

## Perbandingan Arsitektur ResNet50 dan ResNet101 dalam Klasifikasi Kanker Serviks pada Citra *Pap Smear*

Za'imatun Niswati<sup>1</sup>, Rahayuning Hardatin<sup>2</sup>, Meia Noer Muslimah<sup>3</sup>, Siti Nur Hasanah<sup>4</sup>

<sup>1</sup>Department of Informatic, Universitas Indraprasta PGRI, Indonesia

<sup>2,3,4</sup>Department of Computer Science, Institut Pertanian Bogor, Indonesia

### Article Info

#### Article history:

Received June 28, 2021

Revised Sept 15, 2021

Accepted Sep 29, 2021

#### Keywords:

*Kanker serviks*

*Citra pap smear*

*Data SIPaKMeD*

*Convolution Neural Network (CNN)*

*Residual Network (ResNet)*

### ABSTRACT

Cervical cancer is one of the most deadly types of cancer in women. Cervical cancer ranks second for the type of cancer that attacks women the most. Data from the Indonesian Ministry of Health, there are at least 15000 cases of cervical cancer every year in Indonesia. This cancer is a type of tumor that develops in the epithelial tissue of the cervix. In addition to HPV vaccination, cervical cancer detection can also be carried out with a Pap smear test and VIA examination supported by medical image tests such as CT scan, microscopic, and MRI. Pap smear test is a type of test to detect cervical cancer which is quite widely used because the cost of the test is cheaper than the HPV vaccination. This test is carried out by taking samples of uterine cells which are then analyzed for early detection of cervical cancer. Through a pap smear can be found the presence of HPV infection and abnormal cells that can turn into cancer cells. The purpose of this research is to apply the ResNet50 and ResNet101 architectures on pap smear images to identify cervical cancer and evaluate the performance of the ResNet50 and ResNet101 architectures in the classification of cervical cancer on pap smear images. In this study, CNN ResNet50 and ResNet101 were used to classify cervical cancer on pap smear images. This study has created two models to predict the grade of cervical cancer on pap smear images. The ResNet50 architecture gets 91% accuracy while the ResNet101 architecture gets 89%. Although the architecture of ResNet101 is more complex than ResNet50, if viewed from the results of the model evaluation, ResNet101 has worse performance. This is due to the relatively small training data when trained with a large architecture such as ResNet101, not necessarily resulting in better accuracy.

Copyright © 2020 Universitas Indraprasta PGRI.  
All rights reserved.

### Corresponding Author:

Za'imatun Niswati,

Department of Informatic,

Universitas Indraprasta PGRI,

Jl. Nangka No. 58 C, Tanjung Barat, Jagakarsa, Jakarta Selatan.

Email: [zaimatunnis@gmail.com](mailto:zaimatunnis@gmail.com)

## 1. PENDAHULUAN

Kanker serviks adalah salah satu jenis kanker yang paling mematikan pada wanita. Menurut [1] kanker serviks menempati urutan kedua untuk jenis kanker yang paling banyak menyerang wanita. Menurut data Kementerian Kesehatan RI, setidaknya terjadi 15.000 kasus kanker serviks setiap tahunnya di Indonesia. Kanker ini merupakan jenis tumor yang berkembang di jaringan epitel mulut rahim.

Penyakit ini merupakan salah satu jenis kanker yang mungkin dapat dicegah apabila terdeteksi dini dan ditangani dengan benar [2]. Sehingga, upaya pencegahan kanker serviks sangat penting untuk

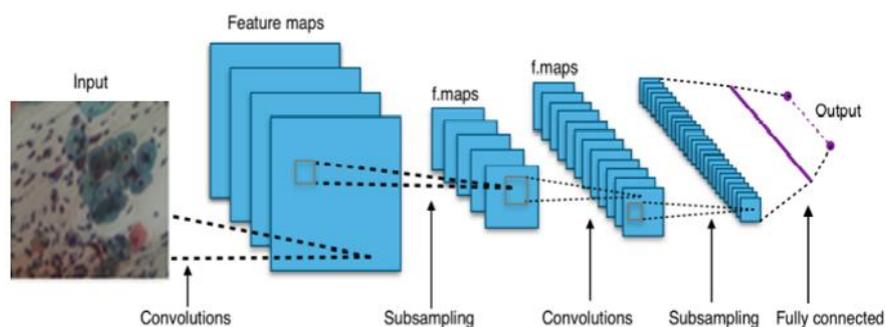
dilakukan dengan cepat. *World Health Organization* (WHO) menyatakan bahwa tindakan pencegahan primer yang dapat dilakukan untuk mencegah terjadinya kanker serviks yaitu dengan melakukan vaksinasi *Human Papilloma Virus* (HPV). Dengan melakukan vaksinasi HPV tubuh akan membentuk antibodi terhadap virus HPV, sehingga tubuh memiliki kekebalan terhadap virus ini.

