Outline:

**BAB II TINJAUAN PUSTAKA**

Bab ini menjelaskan tentang definisi dan teori-teori yang digunakan sebagai landasan penelitian yang berasal dari hasil publikasi dan penelitian dan/atau buku yang relevan.

**REVIEW LITERATUR**

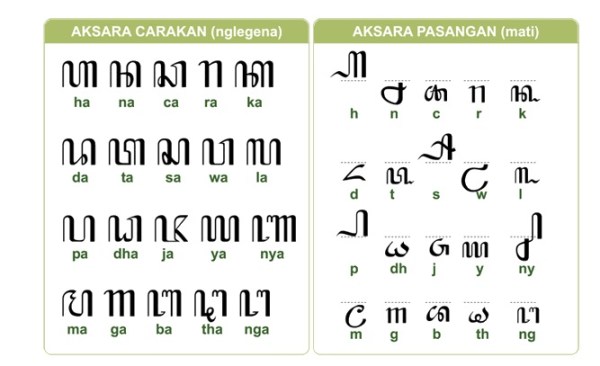
|  |  |  |  |
| --- | --- | --- | --- |
| **Judul** | **Penulis, Tahun** | **Metode** | **Hasil** |
| *Transfer Learning Implementation on Sundanese Script Recognition Using Convolutional Neural Network* | MUHAMMAD KHALIFA U, et al. (2022) | *Transfer learning* dengan *pre-trained* model dari dataset *alphabet*, *devanagari*, *arabic*, dan aksara jawa | Model *transfer learning* terbaik didapat dari *pre-trained* dataset *arabic,* mencapai akurasi 91,86% dan *loss* 0.2814 |
| Javanese Script Text Image Recognition Using Convolutional Neural Networks | Goldy Najma Adli Kesaulya, et al. (2022) | *Transfer learning* dengan *pre-trained* model ResNeXt dengan melakukan *freezing* di 4 layer pertama dari 10 layer model *pre-trained* | Akurasi testing mencapai 98.19% |
| Deep Learning for Image Classification on Very Small Datasets Using Transfer Learning | Menying Shu (2019) | *Fine-tuned pre-trained* model VGG16, VGG19, InceptionV3, InceptionResNetV2 | Akurasi testing mencapai 96% dengan menggunakan *fine-tuned* model dari InceptionResNetV2 |
| A Close Look at Deep Learning with Small Data | L Brigato, et al. (2020) | Review metode yang dapat digunakan untuk menyelesaikan permasalahan *small data*, menggunakan berbagai variasi dari model CNN dan ResNet | *Regularization technique* seperti *data augmentation* dan dropout dapat meningkatkan performa dari *deep learning* |
| A Comprehensive Study on Deep Image Classification with Small Datasets | Gayani Chandrarathne, et al. (2019) | Review metode *transfer learning* dengan *fine-tuned pre-trained* model VGG-16 pada berbagai macam dataset. | *Fine-tuning* (*re-init few layers*)dan *(whole network)* menunjukkan performa yang signifikan dibandingkan dengan *scratch training.* Hasil akurasi yang diperoleh dari *fine-tuning* adalah 91.4%, 95.52% dan 79.6% secara terurut |
| Understanding the Mechanisms of Deep Transfer Learning for Medical Images | Hariharan Ravishankar, et al. (2017) | CNN *transfer learning* CaffeNet | *Transferred* dan *fine-tuned* model dapat menungguli performa *state-of-the-art* *feature engineered pipeline* (Haar) dan menghasilkan akurasi 85% |
| Transfer Learning using CNN for Handwritten Devanagari Character Recognition | Nagender Aneja, et al. (2019) | *Transfer learning* AlexNet, DenseNet, VGG, dan Inception | InceptionV3 menghasilkan akurasi sebesar 99% |
| A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning | Arshia Rehman, et al. (2019) | *Transfer learning* menggunakan *pre-trained* model AlexNet, GoogLeNet, dan VGG16 dari dataset ImageNet | *Fine-tuned* VGG16 menghasilkan akurasi tertinggi sebesar 98.69% |
| A Review of Deep Transfer Learning and Recent Advancements | Mohammadreza Iman, et al. (2023) | Review berbagai metode *deep transfer learning* yang sudah diteliti dalam waktu dekat | (i) *Finetuning*, (ii) *Freezing* CNN *Layers*, dan (iii) *Progressive Learning* adalah teknik yang telah terbukti kemampuan dan efektivitasnya untuk berbagai masalah *machine learning*. |
| Text recognition on images using pre-trained CNN | Afgani Fajar Rizky, et al. (2023) | *Transfer leraning* dengan *augmentation, freeze-layers,* dan *fine-tuning* telah dilakukan dengan menggunakan model AlexNet, VGG, ResNet, dan DenseNet | Akurasi terbaik dihasilkan dari model VGG dengan 0 *freeze layer* dan *augmentation* |
| Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images | Srikanth Tammina (2019) | *Transfer learning* menggunakan *augmentation, fine-tuning,* dan *freeze layers* dengan *pre-trained* model VGG | Hasil terbaik diperoleh dari proses *fine-tuning* dan *augmentation* dengan akurasi 95.40% pada validation |

