

Penerapan Deep Learning untuk Identifikasi Citra Tanaman Obat Anti-inflamasi Menggunakan Algoritma CNN

Triase¹, Melfa Aisyah Hutasuht²

triase@uinsu.ac.id, melfa_aisyah@uinsu.ac.id

¹ Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara

² Program Studi Biologi Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara

Informasi Artikel

Diterima : 22 Okt 2023
Direview : 21 Nov 2023
Disetujui : 20 Des 2023

Kata Kunci

Deep Learning, CNN,
Tanaman Obat, Epoch

Abstrak

Deep Learning bagian kecerdasan buatan dan big data menggunakan Machine Learning solusi paling menjanjikan untuk masalah big data. Machine Learning adalah metode matematis yang diterapkan untuk menciptakan model yang menggambarkan pola-pola dalam data. Tanaman obat adalah jenis tumbuhan yang diketahui memiliki manfaat dalam menjaga kesehatan dan mengobati penyakit. Penggunaan tumbuhan obat sering kali terkait erat dengan pengobatan tradisional, karena sebagian besar pemahaman tentang manfaatnya tidak didasarkan pada uji klinis laboratorium, melainkan berdasarkan pengalaman penggunaan yang telah teruji dari masa ke masa. Untuk memaksimalkan penggunaan tanaman obat dengan mengenali daun tanaman obat ini, maka deteksi citra daun anti-inflamasi akan diuji menggunakan algoritma CNN dengan tahapan scraping data dari bing, preprocessing data secara manual dan pelatihan data dengan 10 epoch berhasil identifikasi dan mengklasifikasikan dengan benar sekitar 16% dari semua instance positif yang ada dari data pengujian dan merupakan tingkat keberhasilan yang rendah, sehingga memiliki kinerja yang tidak memuaskan dalam tugas identifikasi.

Keywords

Deep Learning, CNN,
Medicinal Plants, Epoch

Abstrak

Deep Learning, part of artificial intelligence and big data, uses machine learning, the most promising solution to big data problems. Machine learning is a mathematical method applied to create models that describe patterns in data. Medicinal plants are types of plants that are known to have benefits in maintaining health and treating disease. The use of medicinal plants is often closely related to traditional medicine, because most of the understanding of their benefits is not based on clinical laboratory tests, but rather based on experience of use that has been tested over time. To maximize the use of medicinal plants by recognizing the leaves of these medicinal plants, image detection of anti-inflammatory leaves will be tested using the CNN algorithm with stages of data scraping from Bing, manual data preprocessing and data training with 10 epochs, successfully identifying and classifying approximately 16% correctly. of all existing positive instances of the test data and represents a low success rate, thus having unsatisfactory performance in the identification task.

A. Pendahuluan

Era 4.0 saat ini salah satu yang sangat populer dari perkembangan teknologi adalah inovasi AI (Artificial Intelligence) dan big data menggunakan machine learning. Machine Learning adalah metode matematis yang diterapkan untuk menciptakan model yang menggambarkan pola-pola dalam data. Dengan memanfaatkan data yang telah ada, model dibangun secara tidak langsung untuk kemudian digunakan dalam proses identifikasi.[1]. Banyak sekali Machine Learning telah menyelesaikan permasalahan hidup manusia, contohnya penerapan Machine Learning berbasis AI (Artificial Intelligence) untuk pendeteksi batuk dan bersin[2] dan selain itu juga untuk mendeteksi tanaman obat seperti yang dilakukan pada riset ini.

Tanaman obat atau disebut tanaman herba merupakan sumber bahan dasar obat tradisional maupun modern yang jika digunakan menyebabkan bertambahnya kekebalan tubuh. Tanaman obat adalah jenis tumbuhan yang telah diketahui memiliki manfaat dalam menjaga kesehatan dan mengobati penyakit. Penggunaan tumbuhan obat sering kali terkait erat dengan pengobatan tradisional, karena sebagian besar pemahaman tentang manfaatnya tidak didasarkan pada uji klinis laboratorium, melainkan berdasarkan pengalaman penggunaan yang telah teruji dari masa ke masa[3]. Konsumsi tanaman obat dapat memberikan tindakan preventif bagi tubuh dalam menghadapi berbagai jenis penyakit. Dan memiliki efek samping yang rendah dan berbahaya bagi tubuh jika dikonsumsi dengan takaran yang tepat. Berkembangnya teknologi telah mendorong industri farmasi untuk menggabungkan berbagai tumbuhan yang dijaga kerahasiaannya sebagai sumber bahan obat, termasuk obat antimalaria, obat sariawan, obat anti diare, dan obat anti hipertensi[4].