Selain vaksinasi HPV, deteksi kanker serviks juga dapat dilakukan dengan tes *pap smear* dan pemeriksaan IVA yang didukung dengan tes citra/gambar medis seperti CT scan, mikroskopis dan MRI [3]. Tes *pap smear* merupakan jenis tes untuk mendeteksi kanker serviks yang cukup banyak digunakan karena biaya tesnya yang lebih murah dibandingkan melakukan vaksinasi HPV. Tes ini dilakukan dengan mengambil contoh sel-sel rahim yang kemudian dianalisis untuk mendeteksi dini kanker serviks [4]. Melalui *pap smear* bisa ditemukan adanya infeksi HPV dan sel-sel yang abnormal yang dapat berubah menjadi sel kanker.

Interpretasi visual citra *pap smear* secara manual memiliki banyak keterbatasan, membutuhkan waktu lama dan rawan kesalahan prosedur [5]. Oleh karena itu, analisis secara otomatis dengan bantuan komputer sangat dibutuhkan untuk memudahkan proses interpretasi tersebut. *Computer Aided Diagnosis* (CADx) adalah teknik yang dirancang untuk mengurangi kelalaian pengamatan sehingga dapat mengurangi kemungkinan terjadinya *false negatif* pada hasil tafsiran citra medis oleh dokter. Dalam beberapa tahun terakhir, teknik *deep learning* telah memberikan dampak yang baik pada pemrosesan citra medis dan pengembangan sistem CADx. Untuk mengolah data mentah seperti citra berdimensi 3 dan dapat melakukan ekstraksi fitur secara otomatis untuk tujuan deteksi, segmentasi, dan klasifikasi dapat memanfaatkan *deep learning*. Menurut [6], menyatakan bahwa metode *Convolutional Neural Network* (CNN) dapat mengklasifikasikan citra dan memprediksi kanker serviks pada citra *pap smear*. Pada penelitian ini dilakukan optimasi *hyperparameter* untuk mencari parameter yang optimal dalam melakukan klasifikasi kanker serviks dan diharapkan dapat meningkatkan akurasi.

CNN terdiri atas komponen *convolutional*, *pooling*, *fully connected layers*, dan sebuah *output layer* yang disusun secara bertumpuk [7]. *Convolution layer* memiliki kumpulan *filter* yang ukuran dimensinya lebih kecil dari ukuran dimensi *input*. *Pooling layer* bertujuan untuk mengurangi jumlah parameter dan komputasi yang diperlukan dengan melakukan *down sampling* menggunakan *pooling function*. *Fully connected layer* dalam CNN merupakan *layer* yang setiap neuronnya saling terhubung dengan fungsi aktivasi di *layer* berikutnya [8].

CNN terdiri dari tiga lapisan diantaranya adalah lapisan masukan (*input layer*), lapisan keluaran (*output layer*) dan lapisan tersembunyi (*hidden layer*). Di dalam *hidden layer* terdapat lapisan-lapisan yang tersusun secara bertumpuk. Lapisan-lapisan tersebut diantaranya adalah *convolutional layer*, *pooling layers*, *normalization layer*, *ReLU layer*, *fully Connected layer* dan *loss layer* [9].



Gambar 1. Ilustrasi arsitektur CNN [10]

Peningkatan kedalaman jaringan pada *deep convolutional neural network* tidak selalu menghasilkan akurasi pelatihan yang lebih baik, ada kalanya terdapat penurunan akurasi pelatihan. Hal tersebut disebabkan oleh degradasi yang terjadi pada jaringan karena tidak semua jaringan mudah dioptimisasi. Solusi untuk mengurangi degradasi pada jaringan yang lebih dalam adalah dengan menambahkan pemetaan identitas (*identity mapping*) dan lapisan lainnya yang disalin dari model yang lebih dangkal. Maka dibuatlah jaringan residual atau ResNet. Jaringan ini membiarkan lapisan-lapisan sesuai dengan pemetaan residual. Jaringan ini berupa jaringan *feed forward* dengan koneksi pintas (*shortcut connection*) yang melewati satu atau lebih lapisan. Koneksi pintas ini hanya melakukan pemetaan identitas dan outputnya ditambahkan ke output lapisan yang ditumpuk. Koneksi pintas juga tidak menambahkan parameter atau kompleksitas komputasi. Selain itu, koneksi pintas yang dimiliki oleh

jaringan ResNet seluruh jaringannya juga dapat dilatih oleh *Stochastic Gradient Descent* dan mudah diimplementasikan dengan *library* tanpa memodifikasinya [11].

