

## KLASIFIKASI VIRAL PNEUMONIA MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK DAN SUPPORT VECTOR MACHINE

Naufal Abdillah<sup>1</sup>, Andrey K.W. Hapantenda<sup>1,\*</sup>, Ahmad Habib<sup>1</sup>, Indah Listiowarni<sup>2</sup>

<sup>1</sup>Teknik Informatika, Fakultas Teknik, Universitas 17 Agustus 1945 Surabaya

<sup>2</sup>Teknik Informatika, Fakultas Teknik, Universitas Madura

E-mail: andreyhapantenda@untag-sby.ac.id\*

### ABSTRACT

*Cases of Corona Virus Disease 2019 (Covid-19) have been found in Indonesia since March 2019, for diagnosis in addition to the Real Time - Polymerase Chain Reaction (RT-PCR) examination, a Thorax X-ray is also carried out to detect the presence of Pneumonia. Pneumonia is an inflammation of the air sacs in the lungs that can fill with fluid. Pneumonia can be caused by a bacterial, viral or fungal infection. The proposed method combines the CNN method with SVM for the classification of viral Pneumonia with bacterial Pneumonia. The results of this study are expected to help radiology officers in classifying Pneumonia caused by viruses from the results of Thorax X-rays. The performance of this system is measured using confusion matrix and produces a score of average accuracy 0.85.*

**Keywords:** Convolutional Neural Network, Covid-19, Support vector Machine, Viral Pneumonia.

### ABSTRAK

Kasus Corona Virus Disease 2019 (Covid-19) ditemukan di Indonesia sejak Maret 2019, untuk penegakan diagnose selain pemeriksaan Real Time – Polymerase Chain Reaction (RT-PCR) juga dilakukan pemeriksaan Rontgen Thorax untuk mendeteksi adanya Pneumonia. Pneumonia merupakan peradangan pada kantung udara pada paru-paru yang dapat berisi cairan. Penyebab Pneumonia bisa berasal dari infeksi bakteri, virus ataupun jamur. Metode yang diusulkan menggabungkan metode CNN dengan SVM untuk klasifikasi Pneumonia virus dengan Pneumonia bakteri. Hasil dari penelitian ini diharapkan dapat membantu petugas radiologi dalam mengklasifikasikan Pneumonia yang disebabkan oleh virus dari hasil Rontgen Thorax. Performa sistem ini diukur menggunakan confusion matrix dan menghasilkan skor akurasi rata-rata sebesar 0.85.

**Kata Kunci:** Convolutional Neural Network, Covid-19, Support Vector Machine, Viral Pneumonia

### 1. Pendahuluan

Sebanyak dua puluh persen kematian pada bayi disebabkan oleh Pneumonia (Risksdas, 2019). Pneumonia termasuk salah satu jenis penyakit infeksi saluran pernafasan akut (ISPA) yang disebabkan bakteri, virus, jamur maupun parasit. Pneumonia dapat menyebabkan demam tinggi, sesak nafas dan hipoksia.. Faktor penyebab Pneumonia antara lain kondisi individu, gaya hidup dan perilaku, serta kebersihan lingkungan hidup. Kasus Pnumonia paling tinggi terjadi pada kelompok usia 56-65 tahun, hal ini disebabkan perubahan fisiologi akibat proses degeneratif yang menyebabkan penurunan fungsi Paru

dan daya tahan tubuh. Umumnya, diagnosis Pneumonia dilakukan secara klinis atau fisik gejala oleh dokter. Selain itu, tes lanjutan yang bisa dilakukan adalah Rontgen toraks untuk mendapatkan gambar bagian organ paru-paru yang mengalami kelainan. Rontgen dada pemeriksaan adalah tes pencitraan menggunakan jenis radiasi yang disebut gelombang elektromagnetik. Hasil pemeriksaan akan menghasilkan gambar atau disebut juga gambar dada itu menampilkan penampang organ internal manusia, terutama di rongga dada. Citra dada adalah investigasi yang efisien dalam mengungkap perubahan patologi. Pemeriksaan yang lebih teliti diperlukan untuk setiap perubahan pada

organ di rongga dada, terutama paru-paru. Namun, gambar yang diamati secara manual adalah kurang efektif dalam menentukan diagnosis, sehingga di sebabkan kesulitan dalam analisis visual dari organ atau objek. Ini adalah dipengaruhi oleh beberapa faktor seperti kesamaan visual yang tinggi antara beberapa patologi, *sensor noise*, elektronik gangguan, dan posisi pasien yang bisa berubah visualisasi citra dada yang sebenarnya. Munculnya masalah dalam penalaran kompleks membutuhkan pengamatan yang cermat dan keahlian khusus tentang prinsip-prinsip anatomi, fisiologi, dan patologi. Jadi, sebelumnya butuh waktu lama tenaga medis atau dokter mendiagnosis penyakit yang diderita pasien.

