

IMPROVED ATTENTION MOBILEU-NET UNTUK SEGMENTASI CITRA MEDIS POLIP

Abid Ammar Mahdy¹, Abrar Dwi Fairuz Nadhif² and Muhammad Anang Fathur Rohman³

¹Informatika, Fakultas Teknologi Informasi dan Sains Data, Universitas Sebelas Maret, Indonesia.

²Informatika, Fakultas Teknologi Informasi dan Sains Data, Universitas Sebelas Maret, Indonesia.

³Informatika, Fakultas Teknologi Informasi dan Sains Data, Universitas Sebelas Maret, Indonesia.

¹abidammar1611@student.uns.ac.id, ²abrarnadip123@student.uns.ac.id, ³anangmuhammad245@student.uns.ac.id

Abstrak

The development of deep learning in computer vision has led many researchers to propose methods that not only improve the performance of previous methods but also increase the memory efficiency of the models. One interesting problem is the segmentation of medical polyp images, which has been previously addressed with high-performance methods but with large numbers of parameters and relatively low memory efficiency. This paper proposes the Improved Attention MobileU-Net method for medical polyp image segmentation with significantly fewer parameters than previous methods, resulting in higher memory efficiency. This method is constructed with Sandglass Block and uses an attention framework to achieve performance comparable methods. Test result on the CVC-ClinicDB and Kvasir-Seg datasets show that the proposed method has a small number of parameters, with only 574,579 parameters. The proposed method achieved F1-Score performance of 92.36% and Intersection-over-Union (IoU) value of 86.50% for the CVC-ClinicDB dataset. For the Kvasir-Seg dataset, the proposed method showed an F1-Score performance of 86.59% and IoU of 78.87%. The Improved Attention MobileU-Net outperformed several previous methods, including U-Net, ResU-Net, U-Net++, SFA, and ResU-Net++, with a significantly lower number of parameters.

Keyword: Polyp Image Segmentation, Improved Attention MobileU-Net, Memory Efficiency .

Pendahuluan

Medical Image Segmentation (MIS) merupakan sebuah proses untuk memisahkan objek atau fitur yang diinginkan dari citra medis untuk memudahkan proses diagnosis dan analisis. MIS merupakan hal yang sangat penting karena hasil segmentasi dapat digunakan untuk perawatan pasien, riset klinis, pengembangan teknologi, dan tindak lanjut untuk evaluasi secara berkala terkait kondisi kesehatan pasien (Liu et al., 2020). Hasil segmentasi menjadi hal yang penting bagi dokter untuk memahami sebuah penyakit dan menentukan efektivitas pengobatan. Hal ini memungkinkan dokter lebih memahami pasien dengan representasi citra dan fungsional dari struktur fisiologi yang tersembunyi dari

bagian tubuh seperti tulang, organ, jaringan, dan pembuluh darah untuk pemeriksaan klinis (Al-Masni et al., 2021).

MIS merupakan proses membedakan bagian tertentu pada citra medis 2 dimensi atau 3 dimensi, yang dapat digunakan dokter untuk mempelajari bagian yang diinginkan dari citra medis multi-modal (Ibte haz & Rahman, 2020). MIS merupakan langkah awal yang penting pada sistem *Computer Aided Diagnosis* (CAD) yang sering berperan penting dalam analisis kuantitatif dan kualitatif citra medis (Al-Masni et al., 2021), seperti tumor otak (Stoyanov et al., 2018), inti sel (Zeng et al., 2019), segmentasi polip (Mohapatra et al., 2021), wilayah paru-paru (Nishio et al., 2021), batas sel (Wang & Liu, 2021), citra *ultrasound* payudara (Lou et al., 2021), dan pembuluh darah retina (Zhang et al., 2022).

Terdapat cukup banyak metode yang digunakan untuk MIS contohnya *edge detection*, *clustering*, dan metode berbasis model. Metode *edge detection* menggunakan perubahan intensitas piksel pada citra sebagai garis yang membatasi objek. Metode *clustering* mengelompokkan piksel-piksel pada citra ke dalam beberapa kelompok berdasarkan karakteristik piksel. Metode berbasis model, contohnya seperti *machine learning* menggunakan model matematis untuk mengklasifikasikan piksel yang ada pada citra. Metode berbasis *deep learning* yang merupakan pengembangan dari *machine learning* telah berhasil menjadi metode dengan performa yang lebih baik pada MIS. Metode ini menggunakan *neural network* untuk mempelajari karakteristik dan mengklasifikasikan piksel yang ada pada citra. Dalam beberapa dekade terakhir, perkembangan *neural network* untuk berbagai masalah *computer vision*, khususnya MIS didasarkan pada *Deep Neural Networks* (DNNs), seperti *Fully Convolutional Networks* (FCNs) (Long et al. 2015) dan U-Net (Ronneberger et al. 2015) yang menghasilkan performa lebih baik untuk MIS.

Meskipun metode segmentasi citra medis berkembang dengan pesat, akan tetapi ketidakkonsistenan citra medis, seperti kualitas citra, jenis kamera, dan kondisi pasien menjadi masalah dalam segmentasi citra medis. Oleh karena itu, penelitian lebih lanjut diperlukan untuk mengembangkan metode yang lebih baik. Dalam karya tulis ini diajukan sebuah metode segmentasi citra medis yang bernama *Improved Attention MobileU-Net*. Metode yang diusulkan akan dilatih dan diuji menggunakan beberapa *dataset* sehingga metode yang diajukan dapat digunakan untuk menangani beberapa masalah MIS.

