JESCE

(Journal of Electrical and System Control Engineering)

Available online <http://ojs.uma.ac.id/index.php/jesce>

KLASIFIKASI DAUN TEH GAMBUNG VARIETAS ASSAMICA MENGUNAKAN CONVOLUTIONAL NEURAL NETWORK DENGAN ARSITEKTUR LENET-5

GAMBUNG TEA LEAVES CLASSIFICATION OF ASSAMICA VARIETIES USING CONVOLUTIONAL NEURAL NETWORK WITH LENET-5 ARCHITECTURE

Abdul Hafiez Suherman^{(1)*}, Nur Ibrahim⁽²⁾, Heri Syahrin⁽³⁾, Vitria Puspitasari
Rahadi⁽⁴⁾ & Muhammad Khais Prayoga⁽⁵⁾

1)Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom, Indonesia

2)Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom, Indonesia

3)Pusat Penelitian Teh dan Kina Gamboeng, Bandung, Indonesia

4)Pusat Penelitian Teh dan Kina Gamboeng, Bandung, Indonesia

5)Pusat Penelitian Teh dan Kina Gamboeng, Bandung, Indonesia

Diterima: Agustus 2020; Disetujui: November 2020; Dipublikasi: Februari 2021

*Corresponding Email: abdulhafizsuherman@gmail.com

Abstrak

Indonesia merupakan salah satu pengolahan produk teh gambung terbesar. Produk teh gambung dihasilkan dengan jenis teh yang berbeda. Namun, kualitas system pengolahan produk mengalami penurunan dikarenakan pekebun sulit membedakan jenis daun teh produksi dengan daun teh unggul dan masih menggunakan prosedur pengolahan daun secara manual. Diketahui, daun teh gambung memiliki 11 klon jenis. Daun teh GMB (1-11) merupakan klon unggul jenis teh dari jenis *assamica* maupun jenis *sinensis* dari hasil riset Pusat Penelitian Teh dan Kina (PPTK). Oleh karena itu, diperlukan teknologi pengenalan jenis daun teh sebagai peningkatan kualitas produk. Penelitian ini membuat metode klasifikasi, yaitu dengan menggunakan metode *Convolutional Neural Network* (CNN) sebagai algoritma klasifikasi. Proses klasifikasi data citra daun akan diuji dengan kelas sebanyak 11 jenis daun klon dan jumlah *dataset* diaugmentasi sebesar 4400 data. Arsitektur LeNet-5 akan digunakan pada pengujian model klasifikasi. Proses klasifikasi memperoleh hasil terbaik dengan nilai akurasi sebesar 94.55% dengan parameter *optimizer* Adam dan *learning rate* yang digunakan sebesar 0.001.

Kata Kunci: Arsitektur LeNet-5; CNN; Daun Teh GMB

Abstract

Indonesia is one of the largest processing of gambung tea products. Gambung tea products produced with different types of tea. However, the quality of the product processing system has decreased because farmers are difficult to distinguish the types of production of tea leaves with superior tea leaves and still use manual processing procedures. It is known, gambung tea leaves has 11 types of clones. GMB tea leaves (1-11) are superior tea clones of the type *assamica* and *sinensis* from the research of the Tea and Quinine Research Center (PPTK). Therefore, technology is needed to recognize the types of tea leaves as an increase in product quality. This research makes a classification by using the *Convolutional Neural Network* (CNN) method as a classification algorithm. Leaf image data classification process will be tested to 11 types of leaf clones by the number of augmented datasets is 4400 data. The LeNet-5 architecture will be used in the classification model testing. The classification process is obtained with an accuracy value of 94.55% with

Adam optimizer parameter and the learning rate used of 0.001.

Keywords: *CNN; GMB Tea Leaves; LeNet-5 Architecture*

How to Cite: Suherman, A.H., Ibrahim, N., Syahrin, H., Rahadi, V.P. & Prayoga, M.K. (2021). Klasifikasi Daun Teh Gambung Varietas Assamica Menggunakan Convolutional Neural Network Dengan Arsitektur LeNet-5. *JESCE (Journal of Electrical and System Control Engineering)*. 4 (2): 63-71

PENDAHULUAN

Daun klon gambung (GMB) memiliki 11 klon unggulan teh dari jenis varietas assamica hasil riset dari Pusat Penelitian Teh dan Kina (PPTK). Tujuan dalam penemuan riset tersebut adalah untuk meningkatkan produktivitas kebun.