Beberapa penelitian sebelumnya telah dilakukan menggunakan metode CNN [1], Backpropagation [2], Transformasi Wavelet [3], Teknik Pengolahan Citra[4], namun hanya mengklasifikasikan *Pneumonia* dengan Paru normal. Penelitian Agarap [5] menggabungkan metode CNN dengan SVM untuk mengklasifikasikan citra MNIST dan Fashion MNIST dan menghasilkan akurasi yang tinggi, masing-masing sebesar 99,04% dan 90,72%.

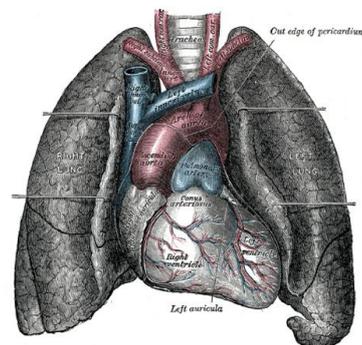
Penelitian ini menggabungkan metode CNN dengan SVM untuk membuat klasifikasi *Pneumonia* virus dengan *Pneumonia* bakteri pada citra digital *Rontgen* thorax. Hasil penelitian ini diharapkan dapat membantu dokter untuk melakukan diagnosa *Pneumonia* disebabkan virus atau bakteri, sehingga membantu dokter untuk memilih terapi yang tepat.

## 2. Tinjauan Pustaka

### 2.1. Paru-Paru

Paru-paru merupakan sepasang organ yang memiliki tekstur kenyal dan berisi udara, dibantu oleh Trakea dalam penghantaran udara. Paru-paru berfungsi sebagai tempat pertukaran oksigen dari udara dengan karbon dioksida dari darah. Paru-paru mengambil oksigen dari udara yang dihirup kemudian masuk ke aliran darah dan didistribusikan ke seluruh bagian sel, ketika

sel bekerja maka dihasilkan gas buangan berupa karbondioksida dilepaskan melalui aliran darah. Organ paru-paru terlibat dalam sintesis, penyimpanan, transformasi dan degradasi zat. Anatomi paru-paru dan jantung dapat dilihat pada Gambar 1.



Gambar 1. Anatomi Paru-Paru

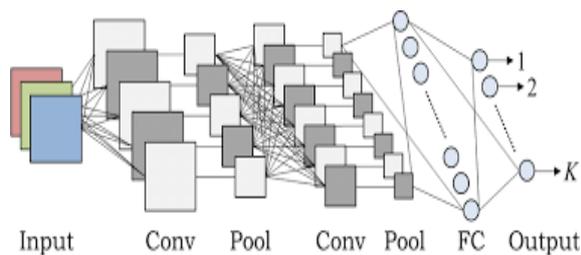
### 2.2. Convolutional Neural Network

Dalam *deep learning*, *Convolutional Neural Network* (CNN), atau ConvNet adalah kelas *deep neural network*, yang paling umum diterapkan untuk menganalisis citra visual. CNN juga dikenal sebagai *shift invariant* atau *space invariant artificial neural networks* (SIANN), berdasarkan arsitektur bobot bersama dan karakteristik invarian terjemahan [6]. Mereka memiliki aplikasi dalam pengenalan gambar dan video, sistem pemberi rekomendasi, klasifikasi gambar, analisis gambar medis, pemrosesan bahasa alami, dan financial time series.

CNN merupakan *multilayer perceptron*. *Multilayer perceptron* biasanya berarti jaringan yang terhubung penuh, yaitu, setiap neuron dalam satu lapisan terhubung ke semua neuron di lapisan berikutnya. "*Fully Connected*" dari jaringan-jaringan ini membuat mereka rentan terhadap *over fitting* data. Cara khas regularisasi termasuk menambahkan beberapa bentuk pengukuran besaran bobot ke *loss function*. CNN mengambil pendekatan berbeda terhadap regularisasi: mereka mengambil keuntungan dari pola hierarkis dalam data dan mengumpulkan pola yang lebih kompleks menggunakan pola yang lebih kecil dan lebih sederhana. Oleh karena itu, pada skala

keterhubungan dan kompleksitas, CNN berada di ekstrem bawah [7].