Improved Attention MobileU-Net menggunakan *improved sandglass block* sebagai block penyusunnya. *Improved sandglass block* merupakan hasil penggabungan dari konsep *sandglass block* (Shao et al., 2020) dan *improved residual network* (Duta et al., 2021). *Sandglass block* (Shao et al., 2020) adalah jenis *bottleneck* yang memiliki performa paling bagus dan efisiensi yang paling tinggi dibanding jenis *bottleneck* yang sering digunakan untuk membuat model yang ringan, yaitu *inverted residual block* (Sandler et al., 2018). *Improved residual network* (Duta et al., 2021) adalah cara baru untuk menyusun *convolution*, *batch normalization*, dan ReLU untuk meningkatkan kinerja dari model. *Improved sandglass block* juga difasilitasi oleh *channel attention* yang terinspirasi oleh *squeeze-excitation block* (Hu et al. 2018). *Channel attention* berfungsi untuk menentukan mana *feature map* yang berisi informasi yang penting dan yang tidak penting (Hu et al. 2018). Hal tersebut membantu model untuk konvergen menuju nilai *global minimum* yang lebih dalam (Hu et al. 2018).

Metode Penelitian (atau Penulisan)

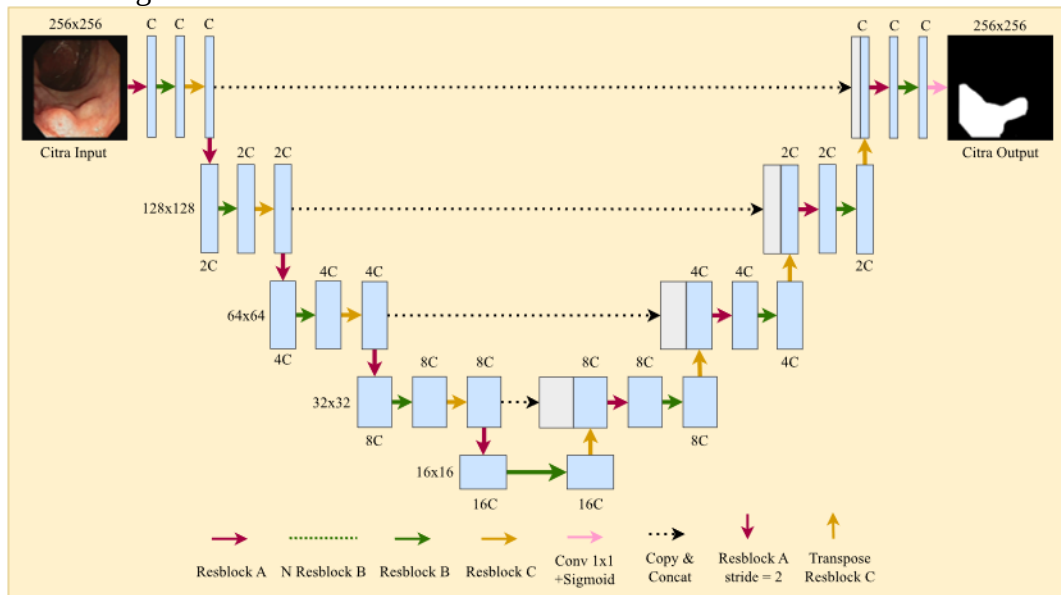
Metodologi penelitian yang diterapkan pada penelitian ini terdiri atas enam fase. Pertama adalah mengumpulkan dan menata dataset, kedua yaitu merancang arsitektur metode yang diajukan, ketiga adalah mengatur hyperparameter, keempat merupakan training metode yang diajukan, dan terakhir adalah melakukan perbandingan dengan metode sebelumnya.

A. Mengumpulkan dan Menata Dataset

Dataset yang digunakan pada riset kali ini adalah citra polip pada dataset CVC-ClinicDB (Bernal et al., 2015) dan Kvasir-Seg (Jha et al., 2020). Dataset CVC-ClinicDB dan Kvasir-Seg secara berturut-turut dapat diakses melalui laman endoscopic vision challenge (<https://polyp.grand-challenge.org/CVCClinicDB/>) dan <https://datasets.simula.no/kvasir-seg/>. Dataset CVC-ClinicDB mengandung 612 citra polip dengan resolusi 288×368 piksel. Sementara itu, Kvasir-Seg memuat 1000 citra polip berdimensi dari 332×487 hingga 1920×1072 piksel. Kedua dataset didistribusikan ke dalam training set, validation set, dan test set dengan rasio secara berturut-turut 80%:10%:10%.

Citra pada dataset memiliki ukuran yang bervariasi. Untuk menyamakan dimensi pada citra, dilakukan resize menjadi ukuran 256×256 . Agar jumlah data citra menjadi lebih banyak dan menghindari potensi terjadinya overfit pada model, augmentasi data diimplementasikan pada training set. Augmentasi yang dilakukan adalah horizontal dan vertical flip, serta rotasi. Total data pada training set setelah augmentasi untuk dataset CVC-ClinicDB dan Kvasir-Seg secara berturut-turut sebesar 1960 dan 3200.