Inflamasi adalah proses fisiologis yang melibatkan respons sistem kekebalan tubuh. Peran utamanya adalah untuk melindungi organisme dari infeksi mikroba, dan dalam beberapa kasus, berfungsi sebagai mekanisme pertahanan fisiologis terhadap penyakit tertentu seperti kanker.[5] Inflamasi adalah respons tubuh terhadap infeksi, iritasi, atau cedera yang menghasilkan tanda-tanda seperti kulit merah, panas, pembengkakan, dan rasa sakit.[6] Cara kerjanya dapat menghancurkan, mengurangi dan melokalisasi agen pencedera ataupun jaringan yang cedera. Tujuan pengobatan inflamasi meringankan rasa nyeri dengan frekuensi sering pada gejala awal, memperlambat dan mencegah penyebaran proses kerusakan jaringan.

Indonesia yang dikenal dengan melimpahnya jenis tanaman obat dengan fungsinya yang berbeda akan tetapi disayangkan masyarakat umum khususnya generasi muda kesulitan dalam mengenali dan mempelajari jenis tanaman obat karena secara morfologi banyak jenis daun tanaman obat yang mirip. Hal itu disebabkan pada umumnya masyarakat hanya mengenali buahnya saja. Untuk memaksimalkan penggunaan tanaman obat dengan mengenali dari daun tanaman obat ini, maka deteksi citra daun anti-inflamasi ini sangat bermanfaat sekali.

Metode yang sangat baik digunakan untuk mendeteksi citra tanaman obat adalah Deep Learning. Deep Learning bagian kecerdasan buatan dan big data menggunakan Machine Learning solusi yang paling menjanjikan untuk masalah big data.[7] karena kemampuannya yang besar dalam mengekstraksi informasi penting dari kumpulan data yang besar. Library Python dirancang khusus agar

dapat menciptakan mesin mampu berpikir, belajar, dan memecahkan masalah. Meskipun deep learning membutuhkan waktu lama untuk melatih model karena jumlah parameter yang besar, namun ia memerlukan waktu yang singkat untuk dijalankan saat pengujian dibandingkan dengan algoritma machine learning lainnya [8]. Pada saat melatih deep learning menggunakan epoch. Sebuah epoch dalam deep learning berfungsi untuk mewakili satu proses lengkap melalui seluruh set data pelatihan selama fase pelatihan dari neural network. Ini adalah ukuran berapa kali algoritma telah melihat seluruh set data pelatihan. Meningkatkan jumlah epoch secara potensial dapat meningkatkan akurasi model, tetapi juga meningkatkan risiko overfitting [9].

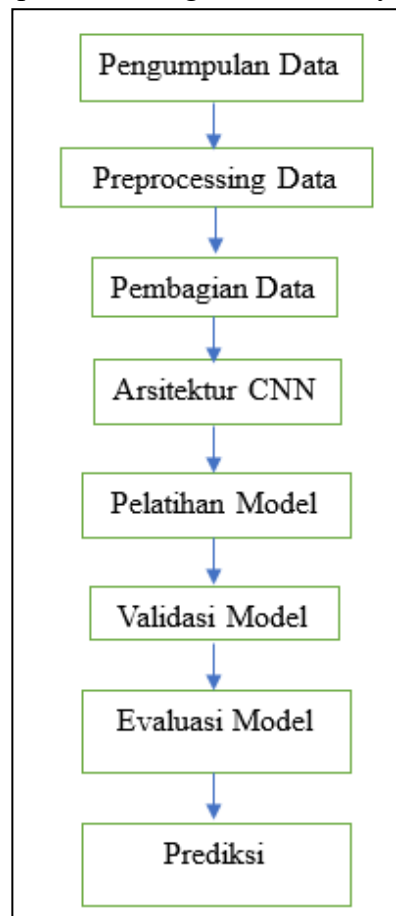
Penelitian ini dikembangkan berdasarkan dari penelitian terdahulu yang pernah diriset oleh Ilahiyah dan Nilogiri[10] yang berjudul "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network". Penelitian tersebut membahas algoritma deep learning untuk CNN dibuat dengan prinsip translation invariance yaitu dapat mengenali objek dalam citra daun pada berbagai macam posisi yang mungkin. Sedangkan riset ini akan melakukan identifikasi citra daun tanaman obat dengan menggunakan CNN. Convolutional Neural Network (CNN) adalah jenis model pembelajaran mendalam yang banyak digunakan untuk tugas-tugas yang melibatkan analisis dan prediksi data, terutama dalam bidang seperti visi komputer dan pengolahan gambar [11]. CNN cocok untuk memproses data dua dimensi dan telah menunjukkan performa yang kuat dalam berbagai aplikasi. CNN memiliki beberapa keunggulan dalam pengolahan gambar, termasuk akurasi tinggi, kinerja yang signifikan dalam deteksi dan lokalisasi objek, kemampuan untuk mendeteksi dan mengklasifikasikan gambar, kemampuan untuk mengidentifikasi pola dan fitur, pembelajaran transfer, pemrosesan real-time, dan penggunaan dataset yang diperkaya[12], [13].

