

KLASIFIKASI POSE MANUSIA BERBASIS *POINT CLOUD* MENGGUNAKAN *DEEP LEARNING*

Muhammad Siddiq^{1*}, Dihin Muriyatmoko¹, Oddy Virgantara Putra¹

¹Program Studi Teknik Informatika, Universitas Darussalam Gontor Ponorogo

Jl. Raya Siman Demangan Siman Ponorogo, Jawa Timur 63471

*Corresponding Author : muh.siddiq25@mhs.unida.gontor.ac.id

ABSTRAK

Dengan menggunakan teknologi untuk mengklasifikasikan pose manusia, pemantauan pekerjaan yang berisiko cedera dapat dilakukan dengan lebih aman, sejalan dengan prinsip mempertahankan keamanan dan privasi yang merupakan bagian dari prinsip-prinsip syariah dalam menjaga jiwa. Namun, dalam pengambilan sampel data tubuh manusia, terdapat risiko pengambilan data aurat yang melanggar prinsip privasi. Melalui penggunaan data point cloud dari LiDAR, bagian tubuh yang menjadi aurat dapat tersamarkan dan menjaga privasi. Meskipun demikian, pose manusia yang dihasilkan belum terlihat dengan jelas. Oleh karena itu, tujuan dari penelitian yang dilakukan ini adalah untuk membuat model klasifikasi pose manusia berbasis voxel point cloud dengan menggunakan deep learning agar dapat mengetahui pose manusia. Dalam penelitian ini, model klasifikasi pose manusia berbasis voxel point cloud dengan menggunakan pendekatan deep learning Conv3D telah berhasil dikembangkan dengan akurasi sebesar 95.76%.

Kata kunci: *Human pose classification; LiDAR; Point cloud data; Deep learning*

ABSTRACT

Utilizing technology to classify human poses enables safer monitoring of injury-prone jobs, in line with principles of upholding safety and privacy integral to Sharia principles in safeguarding lives. However, in sampling human body data, there's a risk of breaching privacy principles by capturing aurat data. Through employing LiDAR point cloud data, aurat areas can be obscured, preserving privacy. Nevertheless, resulting human poses may lack clarity. Hence, the aim of this research is to develop a voxel point cloud-based human pose classification model using deep learning to ascertain human poses. In this study, a voxel point cloud-based human pose classification model utilizing Conv3D deep learning approach has been successfully developed with an accuracy of 95.76%.

Keywords : *Human pose classification; LiDAR; Point cloud data; Deep learning*

PENDAHULUAN

Klasifikasi pose manusia merupakan bagian dari kecerdasan buatan yang tidak hanya menjadi pondasi untuk berbagai aplikasi, seperti navigasi robot dan pengawasan keamanan di tempat kerja, tetapi juga memainkan peran penting dalam pengambilan keputusan yang cepat dan efisien dalam situasi yang memerlukan tindakan responsif. Namun, di tengah perkembangan teknologi ini, muncul tantangan baru terkait dengan masalah sensitivitas privasi yang terkait dengan pengambilan data tubuh manusia. Terutama, di lingkungan kesehatan atau tempat umum, di mana pemantauan kontinu sering diperlukan, pengambilan sampel tubuh manusia dapat mengancam privasi individu dengan mengungkapkan detail anatomis yang sensitif, termasuk lekukan tubuh dan bahkan aurat (Mahanani & Sulistyanta, 2022).

Sebagian besar dari teknologi klasifikasi pose manusia menggunakan data yang mudah dikenali oleh manusia berdasarkan gender dan sebagainya, sehingga penting untuk mempertimbangkan dampaknya terhadap privasi individu. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan pendekatan baru yang mempertimbangkan kepentingan privasi, dan tetap mempertahankan tingkat akurasi dalam klasifikasi pose manusia.

Pemahaman pose manusia adalah asal-usul untuk mengenali niat dan tindakan yang akan dilakukan selanjutnya (Fürst et al., 2020). Dengan melihat pose apa yang sedang dilakukan manusia tersebut, maka dapat dilakukan tindakan sebelum terjadinya hal yang tidak diinginkan. Dengan melakukan pengamatan lewat model klasifikasi, pemantauan terhadap pekerjaan yang memiliki resiko cedera dapat dilakukan dengan cara yang aman (Stenum et al., 2021).