Pada citra digital sudah banyak dilakukan penerapan CNN, yaitu metode CNN-LSTM sudah banyak dilakukan, seperti [12] melakukan deskripsi gambar dengan memperhatikan fitur semantic, penelitian [13] menggunakan CNN-LSTM dimana deskripsi gambar yang dihasilkan cukup baik, penelitian [14] klasifikasi Populasi Ternak menggunakan CNN. Begitu pula dengan penelitian [15] menyimpulkan bahwa akurasi hasil prediksi kata menggunakan metode CNN lebih baik daripada LSTM.

ResNet (*Residual Network*) merupakan jaringan residual yang memiliki jaringan yang dalam. Jaringan terdalam dari ResNet berjumlah 152 lapisan. Jaringan ini 8 kali lebih dalam dari jaringan VGG namun kompleksitasnya masih lebih rendah daripada jaringan VGG. Pada tahun 2015, Jaringan ini berhasil memenangkan juara pertama pada kompetisi ILSVRC (*ImageNet Large Scale Visual Recognition Challenge*) dan *COCO* dalam hal klasifikasi, deteksi dan segmentasi gambar pada dataset *COCO* dan *ImageNet* [11].

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

Gambar 2. Arsitektur dan Jumlah lapisan ResNet [16]

Tujuan penelitian adalah menerapkan arsitektur ResNet50 dan ResNet101 pada citra *pap smear* untuk mengidentifikasi kanker serviks dan mengevaluasi kinerja arsitektur ResNet50 dan ResNet101 dalam klasifikasi kanker serviks pada citra *pap smear*.

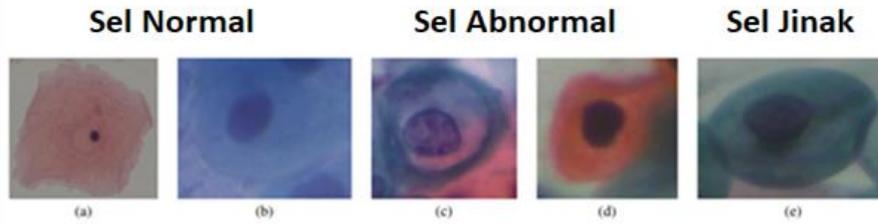
## 2. METODE

Penelitian [6] mengenai penggunaan metode *convolutional neural network* dalam klasifikasi citra menginspirasi penelitian ini untuk mengimplementasi metode tersebut dalam memprediksi kanker serviks pada citra *pap smear*. Metode CNN diimplementasikan pada data citra *pap smear* yang sudah dilakukan segmentasi. Optimasi *hyperparameter* untuk mencari *hyperparameter* yang optimal juga dilakukan pada kasus klasifikasi kanker serviks dan diharapkan dapat membantu dalam meningkatkan kebaikan model CNN. Pada penelitian ini menggunakan CNN ResNet50 dan ResNet101 untuk klasifikasi kanker serviks pada citra *pap smear*.

CNN adalah salah satu model *deep learning* yang dapat mengoptimasi diri sendiri secara adaptif melalui masukan yang diterima oleh neuron [18]. Alasan dibalik penggunaan CNN dibandingkan algoritme lainnya karena penggunaan konsep *weight sharing* sehingga parameter yang digunakan akan lebih sedikit serta menghasilkan generalisasi yang lebih baik [19].

