

Penerapan Arsitektur EfficientNet Untuk Pembuatan Model Algoritma Convolutional Neural Network Pada Klasifikasi Bahasa Isyarat

Amanda Kanaya Putri*, Suroso, Ade Silvia Handayani

Teknik Elektro, Program Studi Teknik Telekomunikasi, Politeknik Negeri Sriwijaya, Palembang, Indonesia

Email: amandakanaya51@gmail.com, osorus11@gmail.com, ade_silvia@polsri.ac.id

Email Penulis Korespondensi: amandakanaya51@gmail.com

Submitted: 17/07/2024; Accepted: 08/09/2024; Published: 09/09/2024

Abstrak—Penelitian ini membahas penerapan arsitektur EfficientNet dalam pengembangan model Convolutional Neural Network (CNN) untuk klasifikasi bahasa isyarat. Bahasa isyarat merupakan metode komunikasi vital bagi komunitas tunarungu, namun pengenalan otomatisnya masih menjadi tantangan dalam bidang computer vision. Salah satu permasalahan utama adalah keterbatasan akurasi dan efisiensi model dalam mengenali variasi kompleks dari bahasa isyarat dalam kondisi dunia nyata. EfficientNet, yang dikenal dengan efisiensi komputasinya, digunakan sebagai backbone untuk membangun model CNN yang mampu mengklasifikasikan pola huruf bahasa isyarat dengan akurasi tinggi namun tetap ringan. Dataset yang digunakan dalam penelitian ini adalah American Sign Language (ASL) dengan teknik augmentasi data untuk meningkatkan variasi dan kualitas dataset. Dataset yang digunakan terdiri dari 14.740 gambar untuk keseluruhan pola huruf bahasa isyarat dari berbagai sudut dan kondisi pencahayaan. Hasil eksperimen menunjukkan bahwa model berbasis EfficientNet yang dikembangkan mencapai akurasi pelatihan dan akurasi validasi sebanyak 98.40% dengan ukuran model dan waktu inferensi yang lebih efisien. Penelitian ini memperlihatkan potensi signifikan dari penggunaan EfficientNet dalam pengembangan sistem klasifikasi bahasa isyarat yang dapat diterapkan pada perangkat dengan sumber daya terbatas, seperti aplikasi mobile dan edge computing. Temuan ini diharapkan dapat meningkatkan aksesibilitas dan inklusi sosial bagi komunitas tunarungu dan tunawicara. Dengan demikian, penelitian ini tidak hanya berkontribusi pada bidang teknologi pengenalan pola, tetapi juga pada upaya peningkatan kualitas hidup bagi penyandang disabilitas komunikasi, melalui pengembangan alat bantu yang efektif dan efisien.

Kata Kunci: Augmentasi Data; Bahasa Isyarat; CNN; Klasifikasi; EfficiencyNet

Abstract—This study discusses the application of EfficientNet architecture in developing a Convolutional Neural Network (CNN) model for sign language classification. Sign language is a vital communication method for the deaf community, but automatic recognition remains a challenge in the field of computer vision. One of the primary issues is the limitation in accuracy and efficiency of models in recognizing complex variations of sign language in real-world conditions. EfficientNet, known for its computational efficiency, is used as a backbone to build a CNN model that can classify sign language letter patterns with high accuracy while remaining lightweight. The dataset used in this study is American Sign Language (ASL) with data augmentation techniques to enhance the variety and quality of the dataset. The dataset comprises 14,740 images of sign language letter patterns from various angles and lighting conditions. Experimental results show that the EfficientNet-based model developed achieves training and validation accuracies of 98.40% with a more efficient model size and inference time. This study demonstrates the significant potential of using EfficientNet in developing sign language classification systems that can be applied to devices with limited resources, such as mobile applications and edge computing. These findings are expected to improve accessibility and social inclusion for the deaf and speech-impaired communities. Thus, this research not only contributes to the field of pattern recognition technology but also to efforts to enhance the quality of life for individuals with communication disabilities through the development of effective and efficient assistive tools.

Keywords: Data Augmentation; Sign Language; CNN; Classification; EfficiencyNet

1. PENDAHULUAN

Penyandang tunawicara dan tunarungu menghadapi hambatan besar dalam komunikasi yang disebabkan oleh berbagai faktor. Masalah pada organ bicara, keterlambatan perkembangan bahasa, atau kerusakan sistem pendengaran dapat signifikan mengurangi kemampuan mereka untuk berkomunikasi secara efektif. Bagi banyak dari mereka, bahasa isyarat menjadi solusi utama dalam menyampaikan pesan dan berinteraksi dengan lingkungan sekitar [1]. Bahasa isyarat menggunakan gerakan tangan, ekspresi wajah, dan bahasa tubuh untuk mengkomunikasikan ide dan emosi, menjadi jembatan vital bagi komunikasi mereka [2]. Namun, kendala besar muncul karena tidak semua individu memahami atau dapat menggunakan bahasa isyarat secara aktif [3]. Hal ini menyebabkan penyandang tunarungu dan tunawicara sering merasa terisolasi dalam masyarakat, mengalami kesulitan dalam mendapatkan informasi, dan menghadapi tantangan besar dalam pendidikan dan pekerjaan [4]. Penelitian ini memiliki urgensi tinggi karena bertujuan mengatasi masalah tersebut melalui pengembangan teknologi pengenalan bahasa isyarat otomatis menggunakan algoritma Convolutional Neural Network.

Kemajuan teknologi, khususnya dalam bidang *Machine Learning* dan *Computer Vision*, telah membuka pintu baru dalam mengatasi tantangan komunikasi yang dihadapi oleh penyandang tunarungu dan tunawicara [5]. Salah satu teknologi yang menjadi sorotan adalah *Convolutional Neural Network* (CNN), sebuah algoritma *Deep Learning* yang terinspirasi dari cara kerja sistem visual biologis [6]. CNN telah terbukti sangat efektif dalam tugas-tugas pengolahan gambar yang kompleks dengan kemampuannya untuk secara otomatis mengekstrak fitur-fitur penting dari data visual, CNN mampu mengenali pola-pola yang rumit dan melakukan klasifikasi dengan tingkat akurasi yang tinggi. Kelebihan ini menjadikannya kandidat ideal untuk diterapkan dalam pengenalan bahasa isyarat [7]. Berbeda dengan metode tradisional yang mengharuskan ekstraksi fitur secara manual, CNN belajar untuk mengenali representasi

hierarkis dari data gambar, sehingga mampu mengenali gerakan tangan dan ekspresi wajah dengan lebih akurat dan andal [8].