Perkembangan teknologi dalam pemrosesan data poin cloud sudah mulai berkembang untuk berbagai keperluan. Ada yang menggunakan data *point cloud* dari LIDAR ini untuk perkembangan self-driving car (Schlager et al., 2020), pemantauan hutan (D. Xu et al., 2021), dan masih banyak lagi.

Klasifikasi pose manusia berbasis *point cloud* dari LIDAR sudah pernah diteliti (Putra et al., 2023). Dimana penelitian tersebut menggunakan metode DOconv berlapis tiga dan menghasilkan akurasi 87,06%.

Kontribusi pada penelitian ini ialah menggunakan metode deep learning dengan menerapkan conv3d sebagai arsitekturnya. Dengan menggunakan ini, kami berhasil mendapatkan akurasi 95.76%.

METODE

Tahapan-tahapan penelitian klasifikasi pose manusia berbasis *point cloud* dengan menggunakan *deep learning* ini dapat dilihat pada blog diagram yang tertera pada gambar 1.



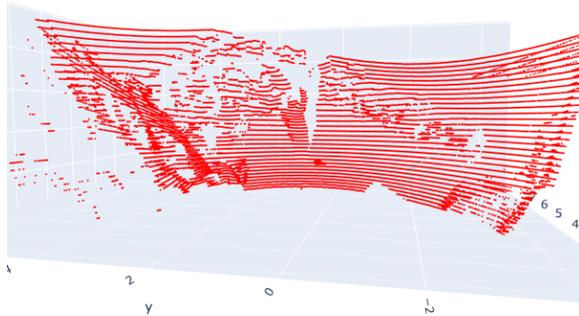
Gambar 1. Blog Diagram

Pada penelitian ini, kami menggunakan LiDAR OUSTER dengan tipe OS1-32 untuk mendapatkan data pose manusia yang berbasis *point cloud*. Alat sensor tersebut dapat mengumpulkan data dalam jangkauan 90 meter, dan bidang pandangnya atau *field of view (FoV)* dengan sudut 45°. Alat sensor tersebut menghasilkan 32x2048 *point cloud* setiap framenya.

Data yang dihasilkan oleh lidar tersebut merupakan data PCAP (*Packet Capture*). Sehingga memerlukan proses lanjutan agar dapat diolah menjadi data yang bisa digunakan dalam klasifikasi.

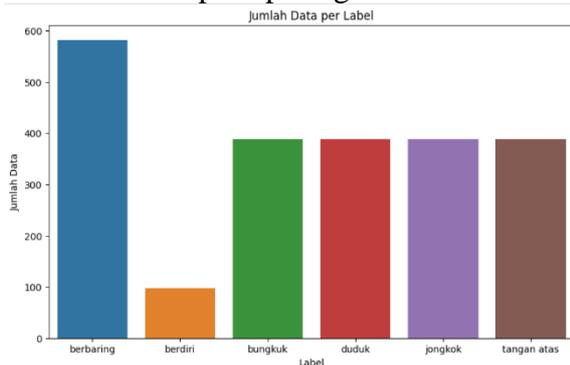
Untuk pengolahan data PCAP tersebut, kami melakukan proses konversi file dengan tipe file PCAP menjadi PCD memerlukan

library dari *ouster* tersebut. *Library ouster* didapat dengan menginstall SDK *Ouster*. Dengan menggunakan SDK dari *Ouster* tersebut, kami berhasil mengekstrak file PCAP menjadi PCD. Adapun hasil ekstraksinya seperti pada gambar 2.