B. Metode Penelitian

Metode Panduan peneliti diperlukan dirancang untuk kredibilitas dan konsistensi dalam penelitian. Untuk itu kerangka metodologi penelitian menggunakan Convolutional Neural Network (CNN) dapat mencakup langkah-langkah sebagai berikut: *scraping data dari bing*

1. Pengumpulan Data: mengumpulkan gambar tanaman obat untuk diidentifikasi. Dataset tanaman obat dikumpulkan melalui scraping data dari bing dan observasi lapangan di galeri tanaman mak ni kota medan.
2. Preprocessing Data: Preprocessing data dilakukan sebelum data diproses di dalam arsitektur CNN, dengan cara membersihkan dataset gambar tanaman obat yang membuat hasil pelatihan untuk tidak overfitting
3. Pembagian Dataset: membagi dataset menjadi tiga subset: pelatihan (training), validasi, dan pengujian (testing). Sebagian besar data digunakan untuk pelatihan, sedikit untuk validasi (untuk penyetelan model), dan yang tersisa untuk pengujian (untuk evaluasi kinerja akhir)
4. Arsitektur CNN: Arsitektur CNN ini mencakup banyak lapisan konvolusi, lapisan pooling, lapisan terhubung sepenuhnya, dan fungsi aktivasi yang digunakan dalam model.
5. Pelatihan Model: Melatih model CNN menggunakan dataset gambar tanaman obat. Proses ini melibatkan penyesuaian parameter model untuk menghasilkan

- hasil yang optimal. Selama pelatihan, menggunakan teknik normalisasi batch, dan optimasi untuk menghindari overfitting.
6. Validasi Model: Menggunakan subset folder validasi untuk memantau kinerja model selama pelatihan. Dilakukan untuk menyesuaikan parameter model dan mencegah overfitting.
 7. Evaluasi Model: Setelah model telah dilatih, evaluasi kinerjanya menggunakan dataset gambar tanaman obat dengan folder pengujian yang terpisah. Menggunakan Metrik evaluasi seperti akurasi, presisi, recall, F1-score, dan kurva ROC untuk melihat hasil evaluasinya
 8. Hasil prediksi: Hasil prediksi menginformasikan yang ditemukan oleh CNN.



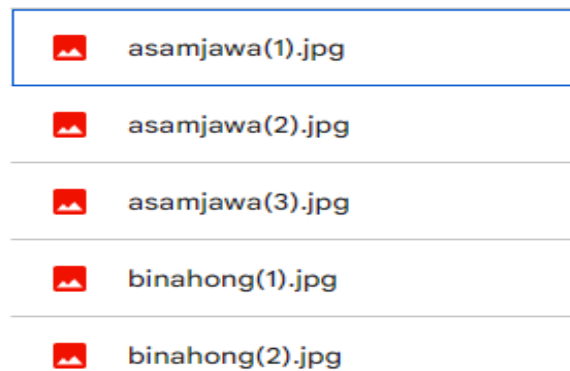
Gambar 1. Kerangka penelitian

C. Hasil dan Pembahasan

Berdasarkan gambar 1. Tahap awal supaya dataset yang dibutuhkan pada analisis algoritma CNN dapat tercapai, yaitu mendapatkan dataset gambar tanaman obat. Dataset ini diperoleh melalui internet dengan teknik scraping dari website bing. Tahapan dalam pengumpulan data dari link bing dengan membuka alamat www.bing.com dan melakukan pencarian data berdasarkan jenis tanaman obat anti-inflamasi yang dibutuhkan. Kemudian data tersebut di download secara realtime menggunakan Bahasa pemrograman pyhon berasal dari google colab yang menjalankan Deep Learning dan Kode Pembelajaran Mesin. Google telah membuat Google Colaboratory yang bersifat online dapat mengunduh kumpulan

