

IMPLEMENTASI DEEP LEARNING DALAM KLASIFIKASI CITRA GAMBAR DENGAN MENGGUNAKAN METODE CNN

Deep Learning Implementation in Image Classification Using CNN Method

Muhammad Imam Dinata¹, Nani Sulistianingsih², Siti Agrippina Alodia Yusuf³

^{1,2,3}Program Sistem Teknologi Informasi, Fakultas Teknik

Universitas Muhammadiyah Mataram

Jl. KH. Ahmad Dahlan No. 1 Pagesangan, Mataram, NTB, Indonesia

email: ¹imam.dinata@ummat.ac.id, ²nani.sulistianingsih@ummat.ac.id, ³siti.agrippina@ummat.ac.id

ABSTRAK

Sepatu biasanya terbuat dari beberapa bahan dan akan dibentuk menjadi berbagai jenis sepatu yang disesuaikan dengan fungsinya. Untuk mengetahui jenis sepatu yang dihasilkan itu baik atau tidak dibutuhkan seorang yang ahli di bidang sepatu. Proses identifikasi sepatu yang baik itu dapat dilihat dari berbagai faktor misalnya: kode sepatu, jahitan sepatu, bagian sol sepatu, logo sepatu, harga, material sepatu, dan lainnya. *Computer Vision* merupakan salah satu cabang ilmu yang memungkinkan mesin komputer dapat melihat, mengidentifikasi dan mengekstrak informasi dari sebuah objek yang ada di sekitar layaknya manusia. *Deep learning*, komputer dapat mengklasifikasikan data gambar atau suara, salah satu metodenya yaitu *Convolutional Neural Network (CNN)*. *CNN* merupakan salah satu penerapan *deep learning*, dimana metode ini berfungsi untuk mengolah citra digital. Dataset yang digunakan dalam penelitian ini diambil di website www.kaggle.com/datasho-es dengan jumlah data sebanyak 10.000 data gambar dan dibagi menjadi 4 kelas yaitu ballet flat, boat, brogue, sneaker dan untuk data validasi sebanyak 2000 data gambar. Dari hasil penelitian ini tingkat akurasi prediksi yang paling tertinggi yaitu untuk sepatu sneakers dengan tingkat akurasi prediksi sebesar 99% sementara untuk tingkat akurasi prediksi terendah yaitu untuk sepatu Brogue sebesar 27%.

Kata Kunci: *Sepatu, Computer Vision, Deep Learning, CNN*

ABSTRACT

Shoes are usually made of several materials and will be formed into various types of shoes according to their function. To find out whether the type of shoes produced are good or not, an expert in the field of shoes is needed. The process of identifying good shoes can be seen from various factors, for example: shoe code, shoe stitching, shoe soles, shoe logos, price, shoe material, and others. *Computer Vision* is a branch of science that allows computer machines to see, identify and extract information from objects around them like humans. *Deep learning*, computers can classify image or sound data, one of the methods is the *Convolutional Neural Network (CNN)*. *CNN* is one of the applications of *deep learning*, where this method functions to process digital images. The dataset used in this research was taken from the website www.kaggle.com/datasho-es with a total of 10,000 image data and divided into 4 classes, namely ballet flat, boat, brogue, sneaker and for validation data as many as 2000 image data. From the results of this study the highest level of prediction accuracy was for sneakers with a prediction accuracy rate of 99% while the lowest prediction accuracy level was for Brogue shoes by 27%.

Keywords: *Shoes, Computer Vision, Deep Learning, CNN*

1. PENDAHULUAN

Salah satu benda yang biasanya digunakan untuk melindungi alas kaki yaitu sepatu. Sepatu biasanya terbuat dari beberapa bahan dan akan dibentuk menjadi berbagai jenis sepatu yang disesuaikan dengan fungsinya. Dalam produksi sepatu, Indonesia menduduki peringkat ke-4 di dunia, mencapai 886 juta dengan berbagai jenis sepatu. Sehingga hal ini merupakan salah satu faktor mendorongnya jumlah industri alas kaki di Indonesia mencapai 18.687 unit usaha yang terdiri dari 18.091 unit usaha kecil, 441 unit usaha menengah, dan 155 unit usaha skala besar [1].