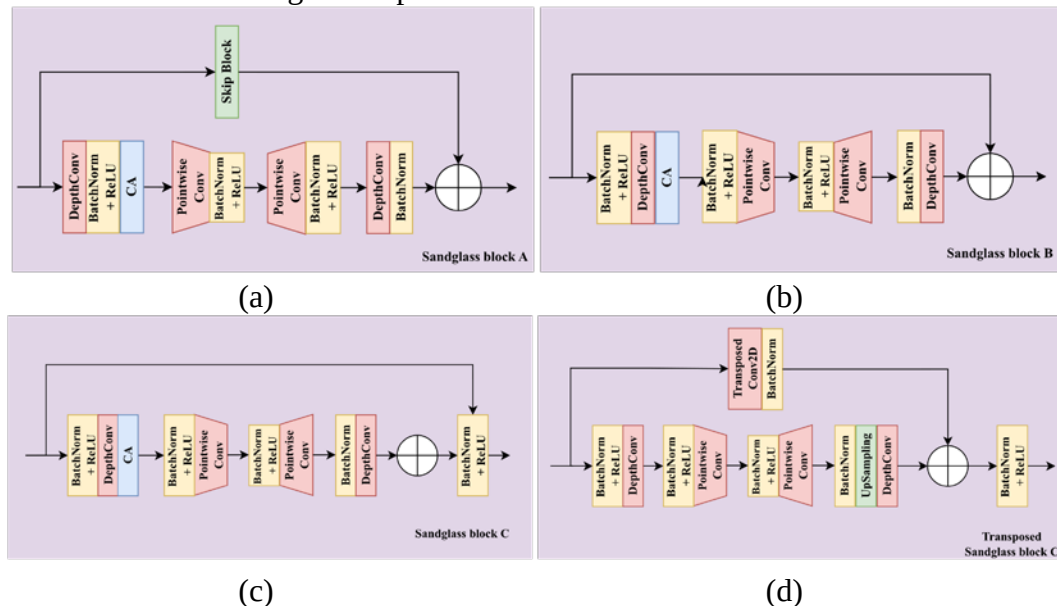
B. Merancang Arsitektur



Gambar 1 Arsitektur Improved Attention MobileU-Net

Metode yang diusulkan adalah Improved Attention MobileU-Net yang menjadikan U-Net sebagai motivasi dalam perancangan arsitektur tersebut. Layaknya U-Net, Improved Attention MobileU-Net tersusun atas encoder, decoder, dan bridge. Level kedalaman pada metode yang diajukan adalah empat, sama seperti pada U-Net. Untuk besaran filter pada setiap tingkat mengikuti teknik yang ada di U-Net, yaitu setiap tingkat lebih besar dua kali lipat dibanding tingkat sebelumnya. Terdapat skip connection yang menyatukan output feature

map hasil upsampling dengan encoder yang setingkat. Demi mencapai performa yang maksimal dengan jumlah parameter yang minimal, dilakukan beberapa modifikasi pada Improved Attention MobileU-Net. Karena fokus penelitian ini adalah mengurangi jumlah parameter, maka diimplementasikan sandglass block sebagai block penyusun Improved Attention MobileU-Net. Model dengan jumlah parameter kecil lazim jika menghasilkan kinerja yang tidak seoptimal model dengan jumlah parameter besar. Mengacu pada persoalan tersebut, diaplikasikan konsep penyusunan block seperti improved ResNet dan ditambahkan channel attention untuk meningkatkan performa model.

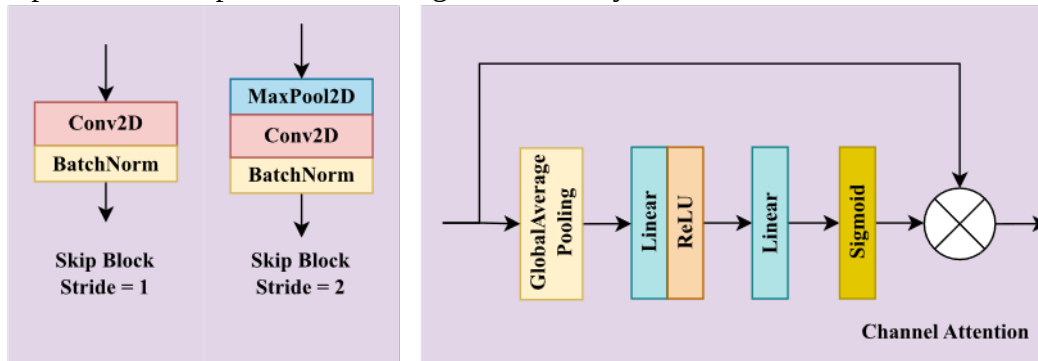


Gambar 2 Improved Sandglass Block. (a) Sandglass Block A, (b) Sandglass Block B, (c) Sandglass Block C, dan (d) Transposed Sandglass Block C

Pembaruan yang dilakukan adalah mengganti semua convolution dengan improved sandglass block pada setiap tingkat. Alih-alih menggunakan dua convolution lalu diikuti dengan pooling layer pada encoder, Improved Attention MobileU-Net menerapkan sandglass block A, B, dan C pada setiap tingkat dengan sandglass block A bertugas untuk mengurangi dimensi feature map. Pengurangan ukuran feature map bisa terjadi ketika nilai pergeseran (stride) bernilai dua di bagian depthwise convolution terakhir. Perbedaan utama dari ketiga sandglass block terlihat pada posisi urutan dari convolution, batch normalization, dan ReLU. Sandglass block A memiliki susunan yang sama dengan versi orisinal. Sementara itu, sandglass block B dan C memiliki susunan yang mirip dengan pre-activation ResNet. Sandglass block C mendapatkan tambahan batch normalization dan ReLU di akhir untuk menstabilkan nilai dari output feature map sebagai persiapan ke tingkat selanjutnya.

Bagian decoder difasilitasi oleh sandglass block C yang dilengkapi upsampling sebelum depthwise convolution terakhir dan projection shortcut yang menggunakan transposed convolution. Pada setiap sandglass block, diletakkan channel attention setelah depthwise convolution awal kecuali pada sandglass block A dan sandglass block C yang bertugas untuk mengubah ukuran feature map. Struktur dari channel attention terinspirasi dari squeeze-and-excitation network. Skip block pada sandglass block A akan ditambahkan ketika stride bernilai dua dan jumlah channel input feature map tidak sama dengan output. Hal

tersebut dilakukan agar shortcut connection tetap bisa dilakukan ketika dimensi input feature map tidak sama dengan keluarannya.



Gambar 3 Kiri: Skip block. Kanan: Channel attention

C. Mengatur Hyperparameter

Langkah selanjutnya adalah mengatur hyperparameter. Hyperparameter adalah nilai yang diatur sebelum dilakukan proses training pada model. Nilai-nilai tersebut akan menentukan bagaimana model akan mempelajari data sehingga mempengaruhi kinerja pada model. Beberapa hyperparameter yang akan dilakukan uji coba antara lain: posisi channel attention, rasio pada channel attention, jumlah sandglass block B pada bridge, dan loss function yang tepat untuk digunakan.