**LANDASAN TEORI**

**AKSARA JAWA**

Aksara Jawa, yang juga dikenal sebagai Hanacaraka dan Dentawyanjana, adalah salah satu aksara tradisional di Indonesia yang berkembang di Pulau Jawa. Aksara ini terutama digunakan untuk menulis bahasa Jawa, tetapi juga digunakan untuk menulis beberapa bahasa daerah lainnya seperti bahasa Sunda, Madura, Sasak dan Melayu serta bahasa historis seperti Sanskerta dan Kawi. Aksara Jawa berasal dari aksara Brahmi India melalui aksara Kawi dan berkerabat dekat dengan aksara Bali. Aksara ini aktif digunakan dalam sastra dan tulisan sehari-hari oleh masyarakat Jawa dari pertengahan abad ke-15 hingga pertengahan abad ke-20 sebelum perlahan-lahan digantikan oleh huruf Latin. Aksara ini masih diajarkan di DI Yogyakarta, Jawa Tengah, Jawa Timur, Cirebon dan Indramayu sebagai bagian dari muatan lokal, tetapi penggunaannya terbatas dalam kehidupan sehari-hari [38, 39].

Aksara Jawa merupakan sistem tulisan abugida yang terdiri dari sekitar 20 aksara dasar (nglegena)



Resource: [Aksara Jawa Lengkap serta Cara Membacanya, Cari Tahu di Sini! (rukita.co)](https://www.rukita.co/stories/aksara-jawa-lengkap/)

**CITRA DIGITAL**

Citra digital terdiri dari sejumlah elemen terbatas, masing-masing memiliki lokasi dan nilai tertentu. Elemen-elemen ini disebut elemen citra, pel, dan piksel. Sebuah gambar digital dapat didefinisikan sebagai fungsi dua dimensi, f(x,y), di mana x dan y adalah koordinat spasial (bidang), dan amplitudo f pada setiap pasangan koordinat (x,y) disebut intensitas atau tingkat abu-abu gambar di titik tersebut. Bidang pengolahan gambar digital mengacu pada pengolahan gambar digital dengan menggunakan komputer digital (Gonzalez & Woods, 2018).

Referensi: Gonzalez, R. C., & Woods, R. E. (2018). Digital Image Processing (4th ed.). Pearson Education Limited.

**PENGOLAHAN CITRA DIGITAL**

image processing, set of computational techniques for analyzing, enhancing, compressing, and reconstructing images. Its main components are importing, in which an image is captured through scanning or digital photography; analysis and manipulation of the image, accomplished using various specialized software applications; and output (e.g., to a printer or monitor). Image processing has extensive applications in many areas, including astronomy, medicine, industrial robotics, and remote sensing by satellites.

“Digital Image Processing” by Rafael C. Gonzalez and Richard E. Woods.