data yang lebih besar langsung dari server ke google drive dengan kecepatan sangat tinggi.[14]. Selain dari website juga dilakukan observasi lapangan dengan memfoto objek menggunakan kamera hp menjadi file jpg. Pengumpulan dataset memperoleh dataset sebanyak 800 tanaman obat anti-inflamasi dengan jenis tanaman obat sebanyak 19 jenis.

Tahapan selanjutnya melakukan preprocessing data dengan cara memilih gambar daun yang terbaik hasilnya disebut dengan preprocessing image, hal itu dilakukan untuk menghindari overfitting. Selanjutnya dataset disimpan dalam bentuk pelabelan dengan memberikan nama pada jenis tanaman obat. Kemudian membuat folder latih dan validasi untuk dapat dilatih menggunakan algoritma CNN.



Gambar 2. Pelabelan Pada Gambar Daun Tanaman Obat

Contoh hasil dari klining data setelah dataset gambar dideteksi oleh Python adalah sebagai berikut:



Gambar 3. Sebelah kanan gambar original dan sebelah kiri gambar setelah diklining

Penggunaan arsitektur CNN (Convolutional Neural Network) adalah data daun tanaman obat melalui proses convulation, maxpolling dan relu. Maka dilakukan pelatihan (fitting) dengan hasil dari arsitektur yang dieksekusi oleh google colab adalah sebagai berikut:

```
CustomModel (Input shape: 32 x 3 x 226 x 226)
```

Layer (type)	Output Shape	Param #	Trainable
Conv2d	32 x 16 x 226 x 226	448	True
Conv2d	32 x 32 x 113 x 113	4640	True
Conv2d	32 x 64 x 56 x 56	18496	True
Conv2d	32 x 128 x 28 x 28	73856	True
Linear	32 x 512	12845568	True
Linear	32 x 19	9747	True

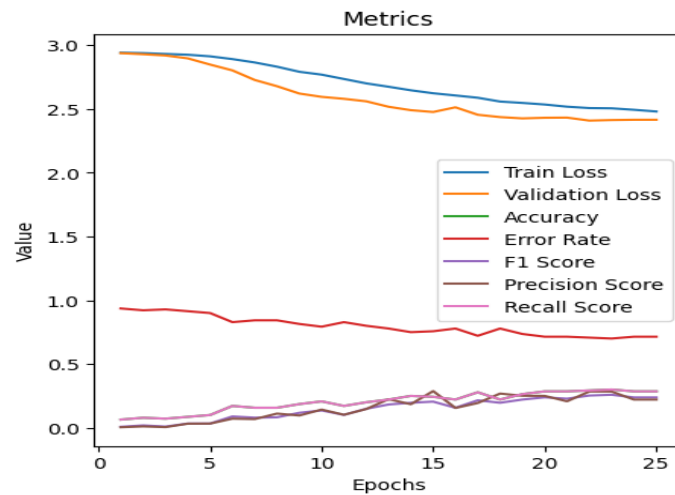
Total params: 12,952,755
 Total trainable params: 12,952,755
 Total non-trainable params: 0

Gambar 4. Arsitektur CNN

Penjelasan dari gambar 4 adalah Input shape "32 x 3 x 226 x 226" mengindikasikan karakteristik dari data masukan yang digunakan dalam jaringan saraf tiruan. Dalam hal ini:

- "32" adalah jumlah sampel (batch size) dalam setiap iterasi saat data dimasukkan ke dalam model.
- "3" mengacu pada jumlah saluran warna atau saluran gambar. Ini mengindikasikan bahwa data masukan adalah gambar berwarna, di mana setiap piksel memiliki tiga saluran warna (misalnya, merah, hijau, biru).
- "226 x 226" menggambarkan dimensi spasial (ruang) gambar, yaitu lebar dan tinggi gambar dalam piksel yang dibaca oleh arsitektur
- Total parameter "12,952,755" mengindikasikan jumlah parameter (bobot dan bias) yang ada dalam model jaringan saraf tiruan. Parameter-parameter ini akan digunakan dalam operasi konvolusi, normalisasi, dan lapisan-lapisan lain dalam model untuk memproses dan mempelajari informasi dari data masukan. Total parameter ini biasanya dipengaruhi oleh arsitektur jaringan dan ukuran setiap lapisan. Semakin banyak parameter, semakin rumit dan kuat kemampuan model, tetapi juga akan memerlukan lebih banyak data dan sumber daya komputasi untuk melatihnya.