D. Training Model

Setelah merancang arsitektur dengan beberapa hyperparameter yang telah disetel, selanjutnya adalah melatih model atau training model. Supaya proses training bisa dilaksanakan, dibutuhkan citra input, citra label, dan model. Citra input yang dimasukkan ke model adalah citra polip berwarna dengan tiga channel yang diambil melalui endoskopi. Model akan menghasilkan sebuah citra tersegmentasi dengan satu channel. Hasil prediksi tersebut akan dicocokkan dengan citra label untuk dikirimkan ke loss function dan mengetahui seberapa baik kinerja model.

E. Melakukan Perbandingan dengan Metode Sebelumnya

Model yang telah melewati proses training dan uji coba akan diuji performanya oleh test set. Evaluasi seberapa baik kinerja model berdasarkan nilai F1-Score, Intersection-over-Union (IoU), dan jumlah parameter. Rumus dari F1-Score dan IoU bisa dilihat secara berturut-turut pada persamaan (1) dan (2).

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \quad (1)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (2)$$

Di mana TP, FP, dan FN secara berturut-turut menyatakan True Positive, False Positive, dan False Negative. True Positive adalah menyatakan seberapa banyak jumlah piksel yang termasuk polip teridentifikasi dengan benar oleh model. False Positive menunjukkan berapa jumlah piksel yang diprediksi polip oleh model, tetapi tidak termasuk polip pada citra label. False Negative merupakan total piksel yang tidak teridentifikasi polip oleh model.

Semakin tinggi nilai dari F1-Score dan IoU, kian membaik pula kinerja pada model. Jumlah parameter yang semakin rendah menandakan efisiensi memori

yang semakin baik. Hasil F1-Score, IoU, dan jumlah parameter akan dibandingkan dengan U-Net (Ronneberger et al., 2015), U-Net++ (Zhoud et al., 2019), ResU-Net (Zhang et al., 2018), ResU-Net++ (Jha et al., 2019), dan SFA (Fang et al., 2019)

Hasil dan Pembahasan

A. Situasi Percobaan

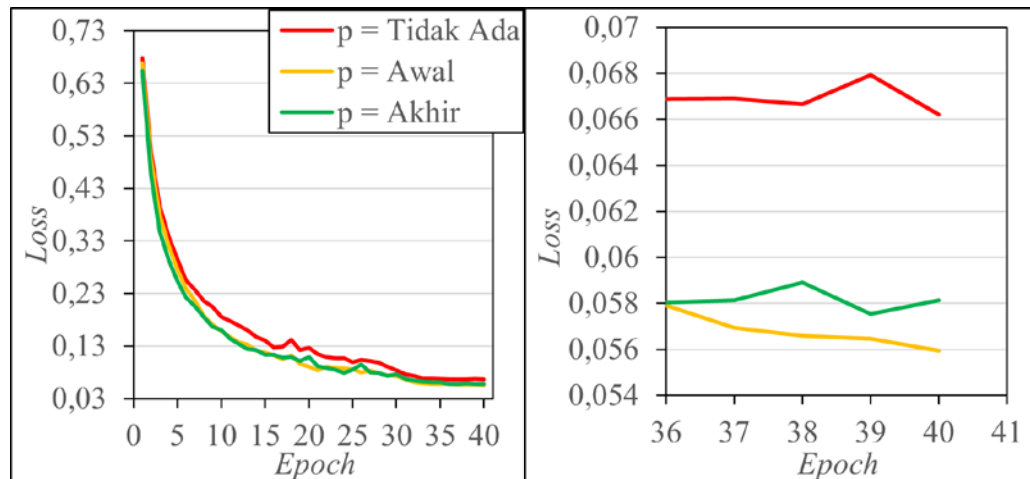
Percobaan ini dilakukan pada *website Kaggle* dengan akselerator TPU v3-8 menggunakan bahasa pemrograman *python* dengan *framework Tensorflow*. Percobaan dilakukan dengan *training set* dari *dataset CVC-ClinicDB* menggunakan *hyperparameter* pada Tabel 1. Nilai dari posisi *attention*, rasio *attention*, jumlah *residual block*, dan *loss function* akan berubah seiring percobaan.

Tabel 1. *Hyperparameter*

<i>Hyperparameter</i>	Nilai
Ukuran <i>Batch</i>	16
Jumlah <i>Epoch</i>	40
<i>Optimizer</i>	<i>Adam</i>
<i>Learning Rate</i>	0.001 (awal) (dibagi 10 setelah <i>epoch</i> 30 setiap 5 <i>epoch</i>)
<i>Activation Function</i>	ReLU
Posisi <i>Attention</i>	Awal
Rasio <i>Attention</i>	2
Jumlah <i>Residual Block</i>	1
<i>Loss Function</i>	<i>Dice Loss</i>

B. Percobaan Posisi *Attention*

Percobaan pertama memiliki tujuan untuk mengetahui bagaimana pengaruh dari posisi *attention* pada *residual block* terhadap performa *neural network* yang diajukan. Percobaan dilakukan dengan nilai posisi *attention* yaitu tidak ada, awal, dan akhir. Hasil percobaan posisi *attention* dapat dilihat pada Gambar 4.

