



SEGMENTASI BAGIAN TUBUH MANUSIA 3D *POINT CLOUD* BERBASIS *SPHERICAL PROJECTION* MENGGUNAKAN SEGNET

Muhammad Akmal^{1*}, Oddy Virgantara Putra², Faisal Reza Pradhana¹

¹Program Studi Teknik Informatika, Universitas Darussalam Gontor, Ponorogo

e-mail: *¹muh.akmal18@mhs.unida.gontor.ac.id, ²oddy@unida.gontor.ac.id,
³faisal@unida.gontor.ac.id

Abstrak

Light Detection and Ranging (LIDAR) adalah sebuah teknologi sensor yang telah berkembang pesat. Teknologi ini sudah banyak diimplementasikan di dalam bidang segmentasi citra dan deteksi objek. Teknologi ini memvisualisasikan hasil pantulan dari objek sebagai 3D *point cloud*. Namun dalam segmentasi 3D *point cloud* memiliki sebuah kendala atau tantangan yaitu struktur data yang tidak teratur, yang mengakibatkan *noise* dan bentuk objek yang menyebabkan saling tumpang tindih, sehingga menghambat akurasi segmentasi dan pengenalan objek. Untuk menyelesaikan masalah ini pendekatan *spherical projection* hadir sebagai jawaban. Dengan mengubah data 3D *point cloud* menjadi citra 2D dapat memudahkan proses segmentasi citra. Dalam melakukan segmentasi citra dari hasil transformasi data 3D menjadi 2D dengan menggunakan pelatihan *deep learning* dan model yang digunakan dalam pelatihan data adalah SegNet yang terkenal efektif dalam melakukan segmentasi citra. Setelah training data, diperoleh akurasi sebesar 59% dengan epoch sebanyak 500. Diharapkan dengan penelitian ini penelitian selanjutnya dapat menyempurnakan dan mengembangkan ilmu pengetahuan pada bidang visi komputer dan kecerdasan buatan.

Kata kunci—Lidar; Point Cloud; Segmentasi; Spherical Projection; SegNet

Abstrak

Light Detection and Ranging (LIDAR) adalah sebuah teknologi sensor yang telah berkembang pesat. Teknologi ini sudah banyak diimplementasikan di dalam bidang segmentasi citra dan deteksi objek. Teknologi ini memvisualisasikan hasil pantulan dari objek sebagai 3D *point cloud*. Namun dalam segmentasi 3D *point cloud* memiliki sebuah kendala atau tantangan yaitu struktur data yang tidak teratur, yang mengakibatkan *noise* dan bentuk objek yang menyebabkan saling tumpang tindih, sehingga menghambat akurasi segmentasi dan pengenalan objek. Untuk menyelesaikan masalah ini pendekatan *spherical projection* hadir sebagai jawaban. Dengan mengubah data 3D *point cloud* menjadi citra 2D dapat memudahkan proses segmentasi citra. Dalam melakukan segmentasi citra dari hasil transformasi data 3D menjadi 2D dengan menggunakan pelatihan *deep learning* dan model yang digunakan dalam pelatihan data adalah SegNet yang terkenal efektif dalam melakukan segmentasi citra. Setelah training data, diperoleh akurasi sebesar 59% dengan epoch sebanyak 500. Diharapkan dengan penelitian ini penelitian selanjutnya dapat menyempurnakan dan mengembangkan ilmu pengetahuan pada bidang visi komputer dan kecerdasan buatan.

Kata kunci—Lidar; Point Cloud; Segmentasi; Spherical Projection; SegNet

I. PENDAHULUAN

Computer Vision merupakan cabang dari *Machine Learning* (Dandil & Cevik,

2019) yang sudah berkembang. Teknologi ini memungkinkan menganalisis data foto atau video (Demir et al., 2020). Perkembangan *computer vision* sangat membantu berbagai bidang (Akagic et al.,



2022). Dikarenakan kesulitannya dan variasi *point cloud* yang tidak teratur, kerapatan yang berbeda, jenis objek yang beragam, dan lainnya, segmentasi *point cloud* merupakan topik penelitian yang sangat hangat diteliti (Grilli et al., 2017). Kegunaan segmentasi dapat bervariasi tergantung pada konteks dan kebutuhannya, berikut fungsi utamanya adalah pengenalan objek, deteksi objek pada citra, dan lainnya (Wolfe et al., 2002).

