

ANALISIS EFEK AUGMENTASI DATASET DAN FINE TUNE PADA ALGORITMA PRE-TRAINED CONVOLUTIONAL NEURAL NETWORK (CNN)

Theopilus Bayu Sasongko^{*1}, Haryoko², Agit Amrullah³

^{1,2,3}Universitas Amikom Yogyakarta, Yogyakarta

Email: ¹theopilus.27@amikom.ac.id, ²haryoko@amikom.ac.id, ³agit@amikom.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 30 September 2023, diterima untuk diterbitkan: 25 Juli 2023)

Abstrak

Saat ini Teknologi deep learning seringkali dikaitkan dengan proses (metode) yang dapat diandalkan Ketika kita memiliki jumlah data yang besar. Dalam *deep learning* CNN (*Convolutional Neural Network*) memainkan peranan yang sangat penting yang sering digunakan untuk menganalisis (klasifikasi ataupun rekognisi) gambar visual. Pada era *deep learning* model seperti Convolutional Neural Network (CNN) yang kompleks seperti saat ini memiliki tantangan-tantangan yang baru seperti gradient vanishing, *overfitting* yang dikarenakan keterbatasan dataset, optimasi parameter hingga keterbatasan perangkat keras. Arsitektur MobileNet dicetuskan pada tahun 2017 oleh Howards, dkk yang merupakan salah satu arsitektur *convolutional neural network* (CNN) yang dapat digunakan untuk mengatasi kebutuhan akan *computing resource* berlebih. Penelitian ini bertujuan mengukur pengaruh teknik fine tune dan augmentasi dataset pada model transfer learning CNN mobilenet, efficientnet, dan nasnetmobile dengan dataset yang sangat kecil. Hasil dari penelitian ini adalah Dari ketiga dataset yang digunakan sebagai dasar pelatihan (*training*) pada model efisien transfer learning (mobilenet, efficientnet, dan nasnetmobile), Teknik augmentasi random erase dan zoom range mendominasi peningkatan akurasi model. Jumlah peningkatan akurasi setelah dilakukan augmentasi random erase ataupun zoom range yang terjadi sekitar 0.03% hingga 0.1%.

Kata kunci: CNN, Transfer learning, Pre-trained Model, Augmentasi dataset, Fine Tuned

EFFECT ANALYSIS OF DATASET AUGMENTATION AND FINE TUNE ON PRE-TRAINED CONVOLUTIONAL NEURAL NETWORK (CNN) ALGORITHM

Abstract

Today deep learning technology is often associated with reliable processes (methods) when we have large amounts of data. In deep learning CNN (Convolutional Neural Network) plays a very important role which is often used to analyze (classify or recognize) visual images. In the era of deep learning models such as the complex Convolutional Neural Network (CNN) as it is today, it has new challenges such as gradient vanishing, *overfitting* due to dataset limitations, parameter optimization to hardware limitations. The MobileNet architecture was coined in 2017 by Howards, et al, which is one of the convolutional neural networks (CNN) architectures that can be used to overcome the need for excessive computing resources. This study aims to measure the effect of fine tune and dataset augmentation techniques on CNN mobilenet, efficientnet, and nasnetmobile transfer learning models with very small datasets. The results of this study are that of the three datasets used as the basis for training in efficient transfer learning models (mobilenet, efficientnet, and nasnetmobile), random erase and zoom range augmentation techniques dominate the increase in model accuracy. The amount of increase in accuracy after random erase or zoom range augmentation that occurs is about 0.03% to 0.1%.

Keywords: CNN, Transfer learning, Pre-trained Model, Dataset augmentation, Fine Tuned

1. PENDAHULUAN

Saat ini Teknologi deep learning seringkali dikaitkan dengan proses (metode) yang dapat diandalkan Ketika kita memiliki jumlah data yang besar. Dalam *deep learning* CNN (Convolutional Neural Network) memainkan peranan yang sangat penting yang sering digunakan untuk menganalisis (klasifikasi ataupun

rekognisi) gambar visual. CNN diciptakan pada tahun 1980-an (Fukushima, 1980), berkat popularitas GPU (Pan & Yang, 2010) merangkul terobosan pada tahun 2000. Pada tahun 2011 CNN berbasis GPU memperoleh kinerja yang tinggi dengan tingkat error hanya 0.35%, 2.53%, dan 19.51% pada rekognisi digit, disusul pengenalan objek 3D (NORB) dan natural images (CIFAR10). Pada era deep model