Selain itu, kemampuan CNN untuk beradaptasi dengan variasi posisi, skala, dan orientasi gerakan tangan menjadikannya solusi yang sangat fleksibel untuk mengatasi kompleksitas dan variasi dalam bahasa isyarat [9]. Pendekatan ini tidak hanya meningkatkan efektivitas dalam mengenali berbagai bentuk dan gaya isyarat, tetapi juga memberikan hasil yang konsisten dan dapat diandalkan. Integrasi model-model terbaru seperti *EfficientNet*, yang menawarkan keseimbangan optimal antara akurasi dan efisiensi komputasi, semakin memperluas kemungkinan penerapan CNN dalam perangkat dengan sumber daya terbatas seperti *smartphone* dan *edge computing* [10].

Dalam beberapa tahun terakhir, sejumlah penelitian telah mengeksplorasi penerapan *EfficientNet* dalam jaringan saraf konvolusional (CNN) untuk klasifikasi bahasa isyarat [11]. Sebagai contoh, studi [12] yang menggunakan *EfficientNet* berhasil mendeteksi alfabet bahasa isyarat Amerika (ASL) secara *real-time*, menunjukkan akurasi dan efisiensi tinggi yang membuatnya cocok untuk perangkat mobile dan *edge computing*. Penelitian lainnya menggunakan *EfficientNetB4* yaitu [13] mencapai akurasi pengenalan gestur tangan yang tinggi, namun tetap dengan waktu inferensi yang cepat. Berbagai tugas pemrosesan gambar, termasuk teknik augmentasi data seperti rotasi dan *flipping* untuk meningkatkan kualitas dataset, yang secara signifikan meningkatkan akurasi klasifikasi. Pengujian dilakukan dengan skenario terbaik, yaitu menggunakan *Cyclical Learning Rate* (CLR) yang membuat laju pembelajaran dinamis selama proses pelatihan, serta menggunakan ukuran inputan yang lebih kecil yaitu 64x64x3 untuk menurunkan kompleksitas. Selain itu, penelitian lain menyoroti kemampuan *EfficientNet* dalam pengenalan visual yang lebih baik dengan mengekstraksi fitur kompleks dari data visual yaitu [14] menunjukkan bahwa model *EfficientNet-Lite 0* yang dikombinasikan dengan fungsi *loss Label Smooth* mencapai akurasi sebesar 94% dalam mengenali 30 alfabet bahasa isyarat Arab. Penelitian lain seperti [15] yang mengeksplorasi penerapan model *EfficientNet* dengan tujuan untuk mengklasifikasikan 35 gerakan ISL menggunakan pendekatan *transfer learning* dengan model *EfficientNet B0*. Penggunaan *EfficientNet* dalam jurnal ini adalah karena *EfficientNet* merupakan model konvolusi yang dirancang untuk mengoptimalkan keseimbangan antara akurasi dan ukuran model. Ini menunjukkan bahwa arsitektur *EfficientNet* yang efisien sangat ideal untuk aplikasi ini tanpa mengorbankan akurasi. Terakhir contoh penelitian sebelumnya yaitu [16] yang menunjukkan Penggunaan *EfficientNet* sebagai backbone arsitektur *Deep Learning* terbukti efektif dalam mengekstraksi fitur-fitur yang relevan untuk klasifikasi bahasa isyarat.

Studi-studi ini secara kolektif menunjukkan potensi besar *EfficientNet* dalam meningkatkan sistem pengenalan bahasa isyarat [17], terutama dalam aplikasi *real-time* pada perangkat dengan sumber daya terbatas. Efisiensi dan kehandalan *EfficientNet* membuatnya menjadi pilihan yang menarik dalam menghadapi tantangan pengenalan bahasa isyarat, terutama ketika diterapkan pada perangkat dengan sumber daya terbatas seperti *smartphone* atau perangkat *edge computing*. Arsitektur ini didesain untuk memberikan keseimbangan optimal antara akurasi dan kecepatan, sehingga sistem dapat beroperasi dengan efisiensi tanpa mengorbankan performa. Dengan memanfaatkan potensi *EfficientNet*, peneliti dan pengembang dapat menciptakan solusi yang lebih cepat dan responsif dalam menginterpretasikan bahasa isyarat, membuka jalan bagi aksesibilitas yang lebih luas bagi penyandang tunarungu dan tunawicara [18]. Implementasi teknologi ini juga mendukung inklusi sosial yang lebih besar, memungkinkan partisipasi yang lebih aktif dalam berbagai aspek kehidupan masyarakat, termasuk pendidikan, pekerjaan, dan interaksi sehari-hari [19].

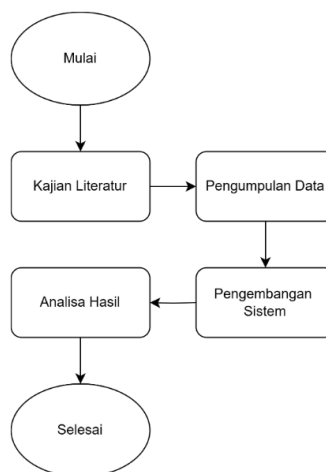
Secara keseluruhan, penggunaan *EfficientNet* dalam konteks pengenalan bahasa isyarat menunjukkan potensi besar untuk meningkatkan kualitas hidup dan integrasi sosial bagi penyandang tunarungu dan tunawicara [20]. Dengan terus mengembangkan dan mengoptimalkan model ini, kita dapat merespons kebutuhan mendesak akan teknologi yang lebih inklusif dan responsif terhadap keberagaman cara berkomunikasi di masyarakat global yang semakin terhubung ini. Perkembangan ini tidak hanya menghadirkan solusi teknologi yang lebih canggih, tetapi juga membuka jalan untuk kemajuan signifikan dalam memperkuat hubungan antarindividu yang berbeda dalam masyarakat modern yang kompleks ini.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini bertujuan untuk mengadaptasi model *Convolutional Neural Network* (CNN) dengan memanfaatkan *EfficientNet* yang telah di-*pre-trained* untuk melakukan klasifikasi bahasa isyarat. Metode yang digunakan adalah *Deep Learning*, yang secara khusus memanfaatkan keunggulan CNN dalam memproses data visual kompleks seperti gerakan tangan dan ekspresi wajah dalam bahasa isyarat.

Dalam penelitian ini, perangkat lunak utama yang digunakan termasuk *Python* sebagai bahasa pemrograman, *TensorFlow* sebagai *framework Deep Learning* yang populer, serta berbagai *library* pendukung seperti *NumPy* untuk operasi numerik yang efisien dan *OpenCV* untuk pengolahan gambar. Pemilihan perangkat keras berfokus pada penggunaan komputer dengan GPU yang mendukung komputasi paralel, memungkinkan percepatan signifikan dalam proses pelatihan dan pengujian model.