Kunci utama dalam mengidentifikasi tanaman adalah karakter pada daun. Dari perbedaan karakter daun, beberapa peneliti telah mengidentifikasi objek daun dengan berbagai macam metode. Salah satunya dengan teknik *Computer Vision* (CV) dalam uji coba aplikasi dan menjadikan daun tanaman sebagai objek dalam mengidentifikasi daun (Kumar, dkk, 2012). Selanjutnya Hall, dkk (2015) telah mengevaluasi karakter data daun flavia dengan membandingkan efektifitas *Convolutional Neural Network* (CNN) dan fitur Hand-Crafted (CFH). Adapun metode *Probabilistic Neural Network* (PNN) digunakan oleh Kadir, dkk (2013) dalam menguji mengklasifikasi ukuran, warna, serta tekstur pada daun. Kalyoncu, dkk (2015) mengusulkan klasifikasi daun dengan menggunakan *Linear Discriminant Classifier* (LDC) sebagai metode terbaru dalam segmentasi ekstraksi fitur daun. Pada metode sebelumnya berbasis *Machine Learning* (ML) dan perlu pengujian lebih lanjut mengenai objek daun. Proses pengenalan fitur pada daun teh GMB memiliki berbagai jenis klon, hal itu dapat dijadikan sebagai tantangan dalam hal klasifikasi serta pengenalan karakteristik daun.

Diketahui, struktur daun bisa dijadikan pembeda antar klon (tanaman teh) dari segi struktur bentuk daun (Rahadi, dkk, 2016). Bisa dikatakan daun klon GMB memiliki keragaman morfologi, namun tingkat kemiripannya sangat tinggi dikarenakan dari proses persilangan yang sama. Dari perbedaan morfologi daun GMB, dapat mempengaruhi kualitas sistem pengolahan produk karena pekebun sulit membedakan jenis daun teh produksi dengan daun teh unggul dan masih menggunakan prosedur pengolahan daun secara manual. Diketahui produktivitas teh juga dipengaruhi oleh faktor iklim dan produk klon yang dihasilkan dapat mencapai 5000kg per tahun (Effend, dkk, 2010). Maka, untuk menghindari produksi teh yang berfluktuasi serta meningkatkan kualitas produk teh diperlukan teknologi identifikasi jenis daun teh GMB.

Pada metode penelitian dimulai dengan melakukan observasi lapangan di perkebunan teh riset PPTK daerah Ciwidey, Bandung. Proses selanjutnya dilakukan pengumpulan objek daun sesuai spesifikasi klon daun GMB lalu dilanjutkan dengan proses pemotretan daun dengan tipe format JPG menggunakan media *smartphone* serta dimasukkan ke dalam data sesuai jenis klon-nya. Data yang telah terkumpul akan dijadikan penelitian lanjutan terhadap klasifikasi dengan menggunakan CNN dan arsitektur yang digunakan hanya berfokus kepada arsitektur LeNet-5.

LANDASAN TEORI

Daun Klon teh seri GMB merupakan 11 klon unggulan teh dari jenis varietas *assamica* hasil riset dari Pusat Penelitian Teh dan Kina (PPTK). Klon seri GMB merupakan klon generasi kedua, karena klon- klon ini diperoleh dari seleksi tanaman F1 hasil persilangan yang melibatkan ketua klon generasi pertama (Sriyadi, dkk 2012). Dalam proses klasifikasi tentu memerlukan informasi identifikasi daun mengenai bentuk, tekstur, warna, serta ukuran dari klon GMB. Setyamidjaja (2000) telah menjelaskan mengenai bentuk daun berjenis *assamica* dengan ciri-ciri ukuran daun sekitar 15cm-20cm, berbentuk lonjong, berbobot, berwarna hijau tua, bergerigi banyak dan terlihat jelas pada ujungnya. Berikut bentuk daun klon GMB 1 hingga GMB 11 pada Gambar 1.

Gambar 1. Bentuk daun klon GMB 1 hingga GMB 11 sesuai urutan.

Deep learning (DL) merupakan bagian dari metode *Machine Learning* (ML) yang memiliki algoritma algoritma abstraksi tingkat tinggi menggunakan

fungsi transformasi non-linier yang ditata berlapis-lapis dan mendalam. Komunitas riset menggunakan metode DL terhadap identifikasi dengan jumlah data yang banyak dan struktur yang rumit dengan algoritma *backpropagation* dan telah menunjukkan proses perhitungan nilai representasi pada tiap lapisan (LeCun, dkk 2015).