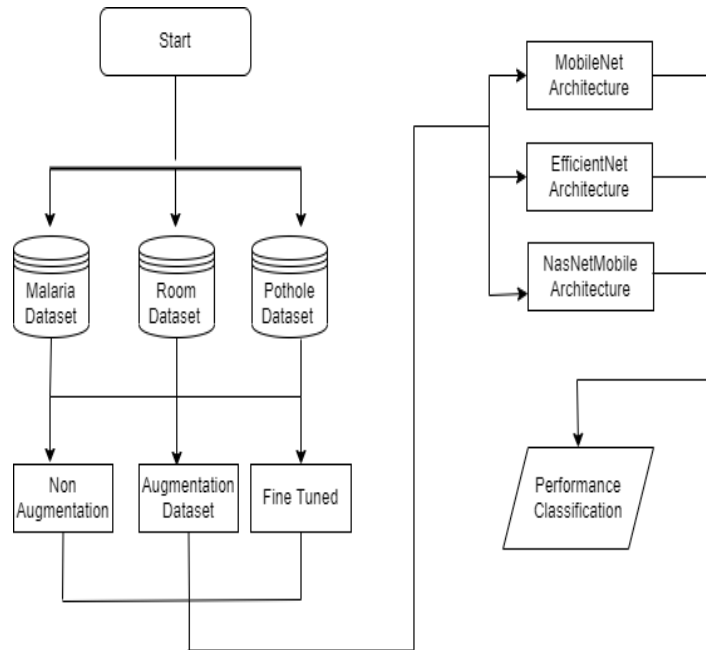
seperti Convolutional Neural Network (CNN) yang kompleks seperti saat ini memiliki tantangan-tantangan yang baru seperti gradient vanishing, overfitting, optimasi parameter hingga keterbatasan perangkat keras. Arsitektur MobileNet dicetuskan pada tahun 2017 oleh Howards, dkk yang merupakan salah satu arsitektur convolutional neural network (CNN) yang dapat digunakan untuk mengatasi kebutuhan akan computing resource berlebih (Howard et al., 2017). Seperti namanya, Mobile, para peneliti dari Google membuat arsitektur CNN yang dapat digunakan untuk ponsel. Arsitektur MobileNets, merupakan salah satu arsitektur convolutional neural network (CNN) yang dapat digunakan untuk mengatasi kebutuhan akan computing resource berlebih. Seperti namanya, Mobile, para peneliti dari Google membuat arsitektur CNN yang dapat digunakan untuk ponsel. EfficientNet merupakan salah satu pre-trained model CNN untuk transfer learning untuk masalah klasifikasi gambar. Model ini dikembangkan oleh Google AI pada tahun 2019 dan bersifat open source. Model EfficientNet terdiri dari 8 model, yaitu B0 hingga B7 yang menandakan jumlah parameter yang semakin banyak dari 5.3 juta hingga 66 juta parameter dan akurasi yang lebih tinggi (Tan & Le, 2019). Arsitektur EfficientNet berfokus pada efisien waktu dan kemampuan komputasi. EfficientNet telah menunjukkan akurasi yang lebih tinggi dengan jumlah parameter yang lebih sedikit dan lebih cepat dibandingkan beberapa pre-trained model lainnya. NASNetMobile merupakan arsitektur NASNet yang lebih kecil dengan jumlah parameter yang menyerupai MobileNet namun dengan performa akurasi yang lebih baik. Arsitektur NASNet terdiri dari 2 blok utama atau sering disebut cell yakni normal cell dan reduction cell susunan layer yang tepat untuk kedua cell tersebut dicari menggunakan recurrent neural network (YILMAZ & DEMİR, 2020). Meskipun MobileNet, EfficientNetB0, NasNetMobile, dan variants arsitektur CNN lain dicetuskan untuk mengatasi keterbatasan hardware akan tetapi dalam implementasi pembentukan modelnya membutuhkan jumlah dataset yang relatif besar. Pengumpulan dataset yang relatif besar membutuhkan biaya dan effort yang lebih. Hal ini akan menjadi kendala jika ketersediaan dataset relatif sangat kecil. Overfitting adalah sebuah kondisi pada model yang disebabkan keterbatasan dataset sehingga jaringan model CNN tidak dapat mempelajari banyak fitur yang dibutuhkan. Banyak Teknik yang dapat diusulkan untuk mengatasi overfitting diantaranya dengan teknik fine tune dengan dropout maupun augmentasi dataset. Tujuan dari fine tuning adalah melatih pre-trained model pada top-layer untuk mengenali feature dari dataset yang kita miliki sedangkan augmentasi dataset merupakan strategi yang memungkinkan kita untuk secara signifikan meningkatkan keragaman data yang tersedia untuk model pelatihan, tanpa benar-benar mengumpulkan