**RESCALE/RESIZE (BILINEAR INTERPOLATION)**

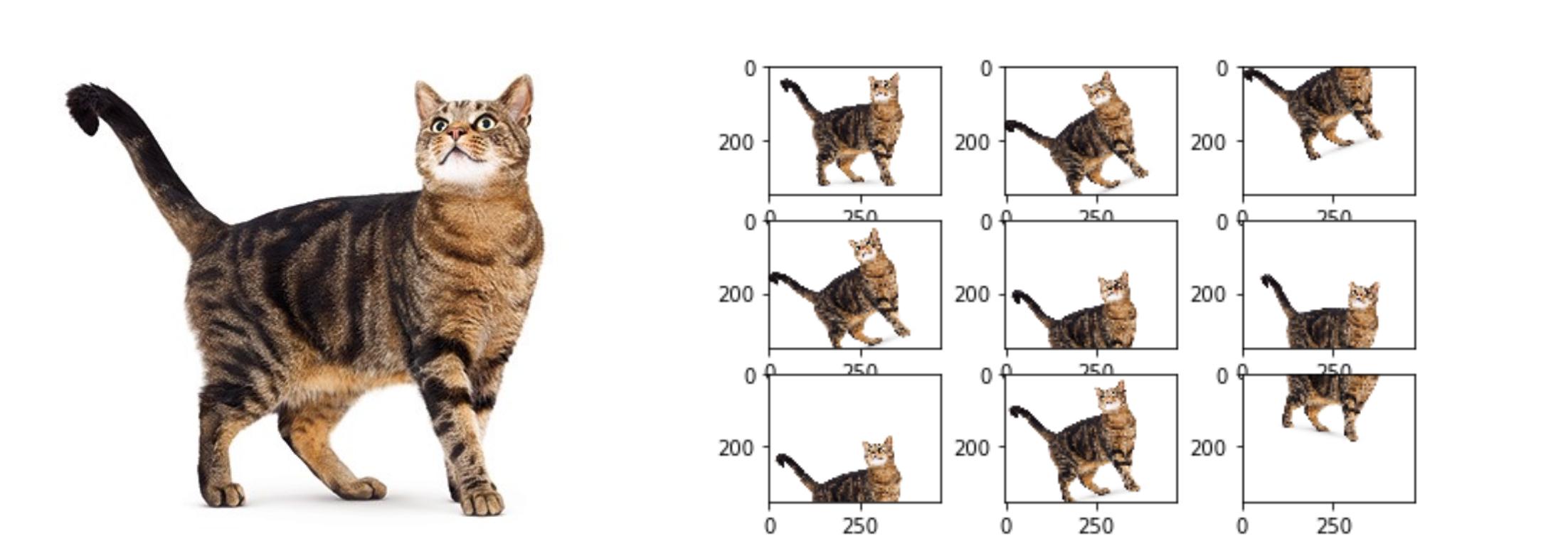
A picture containing diagram, line

Description automatically generated

Source: [Image Processing – Bilinear Interpolation | TheAILearner](https://theailearner.com/2018/12/29/image-processing-bilinear-interpolation/) <https://i0.wp.com/theailearner.com/wp-content/uploads/2018/10/Bilinear_interpolation.png?w=423&ssl=1>

Resizing images is a critical pre-processing step in computer vision. Principally, deep learning models train faster on small images. A larger input image requires the neural network to learn from four times as many pixels, and this increase the training time for the architecture [30].

**DATA AUGMENTATION**



Source: <https://149695847.v2.pressablecdn.com/wp-content/uploads/2020/08/UKwFg.jpg>

[How Data Augmentation Impacts Performance Of Image Classification (analyticsindiamag.com)](https://analyticsindiamag.com/image-data-augmentation-impacts-performance-of-image-classification-with-codes/)

Data augmentation goal is to add new data points to the input space by modifying training images while preserving semantic information and target labels. Thus, it is used to reduce overfitting [29].

Our results confirm the importance of data augmentation in both training and testing and show that it can lead to more performance gains than obtaining new images [29].

**NORMALISASI atau STANDARDISASI**

A picture containing text, font, line, screenshot

Description automatically generated

A picture containing font, white, diagram, line

Description automatically generated

Image normalization or standardization ensures optimal comparisons across data acquisition methods and texture instances. The normalization of pixel values (intensity) is recommended for imaging modalities that do not correspond to absolute physical quantities [31, 32].

