Merkurius : Jurnal Riset Sistem Informasi dan Teknik Informatika
Volume. 2 No. 4 Juli 2024
e-ISSN: 3031-8912; p-ISSN: 3031-8904; Hal. 01-09
DOI: https://doi.org/10.61132/merkurius.v2i4.119

Klasifikasi Suara Instrumen Musik Tiup Menggunakan Metode Convolutional Neural Network

[1] Royan Hisyam Rafliansyah , [2] Basuki Rahmat , [3] Chrystia Aji Putra ,
[1-3] Universitas Pembangunan Nasional "Veteran" Jawa Timur

Alamat: Rungkut Madya No. 1, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur

Abstract. *This research explores the classification of brass instrument sounds using Convolutional Neural Network (CNN) combined with Mel-Frequency Cepstrum Coefficient (MFCC) feature extraction. This research aims to improve the accuracy of brass instrument sound recognition by utilizing CNN's ability to process audio data. Through experiments conducted with different audio durations and variations in CNN model architecture, this study evaluates the impact of dataset separation and model design on classification performance. The results show that dataset duration and CNN model architecture significantly affect classification accuracy, with the highest accuracy achieved in the scenario using 30 seconds of audio duration with an accuracy value of 84%. In addition, experiments varying the number of convolution layers in the CNN model show that the selection of the model architecture plays an important role in classification performance. Overall, this research contributes to advancing the field of audio classification by providing insight into the optimal dataset duration and model architecture for wind instrument speech recognition using CNNs.*

Keywords: *Machine Learning, Convolution Neural Network, MFCC, Brass Instrument, Classification*

Abstrak. Penelitian ini membahas tentang klasifikasi suara alat musik tiup menggunakan Convolutional Neural Network (CNN) yang dikombinasikan dengan ekstraksi fitur Mel-Frequency Cepstrum Coefficient (MFCC). Penelitian ini bertujuan untuk meningkatkan akurasi pengenalan suara alat musik tiup dengan memanfaatkan kemampuan CNN dalam memproses data audio. Melalui eksperimen yang dilakukan dengan durasi audio yang berbeda dan variasi arsitektur model CNN, penelitian ini mengevaluasi dampak dari pemisahan dataset dan desain model terhadap kinerja klasifikasi. Hasil penelitian menunjukkan bahwa durasi dataset dan arsitektur model CNN secara signifikan mempengaruhi akurasi klasifikasi, dengan akurasi tertinggi dicapai pada skenario menggunakan durasi audio 30 detik dengan nilai akurasi sebesar 84%. Selain itu, percobaan dengan memvariasikan jumlah lapisan konvolusi pada model CNN menunjukkan bahwa pemilihan arsitektur model berperan penting dalam performa klasifikasi. Secara keseluruhan, penelitian ini berkontribusi dalam memajukan bidang klasifikasi audio dengan memberikan wawasan tentang durasi dataset yang optimal dan arsitektur model untuk pengenalan suara alat musik tiup menggunakan CNN.

Kata kunci: Machine Learning, Convolutional Neural Network, MFCC, Klasifikasi, Alat Musik Tiup

LATAR BELAKANG

Musik merupakan seni yang tidak dapat terpisahkan dalam kehidupan manusia sejak berabad-abad yang lalu. Penggolongan musik bisa dilakukan berdasarkan berbagai aspek seperti jenis genre, artis, atau pencipta lagu (Mahanta, et al., 2021). Dalam seni musik sendiri memiliki banyak jenis instrumen musik mulai dari instrumen dawai, instrumen tiup, instrumen perkusi, instrumen gesek, dan masih banyak jenis instrumen lainnya. Dengan banyaknya jenis instrumen musik tersebut membuat setiap jenis instrumen memiliki karakteristik yang berbeda-beda saat dimainkan.