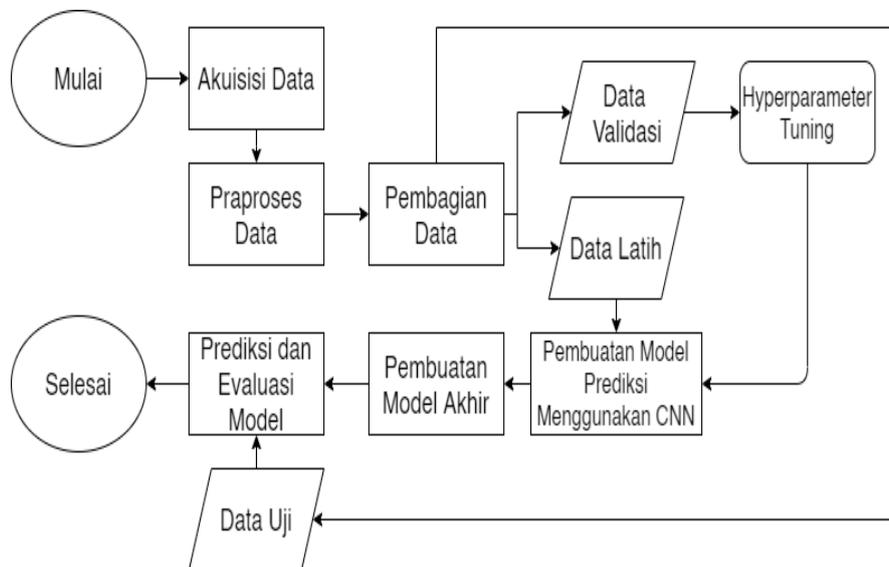
Data yang digunakan untuk pelatihan model merupakan data sel SIPaKMeD <http://www.cs.uoi.gr/~marina/sipakmed.html> yang sudah dilakukan segmentasi dari 5 kelas *superficial-intermediate*, *parabasal*, *koilocytotic*, *dyskeratotic*, dan *metaplastic* [6]. Citra asli yang digunakan adalah 4049 data. Untuk membuat model, dataset dibagi menjadi dataset latih, dataset validasi dan dataset uji dengan perbandingan 80:10:10 dari jumlah total 4049 data. Sehingga jumlah dataset latih berjumlah 3237

data, dataset validasi berjumlah 403 data, dan dataset uji berjumlah 409 data. Model dibuat menggunakan dataset latihan dan dataset validasi. Selain itu juga menggunakan hasil ekstraksi fitur gambar. Kemudian dilakukan penambahan data dengan menggunakan 3 teknik augmentasi (rotasi, vertikal flip, dan horizontal flip).



Gambar 3. Data SIPaKMeD [20]

Penelitian ini terdiri atas 7 (tujuh) tahapan utama, yaitu akuisisi data, praproses data, pembagian data, *hyperparameter tuning*, pembuatan model prediksi menggunakan CNN, pembuatan model akhir, serta prediksi dan evaluasi model. Aktivitas yang dilakukan pada pra proses data yaitu augmentasi data. Tahapan-tahapan yang dilakukan pada penelitian ini dapat dilihat pada Gambar 4.



Gambar 4. Tahapan Penelitian

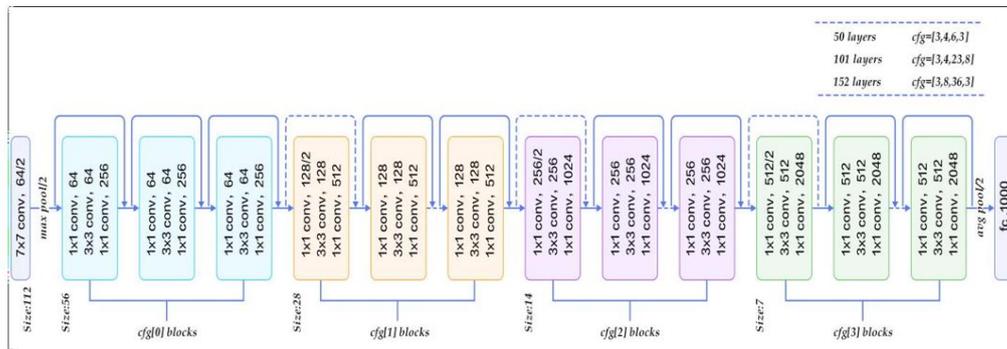
Konsep alur yang akan diterapkan untuk pembuatan deskripsi gambar (*generate image captioning*) kanker serviks pada citra *pap smear* menggunakan metode CNN. Pada penelitian ini, jenis CNN yang digunakan untuk proses ekstraksi fitur gambar pada dataset gambar adalah ResNet50 dan ResNet101.

Proses ekstraksi fitur dataset gambar kanker serviks pada citra *pap smear* dilakukan menggunakan server *Google Colaboratory* dengan pengaturan *runtime GPU* yang terhubung dengan *Google Drive* sebagai media penyimpanan. Penulis melakukan proses ekstraksi fitur gambar dengan mengekstraksi seluruh gambar pada dataset gambar menjadi vektor berukuran tetap. Proses ekstraksi fitur dataset gambar dilakukan menggunakan algoritma *Convolutional Neural Network* yaitu ResNet50 dan ResNet101. Hasil ekstraksi fitur gambar dari ResNet50 akan digunakan untuk membuat Model-1 dan ResNet101 untuk membuat Model-2.