Pada arsitektur DL memiliki *layer* yang berlapis-lapis dan subjek nilai secara keseluruhan atau sebagian besar akan dipelajari. Setiap pemetaan DL dapat memecahkan masalah yang kompleks dan lebih rumit yang terdiri dari beberapa lapisan non-linier dan memiliki kemampuan representasi fitur data secara otomatis. Jaringan saraf pada DL terbentuk dari hirarki sederhana dengan beberapa lapisan. DL menggunakan metode filter, setiap filter berbagi bobot (*weight sharing*) sehingga komputasinya jauh lebih efisien. Salah satu metode yang diunggulkan oleh DL adalah metode CNN dan metode ini sangatlah bagus dalam menemukan fitur yang baik pada citra ke lapisan berikutnya untuk membentuk hipotesis non-linier yang dapat meningkatkan selektivitas dan invarian dari representasi pada sebuah model.

Convolutional Neural Network (CNN) merupakan salah satu metode dalam pengolahan citra dua dimensi. Metode CNN merupakan pengembangan dari *Multi Layer Perception* (MLP) yang terdiri dari *feature learning* konvolusi dan klasifikasi. CNN merupakan salah satu teknik dari DL yang terdiri dari beberapa *layer* dan dirancang untuk implementasi citra dan suara. CNN telah dikembangkan pertama kali oleh

Fukushima, dkk (1982) lalu dimatangkan oleh LeCun, dkk (1998) dengan memperkenalkan arsitektur LeNet pada pengenalan tulisan angka. Berdasarkan arsitektur, CNN terdiri dari *input layer* dan *output layer* sama seperti pada *hidden layer* yang berlapis-lapis. *Hidden layer* pada CNN terdiri dari lapisan konvolusi *convolution layer* yang pada nilai bobot data diproses dengan perkalian atau perkalian titik lainnya. Nilai yang telah dikonvolusi akan dilanjutkan dengan konvolusi tambahan, yaitu *pooling layers*, *fully connected layers*, serta lapisan normalisasi. Meskipun lapisan dalam bahasa sehari-hari disebut sebagai konvolusi, ini hanya berdasarkan konvensi. Secara matematis, teknik konvolusi merupakan produk titik geser atau korelasi silang. Ini memiliki arti penting bagi indeks dalam *array*, karena memengaruhi bagaimana bobot ditentukan pada titik indeks tertentu. Pada proses konvolusi dapat dilihat pada Gambar 2.

Gambar 2. Proses CNN

Arsitektur yang akan diuji dalam pengolahan metode CNN adalah arsitektur LeNet-5. LeNet-5 adalah arsitektur berbasis CNN yang diperkenalkan oleh LeCun, dkk (1998). LeNet-5 memiliki jumlah parameter bebas atau jumlah lapisan yang lebih banyak dari sistem CNN sebelumnya.

Gambar 3. Arsitektur LeNet-5

Arsitektur LeNet-5 pada Gambar 3 terdiri dari 3 *convolutional layer*, 2 *pooling layer*, serta diakhiri dengan 2 *fully connected layer* dan 1 *full connected* dengan klasifikasi fungsi *softmax*. Arsitektur LeNet-5 telah dikembangkan dan dijadikan sebagai arsitektur dasar CNN serta telah diuji oleh para peneliti dengan data citra pengujian yaitu pengenalan tulisan angka (MNIST) dengan dengan kondisi warna *grayscale* (LeCun, dkk 1998), data CIFAR dengan kondisi warna RGB (Xie, dkk 2017), serta data rekognisi rambu lalu lintas. Penelitian tersebut teruji cukup baik dan bisa dijadikan sebagai dasar pengembangan CNN pada data yang akan diuji.