Salah satu instrumen musik yang memiliki karakteristiknya sendiri adalah instrumen musik tiup. Jenis instrumen musik tiup sendiri memiliki banyak jenis seperti, saxophone, klarinet, trumpet, dan masih banyak lainnya. Pada setiap instrumen tiup tersebut memiliki

Received: April 30, 2024; Accepted: Juni 02, 2024; Published: Juli 31, 2024;
* Royan Hisyam Rafliansyah , ---

banyak atribut yang dapat diukur, seperti frekuensi, amplitudo, bentuk gelombang, dan karakteristik lainnya. Karakteristik tersebut dapat digunakan untuk membedakan satu instrumen dengan instrumen lainnya. Klasifikasi yang akurat antara instrumen-instrumen tersebut memiliki nilai penting dalam konteks unik.

Pengklasifikasian suara instrumen musik tiup sebenarnya dapat dilakukan secara manual oleh manusia. Namun, pengenalan tersebut harus memerlukan keahlian dan waktu yang signifikan dalam mempelajari hal tersebut. Pada penelitian ini akan dilakukan Klasifikasi suara instrumen musik tiup menggunakan metode *Convolutional Neural Network* (CNN) dengan menggunakan ekstraksi fitur *Mel-Frequency Cepstrum Coefficient* (MFCC). *Mel-Frequency Cepstrum Coefficient* atau biasa disebut dengan MFCC adalah metode ekstraksi fitur yang populer dibidang teknologi suara (Safitri, et al., 2022). Dalam proses ini, rekaman suara instrumen musik tiup dikonversi menjadi matriks konvolusi, yang dikenal sebagai spektogram, atau dalam bentuk sinyal suara. Selanjutnya, fitur-fitur ini dapat digunakan untuk melakukan klasifikasi suara instrumen musik tiup yang terkandung di dalamnya.

Setelah melakukan ekstraksi fitur akan dilakukan proses klasifikasi menggunakan metode *Convolutional Neural Network* atau biasa dikenal dengan CNN adalah metode *deep learning* yang terutama menggunakan filter seperti matriks ukuran kecil untuk mengekstrak fitur dari kumpulan data yang besar. ConvNet dari CNN dianggap lebih mampu dalam menangkap ketergantungan spasial dan temporal jika dibandingkan dengan model-model lainnya (Su, 2023). CNN dapat digunakan untuk berbagai tugas seperti pengenalan wajah, pengolahan teks menjadi digital, dan pemrosesan bahasa alami, yang diyakini memiliki keunggulan yang signifikan.

Penggunaan metode CNN memiliki beberapa keunggulan utama yang cocok untuk melakukan klasifikasi suara instrumen musik tiup. CNN mampu mengekstrak fitur dari data yang sangat kompleks melalui penggunaan lapisan konvolusi yang memungkinkan model untuk menangkap pola yang lebih halus dalam data suara yang mungkin terlewatkan oleh metode lain. Dengan MFCC sebagai input, CNN dapat lebih efektif mengidentifikasi karakteristik unik dari setiap instrumen musik tiup. Suara memiliki struktur temporal yang kompleks dan CNN sangat efektif dalam menangani data yang memiliki hubungan spasial dan temporal. Selain itu, CNN dapat disesuaikan dengan menambahkan lebih banyak lapisan atau mengubah parameter model untuk meningkatkan kinerja yang memungkinkan untuk mengoptimalkan model erdasarkan kebutuhan spesifik dan kompleksitas data.

Meskipun metode CNN sudah digunakan dalam berbagai aplikasi klasifikasi, implementasinya dalam konteks klasifikasi instrumen musik tiup masih memerlukan penelitian

yang mendalam dan pengembangan lebih lanjut. Perkembangan teknologi pemrosesan sinyal audio dan komputasi telah memberikan peluang baru dalam menggabungkan metode CNN dengan data audio yang sangat rinci untuk meningkatkan akurasi klasifikasi suara instrumen musik tiup.

Pada penelitian yang sejenis yang berjudul "*Instrument Classification Using Different Machine Learning and Deep Learning Methods*". Penelitian tersebut membedakan suara instrumen tunggal dalam musik monofonik atau polifonik. Penelitian tersebut menggunakan dataset sebanyak 5444 file audio. Pada penelitian ini menggunakan enam metode salah satunya *Convolutional Neural Network*. Dari keenam metode tersebut metode CNN mendapatkan nilai akurasi tertinggi dengan nilai 96,82%.