**3. HASIL DAN PEMBAHASAN**

Pada jaringan ResNet 50, 101 dan 152 layer, terdapat modifikasi pada blok bangunan koneksi pintas. Modifikasi tersebut adalah setiap fungsi residual ditumpuk 3 lapisan konvolusi dengan ukuran 1x1, 3x3,dan 1x1. Modifikasi ini disebut *Deeper Bottleneck architecture*, dimana lapisan terakhir 1x1 yang *responsible* untuk mengurangi dan meningkatkan dimensi dengan meninggakan lapisan konvolusi 3x3

input/output yang lebih kecil. Pada ResNet 50,101 dan 152 layer digunakan pintasan identitas yang bebas parameter (*parameter-free identity shortcut*) [11].

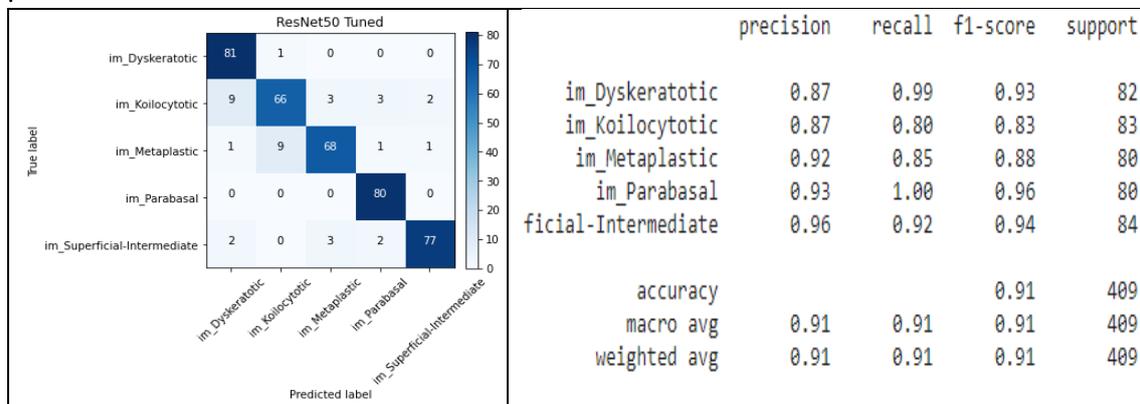


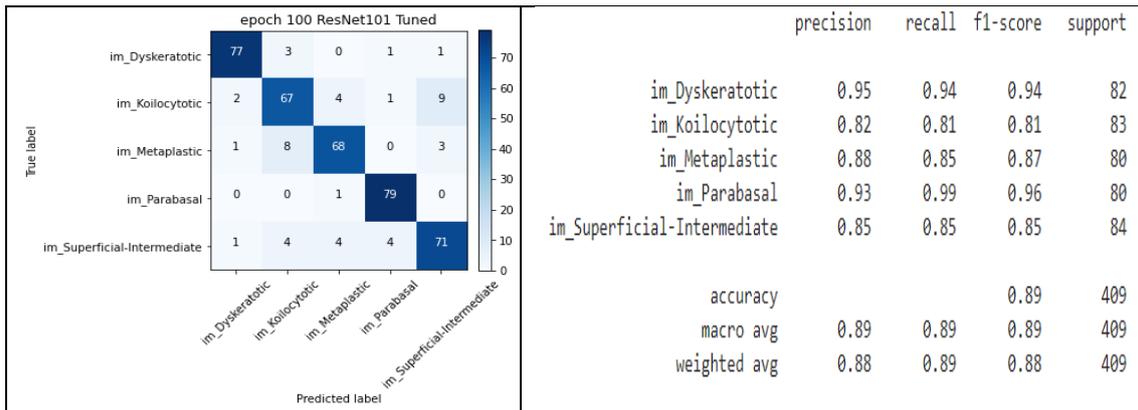
Gambar 5. Arsitektur ResNet [17]

Pada ResNet 50 layer, untuk meningkatkan dimensi menggunakan *projection shortcut* pada pintasan identitas (*identity shortcut*) dan mengubah 2 layer block pada ResNet 34 layer dengan 3 layer *bottleneck block*. Sedangkan pada ResNet 101 dan 152 layer menggunakan lebih dari 3 layer block. Meskipun ResNet 152 layer kedalamannya meningkat secara signifikan, namun masih mempunyai kompleksitas yang lebih rendah dari ResNet50 [11].

Model yang dibuat menggunakan hasil ekstraksi fitur gambar dari ResNet50. Nilai input model diambil dari vektor terakhir dari proses ekstraksi fitur. Model-2 menggunakan hasil ekstraksi fitur gambar dari ResNet101. Nilai input model diambil dari vektor terakhir hasil proses ekstraksi fitur.