ALSO GONZALEZ

**SPLIT DATA**

Pemisahan data melibatkan pengelompokkan dataset menjadi tiga subset: training set, validation set, dan test set. Training set digunakan untuk membangun model, sedangkan validation dan test set berperan sebagai hold-out set. Tidak ada proporsi yang optimal untuk pembagian dataset, namun aturan praktis sebelumnya adalah 70% untuk pelatihan dan 15% untuk validasi dan pengujian. Dalam era big data, proporsi seperti 95% untuk pelatihan dan 2,5%/2,5% untuk validasi/pengujian dapat dipertimbangkan. Dengan menggunakan tiga subset ini, model yang dibangun diharapkan dapat memprediksi dengan baik contoh-contoh yang tidak dilihat oleh algoritma pembelajaran. Validation set digunakan untuk memilih algoritma pembelajaran dan menentukan nilai hiperparameter terbaik, sementara test set digunakan untuk evaluasi model sebelum penggunaannya di lapangan.

REFERENCE: HUNDRED MACHINE LEARNING PAGE[32]

**DEEP LEARNING[16]**

Deep learning adalah cabang dari pembelajaran mesin yang menggunakan jaringan saraf tiruan dengan banyak lapisan tersembunyi untuk memodelkan data yang kompleks (LeCun, Bengio, & Hinton, 2015). Teknik ini memungkinkan komputer untuk mempelajari representasi data yang abstrak dan hierarkis secara otomatis dari data mentah, tanpa perlu pengetahuan domain yang spesifik atau rekayasa fitur yang rumit. Deep learning telah mencapai hasil yang sangat baik dalam berbagai tugas pengenalan pola, termasuk pengenalan suara, pengenalan gambar, dan pemrosesan bahasa alami.

Referensi:

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. doi:10.1038/nature14539

**CONVOLUTIONAL NEURAL NETWORK**

*Convolutional Neural Network* (CNN) adalah jenis *network* khusus yang secara signifikan mengurangi jumlah parameter dalam *deep neural network* dengan banyak unit tanpa kehilangan terlalu banyak kualitas model. CNN telah menemukan aplikasi dalam pengolahan gambar dan teks di mana mereka mengalahkan banyak patokan yang sebelumnya ditetapkan [32].

**ACTIVATION FUNCTIONS**

Untuk memungkinkan jaringan saraf mempelajari batas keputusan yang kompleks, kami menerapkan fungsi aktivasi non-linier pada beberapa lapisannya. Fungsi yang umum digunakan meliputi tanh, ReLU, softmax, dan varian dari fungsi-fungsi tersebut. Secara teknis, setiap neuron menerima sinyal masukan yang merupakan jumlah terbobot dari bobot sinaptik dan nilai aktivasi dari neuron yang terhubung [34].

**RELU ACTIVATION FUNCTIONS**

**A picture containing text, line, screenshot, plot

Description automatically generated**[34]

**SOFTMAX ACTIVATION FUNCTIONS**

Secara umum, pada level terakhir dari FFNN, fungsi softmax diterapkan sebagai batas keputusan. Ini adalah kasus umum, terutama ketika menyelesaikan masalah klasifikasi. Sebaliknya, kita tidak perlu menggunakan fungsi aktivasi sama sekali untuk masalah regresi [34].

Dalam Matematika, fungsi softmax adalah generalisasi dari fungsi logistik, yang “menekan” vektor K-dimensi dari nilai-nilai real sembarang menjadi vektor K-dimensi σ(z) dari nilai-nilai real dalam rentang [0, 1] yang jumlahnya sama dengan 1 [34].

A picture containing font, handwriting, white, diagram

Description automatically generated[34]

**DROPOUT**

Dropout adalah teknik yang digunakan untuk mengurangi overfitting dalam jaringan dengan banyak lapisan dan/atau neuron. Secara umum, lapisan dropout ditempatkan setelah lapisan yang memiliki sejumlah besar neuron yang dapat dilatih [34][42].