Dengan menggabungkan kemampuan ekstraksi fitur dari *Mel Frequency Cepstrum Coefficients* dan kemampuan klasifikasi dari *Convolutional Neuaral Network*, penelitian ini berupaya memberikan solusi inovatif dan efektif dalam menghadapi tantangan pengenalan instrumen musik tiup.

KAJIAN TEORITIS

Dasar acuan penelitian terdahulu merupakan hal yang sangat penting sebagai panduan penulis dalam melakukan penelitian. Terdapat beberapa penelitian terdahulu yang berkaitan dengan topik penelitian ini. Berikut beberapa penelitian terdahulu yang menjadi referensi penulis dalam menyusun penelitian kali ini.

Pada penelitian lain yang dilakukan oleh (Solanki, et al., 2022) yang berjudul "*Music Instrument Recognition Using Deep Convolutional Neural Networks*". Pada penelitian memiliki tujuan untuk mengembangkan sistem pengenalan instrumen musik menggunakan *deep convolutional neural networks*. Penggunaan teknologi ini diharapkan dapat meningkatkan akurasi dalam mengidentifikasi instrumen musik dari sinyal audio yang kompleks. Penelitian ini juga menggunakan *mel spectrogram* untuk merepresentasikan data audio dalm format matriks yang sesuai untuk dilakukan pemrosesan. Penelitian ini juga menggunakan kombinasi fungsi aktivasi ReLu dan *Max Pooling* agar dapat mencapai hasil yang sangat baik dalam pengenalan instrumen musik dari sinyal audio polifonik. Hasil akurasi pada penelitian ini mencapai 92,8%, sistem ini memberikan kontribusi yang signifikan dalam bidang pengambilan informasi musik.

Pada penelitian lain yang berjudul "*Instrument Classification Using Different Machine Learning and Deep Learning Methods*" yang dilakukan oleh (Su, 2023). Penelitian ini akan melakukan klasifikasi alat musik untuk instrumen dengan timbre yang mirip. Dataset yang

digunakan dalam penelitian tersebut mencakup sampel suara dari 20 instrumen. Metode *mel-frequency cepstral coefficients* digunakan untuk ekstraksi fitur yang melibatkan transformasi sinyal audio menjadi spektrum, menerapkan logaritme pada amplitudo, menggunakan skala mel, dan kemudian melakukan transformasi kosinus diskrit untuk mendapatkan fitur MFCC. Penelitian tersebut menggunakan enam metode penelitian salah satunya adalah CNN. Hasil penelitian yang menggunakan metode CNN mendapatkan tingkat akurasi tertinggi, yaitu 96,82%. Peningkatkan hasil akurasi pada metode CNN dapat menambahkan lebih banyak fitur MFCC. Penambahan fitur tersebut meningkatkan akurasi hingga 98,66%.

METODE PENELITIAN

Instrumen Musik Brass

Instrumen musik brass adalah jenis alat musik yang terbuat dari logam yang menghasilkan suara dengan cara ditiupkan melalui bagian-bagian tertentu dari instrumen tersebut. Instrumen-instrumen brass memiliki ciri khas suara yang kuat dan sering digunakan dalam berbagai genre musik, termasuk musik klasik, jazz, dan marching band. Beberapa contoh instrumen musik brass termasuk trumpet, trombone, french horn, euphonium, dan tuba.