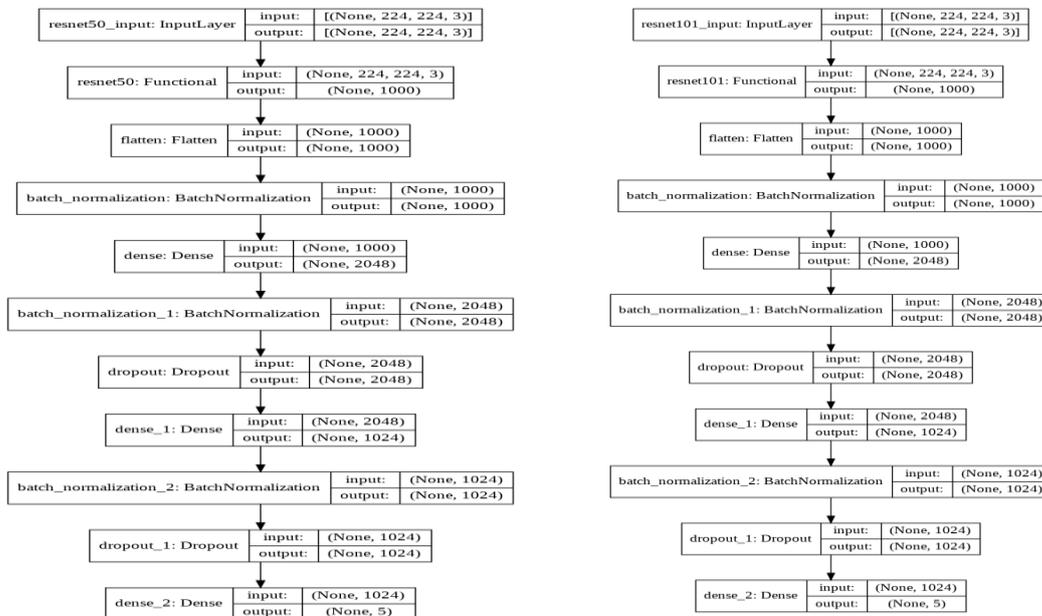
Simulasi dilakukan menggunakan *optimizer Adam* dengan pelatihan setiap model sebanyak 100 *epoch* pada setiap model dengan pengaturan *Early Stopping* yang akan memberhentikan proses pelatihan pada *epoch* tertentu dimana kinerja model membaik pada dataset validasi. Simulasi dilakukan untuk melihat akurasi dari kedua model. Berikut ini merupakan summary model-model yang dihasilkan pada tahapan ini :





Gambar 6. Summary model-1 dan model-2 (sumber : hasil olahan penulis)

Dari hasil simulasi menggunakan *optimizer Adam* diperoleh akurasi 91% untuk ResNet50 dan akurasi 89% untuk ResNet101. Kinerja model-1 lebih baik dibanding kinerja model-2. Berikut ini merupakan arsitektur dari model-model yang dibuat pada tahapan ini :