data baru. Dari latar belakang masalah diatas maka dirumuskan pada penelitian ini yaitu bagaimana menganalisis pengaruh teknik fine tune dan augmentasi dataset pada model transfer learning CNN mobilenet, efficientnet, dan nasnetmobile dengan dataset yang sangat kecil.

2. METODE PENELITIAN

Pada penelitian ini akan dilakukan pengukuran pengaruh teknik fine tune dan augmentasi dataset pada model transfer learning CNN yang memiliki parameter yang relatif kecil (efficient transfer learning) seperti mobilenet, efficientnet, dan nasnetmobile dengan dataset yang sangat kecil. Digunakan tiga dataset binary yang terkelompok menjadi tiga bagian data (kecil, kecil sekali, dan sangat kecil sekali). Ketiga dataset diambil dari sumber public Kaggle dataset (<https://www.kaggle.com/datasets>). Tiga dataset itu yaitu dataset malaria small (<https://www.kaggle.com/datasets/gayathryskb/malaria-small-dataset>) yang terdiri dari 500 citra (data training 176 berlabel parasite dan 154 berlabel uninfected, data validation 44 berlabel parasite dan 42 berlabel uninfected, data testing berjumlah 91 berlabel parasite dan 43 berlabel uninfected), dataset yang kedua (kecil sekali) yaitu messy vs clean room dataset (<https://www.kaggle.com/datasets/cdawn1/messy-vs-clean-room>) yang terdiri dari 222 data citra (data training 96 berlabel clean dan 96 berlabel messy, data validation 10 berlabel clean dan 10 lain berlabel messy, data testing 5 berlabel clean dan 5 lainnya berlabel messy), dataset yang ketiga (sangat kecil sekali) yaitu pothole small dataset (<https://www.kaggle.com/datasets/virenbr11/pothole-detection-small>) yang terdiri dari 56 data citra (data training 25 berlabel pothole dan 15 berlabel plain, data validation 4 berlabel pothole 4 lainnya berlabel plain, data testing 4 berlabel pothole dan 4 lainnya berlabel plain). Seluruh dataset yang ada diambil dari Kaggle dataset. Bagan penelitian seperti pada Gambar 1.

Alur penelitian pada Gambar 1 dimulai dari penentuan dataset. disini digunakan tiga dataset yang berbeda. Dari dataset yang ada satu persatu dilakukan tiga skenario pengujian yaitu tanpa augmentasi, dengan menggunakan augmentasi (rotation, zoom range, shift range, shear, random erase, flip, dan brightness range) dan teknik finetuned dengan maksimal dropout. Setiap skenario yang ada dilakukan model pelatihan pada metode transfer learning yang memiliki parameter yang sedikit (*mobile parameter*) yaitu mobilenet, nasnetmobile, dan efficientnetmobile. Setiap pengujian yang ada kemudian dilakukan pengukuran performansi untuk melihat efek augmentasi dan finetuned pada dataset yang sangat kecil.



Gambar 1. Alur Penelitian dilakukan pengukuran pengaruh teknik fine tune dan augmentasi dataset pada model *transfer learning* CNN

3. HASIL DAN PEMBAHASAN

Pada awal proses yang digunakan untuk *training* (pelatihan) disiapkan terlebih dahulu hyperparameter yang digunakan sebagai dasar bahan training. Pada Tabel 1.