Pada penelitian Yingshen He, Ge Li, Yiting Shao, Jing Wang, Yueru Chen, dan Shan Liu (He et al., 2020). Teknologi kompresi untuk *point cloud* sangat diperlukan untuk penyimpanan dan transmisi. Penelitian ini mengusulkan sebuah kerangka kerja kompresi *point cloud* yang baru berdasarkan *spherical projection* untuk kompresi *point cloud* geometri dan atribut yang bersifat *lossless* dan *lossy*. Hasil penelitian menunjukkan potensi besar dalam kompresi tanpa kehilangan informasi untuk dataset LiDAR dan memiliki kinerja yang baik untuk kompresi, yaitu tidak berkurang informasi dari dataset objek *point cloud*. Pada penelitian yang saya lakukan mengubah data 3D *Point Cloud* menjadi citra 2D dengan *spherical projection* agar tidak mengurangi banyak informasi yang terkandung dalam data 3D, setelah itu melakukan segmentasi anggota tubuh manusia menggunakan model SegNet.

Tugas segmentasi 3D *point cloud* adalah mengambil data *point cloud* dan memisahkannya kedalam wilayah yang memiliki properti yang sama (Nguyen & Le, 2013). *Spherical Projection* adalah teknik pemetaan atau proyeksi yang digunakan untuk mengubah data 3D *point cloud* menjadi citra 2D. SegNet merupakan sebuah arsitektur jaringan saraf yang bertugas dalam segmentasi citra atau pengenalan pola di tingkat piksel. SegNet telah berhasil digunakan dalam berbagai

tugas segmentasi citra dengan mempertahankan akurasi tinggi (Gad et al., 2020). Arsitektur ini telah menunjukkan kinerja yang baik dalam banyak aplikasi dan menjadi salah satu metode yang populer untuk tugas segmentasi objek (Tang et al., 2018). Tugas segmentasi bagian tubuh manusia untuk memisahkan atau mengidentifikasi berbagai bagian tubuh manusia sehingga memungkinkan analisis lebih lanjut tentang struktur dan perilaku manusia (Wen et al., 2012). Hal ini dapat bermanfaat dalam berbagai aplikasi, seperti pengenalan gerakan, analisis postur tubuh, atau pengembangan teknologi realitas virtual (Yang et al., 2020). Karena aplikasinya yang luas dan tantangan yang kompleks (Jalal et al., 2019).

Berbeda dengan penelitian sebelumnya yang hanya menguji pendekatan *spherical projection*, tujuan penelitian ini untuk mengembangkan penelitian sebelumnya, yaitu mengimplementasikan hasil *spherical projection* kedalam segmentasi, dan melihat apakah informasi yang terkandung dalam hasil *spherical projection* tidak banyak berkurang. Dengan menggunakan jaringan SegNet yang populer dalam penelitian segmentasi citra objek. Adapun objek yang digunakan yaitu anggota tubuh manusia.

II. METODOLOGI PENELITIAN

Untuk penelitian ini, peneliti membuat tahapan dan rancangan penelitian yang akan digunakan untuk membangun model segmentasi tersebut;

2.1. Pengumpulan Data

Dalam penelitian ini, dataset yang digunakan adalah hasil dari pemindaian LiDAR yang berbentuk 3D *point cloud*. Dari dataset tersebut memuat file dengan jumlah 21 file, dan setiap file memiliki 2 jenis format file, yaitu .JSON dan .PCAP. Masing-masing file berisikan pose manusia

dengan pose yang berbeda-beda. Dari 21 file tersebut, terdapat file dengan posisi dan pose yang sama tetapi dengan sudut yang berbeda, ada dari sudut depan, belakang, kanan, dan kiri. Berikut rincian dari file tersebut. Pose berbaring terdapat 4 file, pose berdiri terdapat 1 file, pose bungkuk terdapat 4 file, pose duduk terdapat 4 file, pose jongkok terdapat 4 file, dan pose tangan terdapat 4 file.