Timbre Feature

Fitur timbre menggambarkan karakteristik tone suara dengan tingkat halus. Biasanya, fitur ini dihitung untuk segmen sinyal audio dengan durasi 10 hingga 60 detik, dan sering disebut sebagai kualitas tonal yang membedakan suara atau sumber tertentu. Meskipun mungkin kurang bermakna bagi pengguna, fitur ini sangat berguna bagi sistem komputer. Penghitungan fitur timbre didasarkan *short time fourier transform* (STFT) dan diterapkan pada setiap kerangka waktu pendek sari sinyal audio. Contoh fitur timbre yang dignakan dalam klasifikasi adalah sebagai berikut. (Falola, et al., 2021) :

1. *Chroma stft* memiliki peran dalam menentukan kelas nada. Kemudian vektor *chroma* dari setiap frame dikombinasikan untuk menghasilkan rata-rata representatif dan deviasi standar.
2. *Spectral centroid* digunakan untuk mengukur letak dan bentuk spektrum dari titik pusat distribusi energi. Berikut ini rumus dari *spectral centroid*:

   $$f_n = \frac{\sum_k S(k) f(k)}{\sum_k fk} \quad (1)$$

3. *Spectral bandwith* digunakan untuk menentukan resolusi sinyal, dan dapat diungkapkan dengan rumusan seperti berikut:

   $$(\sum_k S(k)(f(k) - f_c)^p)_p^1 \quad (2)$$

4. *Spectral Rolloff* adalah persentase dari total energi spektral yang terletak di bawah frekuensi tertentu. Rata-rata deviasi standar dari *spectral rolloff* dihitung dari semua frame sinyal audio.

5. *Root Mean Square Energy* dihitung secara *frame by frame*. Rata-rata dan standar deviasi kemudian dihitung dari semua frame sinyal audio. Energi yang terkandung dalam sinyal dihitung seperti berikut ini:

   $\sum_{n=1}^{N}|x(n)|^2$ (3)

   RMSE dalam sebuah sinyal dapat dihitung sebagai berikut:

   $\sqrt{\frac{1}{n}\sum_{N}|x(n)|^2}$ (4)

6. *Zero_Crossing Rate* adalah jumlah kali sinyal audio berubah tanda dari positif ke negatif, yang mengindikasi seberapa sering gelombang melintas nilai 0. Rata-rata dan standar deviasi *zero crossing rate* dihitung dari semua frame sinyal audio dan dipilih sebagai fitur yang mewakili.
7. *Mel-Frequency Cepstrum Coefficients* adalah fitur yang berguna untuk menggambarkan karakteristik timbre musik. MFCC dari suatu sinyal audio merupakan set fitur yang relatif kecil, biasanya antara 10 hingga 20, yang secara kompak merepresentasikan bentuk keseluruhan dari amplop spektral. MFCC menggambarkan serangkaian karakteristik jangka pendek dari *spektrum* suara.

Machine Learning

Machine Learning adalah bagian dari bidang kecerdasan buatan yang memanfaatkan bahasa pemrograman untuk mengajarkan komputer agar dapat berprilaku cerdas, mirip dengan manusia (Zer, et al., 2022). Menurut Jo algoritma *Machine learning* atau biasa disebut dengan pembelajaran mesin dapat didefinisikan sebagai paradigma komputasi di mana kapasitas menyelesaikan masalah yang diberikan dibangun oleh contoh-contoh sebelumnya (Jo, 2021). Pada umumnya *Machine Learning* memiliki 3 metode pendekatan, yaitu *supervised learning, unsupervised learning,* dan *reiforcment learning*.

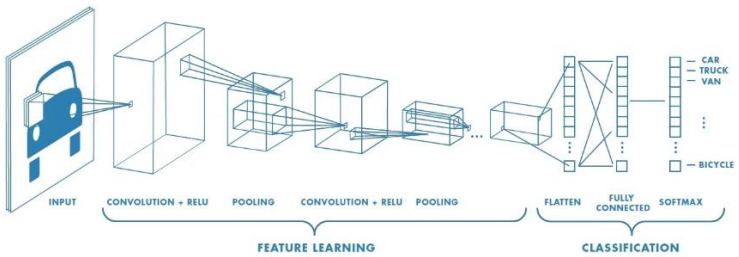
Deep Learning

Deep Learning adalah komponen dari kecerdasan buatan yang pertama kali terinspirasi dari struktur otak manusia. Teknologi ini memungkingkan komputer untuk "mempelajari" data dengan cara yang lebih akurat (Nurhakiki, et al., 2024). Pada *Deep Learning* sendiri dapat mengidentifikasi dan mengenai berbagai jenis data, seperti gambar, video, dan teks, sebagai entitas yang berbeda dengan tingkat ketepatan yang tinggi.