Parameter	Value
Image Size	224 x 224
Batch Size	32
Learning Rate	0.01
Epoch	50

Pada pengukuran pengaruh Teknik *fine tune* dan augmentasi pada model transfer learning CNN yang memiliki parameter yang efisien (*efficient transfer learning*) seperti mobilenet, efficientnet, dan nasnetmobile dengan dataset yang sangat kecil ini memiliki spesifikasi parameter seperti pada Tabel 2.

Tabel 2. Parameter Architecture Model *Mobile Transfer Learning*

Architecture Model	Total Parameters	Trainable Parameters	Non-Trainable Parameters
Mobile-Net	3,491,521	3,469,633	21,888
NasNet-Mobile	4,540,565	4,503,827	36,738
Efficient-Net	4,377,764	4,335,741	42,023

Pada pelatihan model klasifikasi digunakan efisien *transfer learning* dimana parameter yang dimiliki tidak terlalu besar sehingga waktu komputasi tidak memakan waktu yang lama. Pada penelitian ini juga dilakukan pengujian Teknik fine tuning. Teknik fine tuning merupakan sebuah solusi dimana tidak semua *layer* dalam arsitektur *transfer learning* digunakan, sehingga proses komputasi lebih efektif dan efisien.

Tabel 3 merupakan hasil parameter arsitektur transfer learning setelah dilakukan Teknik *fine tuning*.

Tabel 3. Parameter Architecture Model *Mobile Transfer Learning* setelah *Fine Tuning*

Architecture Model after Fine Tuned	Total Parameters	Trainable Parameters	Non-Trainable Parameters
Mobile-Net	3,491,521	2,125,313	1,366,208
NasNet-Mobile	4,540,565	3,460,497	1,080,068
Efficient-Net	4,377,764	1,670,961	2,706,803

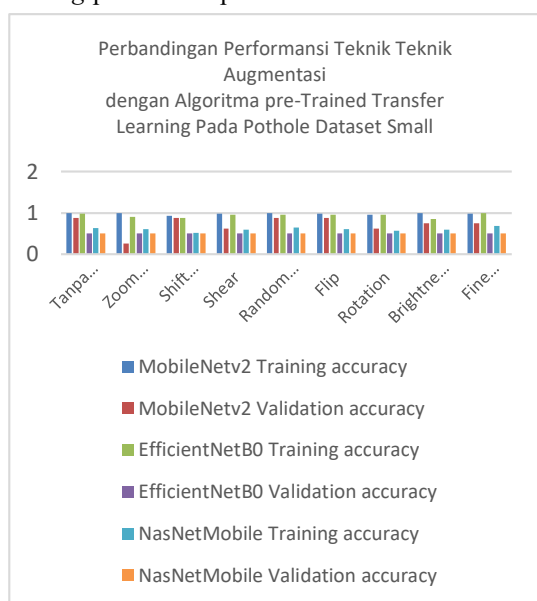
Pada Tabel 3 dapat dilihat bahwa terjadi penurunan jumlah trainable parameters pada ketiga arsitektur model transfer learning yang digunakan. Digunakan dan dibandingkan beberapa metode augmentasi dataset untuk melihat signifikansi pengaruh metode augmentasi terhadap klasifikasi model. Tabel 4 merupakan konfigurasi nilai augmentasi yang digunakan.