Untuk mengetahui jenis sepatu yang dihasilkan itu baik atau tidak dibutuhkan seorang yang ahli di bidang sepatu. Proses identifikasi sepatu yang baik itu dapat dilihat dari berbagai faktor misalnya: kode sepatu, jahitan sepatu, bagian sol sepatu, logo sepatu, harga, material sepatu, dan lainnya. Semakin berkembangnya teknologi proses identifikasi sepatu tersebut, dapat dilakukan dengan memanfaatkan *Computer Vision* dengan cara mengidentifikasi citra digital dari material sepatu. Mengidentifikasi jenis sepatu menggunakan *computer vision* sangat bermanfaat karena dapat mempermudah identifikasi dan mengurangi biaya proses *quality control*.

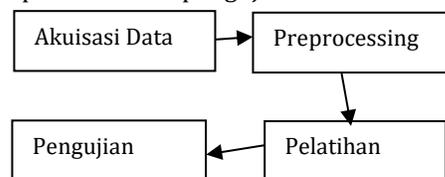
Computer Vision merupakan salah satu cabang ilmu yang memungkinkan mesin komputer dapat melihat, mengidentifikasi dan mengekstrak informasi dari sebuah objek yang ada di sekitar layaknya manusia [2]. Penerapan *Computer Vision* dapat dilakukan dengan salah satu metode kecerdasan buatan yaitu Jaringan Saraf Tiruan (JST). JST merupakan implementasi dari cara otak bekerja yang diterapkan dalam sebuah mesin dengan memanfaatkan sekumpulan data. Perkembangan dari metode JST lebih detail akan menghasilkan cabang ilmu yang lebih baik yaitu *Deep Learning*. *Deep learning*, komputer dapat mengklasifikasikan data gambar atau suara, salah satu metodenya yaitu *Convolutional Neural Network (CNN)*. *CNN* merupakan salah satu penerapan *deep learning*, dimana metode ini berfungsi untuk mengolah citra digital.

Penelitian sebelumnya mengenai implementasi *deep learning* pada identifikasi jenis tumbuhan berdasarkan citra daun menggunakan *convolutional neural network*, yang dilakukan oleh Sarirtul Ilahiyah, Agung Nilogiri. Penelitian ini menggunakan 2000 citra

daun, menghasilkan akurasi 85% [3]. Penelitian lainnya yaitu implementasi *deep learning object detection* rambu K3 pada video menggunakan *CNN* dengan *tensorflow*, penelitian ini dilakukan oleh Syinta Nuri Mashita penelitian ini menggunakan data sebanyak 1500 data citra menghasilkan tingkat akurasi sebesar 50%-97% [4]. Penelitian lainnya dengan mengusulkan arsitektur *Deep Neural Network* Berbasis *ResNet50* untuk mengklasifikasikan kata-kata yang dieja dengan jari. Dataset yang digunakan adalah *American Sign Language Hand gesture* standar yang menghasilkan akurasi sebesar 99,03%. Sementara studi [6] Jaringan Saraf Konvolusional yang Terhubung Padat (*DenseNet*) untuk mengklasifikasikan bahasa isyarat secara real-time menggunakan kamera web dengan akurasi 90,3%. Pengikut penelitian [7]-[11] telah menerapkan model *CNN* untuk pengenalan bahasa isyarat dan diuji menggunakan Dataset Bahasa Isyarat Amerika (ASL), dengan tingkat akurasi 99,92%, 99,85%, 99,3%, 93%, dan 99,91%. Dataset yang digunakan dalam penelitian ini diambil di website www.kaggle.com/datashoes dengan jumlah data sebanyak 10.000 data gambar dan dibagi menjadi 4 kelas yaitu *ballet flat*, *boat*, *brogue*, *sneaker* dan untuk data validasi sebanyak 2000 data gambar dan dibagi menjadi 4 kelas yaitu: *ballet flat*, *boat*, *brogue*, *sneaker*. Permasalahan yang timbul dari penelitian ini yaitu untuk menghasilkan tingkat akurasi yang dilakukan oleh mesin *computer* sehingga diharapkan dari akurasi tersebut, mesin *computer* dapat melakukan prediksi terhadap jenis gambar yang akan dihasilkan.