Gambar 1. Tahapan Penelitian

Pada gambar 1 merupakan metodologi penelitian yang melibatkan beberapa tahap penting. Pertama, dilakukan kajian literatur untuk memahami landasan teori dan penelitian terdahulu terkait penggunaan CNN dalam konteks pengenalan bahasa isyarat. Selanjutnya, dilakukan pengumpulan dan preprocessing data bahasa isyarat yang diperlukan untuk pelatihan model. Proses berikutnya adalah adaptasi model CNN *pre-trained EfficientNet*, yang melibatkan pengaturan ulang lapisan-lapisan akhir untuk menyesuaikan dengan tugas klasifikasi bahasa isyarat.

Penilaian akurasi model dilakukan dengan menggunakan beberapa metrik evaluasi utama, termasuk akurasi, presisi, *recall*, dan *F1-score*. Akurasi mengukur persentase prediksi yang benar dari keseluruhan data, sedangkan presisi dan *recall* memberikan gambaran lebih detail tentang kemampuan model dalam mengidentifikasi kelas-kelas bahasa isyarat tertentu.

Visualisasi metodologi penelitian ini termasuk penggunaan diagram alir untuk mengilustrasikan secara jelas tahapan-tahapan penelitian, mulai dari persiapan data, pelatihan model, evaluasi, hingga analisis hasil. Diagram ini membantu memperjelas proses pengembangan sistem pengenalan bahasa isyarat menggunakan *EfficientNet*. Selain itu, struktur jaringan CNN *pre-trained EfficientNet* yang diadopsi juga dijelaskan secara rinci, menggambarkan bagaimana model ini dapat diimplementasikan secara efektif dalam mengenali berbagai bentuk huruf bahasa isyarat yang berbeda, sehingga memastikan kualitas dan keakuratan dalam pengenalan gerakan tangan yang kompleks.

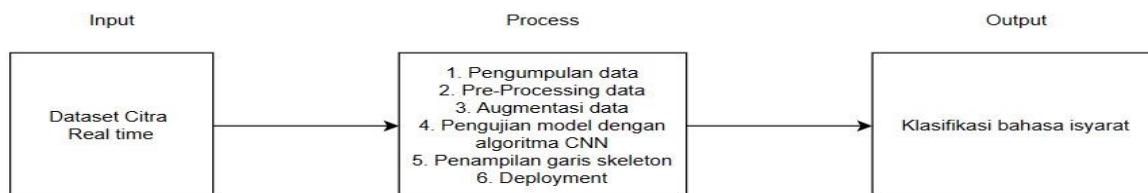
2.2 Kajian Literatur

Untuk memperdalam pemahaman teoritis dalam konteks topik yang dibahas, dilakukan pencarian literatur dari jurnal, buku, dan sumber internet. Tujuannya adalah memperkuat hasil diskusi yang ada dan menemukan literatur yang relevan tentang penerapan metodologi model CNN dengan menggunakan *EfficientNet* yang telah di-*pre-trained* untuk klasifikasi bahasa isyarat. Langkah ini diperlukan guna mengenali teori-teori terbaru dan praktik terbaik dalam penggunaan teknologi ini dalam konteks pengenalan bahasa isyarat secara otomatis.

2.3 Pengumpulan Data

Penelitian dimulai dengan mengakses dataset *American Sign Language (ASL)* yang memuat gambar-gambar tangan menggambarkan isyarat dari huruf A hingga Z dalam alfabet ASL. Dataset ini berisi ribuan gambar yang merepresentasikan posisi tangan yang berbeda untuk setiap huruf, penting untuk pengembangan sistem klasifikasi bahasa isyarat yang komprehensif. Setelah mengunduh dataset, dilakukan pemeriksaan awal untuk memastikan kelengkapan dan kualitas gambar, serta memverifikasi jumlah sampel yang memadai untuk setiap kelas huruf. Hal ini menjamin bahwa dataset ini dapat diandalkan untuk tujuan penelitian yang akurat dan efektif dalam pengenalan bahasa isyarat.

2.4 Pengembangan Sistem



Gambar 2. Blok Diagram Penelitian

Gambar 2 merupakan proses pengembangan sistem yang digunakan pada penelitian ini. Pengembangan sistem adalah langkah perancangan untuk membangun model klasifikasi bahasa isyarat. Mulai dari pengumpulan dataset citra, dilakukan *pre-processing* seperti penyesuaian ukuran gambar dan normalisasi untuk meningkatkan kualitas data. Augmentasi data dilakukan dengan rotasi, *flipping*, *zooming*, dan translasi untuk memperluas variasi dataset. Model *pre-trained EfficientNetB0* disesuaikan parameter untuk klasifikasi bahasa isyarat, dilatih dengan dataset yang diproses untuk optimalisasi parameter. Evaluasi dilakukan dengan data pengujian terpisah untuk memastikan generalisasi yang baik. Terakhir, hasil klasifikasi ditampilkan melalui visualisasi garis skeleton pada gambar isyarat, berperan sebagai alat bantu analisis dan interpretasi output model.

2.5 Analisa Hasil

Analisis kinerja model klasifikasi bahasa isyarat melibatkan evaluasi menggunakan berbagai metrik dan teknik. Pertama, akurasi model digunakan untuk mengukur tingkat pengenalan yang tepat pada data uji. Matriks kebingungan digunakan untuk mengidentifikasi pola klasifikasi yang benar dan salah, memberikan wawasan tambahan tentang performa tiap kelas. Untuk analisis lebih mendalam, presisi, *recall*, dan *F1-score* dihitung untuk menilai keseimbangan antara kepresisian dan *recall* dari setiap kelas. Performa model dievaluasi menggunakan kurva *Receiver Operating Characteristic* (ROC) dan *Area Under the Curve* (AUC) untuk mengevaluasi kemampuan model dalam berbagai threshold klasifikasi. Analisis kesalahan dilakukan untuk memahami contoh yang salah diklasifikasikan, mengidentifikasi pola kesalahan yang perlu dioptimalkan di masa mendatang. Hasil klasifikasi divisualisasikan menggunakan heatmap dari matriks kebingungan dan gambar isyarat dengan garis skeleton, meningkatkan interpretasi performa model secara intuitif.

3. HASIL DAN PEMBAHASAN

Setelah proses pelatihan yang intensif menggunakan algoritma *Convolutional Neural Network* (CNN), model yang telah dibangun siap untuk dievaluasi menggunakan dataset uji. Evaluasi ini bertujuan untuk mengukur seberapa efektif model dalam mengenali dan mengklasifikasikan data yang belum pernah dilihat sebelumnya, yang merupakan simulasi dari situasi dunia nyata di mana model akan digunakan.