Convolutional Neural Network

Convolutional Neural Network atau biasa disebut dengan CNN merupakan salah satu jenis jaringan syaraf tiruan yang sering digunakan untuk memproses data gambar dengan tujuan mendeteksi dan mengenali objeck. CNN dirancang khusus untuk bekerja dengan data dalam bentuk array (Ramadhani, et al., 2023). CNN memanfaatkan sifat alami sinyal, seperti koneksi lokal, pembagian bobot, pooling, dan penggunaan banyak lapisan. Terdapat empat

model utama dalam proses pengolahan data yang digunakan oleh CNN, yaitu *convolutional layer*, *pooling layer*, *dropout layer*, dan *fully connected layer*.



Gambar 1. Arsitektur CNN

Pada Gambar 2.5 dijelaskan arsitektur CNN dimulai dengan tahap konvolusi, dimana kernel dengan ukuran tertentu digunkan untuk melakukan operasi konvolusi. Jumlah kernel yang digunakan bergantung pada jumlah fitur yang ingin dihasilkan. Tahap ini diikuti oleh fungsi aktivasi yang biasanyya mengunakan ReLu (*Rectifier Linier Unit*) sebagai fungsi aktivasi. Selanjutnya, hasil dari fungsi aktivasi akan melewati tahap *pooling*. Pada tahap ini akan diulang hingga beberapa kali hingga peta fitur yang mencakup dihasilkan untuk diteruskan ke *fully connected neural network*, dan *output* akhirnya adalah kelas dari input.

Batch Normalization

Normalisasi adalah proses mengubah skala nilai atribut dari data sehingga nilainya berada dalam rentang tertentu (Purnama, et al., 2022). Namun, *batch normalization* lebih efektif digunakan pada *machine learning* yang dimana proses tersebut memberikan banyak manfaat dalam proses pelatihan *deep neural network* (Chen, et al., 2021). Manfaat yang diberikan oleh proses tersebut, yaitu menghindari masalah peledakan dan lenyapnya gradien, mempercepat proses pelatihan, mengurangi efek inisialisasi mode, dan membuat distribusi *output* lebih merata. Menambahkan *batch normalization* ke *output* setiap lapisan dapat menyesuaikan kembali distribusi *input* ke lapisan saat ini. Dengan demikian, distribusi bobot lapisan sebelumnya tidak akan memengaruhi bobot lapisan selanjutnya.

Confusion Matrix

Confusion Matrix adalah sebuah tata letak tabel yang sering digunakan sebagai alat visual untuk membantu evaluasi model *machine learning*. Matriks ini membandingkan antara label kelas yang diprediksi oleh model dengan label kelas yang sebenarnya dari semua data yang ada (Görtler, et al., 2022). Dalam proses ini terdapat 4 kasus perhitungan, yaitu:

1. *True Positive* (TP) : hasil prediksi yang benar dan sesuai dengan realitas
2. *True Negative* (TN) : hasil prediksi yang benar tetapi tidak sesuai realitas
3. *False Positive* (FP) : hasil prediksi yang salah tetapi sesuai dengan realitas
4. *False Negative* (FN) : hasil prediksi dan realitas salah

Akurasi adalah ukuran yang membandingkan jumlah data yang telah diprediksi dengan benar berdasarkan kelasnya dengan jumlah total semua data. Rumusnya dapat dinyatakan sebagai berikut:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (5)$$

HASIL DAN PEMBAHASAN

Pengumpulan Data

Pada proses pengumpulan data dilakukan pengambilan *sample* secara langsung ke teman-teman marching band. Untuk melakukan proses ini menggunakan alat rekam suara dengan menggunakan handphone. Data yang diambil memiliki durasi 30 detik dengan jumlah *sample* 400 file audio. Selanjutnya, data berdurasi 30 detik tersebut dibagi menjadi 10 detik yang dimana pembagian file tersebut membuat jumlah *sample* menjadi 1200 file audio.

