

Optimasi Identifikasi Adenokarsinoma dari Citra X-Ray dengan Metode *Learning Vector Quantization*

Aziz Septian Amrullah^{*1}, Dadang Iskandar Mulyana²

^{1,2}Teknik Informatika Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika

E-mail: ^{*1}azizseptianamrullah09@gmail.com, ²mahvin2012@gmail.com

Abstrak

Adenokarsinoma adalah jenis kanker yang berkembang dan menyebar di sejumlah organ tubuh manusia, terutama yang memiliki kelenjar dalam, antara lain payudara, paru-paru, kerongkongan, usus besar, pankreas, dan prostat. Karena adenokarsinoma dapat mempengaruhi berbagai organ tubuh dan menyebabkan gejala berkembang secara berbeda, sinar-x adalah satu-satunya metode yang dapat dipercaya untuk mendiagnosis kondisi tersebut. Analisis data sinar-X akan mengungkapkan ada tidaknya adenokarsinoma. Periksa ini untuk mengklasifikasikan hasil rontgen sebagai normal atau kanker. *Learning Vector Quantification (LVQ)* dan ekstraksi ciri merupakan teknik klasifikasi. Upaya yang dilakukan dengan Klasifikasi citra diterapkan sebelum Grayscale, kemudian karakteristik ekstraksi digunakan. First-Rate Kualitas nilai parameter adalah mu, deviasi, skewness, energy, entropy, dan smoothness. 150 gambar x-ray akan diperiksa, dan *Learning Vector Quantization* akan digunakan untuk mengklasifikasikan hasilnya. Pada epoch 100, akurasi penelitian ini mencapai puncaknya, 97,3%, dan learning rate antara 0,1 dan 0,2.

Kata Kunci : Optimasi , Identifikasi , Adenokarsinoma , X-Ray , LVQ

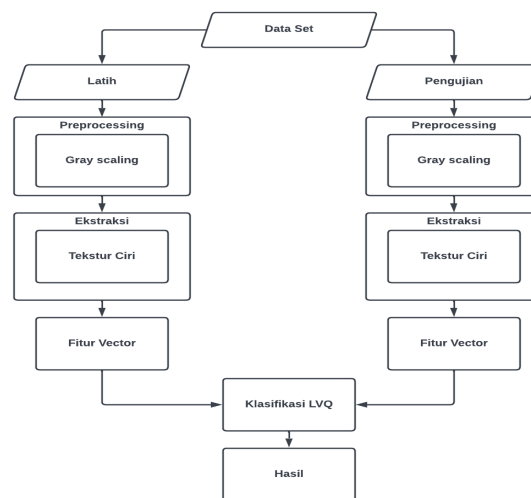
1. PENDAHULUAN

Kanker adalah penyakit berbahaya yang dapat menyerang bagian tubuh mana pun. Neoplasma, kadang-kadang dikenal sebagai tumor ganas, adalah istilah yang digunakan secara luas. Ciri khas kanker adalah pembentukan sel-sel abnormal yang membelah dengan cepat yang biasanya melebihi kapasitas normalnya sebelum bermetastasis ke organ lain. Metastasis adalah penyebab utama kematian akibat kanker. Pada tahun 2020, diperkirakan akan ada 19,3 juta diagnosis kanker baru dan 10 juta kematian terkait kanker di seluruh dunia. Beberapa penelitian menggunakan metode LVQ [1] Pada tahun 2020, Risha Ambar Wati, Hafiz Irsyad, dan M Ezar Al Rivan melakukan makalah berjudul Klasifikasi Pneumonia dengan menggunakan metode Support Vector Machine (SVM) dan ekstraksi Gray Level Occurrence Matrix (GLCM). fungsi. Metode Support Vector Machine menghasilkan tingkat akurasi terbaik yaitu 62,66 persen. [2] Judul penelitian adalah “Jaringan Syaraf Tiruan LVQ Parameter HSV Dalam Penentuan Uang Rupiah Palsu”. Klon LVQ berdasarkan saraf jaringan dapat memberikan akurasi 100 persen selama fase pengujian dan penelitian untuk menilai keabsahan uang kertas, menurut sebuah penelitian yang menggunakan pencahayaan UV untuk mengumpulkan foto. Klasifikasi Inflamasi Biologis Urinaria dan Nefritis dari Pelvis Ginjal yang paling akurat dicapai oleh penelitian Ahmadi Izaaz tahun 2021 menggunakan metode Learning Vector Quantization, dengan akurasi sebesar 86.95652173913 persen. Algoritma HOG dan LVQ diterapkan pada karikatur buah dalam penelitian untuk mengklasifikasikan kematangan buah. Dalam sebuah penelitian berjudul Diagnosis penyakit, peneliti mengklasifikasikan jumlah kematangan buah carica menggunakan 104 dataset, termasuk 80 data pelatihan dan 24 data uji terapan. Mereka menemukan bahwa tingkat presisi tertinggi mereka adalah 91,67 persen [5]. Demam Berdarah Dengue (DBD) hemoglobin, leukosit, trombosit, dan heritrosit digunakan sebagai parameter input dalam metode Learning Vector Quantization. Data tersebut menunjukkan bahwa nilai akurasi terbaik adalah