2. METODE PENELITIAN

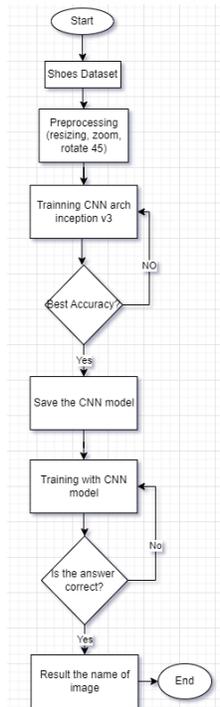
Penelitian ini merupakan penelitian eksperimen kuantitatif untuk mengukur tingkat akurasi dengan menggunakan data gambar sepatu. Secara umum metode yang diusulkan terdiri dari empat tahap yang masing-masing tahap adalah akuisasi data, preprocessing, pelatihan dan pengujian.



Gambar 1. Metode yang diusulkan

Berdasarkan metodologi yang digunakan dalam penelitian ini, sebagaimana di jelaskan dalam

gambar 1, maka tahap yang pertama yaitu akuisasi data, dimana data yang digunakan yaitu data gambar sebanyak 10.000 gambar dibagi menjadi 4 kelas yaitu: ballet flat, boat, brogue, sneaker. Selanjutnya tahap kedua yaitu melakukan proses preproccesing data gambar dengan mengubah ukuran nya menjadi 128x128, mengubah posisi gambar menjadi 45 derajat dan melakukan proses zoom sebanyak 20% dari gambar aslinya. Hal ini dilakukan untuk mengurangi proses waktu pada saat dilakukan pembuatan model training [12]. Selanjutnya pada bagian pelatihan, dalam pembuaan model CNN dengan arsitektur Inception V-3, dimana artistekturnya *inception V-3* akan menghasilkan nilai akurasi lebih dari 78% [13]. Setelah menghasilkan akurasi yang baik, selanjutnya proses terakhir yaitu melakukan proses pengujian terhadap data tersebut, Pengujian ini dilakukan dengan cara menginputkan data gambar secara langsung dengan acak, sealnjutnya akan menampilkan hasil yang sesuai dengan nama gambar yang diinputkannya. Penjelasan mengenai proses pada bagian method dengan menggunakan arsitektur inception V-3 dijelaskan pada gambar 2 dalam bentuk flowchart.



Gambar 2. Flowchart pada proses metode CNN dengan arch Inception V-3

Untuk mengetahui nilai akurasi pada sebuah model, dapat dilakukan dengan menggunakan rumus [14] :

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad [1]$$

Sementara untuk mengetahui dari presisi sebuah model tersebut, dengan mengetahui bahwa model tersebut menghasilkan prediksi yang baik atau tidak, dapat menggunakan rumus [15]:

$$Precision = \frac{TP}{TP + FP} \quad [2]$$

Dan untuk mengetahui nilai dari persentasi akurasi dengan tingkat kebenaran pada klasifikasi yang disebut dengan recall, maka rumus yang digunakan yaitu:

$$Recall = \frac{TN}{TP + FN} \quad [3]$$

Tokenizing berfungsi untuk menghapus karakter-karakter tertentu pada teks (Jaka, 2015), dan terkahir adalah normalisasi berfungsi untuk menghilangkan kata tidak baku menjadi baku (Javed & Kamal, 2018). Kemudian masuk ke tahap pengkelasan dataset dengan vader, kelas yang digunakan akan searah dengan rating google playstore dan berikut daftar kelas yang akan digunakan

