan untuk membangun sistem pencarian gambar berbasis isi atau content-base image retrieval. Namun, banyak gambar yang diunggah tidak diberikan label atau tag sesuai dengan isi dari citra yang diunggah, sehingga sangat sulit untuk dikelola. Untuk dapat mewujudkan sistem pencarian gambar berbasis isi maka setiap obyek pada gambar

harus dikenali terlebih dahulu. Jika pengenalan obyek tersebut dilakukan secara manual maka akan sangat sulit karena akan memakan waktu yang lama dan makna dari setiap orang terhadap suatu gambar berbeda, yang menimbulkan pengenalan obyek yang subyektif. Oleh karena itu dibangunlah sistem penganotasian gambar secara otomatis. Penelitian ini diajukan sebuah metode Convolutional Neural Network untuk menangani sistem penganotasian gambar multilabel. Metode Convolutional Neural Network telah terbukti memiliki performansi yang baik pada kasus klasifikasi gambar, ditunjukkan dengan performansi pada kasus klasifikasi gambar pada ILSVRC yang semakin membaik setiap tahunnya. Performansi tertinggi terhadap data uji pada penelitian ini adalah 81.24%.

III Research Method

3.1 Konsep Dasar Convolutional Neural Network

Convolutional Neural Network (CNN) adalah salah satu jenis neural network yang biasa digunakan pada data image. CNN bisa digunakan untuk mendeteksi dan mengenali object pada sebuah image. Secara garis besar CNN tidak jauh beda dengan neural network biasanya. CNN terdiri dari neuron yang memiliki bobot, bias dan fungsi aktivasi.



Gambar 3.1 Proses Convolutional Neural Network

3.2 Model Convolutional Neural Networks

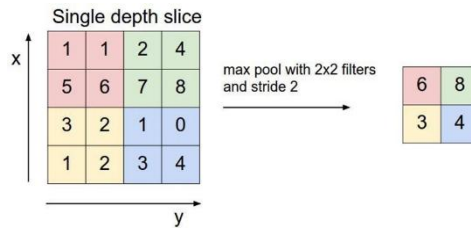
Berikut macam-macam model convolutional neural networks:

1. Convolutional Layer

Convolutional Layer bagian yang melakukan operasi konvolusi yaitu mengkombinasikan linier filter terhadap daerah lokal. Layer ini yang pertama kali menerima gambar yang diinputkan pada arsitektur. Bentuk layer ini adalah sebuah filter dengan panjang (pixel), lebar (pixel) dan tebal sesuai dengan channel image data yang diinputkan. Ketiga filter ini akan bergeser keseluruhan bagian gambar. Pergeseran tersebut akan melakukan operasi “dot” antara input dan nilai dari filter tersebut sehingga akan menghasilkan output yang disebut activation map atau feature map.

2. Pooling Layer

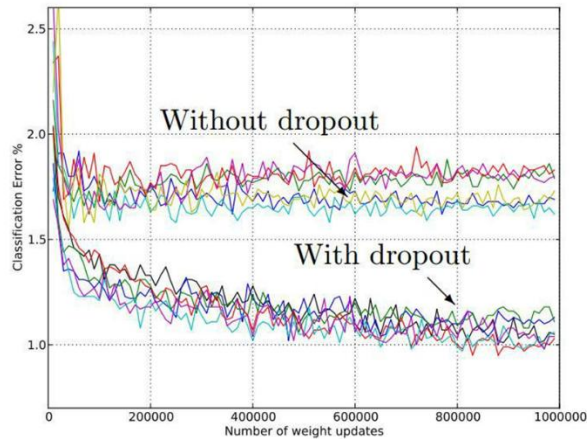
Pooling layer menerima output dari convolution layer, pada layer ini ukuran data citra akan direduksi. Prinsipnya pooling layer terdiri dari filter dengan ukuran tertentu dan stride/langkah kemudian bergeser keseluruh area feature map. Sebagian besar arsitektur CNN, metode Pooling yang digunakan adalah Max pooling. Max pooling membagi output convolution layer menjadi beberapagrid kemudian setia pergeseran filter akan mengambil nilai terbesar dari setiap grid. Tergantung pada panjang langkahnya, gambar yang dihasilkan adalah sebagian kecil dari ukuran aslinya yang berguna untuk mengurangi dimensi data, sehingga mengurangi jumlah parameter pada langkah selanjutnya.