2.2. Spherical Projection

Dalam Segmentasi objek berbasis *spherical projection*, Langkah pertama yaitu mentransformasikan data 3D menjadi citra 2D, dimana posisi setiap titik 3D diproyeksikan ke bidang 2D. Dalam proses ini, banyak keuntungan yang didapat, antara lain; (1) Data dapat direpresentasikan dalam format yang lebih mudah dikelola dan diolah lebih lanjut menggunakan berbagai teknik pemrosesan citra, (2) Dapat memanfaatkan algoritma dan metode citra 2D, (3) Citra 2D yang dihasilkan dari transformasi, dapat diaplikasikan dalam visi komputer.

Transformation, akan dilakukan konversi global cartesian coordinates (x,y,z) menjadi local spherical coordinates (φ,θ,r). Dengan rumus pada Gambar 2.

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c)^2}$$

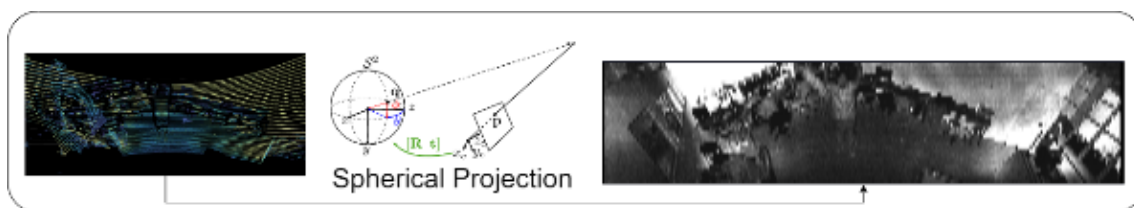
$$\varphi_i = \arccos \frac{|x_i - x_c|}{\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}}$$

$$\theta_i = \arcsin \frac{|z_i - z_c|}{\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c)^2}}$$

Gambar 2. Rumus *Coordinates Transformation*

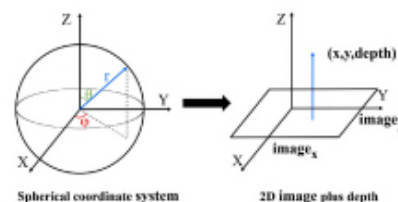
Nilai r_i adalah jarak atau radius dari titik tersebut ke titik pusat. Nilai θ_i dan φ_i masing-masing mewakili sudut horizontal dan sudut vertikal dari titik dalam *spherical coordinates*. Setelah itu, terdapat proses *Range Images Projection*. Pada proses *Range Images Projection*, *Spherical Coordinates* (φ,θ,r) akan dikonversi menjadi 2D *Image plus Depth*.

Dapat dilihat pada Gambar 3 dan Gambar 4 untuk rincian lebih jelas tentang proses *range image*.



Gambar 1. Hasil data *point cloud* setelah melalui proses *spherical projection*

Pada Gambar 1 adalah proses *spherical projection*, pada proses ini data point cloud akan diubah menjadi citra 2D yang akan dilakukan proses segmentasi pada hasil citra 2D tersebut (Cao et al., 2018). Dalam mengolah data *point cloud*, penggunaan sistem koordinat bola akan lebih baik beradaptasi dengan karakteristik dataset, karena titik-titik data point cloud objek berada di permukaan objek dan tidak ada titik pusatnya. Pada proses *Coordinates*



Gambar 3. Proses *Range Image*

$$x_i = \text{round}\left(\frac{\varphi_i}{2\pi} * \text{image}_x\right)$$

$$y_i = \text{round}\left(\frac{\theta_i}{\pi} * \text{image}_y\right)$$

$$\text{depth}_i = \text{round}(r_i)$$

Gambar 4. Rumus *Range Image*

2.3. Pre-processing Data

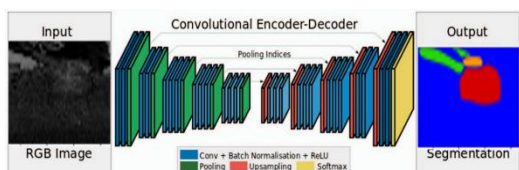
Pada tahapan Preprocessing Data, akan dilakukan pemotongan data gambar range image menjadi gambar yang mudah untuk melakukan masking, lalu resize dan pembagian data.