Tabel 1. Nilai kelas sama dengan rating

kelas	rating
sangat buruk	1
buruk	2
sedang	3
baik	4
sangat baik	5

3. HASIL DAN PEMBAHASAN

3.1 Data Acquisition

Data yang di download dari website kaggle.com, dimana data berisikan mengenai 10.000 data gambar sepatu yang terdiri dari 4 jenis sepatu, dan masing-masing jenis sepatu memiliki data gambar sebanyak 2500 data gambar. Dari data gambar yang berjumlah 10.000 gambar dibagi menjadi dua bagian yaitu 80% data gambar untuk training dan 20% data gambar untuk testing. Jadi total untuk training sebanyak 8000 gambar dan untuk melakukan testing dibutuhkan 2000 data gambar 3.



Gambar 3. Contoh gambar jenis sepatu yang diambil dari kaggle

3.2 Preprocessing

Pada tahap preprocessing, proses perubahan ukuran dengan cara merubah ukuran aslinya menjadi ukuran gambar 128x128 [12]. Langkah ini dilakukan untuk mengurangi waktu pada saat proses pembuatan model training [13]. Selanjutnya dilakukan rotasi pada gambar sebesar 45 dan zoom sebesar 20% dari gambar aslinya, hal ini dilakukan bertujuan untuk meningkatkan akurasi dari gambar dan mempermudah mengenali gambar pada saat melakukan prediksi nantinya [13]. **Gambar.4** menampilkan hasil dari proses preprocessing.



Gambar 4. (a) Gambar asli, (b) resize 128x128, (c) rotate image 45 derajat, (d) zoom 20%

Dari **gambar 4**, bisa dilihat hasil perubahan dimensi gambar sepatu untuk dijadikan sebagai dataset pada saat pembuatan model.

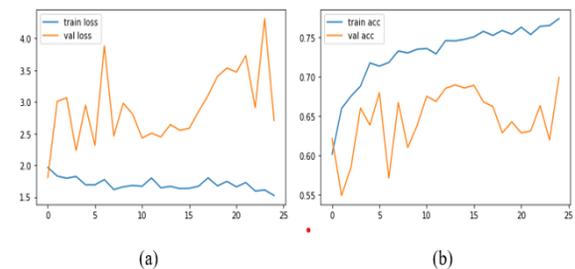
3.3 Training

Pada tahap ini, pembuatan model CNN dengan arsitektur Inception V-3, digunakan untuk proses training pada saat pembuatan model untuk melakukan klasifikasi gambar jenis sepatu. Pembuatan model CNN arsitektur Inception V-3 mengimplementasikan hyperparameter dengan nilai konsistensi dari learning rate, epoch, loss function, dan optimizer. **Table 1** merupakan model CNN arsitektur Inception V-3 yang digunakan dalam penelitian ini.

Table 1. CNN Model Architecture Inception V-3

Layer Type	Description	Size
Input Layer	Input Image	128x128x3
Conv 1d	Convolutional 2D Batch Normalization Activation	32 kernel, 63x63
Conv 2d	Convolutional 2D Batch Normalization Activation	64 kernel, 61x61
.	Max Pooling	Window size 2 x 2
.	.	.
Conv 2d_93	Convolutional 2D Batch Normalization_85 Activation Max Pooling Mixed_10	192 kernel, 2 x 2 192, 2 x 2 320, 2 x 2 Window size 2x2 2048, 2 x 2
Flatten Layer	Flatten	8192
Output Layer	Dense Softmax	4

Dalam proses training model menggunakan perangkat keras dan perangkat lunak dengan spesifikasi yaitu Asus 505 fx laptop, 8 GB DDR4 Ram, processor AMD 7, Nvidia GeoForce 2010 Rtx GPU, 500 SSD. Sistem operasi yang digunakan yaitu Windows 11 Professional dan untuk pembuatan model training dan testing yaitu menggunakan bahasa pemrograman python dengan menggunakan library dari tensorflow. Hasil dari training model yang dihasilkan ada pada **gambar 5**.