Gambar Pooling Layer

3. Dropout Layer

Pada setiap tahap pelatihan, node-node individu dikeluarkan dari internet dengan probabilitas $1-p$ atau disimpan dengan probabilitas p , sehingga jaringan yang berkurang dibiarkan; tepi masuk dan keluar ke node drop-out juga dihapus. Kita membutuhkan Dropout untuk mencegah over-fitting. Lapisan yang terhubung sepenuhnya menempati sebagian besar parameter, dan karenanya, neuron mengembangkan ko-dependensi di antara satu sama lain selama pelatihan yang mengekang daya individu masing-masing neuron yang menyebabkan over-pas dari data pelatihan.



Gambar dropout/without dropout

4. Fully connected layer

Fully connected layer merupakan lapisan dimana semua neuron aktivasi dari lapisan sebelumnya terhubung semua dengan neuron di lapisan selanjutnya. Fully connected layer biasanya digunakan

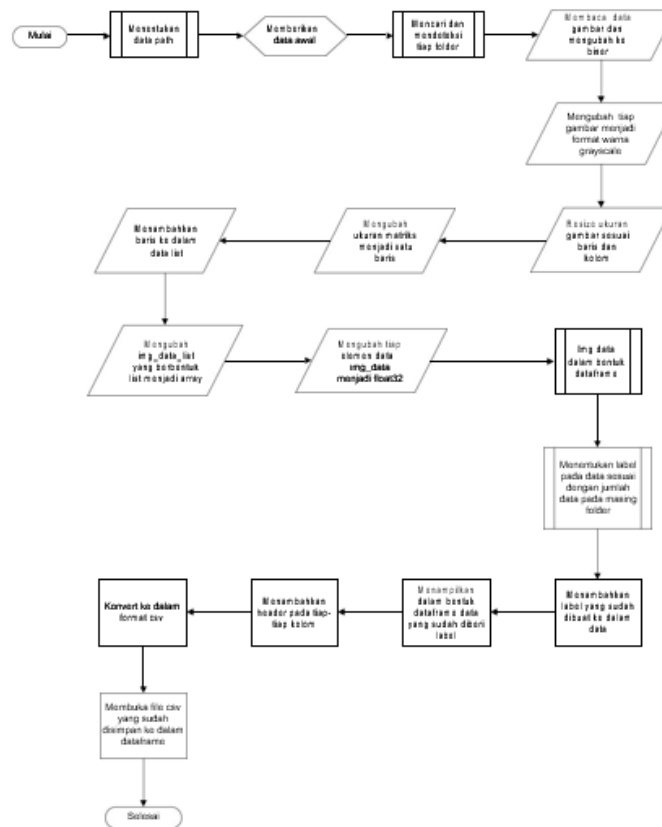
dalam penerapan Multi Layer Perceptron (MLP) dan bertujuan untuk melakukan transformasi pada dimensi data agar dapat diklasifikasikan secara linear.

Perbedaan antara fully connected layer dan convolution layer biasanya adalah neuron di convolution layer terhubung hanya ke daerah tertentu pada input, sedangkan fully connected layer mempunyai neuron yang secara keseluruhan terhubung. Namun, kedua lapisan tersebut masih mengoperasikan produk dot, sehingga fungsinya tidak begitu berbeda

IV Results And Discussion

4.1 Rancangan Sistem

Data yang digunakan pada penelitian ini adalah data-data yang didapatkan dari data training dan data test. Data training terdapat 5216 file gambar yang terdiri dari 1341 file gambar untuk normal dan 3875 file gambar untuk pneumonia sedangkan data test terdapat 624 file gambar yang terdiri dari 234 file gambar untuk normal dan 390 file gambar untuk pneumonia. Jadi total seluruh gambar untuk training dan testing adalah 5840 file. Semua data tersebut dalam format jpg dan jpeg data diambil dari penelitian Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, v2. Dibawah ini adalah flowchart untuk preprocessing:



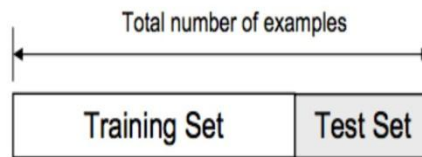
Gambar 4.1 Proses Preprocessing

Hasil image processing ini kita dapatkan dari 2 tabel yaitu tabel test dan tabel train. Tabel test berasal dari csv yang terdiri dari 625 baris dan 784 kolom. Sedangkan tabel train berasal dari csv juga yang terdiri dari 5218 baris dan 784 kolom.