97,14 persen. Metode YOLOV5 yang digunakan oleh mobil-mobil Indonesia dan dijabarkan lebih lanjut dalam judul penelitian Klasifikasi Implementasi Jenis Deteksi Waktu Nyata, dapat mengenali objek dengan tingkat keandalan dan akurasi 90 persen [6]. Sebuah penelitian lain yang mengklasifikasikan foto bunga dahlia menggunakan algoritma Nave Bayes dan Otsu Thresholding menghasilkan tingkat akurasi 98,75 persen [8].

2 METODE PENELITIAN

Dalam tinjauan ini, tahapan-tahapan eksplorasi yang telah diselesaikan terdiri dari tahapan-tahapan studi penulisan yang akan digunakan sebagai sumber perspektif atau referensi dalam konsentrat ini untuk memperkuat masalah dan sebagai premis hipotetis dalam memimpin pemeriksaan. Kajian penulisan diperoleh dari pemahaman jurnal, artikel dan laman-laman di web untuk mendapatkan berbagai referensi yang relevan dengan masalah dalam penelitian ini



Gambar 1. Flow Sederhana pada Metode LVQ

2.1 Alur proses penelitian

Gambar di atas merupakan alur dari sebuah sistem dalam melakukan klasifikasi menggunakan *Learning Vector Quantization*. Proses diawali dengan meng-input citra *chest x-ray*, kemudian *preprocessing* yaitu Citra yang di input akan di konversi ke *grayscale*. Kemudian ekstraksi citra melalui teksur nilai ciri dengan menggunakan parameter μ (intensitas), deviasi, skewness, energi, entropi, smoothness (kehalusan). kemudian nilai menjadi input bagi LVQ untuk mengklasifikasi. Dengan learning rate 0.1 0.2 epoch 100. Ekstraksi ciri memiliki hasil lalu di simpan menjadi kelompok gambar latih pada set

2.2 Data set

Dataset kelompok objek dan atributnya. Selain objek ada jenis nama lainnya point, record, pattern, vector, observasi, case, entitas dan instance yang. Gambar atribut adalah bentuk kejelasan fungsi atau objek dari karakteristik tersebut.[9] Dataset Dapat digunakan sebagai data uji untuk proses Klasifikasi *Learning Vektor Quantization* (LVQ) Kemudian dilakukan latih data untuk mendapatkan model *support*. Setelah model *Learning Vector Quantization* (LVQ) didapatkan selanjutnya dilakukan proses klasifikasi menggunakan LVQ untuk mendapatkan hasil

2.2.1 Grayscale

Gambar skala abu-abu atau redup hanya memiliki satu nilai yang membahas nilai merah, hijau, biru. Tingkat gambar redup memiliki nilai ganda 8 digit atau ukuran 0-255, di mana nilai 0 adalah gelap dan 255 berwarna putih.[10] Perubahan gambar dari nada RGB (Merah, Hijau, Biru) ke gambar gelap ditampilkan dalam Kondisi 1.