2.4. Segmentasi Jaringan Segnet

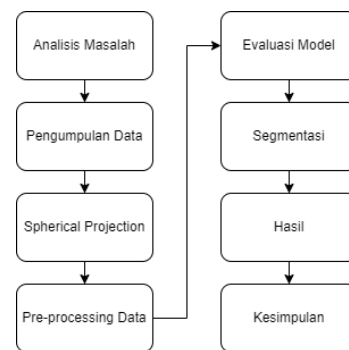
SegNet memiliki 2 tahap (Vianna et al., 2021), pertama adalah encoder, dimana gambar input melewati serangkaian lapisan konvolusi dan max-pooling. Lapisan konvolusi digunakan untuk mengekstraksi fitur-fitur penting dari gambar, sedangkan max-pooling digunakan untuk mengurangi dimensi gambar dan mempertahankan fitur yang paling signifikan.

Tahap kedua, SegNet menggunakan proses decoder untuk merestrukturisasi hasil dari tahap encoder menjadi peta segmentasi piksel. Lapisan ini melibatkan lapisan unpooling, yang menggunakan indeks yang dicatat selama max-pooling untuk mengembalikan dimensi gambar ke ukuran semula.

Komponen penting dalam jaringan SegNet yaitu memiliki convolutional layers, batch normalization layers, activation layers, pooling layers, upsampling layers (Chen & Lu, 2019). Gambar 5, menggambarkan secara jelas proses pada jaringan SegNet.

Gambar 5. Proses jaringan SegNet yang terdapat lapisan *convolutional encoder-decoder*

Penelitian ini akan melewati beberapa tahapan. Antara lain; Analisis masalah, Pengumpulan data, Spherical Projection, Pre-processing Data, Evaluasi Model, Segmentasi, Visualisasi Hasil, dan kesimpulan. Tahapan-tahapan tersebut dapat dilihat pada Gambar 6.

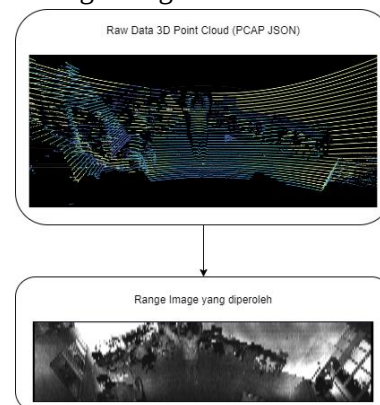


Gambar 6. Rancangan Penelitian yang diajukan

III. HASIL DAN PEMBAHASAN

3.1. Tahap Pengumpulan Data

Raw data yang berupa 3D *Point Cloud* dengan format file PCAP dan JSON, memiliki bentuk yang sulit untuk diolah. Maka *Spherical projection* ada untuk memudahkan proses tersebut. Dalam proses spherical projection, 3D Point Cloud diubah menjadi range image.



Gambar 7. Proses Spherical Projection pada penelitian ini

Pada Gambar 7, adalah bagaimana *range image* diperoleh melalui perangkat lunak Ouster Studio versi 2.0.2 yang dapat membaca file 3D Point Cloud dengan format PCAP dan JSON. Subjek dari data tersebut adalah seorang pria yang mengangkat kedua tangannya dan berputar 360 derajat, dan direkam selama 9 detik.