**A picture containing diagram, line, screenshot, circle

Description automatically generated**[34]

**A picture containing text, screenshot, diagram, font

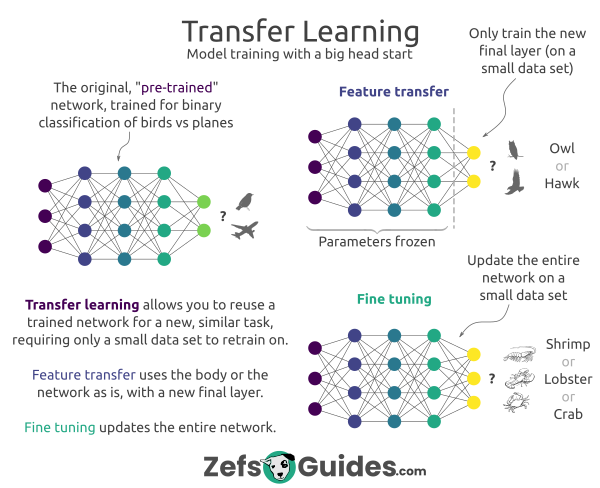
Description automatically generated**[42]

**FREEZEOUT**

***SEMENTARA GA DIMASUKIN KE BAB II***

**TRANSFER LEARNING**

*Transfer learning* adalah peningkatan pembelajaran dalam tugas baru melalui transfer pengetahuan dari tugas terkait yang sudah dipelajari [33]. Pembelajaran transfer terdiri dari mengambil jaringan yang sudah dibangun dan membuat perubahan yang sesuai pada parameter dari berbagai lapisan sehingga dapat menyesuaikan dengan dataset lain [34].



Source: <https://preview.redd.it/4p3j8qej1nn91.png?auto=webp&s=8de0945c56138ede9a0067be28e505a3bca1dc17>

**TRANSFER LEARNING APPROACHES**

**A picture containing text, screenshot, font, diagram

Description automatically generated**

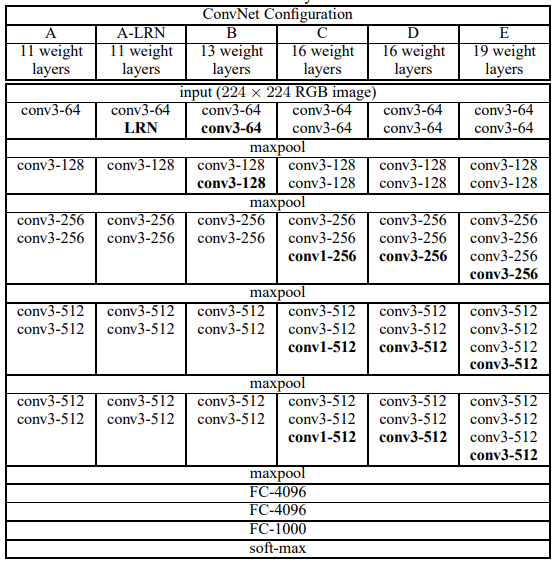
**VGG**

VGG adalah model arsitektur jaringan saraf tiruan yang sangat dalam yang dikembangkan oleh Visual Geometry Group (VGG) di Universitas Oxford (Simonyan & Zisserman, 2014). Model ini terdiri dari beberapa lapisan konvolusi yang diikuti oleh lapisan max-pooling dan beberapa lapisan sepenuhnya terhubung di bagian akhir. Salah satu keunggulan utama dari model VGG adalah penggunaan filter konvolusi berukuran kecil (3x3) di semua lapisannya, yang memungkinkan jaringan untuk memiliki kedalaman yang lebih besar dengan jumlah parameter yang sama dibandingkan dengan jaringan yang menggunakan filter konvolusi berukuran lebih besar. Model VGG mencapai hasil yang sangat baik pada benchmark klasifikasi ILSVRC 2012 dan telah digunakan secara luas sebagai titik awal untuk berbagai tugas pengenalan gambar [40].

Referensi:

- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

**ARSITEKTUR VGG[40]**

****

**INCEPTION**

Inception adalah model arsitektur jaringan saraf tiruan yang digunakan untuk visi komputer. Model ini dirancang untuk meningkatkan kualitas jaringan dengan cara yang efisien secara komputasi dengan menggunakan konvolusi yang difaktorkan dan regularisasi agresif (Szegedy et al., 2015). Inception-v3 adalah salah satu versi dari model Inception yang mencapai kinerja tinggi pada benchmark klasifikasi ILSVRC 2012 dengan biaya komputasi yang relatif rendah dibandingkan dengan arsitektur yang lebih sederhana dan monolitik (Szegedy et al., 2016). Model ini juga menggunakan jumlah parameter yang lebih sedikit dan regularisasi tambahan dengan classifier auxilary yang dinormalisasi batch dan label-smoothing untuk melatih jaringan berkualitas tinggi pada set pelatihan berukuran relatif sedang (Szegedy et al., 2016).