Jaringan konvolusional diilhami oleh proses biologis dalam hal pola konektivitas antara neuron menyerupai organisasi korteks visual hewan. *Neuron* kortikal individu merespons rangsangan hanya di daerah terbatas bidang visual yang dikenal sebagai bidang reseptif. Bidang reseptif dari neuron yang berbeda sebagian tumpang tindih sehingga menutupi seluruh bidang visual, di mana arsitektur CNN ditampilkan pada Gambar 2.



**Gambar 2.** Arsitektur CNN

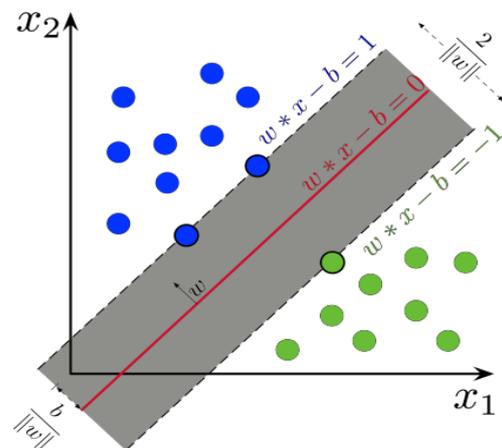
CNN menggunakan pra-pemrosesan yang relatif sedikit dibandingkan dengan algoritma klasifikasi gambar lainnya. Ini berarti bahwa jaringan mempelajari filter yang dalam algoritma tradisional direkayasa secara manual. Kemandirian ini dari pengetahuan sebelumnya dan upaya manusia dalam desain fitur adalah keuntungan utama.

### 2.3 Support Vector Machine

Dalam *machine learning*, *Support Vector Machine* (SVM) adalah model pembelajaran yang diawasi (*supervised learning*) dengan algoritma pembelajaran terkait yang menganalisis data yang digunakan untuk klasifikasi dan analisis regresi. Diberikan serangkaian contoh pelatihan, masing-masing ditandai sebagai milik satu atau yang lain dari dua kategori, algoritma pelatihan SVM membangun model yang memberikan contoh baru untuk satu kategori atau yang lain, menjadikannya sebagai pengelompokan *linear* biner non-probabilistik (walaupun metode seperti sk ala Platt ada untuk menggunakan SVM dalam pengaturan klasifikasi probabilistik). Model

SVM adalah representasi dari contoh-contoh sebagai titik-titik dalam ruang, dipetakan sehingga contoh-contoh kategori yang terpisah dibagi oleh celah yang jelas yang selebar mungkin. Contoh-contoh baru kemudian dipetakan ke dalam ruang yang sama dan diprediksi termasuk dalam kategori berdasarkan sisi kesenjangan di mana mereka jatuh. contoh baru kemudian dipetakan ke dalam ruang yang sama dan diprediksi termasuk dalam kategori berdasarkan sisi kesenjangan di mana mereka jatuh.

Selain melakukan klasifikasi linier, SVM dapat secara efisien melakukan klasifikasi non-linier menggunakan trik kernel, secara implisit memetakan input mereka ke dalam ruang fitur dimensi tinggi. Ilustrasi penggunaan SVM ditampilkan pada Gambar 3.



**Gambar 3.** Hyperplane

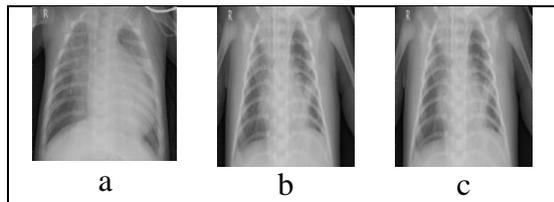
Ketika data tidak diberi label, pembelajaran yang diawasi tidak dimungkinkan, dan diperlukan pendekatan pembelajaran yang tidak diawasi, yang berupaya menemukan pengelompokan alami data ke dalam kelompok, dan kemudian memetakan data baru ke kelompok yang terbentuk ini. Algoritma *support vector clustering*, dibuat oleh Hava Siegelmann dan Vladimir Vapnik, menerapkan statistik *support* vektor, yang dikembangkan dalam algoritma *support vector machine*, untuk mengkategorikan data yang tidak berlabel, dan merupakan salah satu algoritma

pengelompokan yang paling banyak digunakan dalam industri.