Selain itu, file audio berdurasi 30 detik juga dibagi menjadi beberapa bagian yang berdurasi 5 detik. Pembagian tersebut akan membuat jumlah *sample* audio menjadi 2400 file audio. Setelah itu, ada pembagian data audio yang berdurasi 3 detik pembagian ini membuat jumlah data menjadi 4000 file audio. Pembagian file audio ini digunakan untuk skenario uji coba.

Uji Coba Skenario

Pada penelitian ini, proses pengujian dilakukan secara bergantian untuk setiap skenario. Pengukuran akurasi dilakukan dengan menerapkan metode Convolutional Neural Network. Penelitian ini dibagi menjadi dua skenario dimana skenario pertama menggunakan durasi audio yang berbeda. Skenario kedua menggunakan dataset dengan durasi yang sama namun terdapat variasi pada model CNN. Hal ini dilakukan untuk mengetahui perbedaan dataset dan variasi model CNN dapat meningkatkan akurasi yang dicapai. Setiap iterasi uji coba dari model yang dikembangkan akan dievaluasi dengan menggunakan metode confusion matrix. Berikut ini adalah hasil evaluasi untuk setiap skenario uji coba yang dilakukan.

Hasil Uji Coba

1. **Skenario Pertama**

   Pada skenario pertama, dataset yang digunakan memiliki durasi 30 detik. Data berdurasi 30 detik tersebut kemudian dibagi menjadi beberapa segmen kecil menjadi 10 detik, 5 detik, dan 3 detik. Data tersebut kemudian dibagi menjadi data latih dan data uji dimana pembagiannya menjadi 80% data latih dan 20% data uji. Kemudian data tersebut dimasukkan ke dalam proses pelatihan model CNN yang menggunakan 4 layer konvolusi mulai dari 512, 256, 128, dan 64. Berikut adalah hasil dari berbagai pembagian durasi dataset.

Tabel 1. Skenario Pertama

Jenis Dataset	Akurasi (%)
30 Detik	84
10 Detik	80
5 Detik	79
3 Detik	77

Setelah melakukan percobaan, maka akan didapatkan hasil seperti pada tabel 2 hingga 5. Pada percobaan tersebut, nilai akurasi tertinggi didapatkan dengan menggunakan dataset dengan durasi 30 detik yang memiliki nilai akurasi sebesar 84% dan nilai akurasi terendah didapatkan dengan dataset berdurasi 3 detik yang memiliki nilai akurasi sebesar 77%. Pada percobaan tersebut menggunakan 4 layer konvolusi yang sama. Layer konvolusi menggunakan aktivasi ReLu pada setiap layer. Dapat dilihat bahwa hasil percobaan membuktikan bahwa jumlah dataset juga mempengaruhi proses klasifikasi.

2. Skenario Kedua

Pada skenario kedua menggunakan dataset 3 detik. Namun, pada skenario kedua ini menggunakan 2 variasi konvolusi. Percobaan pertama menggunakan 3 lapisan konvolusi. Sedangkan, percobaan kedua menggunakan 2 lapisan konvolusi. Percobaan pertama menggunakan kernel 128, 64, dan 32 sedangkan percobaan kedua menggunakan kernel 64 dan 32. Berikut ini adalah hasil dari percobaan pertama dan kedua.

Tabel 2. Skenario Kedua

Variasi Convolution Layer	Akurasi (%)
3 Lapisan Konvolusi	77
2 Lapisan Konvolusi	78

Dari hasil kedua percobaan tersebut, dapat dilihat bahwa akurasi tertinggi diperoleh 77% dengan menggunakan 3 lapisan konvolusi. Sedangkan ketika menggunakan 2 *convolution layer*, akurasi yang didapatkan adalah 78%. Pada percobaan tersebut, dataset yang digunakan berdurasi 3 detik dengan total data sebanyak 4000 file. Jadi, kedua percobaan ini hanya berbeda pada *convolution layer*-nya saja.