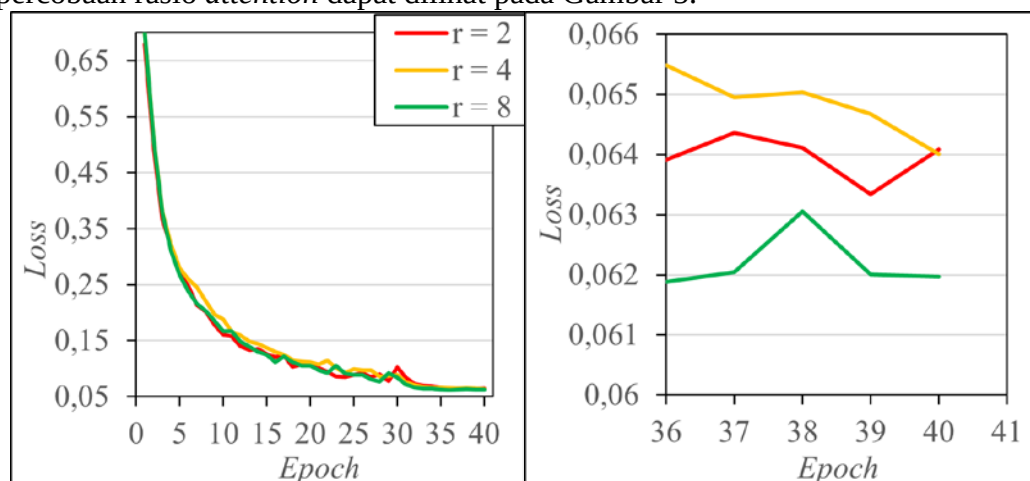


Gambar 4 Hasil Percobaan Posisi Attention (p). Kiri: Semua Epoch. Kanan: 5 Epoch Terakhir

Dari hasil percobaan posisi *attention* pada Gambar 4, posisi *attention* berpengaruh terhadap nilai *loss* dari *neural network* yang diajukan. Setiap posisi *attention* pada *residual block* menghasilkan nilai *loss* yang berbeda-beda. Penggunaan *attention* pada awal *residual block* menghasilkan *neural network* dengan nilai *loss* paling kecil sehingga pada percobaan selanjutnya akan menggunakan posisi *attention* pada awal *residual block*.

C. Percobaan Rasio Attention

Percobaan kedua memiliki tujuan untuk mengetahui bagaimana pengaruh dari rasio *attention* pada *residual block* terhadap performa *neural network* yang diajukan. Percobaan dilakukan dengan nilai rasio *attention* yaitu 2, 4, dan 8. Hasil percobaan rasio *attention* dapat dilihat pada Gambar 5.

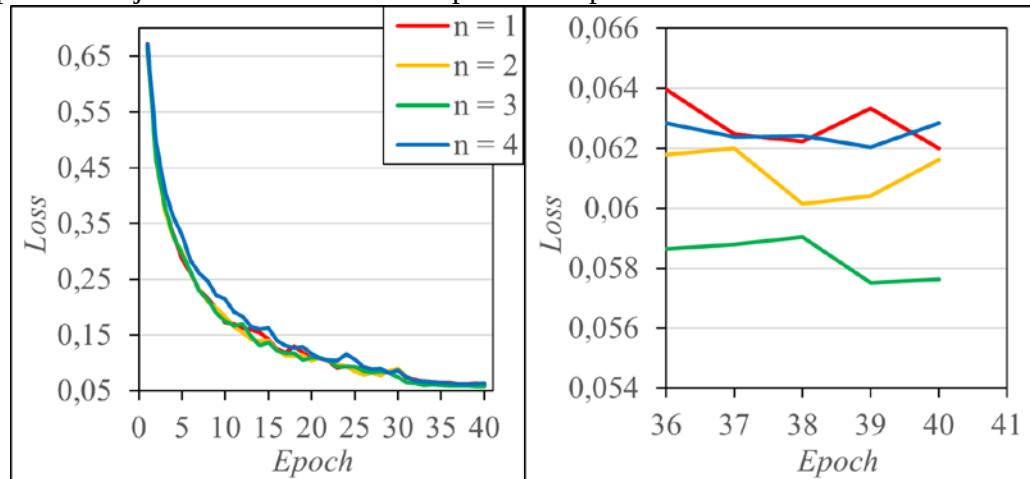


Gambar 5 Hasil Percobaan Rasio Attention (r). Kiri: Semua Epoch. Kanan: 5 Epoch Terakhir

Dari hasil percobaan rasio *attention* pada Gambar 5, rasio *attention* berpengaruh terhadap nilai *loss* dari *neural network* yang diajukan. Setiap rasio *attention* pada *residual block* menghasilkan nilai *loss* yang berbeda-beda. Penggunaan rasio *attention* pada *residual block* sebesar 8 menghasilkan *neural network* dengan nilai *loss* paling kecil sehingga pada percobaan selanjutnya akan menggunakan rasio *attention* pada *residual block* sebesar 8.

D. Percobaan Jumlah Residual Block

Percobaan ketiga memiliki tujuan untuk mengetahui bagaimana pengaruh dari jumlah *residual block* terhadap performa *neural network* yang diajukan. Percobaan dilakukan dengan jumlah *residual block* yaitu 1, 2, 3, dan 4. Hasil percobaan jumlah *residual block* dapat dilihat pada Gambar 6.