### 3. Metode

#### 3.1 Dataset

Pada penelitian ini digunakan 2 macam citra digital *x-ray thorax* yang digunakan yaitu paru-paru *Bacterial Pneumonia*, dan paru-paru *Viral Pneumonia* dan paru-paru Normal. Sebanyak 600 citra digunakan dan dapat dilihat pada Gambar 4. Dataset yang digunakan berasal dari dataset CheXpert dengan 224.316 foto *Rontgent* dari 65.240 pasien[8].



**Gambar 4.** Citra *X-Ray Thorax*: a) Paru-paru *Bacterial Pneumonia* b) Paru-paru *Viral Pneumonia* c) Paru-paru Normal

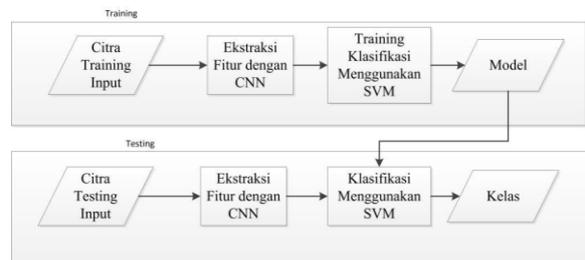
#### 3.2 Rancangan Metode

Tujuan penelitian ini adalah membuat sebuah sistem klasifikasi untuk membedakan

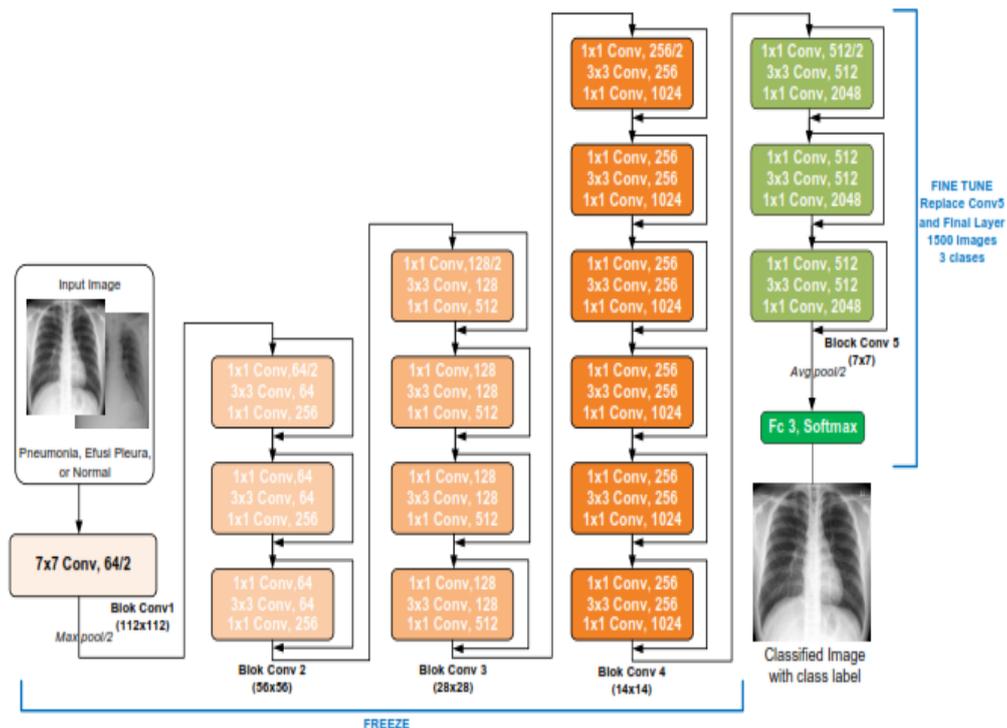
*Pneumonia* yang disebabkan oleh virus dengan *Pneumonia* yang disebabkan oleh bakteri menggunakan metode CNN dan SVM[9].

Tahap pertama adalah proses ekstraksi fitur menggunakan metode CNN, kemudian fitur yang didapat akan diklasifikasi menggunakan SVM, di mana model hasil training akan digunakan untuk melakukan klasifikasi data baru, alur sistem yang diusulkan dapat dilihat pada gambar 5[10].