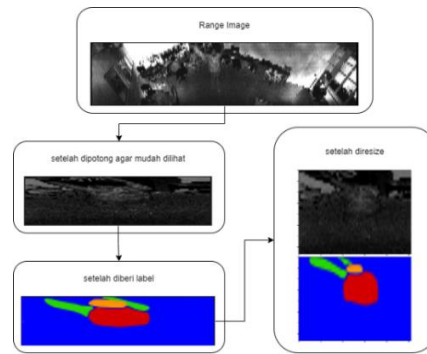
Selama durasi 9 detik, data yang didapat menjadi 71 gambar. Setelah itu data gambar dipotong agar lebih fokus kepada subjek, yaitu seorang pria yang berputar dengan kedua tang diangkat. Gambar dipotong menjadi 126x32, lalu diberi label secara manual, pada Tabel 1 rincian pembagian data yang dipakai.

No	Nama Data	Data Gambar	Label Gambar	Total
1.	Data Train	71	71	142
2.	Data Test	71	71	142
3.	Data Val	71	71	142

Tabel 1. Pembagian data

3.2. Tahap Pre-processing

Sebelum *range image* yang telah melalui proses spherical projection memasuki proses segmentasi, *range image* harus dipotong. Setelah *range image* terpotong menjadi ukuran yang mudah dilihat, memberi label secara manual dengan aplikasi desain. Agar dapat memasuki model yang akan digunakan, dalam SegNet ukuran gambar harus berukuran 192x256 piksel. Maka dari itu, data asli dan data yang telah diberi label akan di *resize* menjadi ukuran yang dibutuhkan agar dapat memberikan kinerja yang maksimal untuk pengujian dan validasi. Gambar 8 menjelaskan secara rinci proses pre-processing.



Gambar 8. Proses Pre-processing

3.3. Tahap Processing

Pada tahap ini, proses pelatihan model dilakukan menggunakan platform Kaggle menggunakan GPU T4x2 yang disediakan. Pada penelitian ini dilakukan pelatihan model SegNet untuk Segmentasi bagian tubuh manusia untuk menentukan kepala, tangan dan badan dari manusia

Tabel 2. Pembagian Parameter

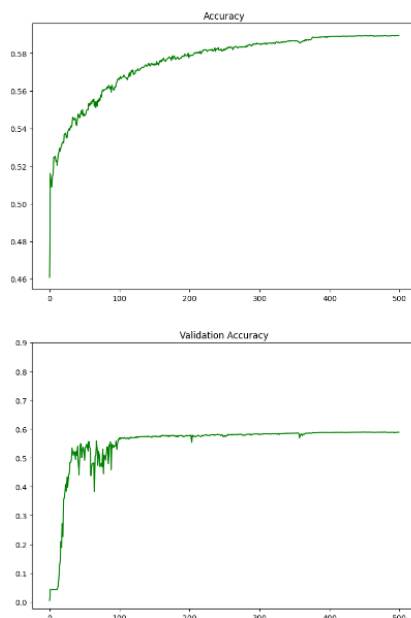
Parameter	Value
Convolutional Layer	26
Pooling Layer	5
Upsampling Layer	5
Activation Function	Relu, Sigmoid
Optimizer	Adam
Loss Function	Categorical Crossentropy
Epoch	100
Batch Size	18
Learning Rate	0.0001

Proses segmentasi dilakukan menggunakan arsitektur model SegNet dengan menggunakan parameter yang sudah dijelaskan di Tabel 2. Pada proses Encoder pada model yang dibuat menggunakan 13 blok dan proses Decoder juga menggunakan 13 blok yang terdiri dari ReLU Layer, Batch Normalization Layer, Convolution Layer, dan juga menggunakan 5 blok Pooling Layer dan 5 blok Upsampling Layer dengan filter kernel 3x3. Optimizer yang digunakan adalah adam dan loss function adalah Categorical