Salah satu metrik utama yang digunakan dalam evaluasi adalah akurasi. Akurasi memberikan gambaran tentang seberapa banyak prediksi model yang benar dibandingkan dengan total prediksi yang dilakukan. Namun, akurasi sendiri tidak cukup untuk sepenuhnya memahami kinerja model, terutama ketika dataset memiliki kelas yang tidak seimbang atau ketika jenis kesalahan yang berbeda memiliki konsekuensi yang berbeda.

Selain akurasi, grafik kurva pembelajaran (*learning curve*) juga sangat penting untuk dianalisis. Kurva pembelajaran menunjukkan bagaimana kinerja model berubah seiring dengan jumlah data pelatihan yang meningkat. Hal ini membantu untuk mengidentifikasi apakah model mungkin mengalami *overfitting* atau *underfitting*. *Overfitting* terjadi ketika model terlalu kompleks sehingga menangkap noise atau detail yang tidak relevan dari data pelatihan, sedangkan *underfitting* terjadi ketika model terlalu sederhana sehingga gagal untuk menangkap pola penting dalam data.

Selain grafik kurva pembelajaran, *confusion matrix* (matriks kebingungan) adalah alat penting lainnya dalam evaluasi kinerja model. *Confusion matrix* menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas. Dari *confusion matrix*, kita dapat menghitung berbagai metrik evaluasi seperti presisi, *recall*, dan *F1-score* untuk setiap kelas. Presisi mengukur seberapa banyak prediksi positif yang benar dari semua prediksi positif yang dibuat oleh model, sedangkan *recall* mengukur seberapa banyak kelas positif yang terdeteksi dengan benar dari semua instance yang sebenarnya positif dalam dataset. Visualisasi grafik kurva pembelajaran dan *confusion matrix* memberikan gambaran yang lebih mendalam tentang kinerja model daripada sekadar angka akurasi. Misalnya, jika kita melihat bahwa akurasi umumnya tinggi tetapi *confusion matrix* menunjukkan *recall* yang rendah untuk kelas tertentu, ini dapat mengindikasikan bahwa model memiliki kesulitan klasifikasi pada kelas tersebut dan mungkin perlu dilakukan penyesuaian lebih lanjut pada arsitektur atau strategi pelatihan.

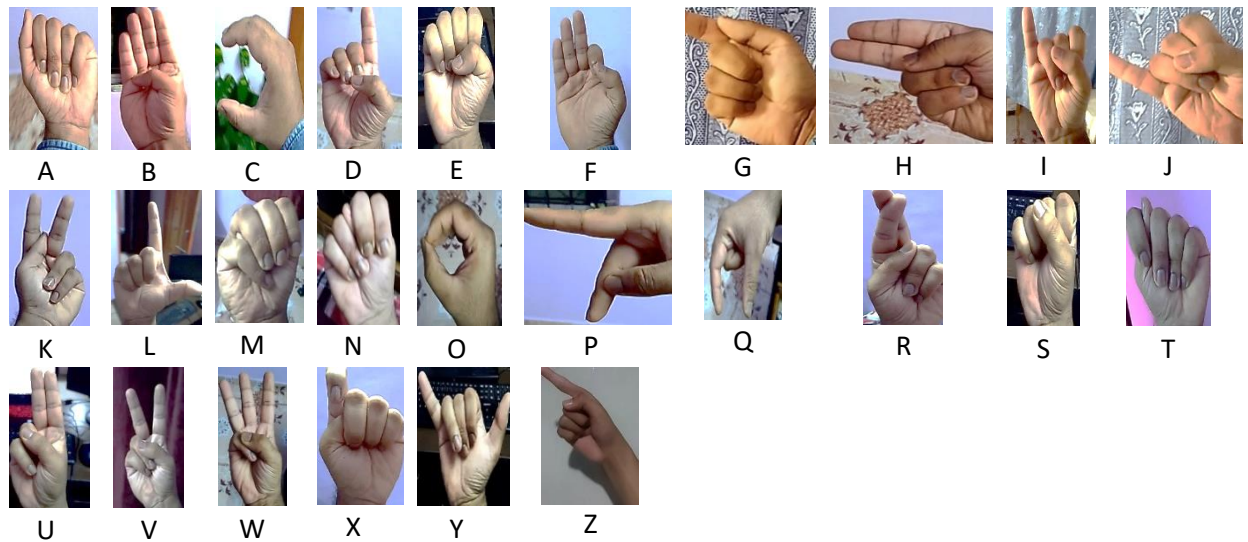
3.1 Hasil Pengambilan Data

Pada percobaan pengembangan sistem pendeteksi bahasa isyarat ini, dataset yang digunakan sangatlah krusial karena akan mempengaruhi kualitas dan akurasi sistem yang dikembangkan. Dataset ini terdiri dari kumpulan gambar yang menampilkan huruf alfabet dalam bahasa isyarat, yang secara total mencakup 26 label klasifikasi yang merepresentasikan huruf A hingga Z. Dengan jumlah total data latih sebanyak 14.740, setiap label memiliki distribusi jumlah data yang berbeda-beda, yang menggambarkan variasi dalam frekuensi kemunculan gestur-gestur tangan tertentu dalam bahasa isyarat.

Dataset ini dipilih setelah mempertimbangkan berbagai koleksi dataset bahasa isyarat yang tersedia di kaggle.com. Kaggle merupakan platform yang terkenal dalam menyediakan dataset untuk berbagai bidang riset, termasuk dalam bidang pengenalan bahasa isyarat. Keputusan untuk menggunakan dataset ini didasarkan pada kualitas visual dari setiap gambar, yang memiliki dimensi sebesar 289x404 piksel. Dimensi yang besar ini memungkinkan representasi yang jelas dan detail terhadap gestur-gestur tangan yang digunakan dalam bahasa isyarat, sehingga memfasilitasi proses ekstraksi fitur yang lebih akurat dan mendalam.

Pentingnya menggunakan dataset yang beragam ini adalah untuk memastikan bahwa sistem yang dikembangkan mampu mengenali dan mengklasifikasikan setiap huruf dalam bahasa isyarat dengan tingkat akurasi yang tinggi. Variasi dalam dataset memungkinkan algoritma pembelajaran mesin untuk melatih model dengan berbagai contoh, sehingga meningkatkan kemampuan generalisasi model terhadap berbagai variasi gestur tangan yang mungkin terjadi dalam kehidupan nyata.

Selain itu, dengan menggunakan dataset yang lebih besar seperti ini, percobaan pengembangan sistem dapat dilakukan dengan cermat untuk mengoptimalkan parameter-model dan teknik pembelajaran mesin yang digunakan. Hal ini penting karena pengenalan bahasa isyarat memerlukan tingkat akurasi yang tinggi untuk memastikan komunikasi yang efektif antara pengguna sistem dan mesin yang digunakan.