METODE PENELITIAN

Pada metode klasifikasi citra GMB dimulai dengan menentukan *dataset* citra dan dilanjutkan dengan proses ekstraksi ciri menggunakan metode CNN dengan arsitektur LeNet-5. Arsitektur yang digunakan berjumlah 6 *layer*, yaitu 3 *convolution layer*, 2 *subsampling layer*, serta 1 *fully connected layer*. *Hyperparameter* yang digunakan pada layer konvolusi yaitu dengan ukuran *kernel size* sebesar (5x5), *stride* berukuran (1x1), dan aktivasi fungsi ReLU. Sedangkan untuk *subsampling* atau *pooling layer*, *hyperparameter* yang digunakan adalah *kernel size* berukuran 2x2, dan *stride* berukuran 2x2. Selanjutnya, dilakukan proses *Fully Connected* (FC) dengan unit sebesar 84 unit serta dilanjutkan dengan FC dengan aktivasi fungsi *softmax* sehingga keluaran sebesar 11 unit sebagai evaluasi nilai jaringan dari hasil klasifikasi. Berikut konfigurasi arsitektur LeNet-5 yang akan digunakan pada Tabel 1.

Tabel 1. Tabel konfigurasi arsitektur LeNet-5

Layer		Feat ure Map	Ker nel Size	Stri de	Activa tion
Inp ut	Image	1	-	-	-
1	Convol ution 1	6	5×5	1	ReLU
2	Maxpo oling 1	6	2×2	2	-
3	Convol ution 2	16	5×5	1	ReLU
4	Maxpo	16	2×2	2	-

oling 2					
5	Convol ution 3	120	5×5	-	ReLU
6	FC	-	84	-	ReLU
Out put	FC	-	11	-	Softma x

Gambar 4. Blok diagram sistem

Metode klasifikasi data citra akan dilakukan dengan beberapa tahapan. Pada Gambar 1 menunjukkan tahapan sistem dan akan dijadikan sebagai metode penelitian. Tahap pertama yaitu akuisisi citra yang merupakan tahap pengambilan data citra klon teh seri GMB dan dilakukan penyesuaian jenis spesifikasi klon GMB ke dalam 11 kelas. Pada data citra GMB diambil dengan menggunakan kamera *smartphone* dengan spesifikasi ukuran sebesar 13, 25, dan 48 *megapixel*. Tahap kedua merupakan proses *preprocessing*, yaitu mengubah ukuran data citra dan melakukan proses data augmentasi. Setelah dilakukan proses *preprocessing*, akan dilanjutkan ke tahap ketiga dengan proses klasifikasi data latih dan data uji. Pada data latih digunakan untuk memperoleh pencapaian uji data citra, sedangkan data uji digunakan untuk menguji performa dan kebenaran atau korelasi data latih dalam model citra tersebut.

HASIL DAN PEMBAHASAN

Data latih dan data uji masing-masing berjumlah 1100 dan 880. Pada data latih akan diaugmentasi sebelum melakukan proses klasifikasi. Ukuran Data citra yang akan diklasifikasi sebesar 64×64 . Klasifikasi pemodelan sistem akan dilatih dengan menggunakan *optimizer* Adam dengan *learning rate* 0.001 dan mencari nilai akurasi, *loss*, dan *F1-score* terbaik terhadap pengaruh dari parameter tersebut.

Tabel 2. Hasil Akurasi validasi dengan data augmentasi dan parameter terbaik.

Model + Augmentasi Data + Optimizer Adam (lr = 0.001)					
Epoch	Valid ation Accur acy	Validat ion Loss	Prec ision	Recal l	F1- Sco re
7	70.45 %	29.55 %	73%	71%	70 %
10	85.90 %	14.1%	87%	87%	86 %
13	94.55 %	5.45%	96%	94%	94 %

Gambar 5. Grafik Peningkatan Akurasi Training dan Akurasi Validasi Data Augmentasi dengan Optimizer Adam dan Learning Rate 0.001

Gambar 6. Grafik penurunan *loss training* dan *loss validasi* data augmentasi dengan *optimizer* Adam dan *learning rate* 0.001.

Tabel 2 menunjukkan bahwa dengan epoch 13 memperoleh nilai akurasi validasi sebesar 94.55%. Nilai *F1-score* diperoleh sebesar 94%, maka tingkat prediksi klasifikasi yang diberikan oleh sistem sudah sangat baik. data citra GMB dengan klon kelas 1 hingga 11 dikatakan sudah baik dalam waktu proses klasifikasi selama 39 detik mampu memperoleh nilai akurasi di atas 90%. Dengan kata lain, proses klasifikasi

model tersebut terhadap arsitektur terbilang sangat optimal. Grafik peningkatan akurasi pada Gambar 5 menunjukkan performansi akurasi validasi dengan sangat baik dilihat perbedaan nilai akurasi *training* dengan akurasi validasi. Pada Gambar 6 menunjukkan *loss training* dan validasi tiap *epoch*. Dapat dilihat bahwa tidak ada pengaruh *overfitting*, dan kurva *loss* validasi selalu mengikuti kurva *loss training*.