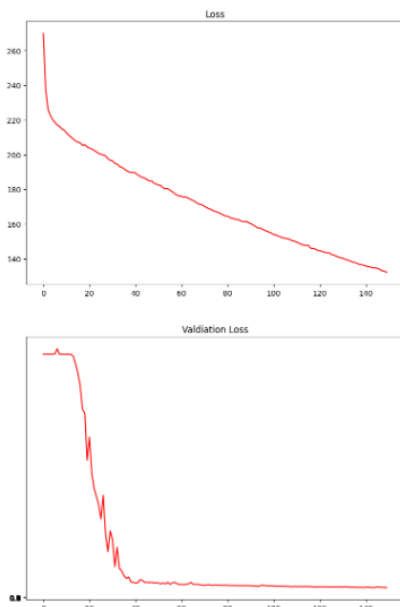
Crossentropy dengan jumlah batch 18. Dengan proses pelatihan sebanyak 500 Epochs.

3.4. Tahap Uji Coba dan Validasi

Akurasi pada pengujian model kali ini terlihat cukup baik dengan mencapai 0.59 . Adapun loss yang didapatkan dari hasil pengujian, penjelasan bisa dilihat pada Gambar 10 dan Gambar 11.

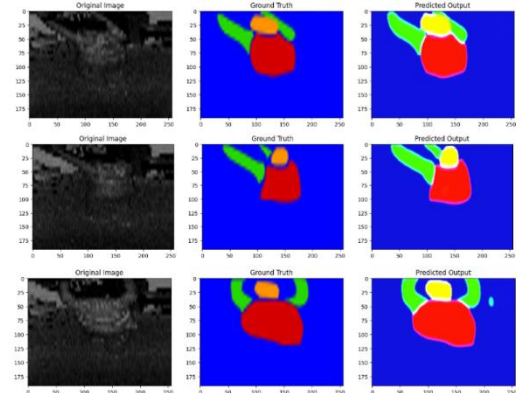


Gambar 10. Grafik Akurasi dan Val Akurasi



Gambar 11. Grafik Loss dan Val Loss

Model *SegNet* yang sudah dirancang sesuai keperluan dalam penelitian ini kemudian dilakukan prediksi pada gambar secara acak. Adapun hasil prediksi gambar bisa dilihat di gambar 12.



Gambar 12. Hasil Segmentasi dan Prediksi Citra

Dari gambar 12 yaitu hasil Segmentasi menggunakan model *SegNet* untuk Segmentasi Anggota tubuh manusia sudah memenuhi tujuan yang diharapkan yaitu bisa membedakan antara kepala, tangan dan badan manusia. Adapun hasil segmentasi, terlihat masih ada yang gagal dan tidak sesuai dengan yang diharapkan, seperti hasil segmentasi yang kurang jelas dan prediksi yang kurang sesuai seperti kepala yang terlalu besar dan badan yang terlalu kecil dari *Ground Truth*.

IV. KESIMPULAN

Pada penelitian ini, dapat dirangkum beberapa kesimpulan yang dapat diambil, sebagai berikut :

1. Dari hasil pelatihan pada model, dapat disimpulkan penggunaan model arsitektur *SegNet* untuk Segmentasi 3D Point Cloud pada daun padi mendapat hasil yang cukup baik dengan melihat akurasi yang diperlihatkan dengan nilai akurasi mencapai 0,59.



2. Terdapat kekurangan yaitu masih kurang jelas hasil segmentasi.
3. Jumlah dataset dan pengambilan citra yang sesuai dengan tingkat keparahan sangat mempengaruhi kinerja model, mulai dari kecepatan proses hingga tingkat akurasi.

V. SARAN

Berdasarkan temuan penelitian ini, peneliti memiliki beberapa saran untuk penelitian selanjutnya yang dapat membantu mengembangkan penelitian ini lebih lanjut, yaitu :

1. Memperluas objek penelitian dengan tidak mengambil 3D Point Cloud anggota tubuh manusia.
2. Memperbanyak pengambilan data untuk bahan pengolahan Segmentasi serta pengambilan citra yang sesuai dengan tingkat keparahan pada penyakit.
3. Mengembangkan hasil model dan mengimplementasikan model yang telah dibuat ke dalam objek seperti *android*.