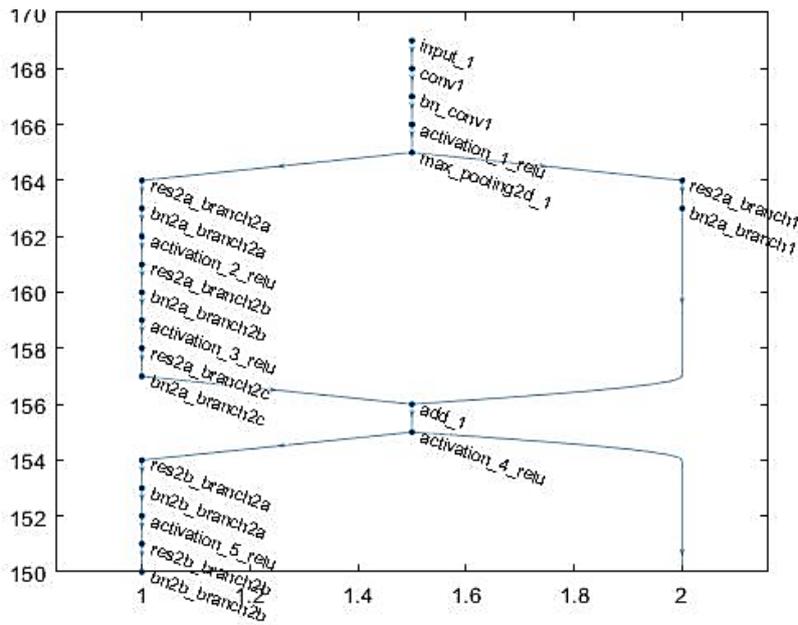
Untuk ekstraksi fitur arsitektur *Convolutional Neural Network* yang digunakan ResNet-50 karena arsitektur ResNet-50 dapat mengurangi reduksi bentuk Gambar 5.



**Gambar 5.** Alur Sistem



**Gambar 6.** Struktur Arsitektur Resnet-50



Gambar 7. First Section of ResNet-50

Metode SVM digunakan pada tahap klasifikasi karena selain melakukan klasifikasi linear, SVM juga dinilai efisien melakukan klasifikasi non-linear dengan trik kernel[11][12].

### 3.3 Arsitektur ResNet-50

Model ResNet-50 terdapat 5 tahap dengan masing-masing tahap memiliki konvolusi dan blok identitas. Pada penelitian ini digunakan model ResNet-50 dengan ImageNet Pretrained untuk efisiensi komputasi[13]. Pada blok konvolusional 5 diganti dengan *fully connected*, dimana arsitektur ResNet-50 yang digunakan pada penelitian ini secara lengkap ditampilkan pada Gambar 6 [14][15].

### 4. Hasil dan Pembahasan

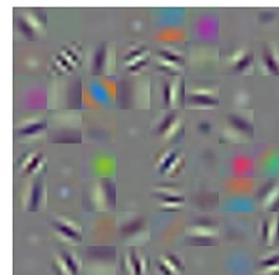
Pada penelitian ini digunakan sejumlah 400 citra X-ray Thorax, di mana masing-masing 100 citra viral *Pneumonia* dan bacterial *Pneumonia* digunakan sebagai data latih dan data tes. Pengukuran performa digunakan *confusion matrix* yang ditampilkan pada Tabel 1.

Tabel 1. Confusion Matrix

		Prediksi		
		Bakteri	Normal	Virus
Aktual	Bakteri	0.90	0.04	0.06
	Normal	0.07	0.79	0.14
	Virus	0.1	0.04	0.86

Hasil uji coba menunjukkan metode yang diusulkan menghasilkan akurasi rata-rata sebesar 0.85, Di mana akurasi untuk klasifikasi *pneumonia* bakteri, paru-paru normal dan paru-paru dengan *Pneumonia* virus masing-masing 0.9, 0.79, dan 0.86.

Untuk CNN pada penelitian ini digunakan arsitektur *Residual Network* (ResNet), *Section* pertama dari resnet dan bobot layer pertama dapat dilihat pada Gambar 7 dan Gambar 8.



Gambar 8. Bobot Layer Convolutional Pertama

Hingga saat ini belum ada penelitian untuk mengklasifikasikan *Pneumonia* yang disebabkan oleh virus atau bakteri.