Gambar 7. Nilai prediksi *confusion matrix* pada tiap kelas GMB.

Apabila dilihat dari penampakan tabel *confusion matrix* pada Gambar 7, data tiap kelas yang diprediksi terhadap data aktual memperoleh nilai yang sangat baik dan sedikit memperoleh kesalahan pada data yang diprediksi.

SIMPULAN

Penelitian ini telah mengusulkan metode klasifikasi daun teh GMB (1-11) menggunakan algoritma CNN dengan arsitektur dasar LeNet-5 dan memperoleh hasil akurasi di atas 90%. Hasil akurasi yang diperoleh sebesar 94.55% pada *epoch* 13 dengan *optimizer*

Adam (*learning rate* 0.001) dan F1-score sebesar 94%.

Adapun beberapa saran peneliti terhadap arsitektur yang telah diuji. Pertama, untuk memperoleh akurasi yang lebih baik lagi dan lebih signifikan apabila jika ditambahkan jumlah *dataset* citra GMB. Kedua, perlu penambahan parameter pada pengujian data citra untuk memperoleh hasil perbandingan yang lebih kompleks.

DAFTAR PUSTAKA

- Kumar, N., Belhumeur, P. N., Biswas, A., Jacobs, D. W., Kress, W. J., Lopez, I. C., & Soares, J. V. 2012. Leafsnap: A computer vision system for automatic plant species identification. In *European conference on computer vision* (pp. 502-516). Springer, Berlin, Heidelberg.
- Hall, D., McCool, C., Dayoub, F., Sunderhauf, N., & Upcroft, B. 2015. Evaluation of features for leaf classification in challenging conditions. In *2015 IEEE Winter Conference on Applications of Computer Vision* (pp. 797-804). IEEE.
- Kadir, A., Nugroho, L. E., Susanto, A., & Santosa, P. I. 2013. Leaf classification using shape, color, and texture features. *arXiv preprint arXiv:1401.4447*.
- Kalyoncu, C., & Toygar, Ö. 2015. Geometric leaf classification. *Computer Vision and Image Understanding*, 133, 102-109.
- Rahadi, V. P., Khomaeni, H. S., Chaidir, L., & Martono, B. 2016. Keragaman dan kekerabatan genetik koleksi plasma nutfah teh berdasarkan karakter morfologi daun dan komponen hasil. *Jurnal Tanaman*

- Industri Dan Penyegar, 3(2): 103-108.
- Effendi, I. D. S., Syakir, M., Yusron, M., Pelaksana, R., Jusniarti, I., Budiharto, A., & Undang-undang, H. C. D. 2010. *Budidaya dan pasca panen teh. Bogor: Pusat Penelitian dan Pengembangan Perkebunan.*
- Sriyadi, B., Suprihatini, R., & Khomaeni, H. S. 2012. The development of high yielding tea clones to increase Indonesian tea production. In *Global Tea Breeding* (pp. 299-308). Springer, Berlin, Heidelberg.
- Setyamidjaja, D. 2000. *Teh Budidaya & Pengolahan Pascapanen.* Kanisius.
- LeCun, Y., Bengio, Y., & Hinton, G. 2015. Deep learning. *nature*, 521(7553): 436-444.
- Fukushima, K., & Miyake, S. 1982. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets* (pp. 267-285). Springer, Berlin, Heidelberg.
- LeCun, Y., Cortes, C., & Burges, C. J. 1998. The MNIST database of handwritten digits, 1998. URL <http://yann.lecun.com/exdb/mnist>, 10(34), 14.
- Xie, Y., Jin, H., & Tsang, E. C. 2017. Improving the lenet with batch normalization and online hard example mining for digits recognition. In *2017 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)* (pp. 149-153). IEEE.
- Xie, L., Wang, J., Wei, Z., Wang, M., & Tian, Q. 2016. *Disturb label: Regularizing cnn on the loss layer.* In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Juni 2016.
- Kingma, D. P., & Ba, J. 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Schulz, H., & Behnke, S. 2012. Deep learning. *KI-Kunstliche Intelligenz*. 26(4): 357-363.