Referensi:

- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

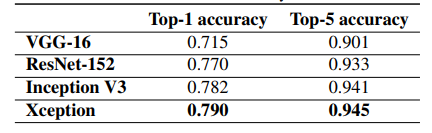
**ARSITEKTUR INCEPTION[41]**

**A picture containing text, screenshot, number, font

Description automatically generated**

**XCEPTION**

Xception adalah arsitektur jaringan saraf konvolusional yang sepenuhnya didasarkan pada lapisan konvolusi terpisah secara mendalam (depthwise separable convolution layers). Arsitektur ini diusulkan sebagai perpanjangan dari arsitektur Inception, di mana modul Inception telah digantikan dengan konvolusi terpisah secara mendalam. Arsitektur Xception adalah tumpukan linier lapisan konvolusi terpisah secara mendalam dengan koneksi residual [35].

****

Benchmark table [35]

**ARSITEKTUR XCEPTION**

**A screenshot of a computer program

Description automatically generated with low confidence**

**CONFUSION MATRIX**

The confusion matrix is a table that summarizes how successful the classification model is at predicting examples belonging to various classes. One axis of the confusion matrix is the label that the model predicted, and the other axis is the actual label. In a binary classification problem, there are two classes [32].

**ALBUMENTATIONS**

Albumentations adalah *library* Python untuk augmentasi gambar yang cepat dan fleksibel. Albumentations secara efisien mengimplementasikan berbagai operasi transformasi gambar yang kaya dan dioptimalkan untuk kinerja, sambil memberikan antarmuka augmentasi gambar yang ringkas namun kuat untuk berbagai tugas visi komputer, termasuk klasifikasi objek, segmentasi, dan deteksi [36][45].

Contoh lebih banyak: [Defining a simple augmentation pipeline for image augmentation - Albumentations Documentation](https://albumentations.ai/docs/examples/example/), [Image augmentation for classification - Albumentations Documentation](https://albumentations.ai/docs/getting_started/image_augmentation/)

**TENSORFLOW**

TensorFlow adalah salah satu *library* yang dapat digunakan untuk implementasi *machine learning*. *Library* ini terkenal dengan penggunaan terkait *Deep Learning* [37][46].

[TensorFlow](https://www.tensorflow.org/)

[API Documentation  |  TensorFlow v2.12.0](https://www.tensorflow.org/api_docs)

BAB I

13: Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application, Urbanization and Regional Imbalances in Indonesia, Rancang Bangun Aplikasi Pembelajaran Aksara Jawa Berbasis Android

14: Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application, Urbanization and Regional Imbalances in Indonesia, Rancang Bangun Aplikasi Pembelajaran Aksara Jawa Berbasis Android

15: [Kurdish Handwritten character recognition using deep learning techniques - ScienceDirect](https://www.sciencedirect.com/science/article/abs/pii/S1567133X22000485)

16: [Sci-Hub | Deep learning. Nature, 521(7553), 436–444 | 10.1038/nature14539](https://sci-hub.se/https:/www.nature.com/articles/nature14539)

17: [(PDF) A Comprehensive Study on Deep Image Classification with Small Datasets (researchgate.net)](https://www.researchgate.net/publication/331728350_A_Comprehensive_Study_on_Deep_Image_Classification_with_Small_Datasets)

18: K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv preprint arXiv:1512.03385, 2015.

19: [Sci-Hub | A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Transactions on Neural Networks and Learning Systems, 1–21 | 10.1109/TNNLS.2021.3084827](https://sci-hub.se/https:/ieeexplore.ieee.org/abstract/document/9451544)

23: J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in Advances in neural information processing systems, pp. 3320-3328, 2014.