Gambar 3. Hasil Pengambilan Dataset

3.2 Hasil Pelatihan Model CNN

Pelatihan menggunakan model jaringan saraf konvolusi (CNN) dipilih berdasarkan arsitektur yang telah ditetapkan sebelumnya untuk pengenalan bahasa isyarat karena CNN memiliki kemampuan unggul dalam mengekstraksi fitur-fitur spasial dari data gambar. Dalam konteks pengenalan gerakan tangan kompleks dan beragam seperti dalam bahasa isyarat, kemampuan CNN untuk secara efektif mengenali pola-pola visual kompleks sangatlah krusial. CNN bekerja dengan cara melakukan operasi konvolusi pada input gambar, yang memungkinkan mereka untuk secara hierarkis memahami fitur-fitur dari tingkat yang rendah hingga tingkat yang lebih abstrak, seperti bentuk tangan dan posisi jari-jari, yang sangat penting dalam pengenalan bahasa isyarat.

Selama proses pelatihan, data yang telah dikumpulkan sebelumnya digunakan untuk menguji dan melatih model CNN ini. Data ini terdiri dari berbagai contoh gerakan tangan yang mencakup sejumlah besar variasi dalam bahasa isyarat, termasuk berbagai pose tangan dan ekspresi wajah yang mungkin menyertai gerakan tersebut. Validasi model menjadi fokus utama selama pelatihan, di mana kita memastikan bahwa model dapat dengan akurat mengklasifikasikan berbagai isyarat tangan yang berbeda. Proses ini melibatkan penggunaan set data yang terpisah untuk pengujian dan validasi, di mana setiap contoh diuji terhadap model yang telah dilatih untuk mengevaluasi tingkat keakuratannya.

Pengukuran performa dan akurasi model CNN yang dilatih merupakan langkah penting dalam pengembangan sistem pengenalan bahasa isyarat yang handal dan efektif. Performa model dievaluasi dengan menggunakan metrik-metrik seperti akurasi (persentase prediksi yang benar), presisi (kemampuan model untuk tidak membuat kesalahan positif), dan *recall* (kemampuan model untuk menemukan semua instansi kelas yang benar). Evaluasi ini membantu dalam menentukan seberapa baik model dapat menggeneralisasi dari data pelatihan ke data baru, yang sangat penting dalam menerapkan teknologi pengenalan bahasa isyarat dalam situasi nyata, di mana variasi gerakan dan kondisi pencahayaan dapat bervariasi secara signifikan.

Selain itu, dalam pengembangan sistem pengenalan bahasa isyarat, ada juga perhatian terhadap optimisasi model untuk kinerja yang lebih baik. Ini melibatkan teknik-teknik seperti *tuning hyperparameter* untuk meningkatkan akurasi dan kecepatan konvergensi selama pelatihan model. Optimisasi ini sering kali melibatkan eksperimen dengan berbagai arsitektur CNN, seperti menyesuaikan jumlah lapisan, ukuran kernel, dan langkah-langkah konvolusi, serta penerapan teknik-teknik seperti regularisasi untuk mengurangi *overfitting*.

Tabel 1. Percobaan Pelatihan Model CNN

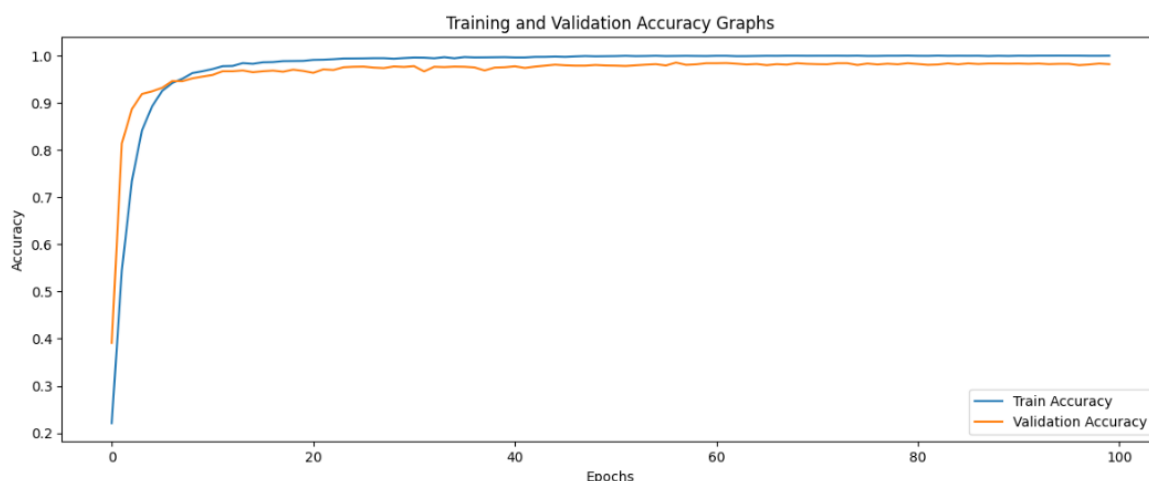
Experiment	Number Of Datasets	Number Of Epochs	Train Loss	Train Accuracy	Validation Loss	Validation accuracy
1	12.000	20	0.5003	0.9866	0.5303	0.9715

Experiment	Number Of Datasets	Number Of Epochs	Train Loss	Train Accuracy	Validation Loss	Validation accuracy
2	12.000	50	0.0879	0.9966	0.1635	0.9799
3	12.000	100	0.0320	0.9994	0.1224	0.9833
4	14.740	20	0.0230	0.0932	0.2465	0.9373
5	14.740	50	0.0145	0.9955	0.1977	0.9842
6	14.740	100	0.0209	0.9994	0.1001	0.9840

Tabel 1 menunjukkan hasil eksperimen pelatihan dan validasi model *Convolutional Neural Network* (CNN) yang dilakukan dalam enam kondisi yang berbeda dengan variasi jumlah dataset. Percobaan yang paling optimal dalam hal akurasi dan *loss* adalah percobaan ke-6, di mana jumlah dataset yang digunakan adalah 14.740. Percobaan ini mencatatkan akurasi pelatihan sebesar 99.94% dan akurasi validasi sebesar 98.40%, dengan nilai *loss* pelatihan sebesar 3.2% dan *loss* validasi sebesar 12.24%. Sebagai perbandingan, percobaan ke-3 menggunakan 12.000 dataset selama 100 *epoch* menghasilkan akurasi validasi yang lebih rendah dibandingkan dengan penggunaan 14.740 dataset, yakni sebesar 98.33%, dengan nilai *loss* yang lebih tinggi. Hasil ini mengindikasikan bahwa peningkatan jumlah dataset dapat signifikan mempengaruhi performa model CNN, dengan menunjukkan peningkatan yang konsisten dalam akurasi dan pengurangan dalam nilai *loss*. Peningkatan jumlah *epoch* dalam pelatihan model *Convolutional Neural Network* (CNN) memainkan peran krusial dalam meningkatkan akurasi dan mengurangi *loss*. Ketika jumlah dataset mencapai 12.000 atau bahkan 14.740, hasil percobaan menunjukkan bahwa peningkatan jumlah *epoch* secara signifikan berkontribusi pada kinerja model. *Epoch* tambahan memungkinkan CNN untuk melakukan lebih banyak iterasi melalui data pelatihan, yang mana proses ini penting untuk memperdalam pemahaman model terhadap pola-pola yang lebih kompleks dan fitur-fitur yang lebih halus dalam dataset.