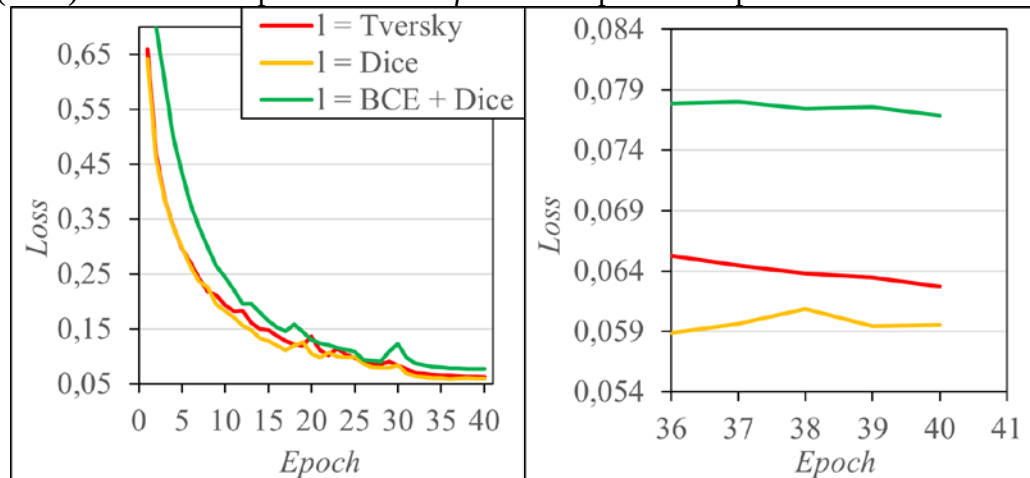


Gambar 6 Hasil Percobaan Jumlah *Residual Block* (n). Kiri: Semua *Epoch*. Kanan: 5 *Epoch* Terakhir.

Dari hasil percobaan jumlah *residual block* pada Gambar 6, jumlah *residual block* berpengaruh terhadap nilai *loss* dari *neural network* yang diajukan. Setiap jumlah *residual block* menghasilkan nilai *loss* yang berbeda-beda. Penggunaan *residual block* sejumlah 3 menghasilkan *neural network* dengan nilai *loss* paling kecil sehingga pada percobaan selanjutnya akan menggunakan *residual block* sejumlah 3.

E. Percobaan *Loss Function*

Percobaan keempat memiliki tujuan untuk mengetahui bagaimana pengaruh dari *loss function* terhadap performa *neural network* yang diajukan. Percobaan dilakukan dengan *loss function* yaitu *tversky*, *dice*, dan *binary cross-entropy* (BCE) + *dice*. Hasil percobaan *loss function* dapat dilihat pada Gambar 7.



Gambar 7 Hasil Percobaan *Loss Function* (l). Kiri: Semua *Epoch*. Kanan: 5 *Epoch* Terakhir

Dari hasil percobaan *loss function* pada Gambar 7, *loss function* berpengaruh terhadap nilai *loss* dari *neural network* yang diajukan. Setiap *loss function* menghasilkan nilai *loss* yang berbeda-beda. Penggunaan *loss function* *dice*

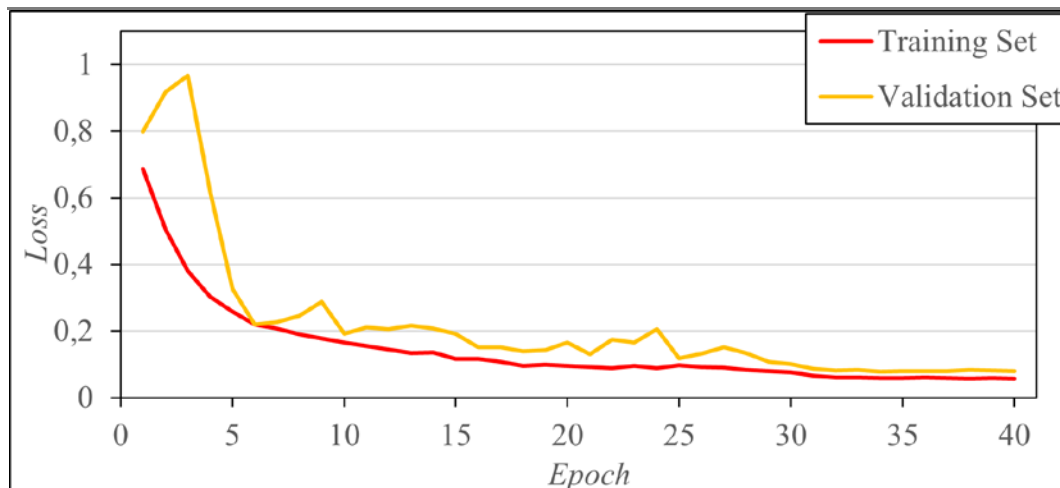
menghasilkan *neural network* dengan nilai *loss* paling kecil sehingga pada percobaan selanjutnya akan menggunakan *loss function dice*.

F. Hasil Percobaan

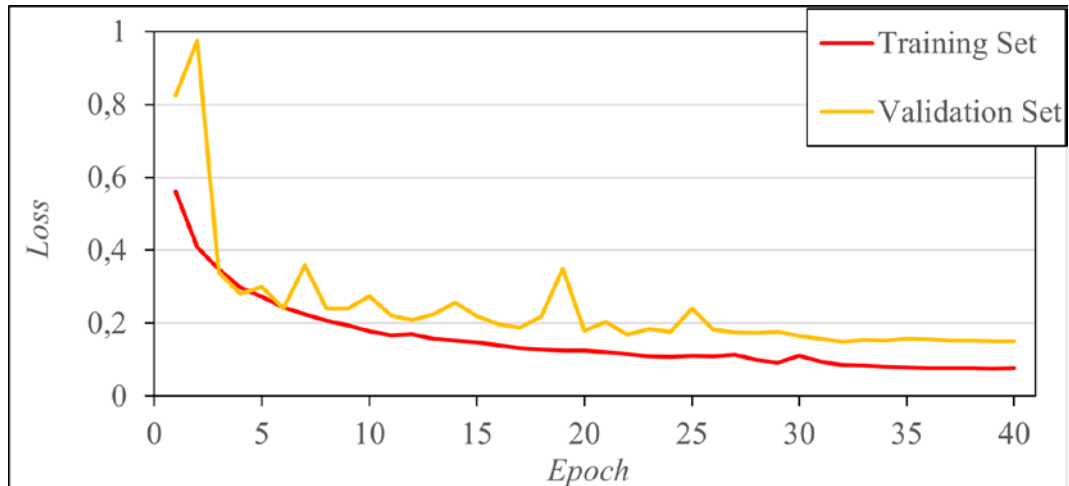
Setelah semua percobaan dilakukan, beberapa nilai *hyperparameter* berubah untuk menghasilkan performa terbaik dari *neural network* yang diajukan. *Hyperparameter* yang berubah yaitu rasio *attention* dan jumlah *residual block*. *Hyperparameter* dengan performa terbaik yang dijabarkan pada Tabel 2 menghasilkan *neural network* dengan parameter sejumlah 574.579 yang akan dilatih menggunakan *training set* dan *validation set* dari dataset *CVC-ClinicDB* dan dataset *Kvasir-Seg*.