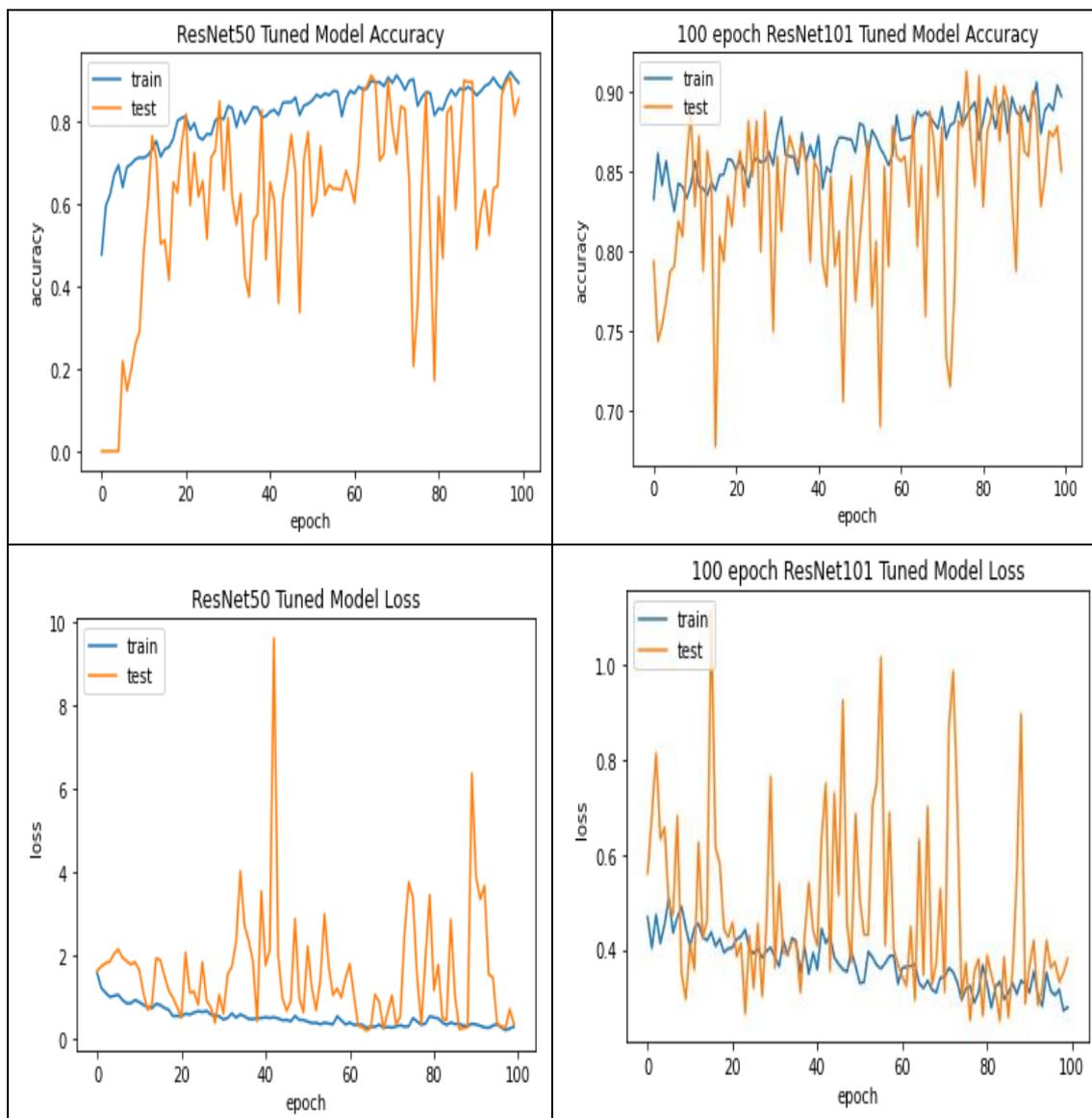
Gambar 7. Arsitektur Model Dengan ResNet50 dan ResNet101 (sumber : hasil olahan penulis)

Untuk validasi, dilakukan penilaian terhadap kinerja dari model- model yang dihasilkan dengan membandingkan nilai *loss* dan akurasi pada masing-masing model. Setelah itu dilakukan pengujian masing-masing model menggunakan data uji untuk mendapatkan deskripsi prediksi dengan deksripsi aktual. Tingkat akurasi dari deskripsi prediksi dan aktual akan dinilai menggunakan skor BLEU.

Dari kedua model yang dihasilkan dalam penelitian ini, dihitung nilai *loss* dan akurasi pada dataset latih dan dataset validasi dengan pengujian model sebanyak 100 *epoch* dengan pengaturan *Early Stopping*.

Berikut ini merupakan hasil nilai *loss* dan akurasi dari kedua model :

ResNet50	ResNet101
----------	-----------



Gambar 8. Grafik Akurasi dan Loss (sumber : hasil olahan penulis)

Model ResNet50 menghasilkan nilai akurasi 0.91 dan sedikit berfluktuasi untuk data testing. Sedangkan pada model ResNet101 menghasilkan nilai akurasi 0.89 dengan fluktuasi yang lebih banyak. Meskipun arsitektur ResNet101 lebih kompleks dari ResNet50 namun jika dilihat dari hasil evaluasi model, ResNet101 memiliki performa yang lebih buruk. Hal ini diakibatkan oleh data latih yang relatif kecil apabila dilatih dengan arsitektur yang besar seperti ResNet-101, belum tentu menghasilkan akurasi yang lebih baik.

Hasil model prediksi dievaluasi dengan melihat metrik akurasi dan nilai *categorical cross entropy loss function*. Nilai akurasi dihitung berdasarkan proporsi antara *true positive* dan *true negative* dengan keseluruhan hasil prediksi. Kesalahan diagnosa dalam dunia medis merupakan hal yang berakibat sangat fatal, sehingga model arsitektur ResNet50 lebih direkomendasikan untuk klasifikasi citra *pap smear*.

#### 4. PENUTUP

Penelitian ini telah membuat dua model untuk memprediksi kelas kanker serviks pada citra *pap smear*. Arsitektur ResNet50 mendapatkan akurasi 91% sedangkan arsitektur ResNet101 89%. Berdasarkan hasil loss dan akurasi pada kedua model serta grafik yang sudah dipaparkan diatas menunjukkan bahwa kinerja model yang menggunakan hasil ekstraksi fitur gambar dengan ResNet50

(Model-1) lebih baik daripada kinerja model yang menggunakan ekstraksi fitur gambar dengan ResNet101 (Model-2). Hal tersebut ditunjukkan oleh akurasi pada data latih dan data validasi Model-1 yang lebih besar daripada Model-2 dan *loss* pada Model-1 lebih kecil daripada *loss* Model-2.