Secara khusus, dengan meningkatnya jumlah *epoch*, CNN memiliki kesempatan lebih banyak untuk menyesuaikan bobot dan biasnya terhadap data yang diberikan. Ini berarti model dapat memperbaiki representasi internalnya dari data, yang pada gilirannya meningkatkan kemampuan model untuk membuat prediksi yang lebih akurat. Proses ini tidak hanya meningkatkan performa saat dilakukan pengujian, tetapi juga memperkuat generalisasi model terhadap data baru yang belum pernah dilihat sebelumnya. Lebih lanjut, peningkatan *epoch* sering kali juga berdampak positif terhadap kemampuan model untuk mengatasi *overfitting*. Meskipun risiko *overfitting* meningkat dengan peningkatan *epoch*, teknik seperti regularisasi dan dropout dapat diterapkan untuk mengurangi dampaknya. Dengan demikian, peningkatan jumlah *epoch* bukan hanya tentang meningkatkan kinerja model pada data pelatihan, tetapi juga tentang mempertahankan kemampuan model untuk menggeneralisasi pola-pola yang relevan dalam data yang baru.

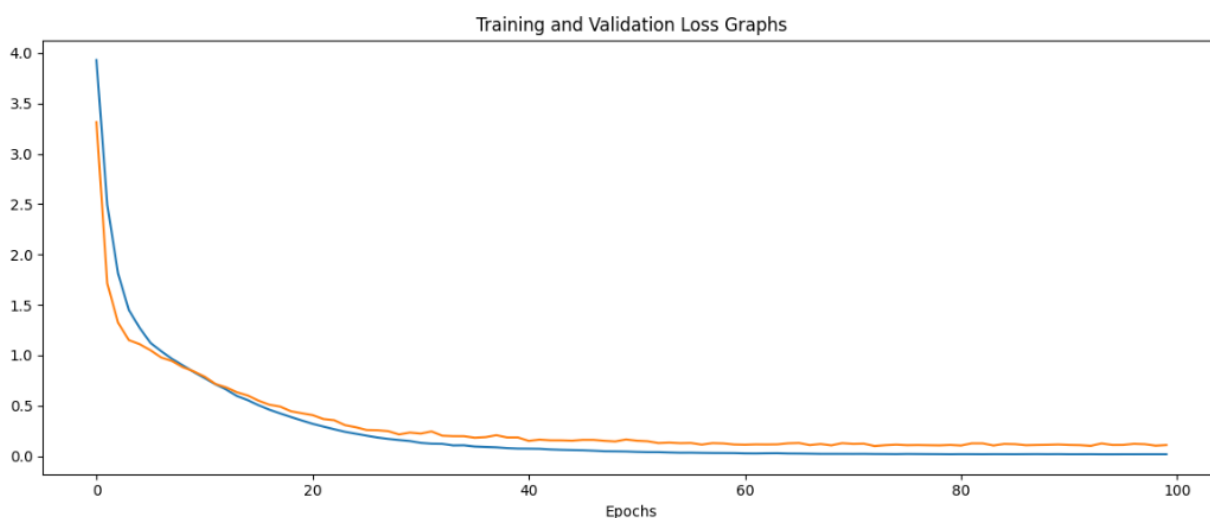
Perbandingan akurasi antara data latih (*train accuracy*) dan data validasi (*validation accuracy*) selama 100 *epochs* memperlihatkan pola yang menarik dalam evaluasi model bahasa isyarat. Grafik pada Gambar 4 menggambarkan tren ini dengan jelas: garis biru mengindikasikan akurasi pada data latih, sedangkan garis oranye mencerminkan akurasi pada data validasi yang tidak digunakan dalam proses pelatihan. Pada awalnya, akurasi pada data validasi meningkat secara konsisten seiring dengan model mempelajari pola-pola dalam dataset. Namun, setelah mencapai titik tertentu, akurasi validasi mulai menetap di sekitar 0.96, menunjukkan bahwa model mampu menggeneralisasi pola-pola umum dengan baik namun mungkin tidak mampu mengenali pola-pola yang lebih kompleks atau spesifik. Di sisi lain, akurasi pada data latih meningkat pesat pada tahap awal pelatihan, mencapai tingkat yang sangat tinggi sekitar 0.99. Namun, setelah beberapa *epoch*, peningkatan ini melambat dan akhirnya stabil, menandakan bahwa model sudah "menghafal" pola-pola dalam data latih dengan baik.



Gambar 4. Grafik Akurasi Training dan Validasi

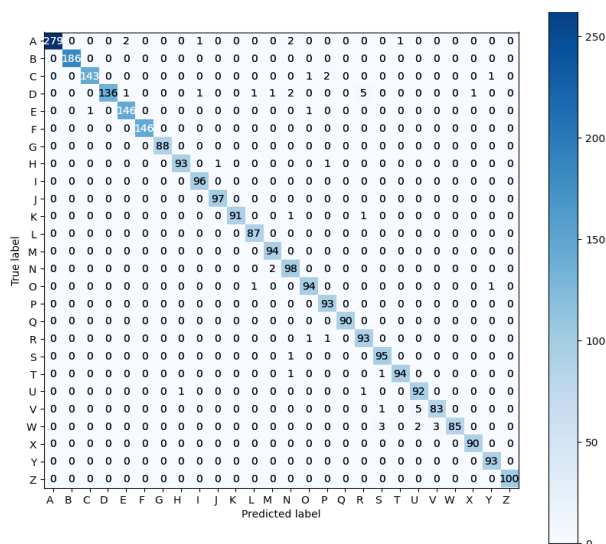
Grafik yang menggambarkan perubahan nilai *loss* selama pelatihan model bahasa isyarat, seperti yang ditunjukkan pada Gambar 5, memberikan *insight* yang mendalam tentang kinerja dan stabilisasi model. Garis biru mewakili perubahan *loss* pada data latih, menunjukkan penurunan yang signifikan awalnya dan stabilisasi di sekitar nilai 0.0 setelah beberapa *epoch*. Ini mengindikasikan bahwa model efektif dalam mengurangi kesalahan dalam memprediksi label dari data latih. Garis oranye, yang mewakili *loss* pada data validasi, juga menunjukkan penurunan yang cukup signifikan awalnya, mencerminkan kemampuan model untuk menggeneralisasi dengan baik terhadap data yang tidak terlihat sebelumnya.