Gambar 2. Hasil ekstraksi PCAP menjadi PCD

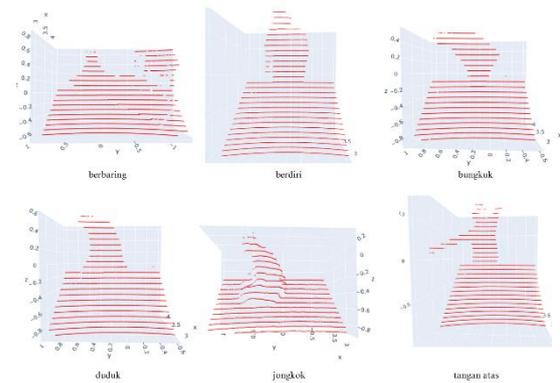
Ekstraksi file itu menghasilkan 98 PCD setiap data PCAP. Kami mengambil data pose dengan 6 pose dan beberapa sudut pandang yang berbeda. Sudut pandang pada berbaring ada 5, berdiri 1, bungkuk 4, duduk 4, jongkok 4, dan tangan atas ada 4 sudut pandang. Sehingga menghasilkan sebaran data seperti pada gambar 3.



Gambar 3. Sebaran data

Setelah berhasil untuk ekstraksi file PCAP menjadi PCD. Maka tahap selanjutnya ialah mengambil objek yang penting untuk klasifikasi dengan cara *crop* PCD. Pada tahap tersebut sudah termasuk *preprocessing*.

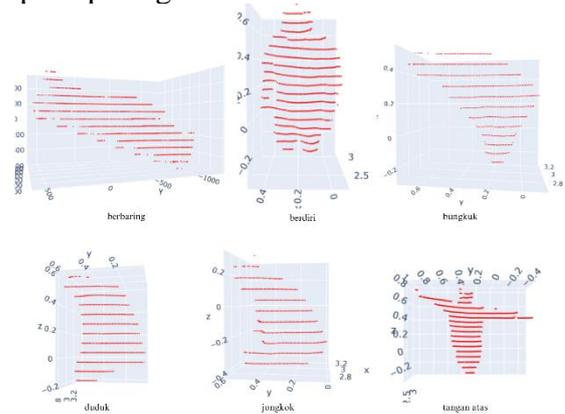
Letak objek berada pada ditengah-tengah ruangan. Dengan menggunakan bantuan *library* *Open3D* dan *numpy*, kami berhasil menentukan *bounding box* yang sesuai (Szutor, 2020). Sehingga mendapatkan hasil seperti gambar 4.



Gambar 4. hasil crop PCD

Crop PCD berhasil, namun hasil *crop* PCD tersebut mengambil bayangan yang mengganggu seperti *ground point* atau biasa disebut tanah datar. Dikarenakan tanah datar tersebut bisa ikut terklasifikasi, maka perlu dihilangkan.

Dengan menggunakan metode *RANSAC* yang menggunakan *threshod* 0,2 (Wu et al., 2020). Maka dihasilkanlah seperti pada gambar 5.



Gambar 5. Hasil Remove Ground Point

Setelah melalui berbagai tahap untuk perubahan file dari PCAP menjadi PCD. Maka tahap selanjutnya ialah mempersiapkan data tersebut agar bisa masuk kedalam arsitektur *Conv3D* dengan cara mengekstraksi fitur.

Dalam proses ekstraksi fitur, kami mengisi atribut *points* dari objek *pcd* dengan *points* yang diberikan sebagai parameter. Langkah ini dilakukan untuk mengkonversi array *numpy* menjadi tipe data yang dapat digunakan oleh *Open3D* (Zhou et al., 2018). Ekstraksi fitur dilakukan untuk mengubah representasi data *point cloud* menjadi vektor

fitur yang dapat digunakan oleh model klasifikasi.

Namun, sebelum dimasukkan kedalam model klasifikasi. Kami membaginya terlebih dahulu 80% untuk data *train* dan 20% untuk data *testing*. Setelah itu label di *encode* menjadi numerik agar memudahkan model dalam klasifikasi dengan hasil encode seperti pada tabel 1.

Tabel 1. Label Encode

Label	Label Encode
berbaring	0
berdiri	1
bungkuk	2
duduk	3
jongkok	4
tangan atas	5

HASIL DAN PEMBAHASAN

Model Conv3D yang telah dilatih berhasil mencapai akurasi sebesar 96,6% pada data uji. Selain itu, kami juga melakukan analisis kelas-wise accuracy untuk memahami kinerja model dalam mengklasifikasikan setiap kelas pose manusia. semoga dapat menjadi fokus untuk peningkatan kinerja model di masa depan.