KESIMPULAN DAN SARAN

Pada penelitian ini menunjukkan bahwa implementasi Convolutional Neural Network (CNN) efektif dalam klasifikasi suara instrumen musik tiup, dengan hasil akurasi prediksi yang bervariasi berdasarkan skenario uji yang berbeda. Dari enam skenario uji yang dilakukan, penggunaan dataset dengan durasi yang lebih panjang cenderung menghasilkan akurasi yang lebih tinggi, seperti yang terlihat pada skenario pertama dengan dataset 30 detik yang mencapai akurasi 84%. Sebaliknya, pengurangan durasi dataset cenderung menurunkan akurasi, dengan skenario uji ketiga dan keempat masing-masing mencatat akurasi 79% dan 77%. Selain itu, variasi dalam arsitektur model CNN, seperti penambahan jumlah lapisan konvolusi, tidak

selalu memberikan peningkatan yang signifikan dalam akurasi prediksi. Misalnya, skenario kelima dengan tiga lapisan konvolusi mencapai akurasi 77%, sementara skenario keenam dengan dua lapisan konvolusi justru sedikit lebih tinggi, yaitu 78%. Hal ini mengindikasikan bahwa pengolahan data dan konfigurasi lapisan konvolusi adalah faktor penting yang mempengaruhi kinerja model, namun penyesuaian arsitektur CNN perlu dipertimbangkan dengan cermat agar dapat menghasilkan hasil yang optimal.

DAFTAR REFERENSI

Chen, H., Wang, Y., & Fan, C. (2021, March). A convolutional autoencoder-based approach with batch normalization for energy disaggregation. The Journal of Supercomputing, 77, 2961-2978.

Falola, P. B., & Akinola, S. O. (2021, August). Music genre classification using 1D convolutional neural network. International Journal on Human Computing Studies, 3(6), 3-21.

Görtler, J., Hohman, F., Moritz, D., Wongsuphasawat, K., Ren, D., Nair, R., ... Patel, K. (2022). Neo: Generalizing confusion matrix visualization to hierarchical and multi-output labels. CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 1-13.

Jo, T. (2021). Machine learning foundations: Supervised, unsupervised, and advanced learning. Cham: Springer International Publishing.

Mahanta, S. K., Khilji, A. F. U. R., & Parkay, P. (2021). Deep neural network for musical instrument recognition using MFCCs. Computacion y Sistemas, 25(2), 351-360.

Nurhakiki, J., & Yahfizham. (2024). Studi kepustakaan: Pengenalan 4 algoritma pada pembelajaran deep learning beserta implikasinya. Jurnal Pendidikan Berkarakter, 2(1), 270-281.

Purnama, J. J., & Rahayu, S. (2022, June). Klasifikasi konsumsi energi industri baja menggunakan teknik data mining. Jurnal Teknoinfo, 16(2), 395-407.

Ramadhani, F., Satria, A., & Salamah. (2023, December). Implementasi algoritma convolutional neural network dalam mengidentifikasi dini penyakit pada mata katarak. Jurnal Teknik Informatika, 167-175.

Safitri, P. W., & Karyawati, A. E. (2022, November 1). Kombinasi metode MFCC dan KNN dalam pengenalan emosi manusia melalui ucapan. Jurnal Nasional Teknologi Informasi dan Aplikasinya, 1, 133-140.

Solanki, A., & Pandey, S. (2022). Music instrument recognition using deep convolutional neural networks. International Journal of Information Technology, 1659-1668.

Su, Y. (2023). Instrument classification using different machine learning and deep learning methods. Highlights in Science, Engineering and Technology, 136-142.

Zer, P., Hayadi, B., & Damanik, A. (2022, August 15). Pendekatan machine learning menggunakan algoritma C4.5 berbasis PSO dalam analisa pemahaman pemrograman berbasis website. Jurnal Informatika dan Teknik Elektro Terapan, 10(3), 150-156.