Langkah selanjutnya adalah pelatihan model dengan melakukan 25 epoch yaitu hasilnya dapat dilihat pada gambar grafik berikut:



Gambar 6. Pelatihan dengan 25 Epoch

Penjelasan dari gambar 6 uji coba dengan 25 epoch, namun hanya mendapatkan accuracy paling tinggi sebanyak 26% dari data pengujian dan merupakan tingkat keberhasilan yang rendah. Model dievaluasi menggunakan confusion metrics, dengan hasil dapat dilihat pada gambar berikut :

Daun Asam Jawa	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	
Daun Binahong	0	3	1	0	1	1	0	0	0	1	0	1	0	0	0	0	0	0	
Daun Jarak	0	1	3	0	0	0	2	0	0	1	1	0	1	0	0	0	0	0	
Daun Jeruk Purut	0	1	1	0	0	1	0	0	0	0	1	0	1	0	0	2	0	0	
Daun Jintan	0	0	0	0	1	1	0	0	0	0	1	0	3	0	0	0	0	2	
Daun Kari	0	1	0	0	0	1	0	0	0	1	3	0	0	0	0	1	0	0	
Daun Kelor	0	1	3	0	0	1	0	1	0	0	0	0	0	0	2	0	0	0	
Daun Kemangi	0	0	1	0	0	0	0	2	0	0	0	0	4	0	1	0	0	0	
Daun Kunyit	0	0	3	0	0	0	0	0	0	1	0	0	2	0	1	0	0	0	
Daun Lidah Buaya	0	0	0	0	0	0	1	0	0	5	0	0	1	0	1	0	0	0	
Daun Lidah Mertua	0	1	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	
Daun Mangga	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	3	0	1	
Daun Mint	0	0	1	1	1	0	0	0	1	0	0	0	4	0	1	0	0	1	
Daun Pandan	0	0	0	0	0	0	0	1	1	1	0	0	2	0	3	0	0	0	
Daun Pepaya	0	0	0	0	0	0	0	0	0	0	0	0	2	0	3	2	1	0	
Daun Salam	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	5	0	0	
Daun Singkong	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	2	2	
Daun Teh Hijau	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	2	
Daun sambiloto	0	0	0	0	0	0	0	2	0	1	0	0	1	0	0	0	0	2	
	Daun Asam Jawa	Daun Binahong	Daun Jarak	Daun Jeruk Purut	Daun Jintan	Daun Kari	Daun Kelor	Daun Kemangi	Daun Kunyit	Daun Lidah Buaya	Daun Lidah Mertua	Daun Mangga	Daun Mint	Daun Pandan	Daun Pepaya	Daun Salam	Daun Singkong	Daun Teh Hijau	Daun sambiloto
Actual																			
								</											

Gambar 7. Confusion Matrix

Berdasarkan gambar 7 bahwasannya banyak jenis daun yang tidak terdeteksi kemiripannya, Hasil confusion matriksnya dari data validasi yang salah mengenali adalah daun binahong bernilai 1 dikenali sebagai daun teh hijau dan dikenali sebagai daun lidah mertua, sebagai daun kelor, daun kari, daun jeruk perut dan daun ajrak bernilai 1 dikenali sebagai daun kari dan beberapa daun yang lainnya.

D. Simpulan

Berdasarkan hasil dari penggunaan arsitektur CNN dari tahap preprocessing dan pelatihan model dan evaluasi maka arsitektur CNN yang diteliti menghasilkan nilai akurasi rendah karena tidak didampingi menggunakan model arsitektur yang lebih baik, sehingga kemungkinan untuk menghasilkan akurasi yang tinggi diperlukan epoch lebih dari 25 kali sehingga, sesuai dengan teori CNN bahwa CNN merupakan arsitektur dengan akurasi yang tinggi dapat tercapai, kemudian dalam melakukan testing pada riset ini dataset daun diuji satu persatu tidak menggunakan teknik otomatis atau split sehingga ini juga mempengaruhi hasil akurasi yang rendah.