Pembuatan Arsitektur Model

Model yang kami gunakan adalah *Convolutional Neural Networks* 3D (Conv3D) yang merupakan jenis arsitektur jaringan saraf tiruan yang digunakan untuk tugas-tugas pemrosesan citra dalam tiga dimensi. Pada penelitian ini, kami menggunakan Conv3D untuk melakukan klasifikasi pose manusia berdasarkan data *point cloud* (C. Xu et al., 2022).

Setelah data dipersiapkan, kami membangun arsitektur Conv3D. Arsitektur ini terdiri dari beberapa lapisan Conv3D, Flatten, Dense, dan Dropout. Pada lapisan Conv3D, kami menggunakan 32 filter dengan kernel konvolusi berukuran (1, 1, 3) dan fungsi aktivasi ReLU. Alasan kami menggunakan kernel konvolusi 1,1,3 karena shape data sebelum di reshape adalah (1792, 3).

Kami juga memasukkan lapisan Flatten untuk meratakan output dari lapisan sebelumnya menjadi vektor satu dimensi, kemudian lapisan Dense dengan 128 unit dan

fungsi aktivasi ReLU, serta lapisan Dropout dengan tingkat dropout sebesar 0.5 untuk mencegah overfitting.

Selanjutnya, kami menambahkan lapisan Dense terakhir dengan jumlah unit sesuai dengan jumlah kelas output (dalam kasus ini, jumlah kelas adalah 6 untuk klasifikasi pose manusia) dan menggunakan fungsi aktivasi softmax. Kami juga melakukan kompilasi model dengan loss function *categorical_crossentropy*, optimizer Adam dengan learning rate 0.001, dan metrik akurasi (accuracy).

Pada model ini, dilakukan dengan tahapan-tahapan layer yang dibutuhkan. Maka arsitektur CNN yang telah kami buat dapat dilihat jelasnya pada tabel 2.

Tabel 2. Arsitektur CNN yang kami gunakan

Nomor Layer	Tipe	Fungsi	Bentuk output	Kernel	Filter
1	Input	-	(1,1,3,1)	-	-
2	Conv3D	ReLU	(1,1,1,32)	(1,1,3)	32
3	Flatten	-	32	-	-
4	dense	ReLU	128	-	-
5	dropout	-	128	-	-
6	dense	Softmax	6	-	-

Dengan demikian, kami telah membuat Conv3D sebagai bagian dari arsitektur CNN yang telah kami desain dan mengintegrasikannya dengan data *point cloud* yang telah kami siapkan, maka tahap selanjutnya adalah pelatihan dan evaluasi model untuk mencapai tujuan klasifikasi pose manusia.

Training Model

Dalam proses *training* model, kami menggunakan kaggle dalam pengerjaan penelitian ini. Dengan menggunakan GPU P100, waktu untuk *training* dan *testing* tidak terasa lama.

Model klasifikasi pose manusia dilatih menggunakan metode fit dengan menggunakan dataset pelatihan (X_{train} , $y_{train_encoded}$). Pada metode fit, kami menentukan parameter-parameter seperti *batch size* (ukuran batch yang digunakan dalam setiap iterasi), *epochs* (jumlah iterasi pelatihan), dan *validation* data (dataset pengujian). Selama proses pelatihan, akan ditampilkan progres pelatihan setiap epoch, termasuk nilai loss dan akurasi untuk data

pelatihan (*training loss* dan *training accuracy*) serta untuk data validasi (*validation loss* dan *validation accuracy*).

Dalam penelitian ini kami menggunakan batch size 32 dan epochs sebanyak 2500.