Tabel 2. *Hyperparameter* Performa Terbaik

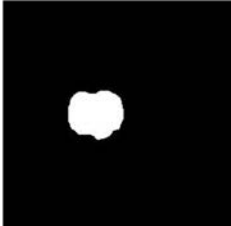
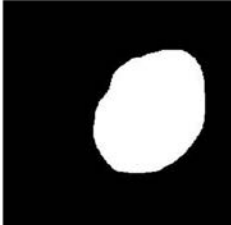
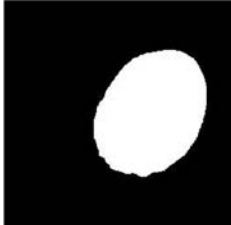
<i>Hyperparameter</i>	Nilai
Ukuran <i>Batch</i>	16
Jumlah <i>Epoch</i>	40
<i>Optimizer</i>	<i>Adam</i>
<i>Learning Rate</i>	0.001 (awal) (dibagi 10 setelah <i>epoch</i> 30 setiap 5 <i>epoch</i>)
<i>Activation Function</i>	ReLU
Posisi <i>Attention</i>	Awal
Rasio <i>Attention</i>	8
Jumlah <i>Residual Block</i>	3
<i>Loss Function</i>	<i>Dice Loss</i>



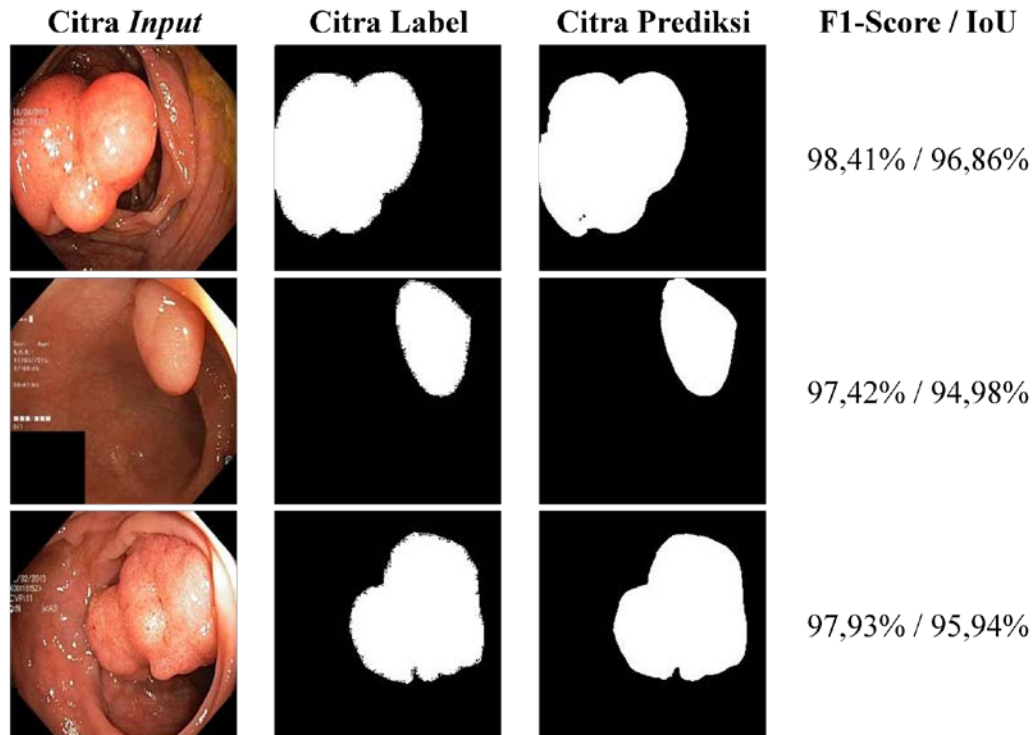
Gambar 8 Hasil Pelatihan Dataset *CVC-ClinicDB*

Gambar 9 Hasil Pelatihan *Dataset Kvasir-Seg*

Gambar 8 menunjukkan hasil pelatihan menggunakan *dataset CVC-ClinicDB*, sedangkan Gambar 9 menunjukkan hasil pelatihan menggunakan *dataset Kvasir-Seg*. *Neural network* yang telah dilatih kemudian digunakan untuk melakukan pengujian menggunakan *test set* dari *dataset CVC-ClinicDB* dan *dataset Kvasir-Seg* untuk mendapatkan nilai *F1-Score* dan *IoU*. Selain mendapatkan *F1-Score* dan *IoU*, dari pengujian juga didapatkan citra prediksi dari masing-masing *dataset*. Contoh citra pengujian dapat dilihat pada Gambar 10 dan Gambar 11. Hasil pengujian dijabarkan pada Tabel 3.

Citra Input	Citra Label	Citra Prediksi	F1-Score / IoU
			97,25% / 94,64%
			97,20% / 94,56%
			98,24% / 96,53%

Gambar 10 Citra Hasil Pengujian *Dataset CVC-ClinicDB*



Gambar 11 Citra Hasil Pengujian Dataset Kvasir-Seg

Tabel 3. Hasil Pengujian

Dataset	F1-Score	IoU
CVC-ClinicDB	92,36%	86,50%
Kvasir-Seg	86,59%	78,87%

G. Perbandingan dengan Metode Sebelumnya

F1-Score dan *IoU* hasil pengujian digunakan untuk perbandingan dengan metode sebelumnya untuk melihat perbandingan performa dengan metode sebelumnya. Performa dari *Improved Attention MobileU-Net* akan dibandingkan dengan metode U-Net, ResU-Net, U-Net++, SFA, ResU-Net++. Perbandingan dengan metode sebelumnya dapat dilihat pada tabel 4.