Gambar 5. Graphics of the Training Model Proses: (a) Loss of Training and Validation (b) Accuracy of Training and Validation

Dapat dilihat dari Fig 5 diatas, proses training dilakukan sebanyak 25 epochs, pada saat 1 epoch tingkat akurasi training sebesar 60% dan akurasi validasi sebesar 63.01%, sementara untuk training loss sebesar 20% dan validasi loss sebesar 17,94%. Dan di 10 epoch, tingkat akurasi training yaitu 77% dan tingkat akurasi validasi sebesar 66%, sementara untuk training loss sebesar 17,4% dan untuk validasi loss sebesar 26%. Dan yang terakhir untuk 25 epochs tingkat training sebesar 78.56% dan untuk tingkat validasi sebesar 68,75%. Sementara untuk training loss sebesar 16,15% dan untuk validasi loss sebesar 27,73%.

3.4. Testing

Dari hasil proses pembuatan model training, tahap selanjutnya yaitu melakukan pengujian pada model tersebut. Proses yang dilakukan hampir sama dengan pada proses training, hanya saja letak perbedaannya yaitu iterasi backpropagation out backward pass atau backpropagation tidak mengubah bobot atau bobot model seperti pada proses training. [18]. Hasil yang dihasilkan pada proses testing, dapat dilihat dari tabel 2.

Table 2. Hasil testing pada model

No	Ballet Flat	Boat	Brogue	Sneaker
Ballet Flat	44%	99%	49%	14%
Boat	31%	36%	100%	24%
Brogue	11%	25%	27%	27%
Sneakers	42%	22%	99%	99%

Dari **table 2.** Diatas dapat disimpulkan bahwa tingkat akurasi prediksi yang paling tertinggi yaitu untuk sepatu sneakers dengan tingkat akurasi prediksi sebesar 99% sementara untuk tingkat akurasi prediksi terendah yaitu untuk sepatu Brogue sebesar 27%.

4. Kesimpulan

Dari penelitian ini, data gambar sepatu dengan model CNN arsitektur Inception V-3 berhasil untuk dibuat dengan hasil akurasi sebesar 68.75% dan untuk validasi loss sebesar 27.73%. Diharapkan untuk penelitian selanjutnya dapat menambahkan variasi jenis sepatu dengan jumlah data yang seimbang dan meningkatkan akurasi dengan menambahkan parameter di dalam pembuatan arsitektur modelnya.