**Tabel 2.** *Confusion Matrix* Arsitektur Model Resnet-50

		Correct Value		
		Normal	<i>Pneumonia</i>	Class Precision
Prediksi	Normal	215	5	97.7 %
	<i>Pneuomonia</i>	21	199	90.5 %
	Class Recall	91.1%	97.5%	

Pada Tabel 2 dan 3 ditampilkan *confusion matrix* klasifikasi paru normal dengan paru *Pneumonia* menggunakan arsitektur Resnet-50 dan Alexnet.

**Tabel 3.** *Confusion Matrix* Arsitektur Model AlexNet

		Correct Value		
		Normal	<i>Pneumonia</i>	Class Precision
Prediksi	Normal	209	11	95%
	<i>Pneuomonia</i>	3	217	98.6%
	Class Recall	98.6%	95.2%	

Dengan rata-rata skor presisi masing-masing 94.1% dan 96.8%, dan skor *recall* masing-masing 94.1% dan 96.9%

## 5. Penutup

Berdasar *confusion matrix* hasil uji coba, didapat akurasi rata-rata sebesar 0.85, sehingga dapat disimpulkan metode yang diusulkan mempunyai akurasi yang cukup tinggi. Untuk penelitian selanjutnya bisa ditambahkan tahapan *preprocessing* sebelum tahap ekstraksi fitur dan menggunakan arsitektur CNN yang lain untuk ekstraksi fitur, dan menggunakan metode klasifikasi yang lain.

## 6. Daftar Pustaka

[1] M. Suyuti and E. Setyati, "Pneumonia

Classification of Thorax Images using Convolutional Neural Networks," *Jurnal INFORM*, vol. 5, no. 2, p. 62, 2020.

- [2] J. Estomihi, "Klasifikasi Penyakit Pneumonia Dari Citra X-Ray Menggunakan Backpropagation," pp. 1–74, 2019.
- [3] N. Singh, "Wavelet Transform Based Pneumonia Classification of Chest X-Ray Images," in *Int. Conf. Comput. Power Commun. Technol.*, 2019, pp. 540–545.
- [4] A. Sharma, D. Raju, and S. Ranjan, "Detection of pneumonia clouds in chest X-ray using image processing approach," in *Nirma Univ. Int. Conf. Eng. NUiCONE*, 2018, pp. 1–4.
- [5] A. F. Agarap, "An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification," pp. 5–8, 2017.
- [6] S. Albawi, T. . Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Int. Conf. Eng. Technol. ICET 2017*, 2018, pp. 1–6.
- [7] R. Poojary and A. Pai, "Comparative Study of Model Optimization Techniques in Fine-Tuned CNN Models," in *International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, 2019, pp. 1–4.
- [8] J. Irvin *et al.*, "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, pp. 590–597, 2019.
- [9] A. Hapantenda and F. X. Ferdinandus, "Sistem Otomatis Untuk Cropping Region of Interest Sel Darah Putih Pada Citra Leukemia Limfoblast Akut," 2017.
- [10] A. K. W. Hapantenda, F. X. Ferdinandus, and R. A. Harianto, "Deteksi Jumlah

- Leukosit Bersentuhan Pada Citra Mikroskopis Leukemia Limfoblastik Akut Menggunakan Multiple K-Means Clustering,” *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, pp. 57–62, 2018.
- [11] A. K. W. Hapantenda, A. Januanto, and I. Listiowarni, “Studi Independen Komparasi Segmentasi Sel darah Putih Menggunakan Ruang Warna HSV Dengan CIE-L\*a\*b,” *Konvergensi*, vol. 15, no. 2, pp. 97–104, 2019.
- [12] A. Januanto, A. . Septiyanto, and A. K. W. Hapantenda, “Penentuan Rute Optimal Pada Distribusi Barang Menggunakan Algoritma Greedy (Studi Kasus : Ud.Xyz),” *Konvergensi*, vol. 17, no. 1, 2021.
- [13] S. Sedai, D. Mahapatra, Z. Ge, R. Chakravorty, and R. Garnavi, “Deep multiscale convolutional feature learning for weakly supervised localization of chest pathologies in X-ray images,” *LNCS. Springer*, vol. 11046, 2018.
- [14] S. Basiriyah, I. Listiowarni, and A. K. W. Hapantenda, “Analisis Penerapan Game-Based Student Response System Pada Flipped Classroom Biologi Sman 5 Pamekasan,” *Konvergensi*, vol. 16, no. 2, pp. 62–69, 2020.
- [15] A. Habib and A. Kartika W. H., “Development of an Online Sales Information System for SMEs Using Incremental Methods,” *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 4, no. 1, pp. 51–62, 2020.