VI. DAFTAR PUSTAKA

- Akagic, A., Krivic, S., Dizdar, H., & Velagic, J. (2022). Computer Vision with 3D Point Cloud Data: Methods, Datasets and Challenges. *2022 28th International Conference on Information, Communication and Automation Technologies, ICAT 2022 - Proceedings, June*.
<https://doi.org/10.1109/ICAT54566.2022.9811120>
- Cao, Z., Huang, Q., & Karthik, R. (2018). 3D object classification via spherical projections. *Proceedings - 2017 International Conference on 3D Vision, 3DV 2017*, 566–574.
<https://doi.org/10.1109/3DV.2017.00070>
- Chen, H., & Lu, S. (2019). Building Extraction from Remote Sensing Images Using SegNet. *2019 IEEE 4th International Conference on Image, Vision and Computing, ICIVC 2019*, 227–230.
<https://doi.org/10.1109/ICIVC47709.2019.8981046>
- Dandil, E., & Cevik, K. K. (2019). Computer Vision Based Distance Measurement System using Stereo Camera View. *3rd International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2019 - Proceedings*, 6–9.
<https://doi.org/10.1109/ISMSIT.2019.8932817>
- Demir, E., Yildiz, K., Demir, O., & Ulku, E. E. (2020). Computer Vision Based Intelligent 3D CNC Machines. *Proceedings - 2020 Innovations in Intelligent Systems and Applications Conference, ASYU 2020*.
<https://doi.org/10.1109/ASYU50717.2020.9259826>
- Gad, G. M., Annaby, A. M., Negied, N. K., & Darweesh, M. S. (2020). Real-Time Lane Instance Segmentation Using SegNet and Image Processing. *2nd Novel Intelligent and Leading Emerging Sciences Conference, NILES 2020*, 253–258.
<https://doi.org/10.1109/NILES50944.2020.9257977>
- Grilli, E., Menna, F., & Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(2W3), 339–344.
<https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017>
- He, Y., Li, G., Shao, Y., Wang, J., Chen, Y., & Liu, S. (2020). A point cloud compression framework via spherical projection. *2020 IEEE International Conference on Visual Communications and Image Processing, VCIP 2020, December*, 62–65.
<https://doi.org/10.1109/VCIP49819.2020.9301809>
- Jalal, A., Nadeem, A., & Bobasu, S. (2019). Human Body Parts Estimation and Detection for Physical Sports Movements. *2019 2nd International Conference on Communication, Computing and Digital Systems, C-*



- CODE 2019, 104–109.
<https://doi.org/10.1109/C-CODE.2019.8680993>
- Nguyen, A., & Le, B. (2013). 3D Point Cloud Segmentation. *2013 6th IEEE Conference on Robotics*, 1–6.
- Tang, J., Li, J., & Xu, X. (2018). Segnet-based gland segmentation from colon cancer histology images. *Proceedings - 2018 33rd Youth Academic Annual Conference of Chinese Association of Automation, YAC 2018*, 1078–1082.
<https://doi.org/10.1109/YAC.2018.8406531>
- Vianna, P. O., Farias, R., & Pereira, W. C. A. (2021). Performance of the SegNet in the Segmentation of Breast Ultrasound Lesions. *Pan American Health Care Exchanges, PAHCE, 2021-May*.
<https://doi.org/10.1109/GMEPE/PAHCE50215.2021.9434877>
- Wen, Z., Yan, Y., & Cui, H. (2012). Study on segmentation of 3D human body based on point cloud data. *Proceedings - 2012 International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2012*, 657–660.
<https://doi.org/10.1109/ISdea.2012.676>
- Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research*, 42(28), 2985–3004.
[https://doi.org/10.1016/S0042-6989\(02\)00388-7](https://doi.org/10.1016/S0042-6989(02)00388-7)
- Yang, Z., Li, Y., Yang, L., Zhang, N., & Luo, J. (2020). Weakly supervised body part segmentation with pose based part priors. *Proceedings - International Conference on Pattern Recognition*, 286–293.
<https://doi.org/10.1109/ICPR48806.2021.9412887>