DAFTAR PUSTAKA

- [1] Pikiran rakyat. Info produksi sepatu di Indonesia, <https://kemenperin.go.id/artikel/20540/Pr-oduksi-Sepatu-di-Indonesia-Urutan-Keempat-di-Dunia> diakses pada tanggal 21 Desember 2022
- [2] Jimmy Feriawan, Daniel Swanjaya. 2020. Perbandingan Arsitektur Visual Geometry Group dan MobileNet Pada Pengenalan Jenis Kayu. Seminar Nasional Inovasi Teknologi e-ISSN: 2549-7952
- [3] Sarirotul Ilahiyah , Agung Nilogiri. 2018. Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network. JUSTINDO (Jurnal Sistem & Teknologi Informasi Indonesia), Vol. 3, No. 2, Agustus 2018
- [4] Syinta Nuri Mashita. 2020. Implementasi Deep Learning Object Detection Rambu K3 Pada Video Menggunakan Metode Convolutional Neural Network (CNN) Dengan Tensorflow. <https://dSPACE.uisi.ac.id/123456789/28781> diakses pada tanggal 21 Desember 2022
- [5] P. Rathi, R. K. Gupta, S. Agarwal, A. Shukla, and R. Tiwari, "Sign Language Recognition Using ResNet50 Deep Neural Network Architecture Pulkit," Next Gener. Comput. Technol. 2019 Sign, pp. 1–7, 2019, <http://dx.doi.org/10.2139/ssrn.3545064>.
- [6] R. Daroya, D. Peralta, and P. Naval, "Alphabet Sign Language Image Classification Using Deep Learning," IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON, pp. 646–650, 2019, <https://doi.org/10.1109/TENCON.2018.8650241>.
- [7] M. M. Rahman, M. S. Islam, M. H. Rahman, R. Sassi, M. W. Rivolta, and M. Aktaruzzaman, "A new benchmark on american sign language recognition using convolutional neural network," in 2019 International Conference on Sustainable Technologies for Industry 4.0, STI 2019, pp. 1-6, 2019, <https://doi.org/10.1109/STI47673.2019.9067974>.
- [8] M. R. M. Bastwesy, N. M. ElShennawy, and M. T. F. Saidahmed, "Deep Learning Sign Language Recognition System Based on Wi-Fi CSI," Int. J. Intell. Syst. Appl., vol. 12, no. 6, pp. 33–45, 2020, <https://doi.org/10.5815/ijisa.2020.06.03>.
- [9] A. Abdulhussein and F. Raheem, "Hand Gesture Recognition of Static Letters American Sign Language (ASL) Using Deep Learning," Eng. Technol. J., vol. 38, no. 6, pp.

- 926–937, 2020,
<https://doi.org/10.30684/etj.v38i6a.533>.
- [10] R. S. Sabeenian, S. Sai Bharathwaj, and M. Mohamed Aadhil, “Sign language recognition using deep learning and computer vision,” *J. Adv. Res. Dyn. Control Syst.*, vol. 12, no. 5, pp. 964–968, 2020,
<https://doi.org/10.5373/JARDCS/V12SP5/20201842>.
- [11] K. H. Rawf, “Effective Kurdish Sign Language Detection and Classification Using Convolutional Neural Networks,” 2022,
<https://doi.org/10.21203/rs.3.rs-1965056/v1>.
- [12] P. Thévenaz, T. Blu, and M. Unser, “Image interpolation and resampling,” *Handb. Med. Image Process. Anal.*, pp. 465–493, 2009,
<https://doi.org/10.1016/B978-012373904-9.50037-4>.
- [13] inception v3.cloud google,
<https://cloud.google.com/tpu/docs/inception-v3-advanced>, diakses pada tanggal 29 januari 2023
- [14] L. K. S. Tolentino, R. O. S. Juan, A. C. Thio-ac, M. A. B. Pamahoy, J. R. R. Forteza, and X. J. O. Garcia, “Static Sign Language Recognition Using Deep Learning,” *Int. J. Mach. Learn. Comput.*, vol. 9, no. 6, pp. 821–827, 2019,
<https://doi.org/10.18178/ijmlc.2019.9.6.879>.
- [15] M. J. Cheok, Z. Omar, and M. H. Jaward, “A review of hand gesture and sign language recognition techniques,” *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 1, pp. 131–153, 2019,
<https://doi.org/10.1007/s13042-017-0705-5>.
- [16] R. Rastgoo, K. Kiani, and S. Escalera, “Hand sign language recognition using multi-view hand skeleton,” *Expert Syst. Appl.*, vol. 150, p. 113336, 2020,
<https://doi.org/10.1016/j.eswa.2020.113336>.
- [17] Anonymus“Shoes dataset”,2022,Kaggle.<https://www.kaggle.com/datasets/utkarshsaxenadn/shoes-classification-dataset-13k-images> (accessed 10 Januari 2023)
- [18] T. R. Gadekallu et al., “Hand gesture classification using a novel CNN-crow search algorithm,” *Complex Intell. Syst.*, vol. 7, no. 4, pp. 1855–1868, 2021,
<https://doi.org/10.1007/s40747-021-00324-x>
-