Tabel 4. Metode Augmentasi Dataset

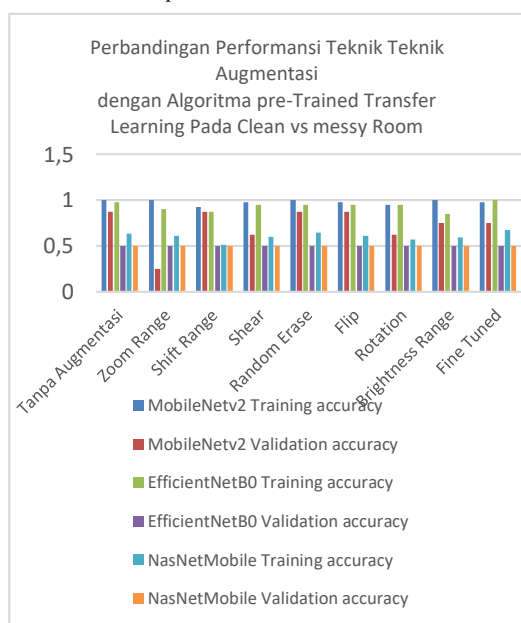
Parameter	Keterangan
Zoom Range	Metode Zoom Range digunakan untuk memperbesar gambar dengan menambah beberapa pixel disekitar gambar. Nilai zoom range yang digunakan adalah 0.15
Shift Range	Metode Shift Range merupakan metode augmentasi dataset dengan menggeser pixel tanpa mengubah dimensi dari citra. Beberapa pixel citra akan terpotong dan akan ada area baru dimana pixel baru citra akan ditambahkan untuk mendapatkan dimensi citra yang sama. Nilai Shift Range yang digunakan adalah W=0.2, H=0.2
Shear	Metode Shear berarti melakukan penggeseran dan mengakibatkan citra akan terdistorsi sepanjang sumbu x maupun y. Biasanya digunakan memperbaiki sudut sehingga persepsi computer lebih mudah memahami. Digunakan shear range sebesar 0.15
Random Erase	Random Erase merupakan metode augmentasi dengan memilih secara acak wilayah persegi Panjang dalam citra dan menghapus pixel yang ada dengan nilai acak. Random erase membuat model tahan terhadap pengaruh

Parameter	Keterangan
	oklusi. Digunakan nilai random erase sebesar $v_l=0$, $v_h=1$
Flip	Teknik Flipping merupakan Teknik sederhana yang mampu mengurangi overfitting. Digunakan vertical flip dan horizontal flip
Rotation	Teknik rotation range merupakan Teknik augmentasi dengan transformasi memutar gambar. Digunakan rotation range sebesar 20
Brightness Range	Teknik brightness range merupakan Teknik augmentasi dengan mengubah pencahayaan dari citra. Digunakan brightness range sebesar 0.2 dan 1.0

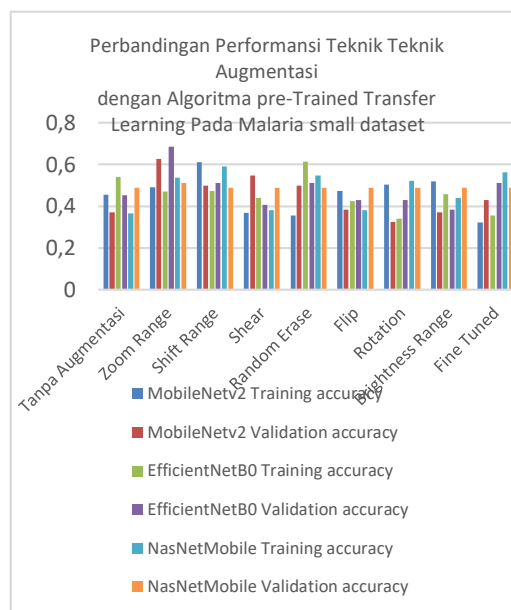
Pengujian dilakukan dengan membandingkan beberapa Teknik augmentasi pada beberapa dataset yang sangat kecil seperti pada Gambar 2 dengan menggunakan beberapa model pembelajaran *transfer learning* pada beberapa *dataset*.



Gambar 2. Performansi Teknik augmentasi dengan pre-trained pada Pothole small dataset



Gambar 3. Performansi Teknik augmentasi dengan pre-trained pada clean vs messy dataset



Gambar 4. Performansi Teknik augmentasi dengan pre-trained pada malaria small dataset

Pada Gambar 2-4 dapat disimpulkan bahwa Random Erase dan Zoom Range terbukti mendominasi dapat meningkatkan akurasi (training dan validation set) pada semua arsitektur efficient transfer learning di semua dataset small. Jumlah peningkatan akurasi setelah dilakukan augmentasi random erase ataupun zoom range yang terjadi sekitar 0.03% hingga 0.1%.