#### DAFTAR PUSTAKA

- [1] Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I and F., "Global Cancer Observatory: Cancer Today.," International Agency for Research on Cancer. 2018.
- [2] Rahaman Md, Li C, Wu X, Yao Y, Hu Z, Jiang T, Li X, "A Survey for Cervical Cytopathology Image Analysis Using Deep Learning.," IEEE Acces, 2020, doi:10.1109/ACCESS.2020.2983186.
- [3] Akbar H, Sandfreni. "Klasifikasi Kanker Serviks Menggunakan Model Convolutional Neural Network (Alexnet).," *J. Inform. dan Komputer.*, 2021, doi: 10.33387/jiko.
- [4] *Badan Penyelenggara Jaminan Sosial Kesehatan Republik Indonesia*. 2021.( <https://bpjs-kesehatan.go.id/bpjs/index.php/post/read/2014/262/Jangan-Khawatir-BPJS-Kesehatan-Menjamin-Deteksi-Sebelum-Kanker-Serviks-Menyerang/berita>)
- [5] Kurniawan R, Sasmito Kartikaning DE, Suryani F "Klasifikasi Sel Serviks Menggunakan Analisis Fitur Nuclei pada Citra Pap Smear.," 2013. *Seminar Nasional Informatika Medis (SNIMed) IV*.
- [6] Plissiti M, Dimitrakopoulos P, Sfikas G, Nikou C, Krikoni O, "Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images.," pp. 3144-3148., 2018, doi: 10.1109/ICIP.2018.8451588.
- [7] Q. Liao, Y. Ding, Z. L. Jiang, X. Wang, C. Zhang, and Q. Zhang, "Multi-task deep convolutional neural network for cancer diagnosis," *Neurocomputing*, vol. 348, pp. 66–73, 2019, doi: 10.1016/j.neucom.2018.06.084.
- [8] Agrawal H, Kalantri M, "Comparative Analysis of different Convolutional NeuralNetwork Algorithm for Image Classification.," *ISRAJET.*, vol. 8(9)., 2020.
- [9] Suyanto, *Machine Learning Tingkat Dasar dan Lanjut*. Bandung: Penerbit Informatika., 2018.
- [10] "Convolutional neural network". [https://en.wikipedia.org/wiki/Convolutional\\_neural\\_network](https://en.wikipedia.org/wiki/Convolutional_neural_network)
- [11] He, K., Zhang, X., Ren, S., & Sun, "Deep Residual Learning for Image Recognition Kaiming.," *IEEE*, vol. 32(5), 428, 2015, <https://doi.org/10.1246/cl.2003.428>.
- [12] M. Soh, "Learning CNN-LSTM Architectures for Image Caption Generation. Nips, (c), 1–9.," 2016. <https://cs224d.stanford.edu/reports/msoh.pdf>.
- [13] P. Arnav, A., Jang, H., & Maloo, *Image Captioning Using Deep Learning*. 2017.
- [14] A. Primawati, I. Mutia, "Analisis Klasifikasi Populasi Ternak Kambing dan Domba dengan Model Convolutional Neural Network," *Fakt. Exacta*, vol. 14, no. 1, pp. 22–33, 2021, [Online]. Available: [http://journal.lppmunindra.ac.id/index.php/Faktor\\_Exacta](http://journal.lppmunindra.ac.id/index.php/Faktor_Exacta).
- [15] Aneja, J., Deshpande, A., & Schwing, "Convolutional Image Captioning. Proceedings," 2018, <https://doi.org/10.1109/CVPR.2018.00583>.
- [16] "ResNet." [https://pytorch.org/hub/pytorch\\_vision\\_resnet/](https://pytorch.org/hub/pytorch_vision_resnet/).
- [17] S. Das, "CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more," [Online]. Available: <https://medium.com/analytics-vidhya/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5>.
- [18] O'Shea K, Nash R. *An Introduction to Convolutional Neural Networks*. ArXiv e-prints., 2015.
- [19] Arel I, Rose D C, "Deep machine learning-a new frontier in artificial intelligence research [research frontier].," *IEEE Comput. Intell. Mag.*, vol. 5(4): 13–1, 2010.
- [20] "SIPaKMeD Database," [Online]. Available: <https://www.cs.uoi.gr/~marina/sipakmed.html>.