Perbedaan yang minimal antara *loss* pada data latih dan validasi mengindikasikan bahwa model tidak mengalami *overfitting*, sementara stabilisasi kedua garis *loss* setelah beberapa *epoch* menunjukkan bahwa model telah mencapai konvergensi optimal di mana peningkatan tambahan dalam pelatihan tidak signifikan meningkatkan kinerja model. Analisis terhadap grafik ini memberikan pemahaman yang mendalam tentang kemampuan model dalam memahami pola-pola bahasa isyarat dan relevansinya dalam aplikasi praktis untuk pengenalan bahasa isyarat secara *real-time*. Dengan demikian, pemantauan terhadap perubahan *loss* selama proses pelatihan tidak hanya membantu dalam evaluasi kualitas model, tetapi juga mendukung pengembangan teknik dan strategi yang lebih baik dalam implementasi sistem pengenalan bahasa isyarat yang efisien dan handal.



Gambar 5. Grafik *Loss* Training dan Validasi

Model yang digunakan telah menunjukkan performa yang sangat baik dalam validasi, yang tercermin dari tingkat prediksi yang tepat. Grafik atau tabel yang menampilkan matriks kebingungan (*confusion matrix*) dapat mengilustrasikan dengan jelas distribusi prediksi yang benar (diagonal biru tua) untuk setiap kelas. Ini memberikan indikasi kuat bahwa model mampu mengenali pola-pola yang diperlukan untuk klasifikasi dengan akurasi yang tinggi. Meskipun secara umum performa model sangat baik, masih ada beberapa contoh di mana kesalahan prediksi terjadi. Analisis lebih lanjut terhadap kesalahan ini dapat memberikan wawasan yang berharga untuk meningkatkan model, seperti mengidentifikasi kelas-kelas yang sering salah diprediksi atau menemukan pola kesalahan yang berulang.



Gambar 6. Confussion Matrix Pelatihan

Dengan menganalisis pola kesalahan klasifikasi, dapat dilakukan strategi perbaikan lebih lanjut pada model. Salah satu pendekatan yang dapat diambil adalah dengan meningkatkan fitur yang membedakan antara kelas-kelas yang sulit diidentifikasi. Ini dapat dilakukan dengan menambahkan fitur-fitur baru yang lebih relevan atau dengan mengubah cara fitur-fitur tersebut diekstraksi atau diproses sebelum dimasukkan ke dalam model. Peningkatan seperti ini tidak hanya berpotensi meningkatkan akurasi prediksi secara keseluruhan, tetapi juga dapat mengurangi kemungkinan kesalahan dalam klasifikasi pada kasus-kasus tertentu. Selain itu, penggunaan teknik validasi seperti *cross-validation* dapat memvalidasi kehandalan performa model di luar data pelatihan yang digunakan. *Cross-validation* membagi data menjadi subset yang saling eksklusif, di mana masing-masing subset digunakan sebagai data uji dan pelatihan secara bergantian. Hal ini membantu mengurangi risiko *overfitting* dan memberikan perkiraan yang lebih konsisten tentang performa aktual model di berbagai kondisi data.

Implikasi dari performa yang tinggi dalam pengenalan pola klasifikasi ini sangat signifikan dalam berbagai aplikasi dunia nyata. Misalnya, dalam konteks pengenalan bahasa isyarat, keakuratan yang tinggi dalam mengenali gerakan tangan dan ekspresi adalah kunci untuk memfasilitasi komunikasi yang efektif bagi individu yang menggunakan bahasa isyarat sebagai bahasa utama mereka. Dengan demikian, pemahaman yang mendalam tentang evaluasi performa model dan langkah-langkah untuk memperbaikinya tidak hanya penting untuk penelitian dan pengembangan model ini, tetapi juga memiliki dampak yang positif dalam meningkatkan teknologi untuk kepentingan sosial dan komunitas yang lebih luas.

4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa model *EfficientNet*, terutama *EfficientNetB0*, terbukti sangat efektif dalam pembuatan model Jaringan Saraf Tiruan (CNN) untuk klasifikasi bahasa isyarat. Keefektifan model ini memungkinkan pengembangan sistem yang dapat mengatasi masalah komunikasi yang dihadapi oleh komunitas tuna rungu dan tuna wicara. Dalam konteks ini, penggunaan *EfficientNet* dalam mengolah data bahasa isyarat telah terbukti mampu menghasilkan model yang tidak hanya akurat tetapi juga efisien dalam penggunaan sumber daya. Pada tahap pengumpulan data, penggunaan dataset *American Sign Language* (ASL) yang telah di-augmentasi adalah langkah penting. Augmentasi data dengan teknik rotasi, penskalaan, dan pembalikan memperluas variasi dan kualitas dataset, yang pada gilirannya meningkatkan kemampuan model untuk mengenali berbagai pose dan ekspresi dalam bahasa isyarat. Variasi ini tidak hanya memperkaya latihan model tetapi juga meningkatkan kemampuan generalisasi model terhadap data baru yang belum pernah dilihat sebelumnya. Dalam proses pelatihan, model yang dikembangkan mencapai tingkat akurasi yang sangat tinggi, dengan akurasi pelatihan mencapai 99,94% dan akurasi validasi sebesar 98,40%. Keakuratan ini adalah hasil langsung dari kemampuan *EfficientNet* dalam mengekstraksi dan mempelajari fitur-fitur penting dari dataset bahasa isyarat yang kompleks. Dengan demikian, *EfficientNet* tidak hanya berhasil dalam mengklasifikasikan isyarat tangan dengan akurasi tinggi tetapi juga menunjukkan potensi besar dalam pengembangan sistem klasifikasi bahasa isyarat yang dapat diandalkan. Penggunaan *EfficientNet* juga memperhatikan aspek efisiensi model. Dalam implementasi pada perangkat mobile atau embedded yang memiliki keterbatasan daya komputasi, efisiensi model menjadi krusial. *EfficientNet* terkenal karena desainnya yang mengoptimalkan penggunaan parameter dan komputasi, sehingga cocok untuk aplikasi pada platform dengan sumber daya terbatas tanpa mengorbankan performa.