Tabel 4. Perbandingan dengan Metode Sebelumnya

Metode	CVC-ClinicDB		Kvasir-Seg		Parameter
	<i>F1-Score</i>	<i>IoU</i>	<i>F1-Score</i>	<i>IoU</i>	
U-Net	82,3%	75,5%	81,8%	74,6%	43,93 juta
ResU-Net	77,9%	n/a	79,1%	n/a	8,06 juta
U-Net++	79,4%	72,9%	82,1%	74,3%	9,04 juta
SFA	70,0%	60,7%	72,3%	61,1%	n/a
ResU-Net++	79,6%	79,6%	81,3%	79,3%	4,06 juta
Metode yang diusulkan	92,36%	86,50%	86,59%	78,87%	0,58 juta

(Improved
Attention
MobileU-
Net)

Kesimpulan

Pada penelitian ini, diajukan sebuah arsitektur *convolutional neural network* dengan kinerja yang baik dan efisiensi memori yang tinggi, yaitu *Improved Attention MobileU-Net*. Jumlah parameter *Improved Attention MobileU-Net* relatif kecil karena menggunakan *sandglass block* sebagai penyusunnya. Demi mendorong performa model menuju titik optimalnya, diterapkan susunan *improved ResNet* dan *channel attention* pada *sandglass block*. Penggabungan tersebut menghasilkan *improved sandglass block* yang membuat metode yang diajukan mampu melampaui kinerja metode sebelumnya. *F1-Score* dan *IoU* yang dihasilkan oleh metode yang diajukan pada *dataset CVC-ClinicDB* secara berturut-turut sebesar 92,36% dan 86,50%. Sementara itu, untuk *dataset Kvasir-Seg* memberikan *F1-Score* dengan nilai 86,59% dan 78,87% untuk *IoU*. Jumlah parameter yang kecil, yaitu 0,58 juta, *Improved Attention MobileU-Net* menjadi model dengan efisiensi memori paling tinggi dibandingkan dengan metode sebelumnya.

Daftar Pustaka

- Al-Masni, M.A. and Kim, D.H., 2021. CMM-Net: Contextual multi-scale multi-level network for efficient biomedical image segmentation. *Scientific reports*, 11(1), pp.1-18.
- Duta, I.C., Liu, L., Zhu, F. and Shao, L., 2021, January. Improved residual networks for image and video recognition. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 9415-9422). IEEE.
- Fang, Y., Chen, C., Yuan, Y. and Tong, K.Y., 2019. Selective feature aggregation network with area-boundary constraints for polyp segmentation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22* (pp. 302-310). Springer International Publishing.
- Hu, J., Shen, L. and Sun, G., 2018. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).
- Ibtehaz, N. and Rahman, M.S., 2020. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural networks*, 121, pp.74-87.
- Jha, D., Smedsrud, P.H., Riegler, M.A., Johansen, D., De Lange, T., Halvorsen, P. and Johansen, H.D., 2019, December. Resunet++: An advanced architecture for medical image segmentation. In *2019 IEEE International Symposium on Multimedia (ISM)* (pp. 225-2255). IEEE.
- Liu, L., Cheng, J., Quan, Q., Wu, F.X., Wang, Y.P. and Wang, J., 2020. A survey on U-shaped networks in medical image segmentations. *Neurocomputing*, 409, pp.244-258.

- Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).
- Lou, A., Guan, S. and Loew, M., 2023. Cfpnet-m: A light-weight encoder-decoder based network for multimodal biomedical image real-time segmentation. *Computers in Biology and Medicine*, 154, p.106579.
- Mohapatra, S., Swarnkar, T. and Das, J., 2021. Deep convolutional neural network in medical image processing. In Handbook of deep learning in biomedical engineering (pp. 25-60). Academic Press.
- Nishio, M., Fujimoto, K. and Togashi, K., 2021. Lung segmentation on chest X-ray images in patients with severe abnormal findings using deep learning. *International Journal of Imaging Systems and Technology*, 31(2), pp.1002-1008.
- Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520).
- Stoyanov, D., Taylor, Z., Carneiro, G., Syeda-Mahmood, T., Martel, A., Maier-Hein, L., Tavares, J.M.R., Bradley, A., Papa, J.P., Belagiannis, V. and Nascimento, J.C. eds., 2018. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings (Vol. 11045). Springer.
- Wang, D. and Liu, Y., 2021. An Improved Neural Network Based on UNet for Surface Defect Segmentation. In 3D Imaging Technologies—Multidimensional Signal Processing and Deep Learning: Methods, Algorithms and Applications, Volume 2 (pp. 27-33). Springer Singapore.
- Zeng, Z., Xie, W., Zhang, Y. and Lu, Y., 2019. RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images. *Ieee Access*, 7, pp.21420-21428.
- Zhang, Y., He, M., Chen, Z., Hu, K., Li, X. and Gao, X., 2022. Bridge-Net: Context-involved U-net with patch-based loss weight mapping for retinal blood vessel segmentation. *Expert Systems with Applications*, 195, p.116526.
- Zhang, Z., Liu, Q. and Wang, Y., 2018. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), pp.749-753.
- Zhou, D., Hou, Q., Chen, Y., Feng, J. and Yan, S., 2020. Rethinking bottleneck structure for efficient mobile network design. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16 (pp. 680-697). Springer International Publishing.
- Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N. and Liang, J., 2018. Unet++: A nested u-net architecture for medical image segmentation. In *Deep*

Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4 (pp. 3-11). Springer International Publishing.