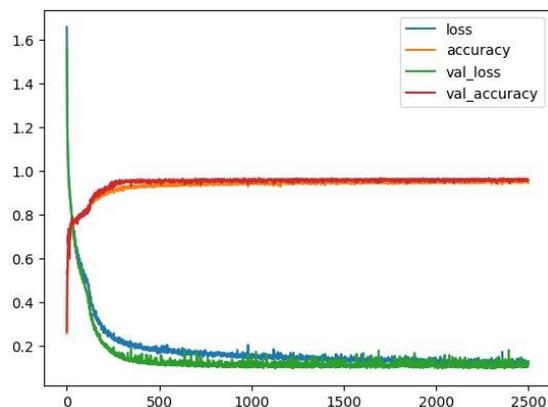
Evaluasi Model

Setelah model selesai dilatih, kami melakukan evaluasi kinerja model menggunakan metode evaluate dengan dataset pengujian ($X_{test}, y_{test_encoded}$). Metode ini mengembalikan dua nilai, yaitu loss dan accuracy dari dataset pengujian.

Dengan langkah-langkah ini, kami dapat mengevaluasi kinerja model pada dataset pengujian untuk memastikan bahwa model dapat menggeneralisasi dengan baik pada data baru. Pada contoh output yang diberikan, kami dapat melihat bahwa model mencapai akurasi sebesar 95.76% pada dataset pengujian setelah 2500 epochs.

Dengan menggunakan pustaka pandas, kami membuat DataFrame dari history yang berisi informasi tentang loss dan akurasi pada setiap epoch selama pelatihan model. Selanjutnya, kami melakukan plotting untuk menampilkan grafik loss dan akurasi pada setiap epoch selama pelatihan.

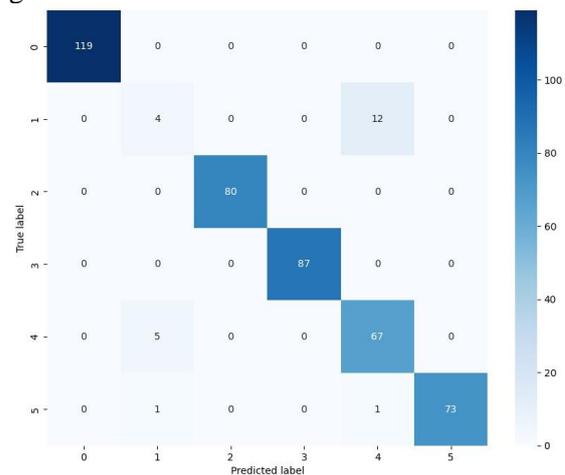
Plot ini membantu kami memahami tren performa model selama pelatihan, apakah ada peningkatan atau penurunan loss dan akurasi. Untuk grafiknya dapat dilihat pada gambar 6.



Gambar 6. Grafik peningkatan dan penurunan dari loss dan akurasi

Setelah itu, kami menggunakan confusion matrix untuk mengevaluasi kinerja model dengan melihat seberapa baik model dapat mengklasifikasikan data ke setiap kelas. Confusion matrix membantu kami melihat

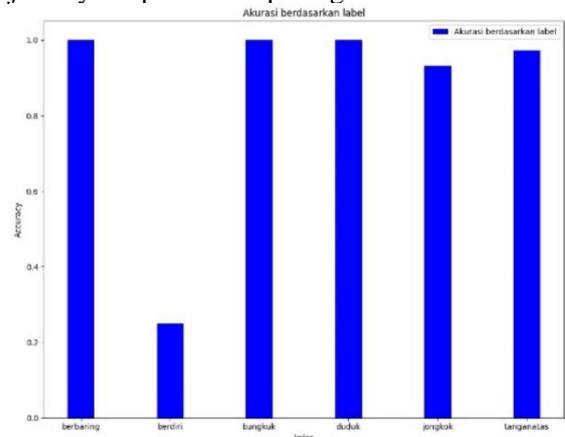
seberapa baik model dapat membedakan antara kelas-kelas yang berbeda. Contohnya ada pada gambar 7.



Gambar 7. Confusion matrix

Setelah mendapatkan confusion matrix, kami melakukan perhitungan akurasi berdasarkan kelas. Kami melakukan iterasi untuk setiap kelas dan dihitung akurasi berdasarkan jumlah prediksi yang benar dibagi dengan jumlah total prediksi untuk kelas tersebut.