Pada penelitian selanjutnya disarankan untuk menambah jumlah dataset daunnya dan menggunakan arsitektur pendukung CNN serta melakukan epoch lebih banyak atau menggunakan framework yang lebih baik lagi.

E. Ucapan Terima Kasih

Kami ingin mengucapkan terima kasih kepada semua individu khususnya mahasiswa kami yang telah membantu penelitian Dzira Faza Zarasky dan lembaga yang telah memberikan bantuan dalam penyusunan artikel ini. Artikel ini dibiayai oleh Dana Penelitian dari BOPTN Universitas Islam Negeri Sumatera Utara Medan pada tahun 2022.

F. Referensi

- [1] K. Triase, sriani, "Indonesian Journal of Computer Science," vol. 11, no. 1, pp. 566-576, 2022.
- [2] S. B. Mulia, N. Wisma Nugraha, M. H. Robbani, T. Otomasi, M. & Mekatronika, and P. Manufaktur Bandung, "Implementasi Machine Learning Untuk Identifikasi Orang Batuk/Bersin," *J. Energy Electr. Eng.*, vol. 81, no. 2, pp. 81-86, 2023.
- [3] D. Sahila, T. I. Wahyuni, and Z. Arifin, "Sosialisasi Pemanfaatan Tanaman Obat Keluarga Guna Meningkatkan Upaya Kesehatan Mandiri Warga Di Dusun Sangiang Desa Kumbang Kecamatan Masbagik Lombok Timur," pp. 4-6, 2023.
- [4] D. Nita Utami, D. Rosanti, and T. Kartika, "Karakteristik Morfologi Jenis-Jenis Tanaman Obat Di Kelurahan Prabujaya Kecamatan Prabumulih Timur Kota Prabumulih," *Indobiosains*, vol. 5, no. 2, pp. 56-65, 2023, doi: 10.31851/indobiosains.v5i2.9153.
- [5] A. Bouyahya *et al.*, "Anti-inflammatory and analgesic properties of Moroccan medicinal plants: Phytochemistry, in vitro and in vivo investigations, mechanism insights, clinical evidences and perspectives," *J. Pharm. Anal.*, vol. 12, no. 1, pp. 35-57, 2022, doi: 10.1016/j.jpha.2021.07.004.
- [6] E. Emelda, R. Nugraeni, and K. Damayanti, "Review: Exploration of

- Indonesian Herbal Plants for Anti Inflammatory," *INPHARMED J. (Indonesian Pharm. Nat. Med. Journal)*, vol. 6, no. 2, p. 58, 2023, doi: 10.21927/inpharmed.v6i2.1938.
- [7] M. Saber and E. M.El-Kenawy, "Design and implementation of accurate frequency estimator depend on deep learning," *Int. J. Eng. Technol.*, vol. 9, no. 2, p. 367, 2020, doi: 10.14419/ijet.v9i2.30473.
- [8] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 6, pp. 1–20, 2021, doi: 10.1007/s42979-021-00815-1.
- [9] S. Lim, "Estimation of gender and age using CNN-based face recognition algorithm," *Int. J. Adv. Smart Converg.*, vol. 9, no. 2, pp. 203–211, 2020, [Online]. Available: <http://dx.doi.org/10.7236/IJASC.2020.9.2.203>
- [10] S. Ilahiyah and A. Nilogiri, "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network," pp. 49–56, 2000.
- [11] A. R. F. Dewandra, A. P. Wibawa, U. Pujiyanto, A. B. P. Utama, and A. Nafalski, "Journal Unique Visitors Forecasting Based on Multivariate Attributes Using CNN," *Int. J. Artif. Intell. Res.*, vol. 6, no. 2, 2022, doi: 10.29099/ijair.v6i1.274.
- [12] A. Desiani, M. Erwin, B. Suprihatin, S. Yahdin, A. I. Putri, and F. R. Husein, "Bi-path Architecture of CNN Segmentation and Classification Method for Cervical Cancer Disorders Based on Pap-smear Images," *IAENG Int. J. Comput. Sci.*, vol. 48, no. 3, pp. 1–9, 2021.
- [13] Z. Tasnim *et al.*, "Deep Learning Predictive Model for Colon Cancer Patient using CNN-based Classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 8, pp. 687–696, 2021, doi: 10.14569/IJACSA.2021.0120880.
- [14] P. Kanani and M. Padole, "Deep learning to detect skin cancer using google colab," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 2176–2183, 2019, doi: 10.35940/ijeat.F8587.088619.