4.2 Pembahasan

4.2.1 Train / Test Split

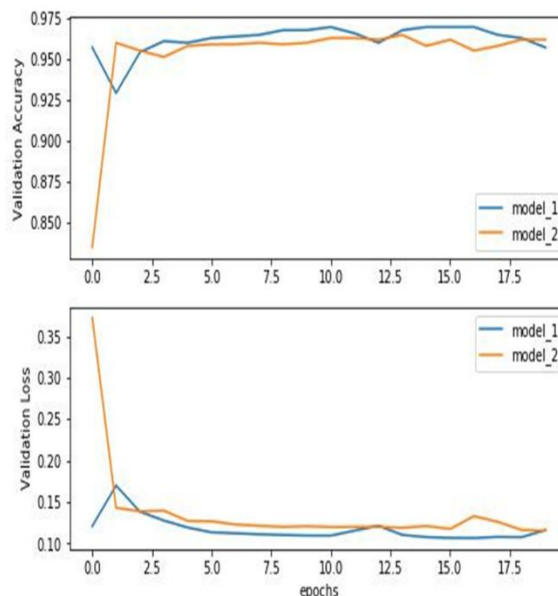
Data yang di gunakan biasanya dibagi menjadi data pelatihan dan data uji. Set pelatihan berisi output yang diketahui dan model belajar pada data ini untuk digeneralisasikan ke data lain di kemudian hari. Kami memiliki dataset uji (atau subset) untuk menguji prediksi model kami pada subset ini.



Gambar 4. 1 Train / Test Split

4.2.2 Training Model

Training dilakukan menggunakan 2 model yang berbeda yaitu model 1 yang dapat dilihat di gambar 3.6 dan model 2 yang dapat dilihat pada gambar 3.8, kedua model tersebut memiliki jumlah layer dan filter yang sama namun perbedaan mendasar adalah pada metode pooling yang digunakan masing masing model tersebut. Pada model 1 menggunakan max-pooling dan pada model 2 menggunakan average pooling, training dilakukan pada 5216 gambar dengan label pneumonia dan normal. Dari proses training tersebut diharapkan dapat menghasilkan bobot yang tepat untuk dapat memprediksi penyakit pneumonia. Pada training model akan menghasilkan grafik validasi akurasi dan validasi dengan cara epochs seperti gambar dibawah:



Gambar 4. 2 Training Model

4.2.3 Testing Model

Pada testing model kita mendapatkan test loss dan test akurasi, data yang digunakan pada proses testing merupakan data yang berbeda dengan data yang digunakan pada proses training proses testing tidak menggunakan data sebanyak proses training, selain itu testing digunakan untuk mengevaluasi bobot yang dihasilkan pada proses training sehingga proses testing tidak kan mengganti nilai bobot seperti pada proses training. Hasil dari proses testing dapat dilihat pada gambar 4.3

```
In [72]: score = cnn_model_1.evaluate(x_test, y_test, verbose=1)
print('Test Loss : {:.4f}'.format(score[0]))
print('Test Acc : {:.4f}'.format(score[1]))

624/624 [=====] - 0s 359us/step
Test Loss : 0.7369
Test Acc : 0.7756

In [73]: score = cnn_model_2.evaluate(x_test, y_test, verbose=1)
print('Test Loss : {:.4f}'.format(score[0]))
print('Test Acc : {:.4f}'.format(score[1]))

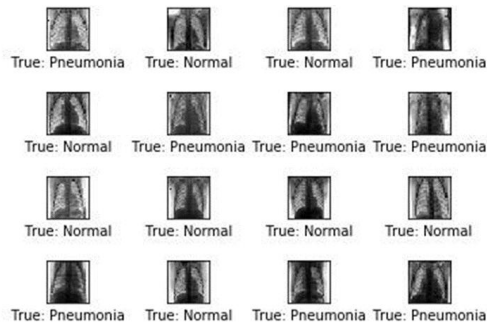
624/624 [=====] - 0s 571us/step
Test Loss : 0.7859
Test Acc : 0.7676
```