4. KESIMPULAN DAN SARAN

Pada Penelitian yang telah dilakukan mengenai pengukuran pengaruh teknik *fine tune* dan *augmentasi dataset* pada model *transfer learning* CNN *mobilenet*, *efficientnet*, dan *nasnetmobile* dengan *dataset* yang sangat kecil. dapat ditarik beberapa kesimpulan dan saran diantaranya adalah:

1. Dilakukan penelitian mengenai pengukuran pengaruh teknik *fine tune* dan *augmentasi dataset* pada model *transfer learning* CNN *mobilenet*, *efficientnet*, dan *nasnetmobile* dengan *dataset* yang sangat kecil.
2. Dari ketiga dataset yang digunakan sebagai dasar pelatihan (*training*) pada model efisien transfer learning (*mobilenet*, *efficientnet*, dan *nasnetmobile*), Teknik augmentasi random erase dan zoom range mendominasi peningkatan akurasi model.
3. Jumlah peningkatan akurasi setelah dilakukan augmentasi random erase ataupun zoom range yang terjadi sekitar 0.03% hingga 0.1%.

Pada penelitian ini memiliki saran saran yang digunakan sebagai acuan pengembangan penelitian terkait yang dapat dikembangkan diantaranya adalah:

1. Melakukan pengujian ke lebih banyak metode augmentasi dataset
2. Melakukan otomatisasi pencarian range nilai pada metode augmentasi yang digunakan.

3. Melakukan otomatisasi pemilihan teknik augmentasi dataset
4. Untuk mengetahui kehandalan model maka pada penelitian yang akan datang perlu dilakukan justifikasi model yang telah terbentuk dengan dataset yang lainnya.

DAFTAR PUSTAKA

- FUKUSHIMA, K., 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202. <https://doi.org/10.1007/BF00344251>
- HOWARD, A. G., ZHU, M., CHEN, B., KALENICHENKO, D., WANG, W., WEYAND, T., ANDREETTO, M., & ADAM, H., 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. <http://arxiv.org/abs/1704.04861>
- HUSSAIN, Z., GIMENEZ, F., YI, D., & RUBIN, D., 2017. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*, 2017, 979–984.
- JIANG, M. T.-J., WU, S.-H., CHEN, Y.-K., GU, Z.-X., CHIANG, C.-J., WU, Y.-C., HUANG, Y.-C., CHIU, C.-H., SHAW, S.-R., & DAY, M.-Y., 2020. Fine-tuning techniques and data augmentation on transformer-based models for conversational texts and noisy user-generated content. 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 919–925. <https://doi.org/10.1109/ASONAM49781.2020.9381329>
- PAN, S. J., & YANG, Q., 2010. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
- RADENOVIC, F., TOLIAS, G., & CHUM, O., 2019. Fine-Tuning CNN Image Retrieval with No Human Annotation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7), 1655–1668. <https://doi.org/10.1109/TPAMI.2018.2846566>
- SHIJIE, J., PING, W., PEIYI, J., & SIPING, H., 2017. Research on data augmentation for image classification based on convolution neural networks. 2017 Chinese Automation Congress (CAC), 4165–4170. <https://doi.org/10.1109/CAC.2017.8243510>
- TAN, M., & LE, Q. V., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. 36th International Conference on Machine Learning, ICML 2019, 2019-June, 10691–10700.
- TAORMINA, V., CASCIO, D., ABBENE, L., & RASO, G., 2020. Performance of Fine-Tuning Convolutional Neural Networks for HEp-2 Image Classification. *Applied Sciences*, 10(19). <https://doi.org/10.3390/app10196940>
- YILMAZ, F., & DEMİR, A., 2020. Cutting Effect on Classification Using Nasnet Architecture. 2020 Medical Technologies Congress (TIPTEKNO), 1–3. <https://doi.org/10.1109/TIPTEKNO50054.2020.9299313>
- ZHONG, Z., ZHENG, L., KANG, G., LI, S., & YANG, Y., 2020. Random erasing data augmentation. *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, 13001–13008. <https://doi.org/10.1609/aaai.v34i07.7000>

Halaman ini sengaja dikosongkan