Hasilnya dicetak untuk dianalisis lebih lanjut. Analisis ini membantu kami memahami kinerja model pada setiap kelas secara lebih terperinci. Adapun untuk akurasi berdasarkan kelas adalah 100% (berbaring), 56,25% (berdiri) 98,75% (bungkuk), 100% (duduk), 93,05% (jongkok), 100% (tangan atas). Atau lebih jelasnya dapat dilihat pada gambar 8.



Gambar 8. Akurasi berdasarkan label

SIMPULAN DAN SARAN

Dalam penelitian ini, kami berhasil melakukan klasifikasi pose manusia menggunakan data *point cloud* yang telah diproses melalui voxelisasi, serta menerapkan

pendekatan deep learning dengan *Convolutional Neural Networks* 3D (Conv3D).

Model yang dikembangkan mampu mengklasifikasikan pose manusia dengan akurasi yang baik, namun masih terdapat potensi untuk peningkatan kinerja dengan mengoptimalkan arsitektur model dan proses pelatihan. Penelitian ini memberikan kontribusi dalam pengembangan sistem pengenalan pose manusia berbasis data tiga dimensi.

DAFTAR PUSTAKA

- Fürst, M., Gupta, S. T. P., Schuster, R., Wasenmüller, O., & Stricker, D. (2020). *HPERL: 3D Human Pose Estimation from RGB and LiDAR*. <http://arxiv.org/abs/2010.08221>
- Mahanani, A., & Sulistyanta, '. (2022). Tinjauan Yuridis Tindak Pidana Pengancaman Dengan Foto Rekayasa Bermuatan Pornografi (Studi Putusan Nomor: 125/Pid.Sus/2019/Pn Bjn). *Recidive : Jurnal Hukum Pidana Dan Penanggulangan Kejahatan*, 11(1), 36. <https://doi.org/10.20961/recidive.v11i1.67433>
- Putra, O. V., Riandini, Yuniarno, E. M., & Purnomo, M. H. (2023). Depthwise Over-Parameterized CNN for Voxel Human Pose Classification. *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 54–59. <https://doi.org/10.1109/ISITIA59021.2023.10221054>
- Schlager, B., Muckenhuber, S., Schmidt, S., Holzer, H., Rott, R., Maier, F. M., Saad, K., Kirchengast, M., Stettinger, G., Watzenig, D., & Ruebsam, J. (2020). State-of-the-Art Sensor Models for Virtual Testing of Advanced Driver Assistance Systems/Autonomous Driving Functions. *SAE International Journal of Connected and Automated Vehicles*, 3(3), 233–261. <https://doi.org/10.4271/12-03-03-0018>
- Stenum, J., Cherry-Allen, K. M., Pyles, C. O., Reetzke, R. D., Vignos, M. F., & Roemmich, R. T. (2021). Applications of pose estimation in human health and performance across the lifespan. In *Sensors* (Vol. 21, Issue 21). MDPI. <https://doi.org/10.3390/s21217315>
- Szutor, P. (2020). *FFT Based Airborne LIDAR Classification with Open3D and Numpy/Scipy*. <http://ceur-ws.org>
- Wu, Z., Zhuo, G., & Xue, F. (2020). Self-Supervised Monocular Depth Estimation Scale Recovery using RANSAC Outlier Removal. *2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 97–102. <https://doi.org/10.1109/CVCI51460.2020.9338538>
- Xu, C., Yang, S., Galanti, T., Wu, B., Yue, X., Zhai, B., Zhan, W., Vajda, P., Keutzer, K., & Tomizuka, M. (2022). Image2Point: 3D Point-Cloud Understanding with 2D Image Pretrained Models. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13697 LNCS, 638–656. https://doi.org/10.1007/978-3-031-19836-6_36
- Xu, D., Wang, H., Xu, W., Luan, Z., & Xu, X. (2021). LiDAR applications to estimate forest biomass at individual tree scale: Opportunities, challenges and future perspectives. *Forests*, 12(5), 1–19. <https://doi.org/10.3390/f12050550>
- Zhou, Q.-Y., Park, J., & Koltun, V. (2018). *Open3D: A Modern Library for 3D Data Processing*. <http://arxiv.org/abs/1801.09847>