Gambar 4. 3 Testing Model

Dari gambar 4.3 dapat dilihat hasil dari testing dengan menggunakan data test, didapatkan nilai akurasi untuk model 1 sebesar 0.7756 atau 77.56% dan untuk model 2 sebesar 0.7676 atau 76.76 % dari hasil tersebut dapat dilihat bahwa model 1 memiliki akurasi yang lebih baik ketimbang model 2 namun perbedaannya tidak signifikan. Sedangkan untuk perbandingan nilai *loss* untuk kedua model tersebut adalah 0.7369 atau 73.69 % untuk model 1 dan 0.7859 atau 78.59% untuk model 2, lagi-lagi dapat dilihat perbandingan kedua model tersebut yaitu nilai *loss* dari model 2 lebih besar namun sama seperti akurasi perbedaannya tidak signifikan

4.2.4 Random Sampiling Tes

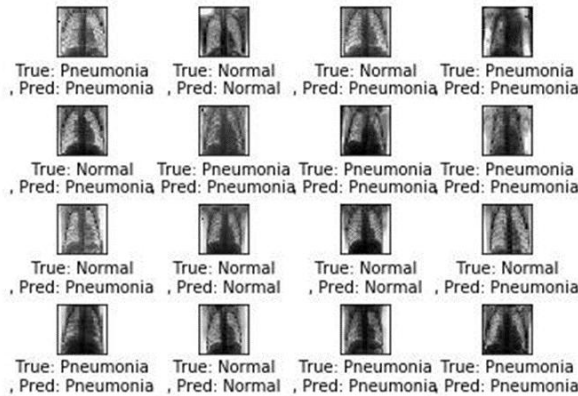
Dibawah ini adalah data *first images* dari *test-set*. Dimana data ini belum ada prediksi dan data ini terdiri dari 8 data pneumonia dan 8 data normal.



Gambar 4. 4 First Images

A. Model 1

Pada model 1 dilakukan random sample sebanyak 16 foto dari data test dengan label masing-masing, yang menghasilkan 12 file foto dengan tepat dapat diprediksi dengan benar sesuai dengan label datanya dan 4 diprediksi tidak tepat atau salah, bisa dilihat di gambar bawah berikut ini:

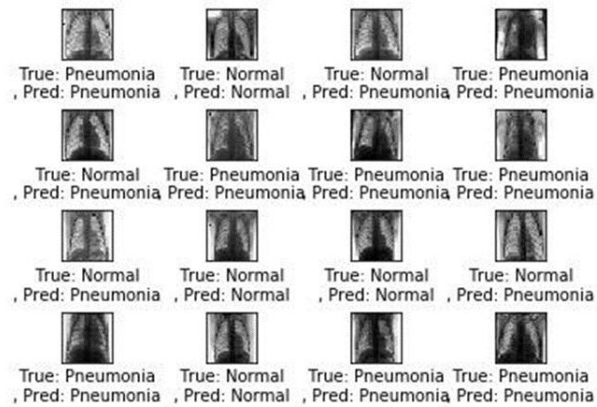


Gambar Model 1

Sehingga dapat diperoleh nilai akurasi dari eksperimen tersebut sebesar 75%.

B. Model 2

Pada model 1 kita melakukan random sample yang menghasilkan 12 file foto kesamaan pada true dan prediksi dan 4 yang tidak tepat antara true dan prediksi, bisa dilihat di gambar bawah berikut ini:



Gambar Model 2

Dari semua rangkaian yang kita lakukan mulai dari training data yang akan menghasilkan grafik validasi akurasi dan validasi loss dengan cara epochs lalu kita melakukan testing data sehingga kita mendapatkan test loss dan test akurasi. Setelah itu semua kita melakukan random simple testing dari data first images dari test-set. Dimana data ini belum ada prediksi dan data ini terdiri dari 8 data pneumonia dan 8 data normal. Sehingga kita memperoleh data untuk model 1 yakni 12 file foto kesamaan pada true dan prediksi dan 4 yang tidak tepat antara true dan prediksi sedangkan untuk model 2 sama hasilnya dengan model 1 yakni 12 file foto kesamaan pada true dan prediksi dan 4 yang tidak tepat antara true dan prediksi.