24: J. Deng, W. Dong, R. Socher, L. J. Li, L. Kai, and F.-F. Li, "Image{N}et: {A} large-scale hierarchical image database," in Proceedings of the Computer Vision and Pattern Recognition (CVPR), pp. 248-255, 2009

25: [Javanese Script Text Image Recognition Using Convolutional Neural Networks | IEEE Conference Publication | IEEE Xplore](https://ieeexplore.ieee.org/abstract/document/9888527)

26: [TRANSFER LEARNING IMPLEMENTATION ON SUNDANESE SCRIPT RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK (ugm.ac.id)](http://etd.repository.ugm.ac.id/penelitian/detail/213204)

BAB II

27: [A survey of transfer learning | Journal of Big Data | Full Text (springeropen.com)](https://journalofbigdata.springeropen.com/articles/10.1186/s40537-016-0043-6)

28: [A Comprehensive Survey on Transfer Learning | IEEE Journals & Magazine | IEEE Xplore](https://ieeexplore.ieee.org/abstract/document/9134370)

29: [Our results confirm the importance of data augmentation in both training and testing and show that it can lead to more performance gains than obtaining new images. - Consensus](https://consensus.app/details/results-confirm-importance-data-augmentation-training-perez/7309cabc177d5f9a85722a741270e505/)

30: [Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance | SpringerLink](https://link.springer.com/chapter/10.1007/978-3-030-95498-7_2#:~:text=Resizing%20images%20is%20a%20critical,training%20time%20for%20the%20architecture.)

31: [Fundamentals of Texture Processing for Biomedical Image Analysis: A General Definition and Problem Formulation - ScienceDirect](https://www.sciencedirect.com/science/article/abs/pii/B9780128121337000016)

32: [The Hundred-Page Machine Learning Book by Andriy Burkov (themlbook.com)](https://themlbook.com/)

33: [torrey.handbook09.pdf (wisc.edu)](https://ftp.cs.wisc.edu/machine-learning/shavlik-group/torrey.handbook09.pdf) - Transfer Learning Lisa Torrey and Jude Shavlik, [Transfer Learning (wisc.edu)](https://pages.cs.wisc.edu/~shavlik/abstracts/torrey.handbook09.abstract.html)

34: buku: Expert Insight Deep Learning with Tensorflow

35: [Xception: Deep Learning With Depthwise Separable Convolutions (thecvf.com)](https://openaccess.thecvf.com/content_cvpr_2017/papers/Chollet_Xception_Deep_Learning_CVPR_2017_paper.pdf)

36: [Albumentations: fast and flexible image augmentations](https://albumentations.ai/)

37: [TensorFlow](https://www.tensorflow.org/)

38: Poerwadarminta, W.J.S (1939). *[Baoesastra Djawa](https://www.sastra.org/katalog/judul?ti_id=75)* (dalam bahasa Jawa). Batavia: J.B. Wolters. [ISBN](https://id.wikipedia.org/wiki/International_Standard_Book_Number) [0834803496](https://id.wikipedia.org/wiki/Istimewa:Sumber_buku/0834803496)

39: *Everson, Michael (6 Maret 2008).*["Proposal for encoding the Javanese script in the UCS"](http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3319.pdf)*(PDF). ISO/IEC JTC1/SC2/WG2. Unicode (N3319R3).*

40: [[1409.1556] Very Deep Convolutional Networks for Large-Scale Image Recognition (arxiv.org)](https://arxiv.org/abs/1409.1556)

41: [CVPR 2016 Open Access Repository (cv-foundation.org)](https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html)

42: buku: Understanding Deep Learning Chitta Ranjan 2020

43: [LayerOut: Freezing Layers in Deep Neural Networks | SpringerLink](https://link.springer.com/article/10.1007/s42979-020-00312-x)

44: [Technologies | Free Full-Text | A Review of Deep Transfer Learning and Recent Advancements (mdpi.com)](https://www.mdpi.com/2227-7080/11/2/40)

45: [Information | Free Full-Text | Albumentations: Fast and Flexible Image Augmentations (mdpi.com)](https://www.mdpi.com/2078-2489/11/2/125)

46: [Citing TensorFlow](https://www.tensorflow.org/about/bib)

# References

|  |  |
| --- | --- |
| [1] | A. R. G., A. Tandra, I. Susanto, J. Harefa and A. Chowanda, "Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application," *Procedia Computer Science,* pp. 499-505, 2019. |