Dengan demikian, hasil penelitian ini tidak hanya mengonfirmasi keunggulan teknologi *EfficientNet* dalam konteks pengenalan bahasa isyarat tetapi juga mengilustrasikan dampaknya dalam meningkatkan aksesibilitas dan komunikasi bagi individu dengan kebutuhan khusus. Dalam pandangan yang lebih luas, penelitian ini memberikan landasan kuat untuk pengembangan lebih lanjut dalam bidang pengenalan bahasa isyarat menggunakan teknologi AI. Fokus pada penerapan model yang efisien dan akurat adalah langkah penting menuju tujuan meningkatkan interaksi dan integrasi sosial bagi komunitas bahasa isyarat global. Dengan terus berkembangnya teknologi dan metodologi dalam bidang ini, masa depan sistem pengenalan bahasa isyarat dapat semakin memberikan manfaat signifikan bagi masyarakat yang membutuhkan.

REFERENCES

- [1] S. Mutiara et al., "Cracteristics And Models Of Guidance Or Islamic Education For Childrend With Disabilities In The Lubuk Lintang Sub-District Community Gang Macang Besar RT 07 RW 03." 2023, doi: <https://doi.org/10.55583/jkip.v4i1.591>.
- [2] I. Sari and E. Altirika, "Sistem Pengembangan Bahasa Isyarat Untuk Berkomunikasi dengan Penyandang Disabilitas (Tunarungu)," 2023
- [3] Mariah Ulfah dan Siti Ubaidah, "Penerapan Bahasa Isyarat dalam Pembelajaran bagi Anak Berkebutuhan Khusus Tuna Rungu," JDSR, Jambi, 2023.
- [4] S. Siallagan et al., "Peran Guru dan Orang Tua dalam Mendukung Pembelajaran Anak Tuna Rungu di Sekolah Luar Biasa (SLB) PGRI Kamal, Bangkalan", 2024, doi: 10.62383/dilan.v1i3.475.
- [5] A. Saifudin, "PERAN KECERDASAN BUATAN DALAM MENINGKATKAN PEMBELAJARAN SISWA TUNARUNGU DI ERA DIGITAL," 2024

- [6] I. Suhardin, A. Patombongi, A. Muhammad Islah, and S. H. Catur Sakti Kendari Jl Abdullah, "MENGIDENTIFIKASI JENIS TANAMAN BERDASARKAN CITRA DAUN MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK," vol. 6, no. 2, 2021, doi: <https://doi.org/10.51876/simtek.v6i2.101>.
- [7] J. Homepage, S. Sahibu, I. Taufik, and P. Studi Sistem Komputer, "Implementation of the Convolutional Neural Network Algorithm for Classifying Types of Organic and Non-Organic Waste," vol. 4, no. 3, pp. 840–852, 2024, doi: [10.57152/malcom.v7i2.1346](https://doi.org/10.57152/malcom.v7i2.1346).
- [8] I. Maulana, N. Khairunisa, R. Mufidah Informatika, U. H. Singaperbangsa Karawang Jl Ronggo Waluyo, T. Timur, and J. Barat, "DETEKSI BENTUK WAJAH MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN)," Karawang, 2023. doi: <https://doi.org/10.36040/jati.v7i6.8171>.
- [9] G. Abdillah and R. Ilyas, "Deteksi Objek Bahasa Isyarat Huruf Bisindo Menggunakan SSD-Mobilenet," 2024. doi: <https://doi.org/10.30645/kesatria.v5i1.295.g292>.
- [10] S. Arnandito and T. B. Sasongko, "Comparison of EfficientNetB7 and MobileNetV2 in Herbal Plant Species Classification Using Convolutional Neural Networks," Yogyakarta, 2024. doi: <https://doi.org/10.30871/jaic.v8i1.7927>.
- [11] A. C. Milano, "KLASIFIKASI PENYAKIT DAUN PADI MENGGUNAKAN MODEL DEEP LEARNING EFFICIENTNET-B6," Jurnal Informatika dan Teknik Elektro Terapan, vol. 12, no. 1, Jan. 2024, doi: [10.23960/jitet.v12i1.3855](https://doi.org/10.23960/jitet.v12i1.3855).
- [12] B. Alsharif, A. S. Altaher, A. Altaher, M. Ilyas, and E. Alalwany, "Deep Learning Technology to Recognize American Sign Language Alphabet," Sensors, vol. 23, no. 18, Sep. 2023, doi: [10.3390/s23187970](https://doi.org/10.3390/s23187970).
- [13] B. Kusuma Wardana, E. Rachmawati, T. Agung, and B. Wirayuda, "PENGENALAN GESTUR TANGAN STATIS MENGGUNAKAN CNN DENGAN ARSITEKTUR EFFICIENT-NET B4," Bandung, 2021.
- [14] B. Y. Al-Khuraym and M. M. Ben Ismail, "Arabic Sign Language Recognition using Lightweight CNN-based Architecture," International Journal of Advanced Computer Science and Applications, vol. 13, no. 4, pp. 319–328, 2022, doi: [10.14569/IJACSA.2022.0130438](https://doi.org/10.14569/IJACSA.2022.0130438).
- [15] T. S. Arthi and Vinotha D, "EfficientNet Indian Sign Language Recognition Using Pretrained Weights," 2024.
- [16] F. Alrowais, S. S. Alotaibi, S. Dhahbi, R. Marzouk, A. Mohamed, and A. M. Hilal, "Sign Language Recognition and Classification Model to Enhance Quality of Disabled People," Computers, Materials and Continua, vol. 73, no. 2, pp. 3419–3432, 2022, doi: [10.32604/cmc.2022.029438](https://doi.org/10.32604/cmc.2022.029438).
- [17] C. Umam Wiranda, "Pengembangan Aplikasi Mobile Android untuk Deteksi Otomatis Mata Katarak Menggunakan CNN dan Tensorflow," Jurnal Kendali Teknik dan Sains, vol. 2, no. 3, pp. 128–140, 2024, doi: [10.59581/jkts-widyakarya.v2i3.3722](https://doi.org/10.59581/jkts-widyakarya.v2i3.3722).
- [18] M. Tan and Q. V Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.", 2019
- [19] M. M. Islam, Md. A. Talukder, M. A. Uddin, A. Akhter, and M. Khalid, "BrainNet: Precision Brain Tumor Classification with Optimized EfficientNet Architecture," International Journal of Intelligent Systems, vol. 2024, no. 1, Jan. 2024, doi: [10.1155/2024/3583612](https://doi.org/10.1155/2024/3583612).
- [20] S. Gnanapriya and K. Rahimunnisa, "A Hybrid Deep Learning Model for Real Time Hand Gestures Recognition," Intelligent Automation and Soft Computing, vol. 36, no. 1, pp. 1105–1119, 2023, doi: [10.32604/iasc.2023.032832](https://doi.org/10.32604/iasc.2023.032832).