Jadi, dapat disimpulkan dari semua rangkaian itu kita dapat memperoleh hasil prediksi yang benar sekitar 75% dari data yang sesuai dengan hasil prediksi pada sampel acak

V Conclusion

Berdasarkan penulisan dan penelitian yang telah penulis uraikan, mengenai prediksi pneumonia dengan menggunakan algoritma convolutional neural networks, maka didapat dibuat kesimpulan yaitu :

1. Metode CNN dapat melakukan proses prediksi gambar khususnya klasifikasi dengan tingkat akurasi yang cukup tinggi.
2. Setelah dilakukan perbandingan dengan kedua model yang dirancang pada bab sebelumnya didapatkan hasil diatas 75% baik dengan evaluasi menggunakan data testing dan random sample, namun kendala yang ditemukan karena nilai akurasi masih dibawah angka 85% sehingga masih harus dilakukan penelitian kembali untuk meningkatkan nilai akurasi dari model tersebut

VI ACKNOWLEDGMENT

Penelitian ini didanai oleh Ristekdikti melalui Program Penelitian Dosen Pemula. Terima kasih banyak disampaikan ke Ristekdikti

VII References

- [1] Ahmad, A. (2017). *Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning*. Jurnal Teknologi Indonesia
- [2] Alodokter. (2018, Juli 20). Retrieved from Alodokter.
- [3] Ashby, B., & Turkington, C. (2007). *The encyclopedia of infectious diseases* (edisi ke-3rd). New York: Facts on File.
- [4] Danukusumo, K. P. (2017). Implementasi Deep Learning Menggunakan Convolutional Neural Network Untuk Klasifikasi Citra Candi Berbasis GPU. Yogyakarta : Universitas Atma Jaya.
- [5] E. P, I. W. (2016). *Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN)*. Surabaya : Institut Teknologi Sepuluh Maret.
- [6] Indonesia, U. (2013). UNICEF/WHO: New plan to address pneumonia and diarrhoea could save 2 million children a year.
- [7] J. Han, M. K. (2006). *Data Mining: Concepts and Techniques*. Morgan Kaufmann. Kusumaningrum, T. F. (2018). IMPLEMENTASI CONVOLUTION NEURAL NETWORK (CNN) UNTUK KLASIFIKASI JAMUR KONSUMSI DI INDONESIA MENGGUNAKAN KERAS.
- [8] Kusumaningrum, T. F. (2018). IMPLEMENTASI CONVOLUTION NEURAL NETWORK (CNN) UNTUK KLASIFIKASI JAMUR KONSUMSI DI INDONESIA MENGGUNAKAN KERTAS.YOGYAKARTA: UNIVERSITAS ISLAM INDONESIA YOGYAKARTA.
- [9] Leach, R. E. (2009). *Acute and Critical Care Medicine at a Glance*. pp. Wiley-Blackwell. ISBN 1-4051-6139-6.
- [10] M. J. Zaki, W. M. (2014). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge: University Press.
- [11] McLuckie, A. (2009). *Respiratory disease and its management*. New Yor: Springer. Nasichuddin, M. A. (2017). Implementasi CNN Untuk Klasifikasi Teks Menggunakan Tensorflow.
- [12] Pommerville, J. C. (2010). *Alcama's Fundamentals of Microbiology*. Sudbury MA:
- [13] Jones & Bartlett. Pommerville, J. C. (2010). *Alcama's Fundamentals of Microbiology*. Sudbury MA: Jones & Bartlett.
- [14] Putra, D. (2010). *Pengolahan Citra Digital*. Yogyakarta: Andi.
- [15] Rismiyati. (2016). *Implementasi Convolutional Neural Network untuk Sortasi Salak Ekspor Berbasis Citra Digital*. Yogyakarta: Universitas Gadjah Mada.
- [16] Sena, S. (2017). *Pengenalan Deep Learning*.
- [17] Suhartono, D. (2012). *Dasar Pemahaman Neural Network*.
- [18] Sutojo, T. E. (2017). *Teori Pengolahan Citra Digital*. Semarang : Universitas Dian Nuswantoro.
- [19] Tintinalli, J. E. (2010). *Emergency Medicine: A Comprehensive Study Guide*
- [20] (Emergency Medicine (Tintinalli)). New York: McGraw-Hill Companies.
- [21] Utami, E. d. (2004). *Logika, Algoritma dan Impelementasinya dalam Bahasa Python di GNU/Linux*. Yogyakarta : Andi Offset.
- [22] Z, H. (2006). "Pneumonia: update on diagnosis and management".