$$gray = (0.299 \times R) + (0.587 \times G) + (0.114 \times B) \quad (1)$$

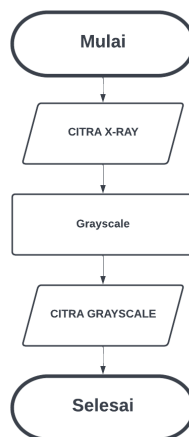
Pengertian:

R = warna merah pada lapisan gambar

G = warna hijau pada lapisan gambar

B = warna biru pada lapisan gambar

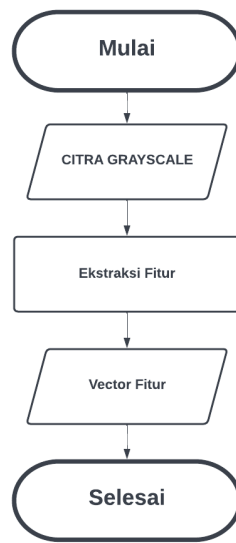
Pada tahap ini, dilakukan rencana perubahan gambar x-ray ke skala abu-abu. Perubahan terfokus pada satu lapisan sehingga lebih optimal. Adapun rencana prosesnya dimodelkan pada *flowchart* berikut :



Gambar 2. *Flowchart Citra Grayscale*

2.2.2 Fitur Textur

Nilai tekstur ekstraksi ialah Fitur salah satu terpenting pada objek gambar. Fitur ini mempunyai keterangan tentang urutan bagan pada citra suatu gambar.[11] Tekstur Nilai Pada tahap ini, proses ekstraksi komponen diselesaikan dari gambar skala abu-abu dengan permukaan. Ekstraksi komponen permukaan selesai untuk mengambil atribut gambar, dalam klasifikasi ini fitur ini sangat berperan untuk menentukan. Pergerakan menuju melakukan ekstraksi komponen tekstur ditampilkan dalam diagram alur[12] berikut Gambar 3:



Gambar 3. Flowchart Ekstraksi Fitur

2.2.3 Ekstraksi Ciri

Nilai tekstur Ciri yang sudah diekstrak digunakan sebagai parameter atau nilai masukan untuk membedakan antara objek satu dengan objek yang lain pada tahapan identifikasi dan klasifikasi.[13]

$$m = \sum_{i=0}^{L-1} i \cdot p(i) \quad (2)$$

Dalam hal ini, i adalah aras keabuan pada citra f dan $p(i)$ menyatakan probabilitas kemunculan i dan L menyatakan nilai aras keabuan tertinggi. Rumus di atas akan menghasilkan rerata kecerahan objek.

Fitur kedua berupa deviasi standar. Perhitungannya sebagai berikut:

$$\sigma = \sqrt{\sum_{i=1}^{L-1} (i - m)^2 p(i)} \quad (3)$$

Dalam hal ini, σ^2 dinamakan varians atau momen orde dua ternormalisasi karena $p(i)$ merupakan fungsi peluang. Fitur ini memberikan ukuran kekontrasan.

Fitur skewness merupakan ukuran ketidaksimetrisan terhadap rerata intensitas. Definisinya :

$$skewness = \frac{\sum_{i=1}^{L-1} (i - m)^3 p(i)}{\sigma^3} \quad (4)$$

Skewness sering disebut sebagai momen orde tiga ternormalisasi. Nilai negatif menyatakan bahwa distribusi kecerahan condong ke kiri terhadap rerata dan nilai positif menyatakan bahwa distribusi kecerahan condong ke kanan terhadap rerata. Dalam praktik, nilai skewness dibagi dengan $(L-1)^2$ supaya ternormalisasi.

Deskriptor energi adalah ukuran yang menyatakan distribusi intensitas piksel terhadap jangkauan aras keabuan. Definisinya sebagai berikut:

$$energi = \sum_{i=0}^{L-1} [p(i)]^2 \quad (5)$$

Citra yang seragam dengan satu nilai aras keabuan akan memiliki nilai energi yang maksimum, yaitu sebesar 1. Secara umum, citra dengan sedikit aras keabuan akan memiliki energi yang lebih tinggi daripada yang memiliki banyak nilai aras keabuan. Energi sering disebut sebagai keseragaman.

Entropi mengindikasikan kompleksitas citra. Perhitungannya sebagai berikut:

$$entropi = - \sum_{i=0}^{L-1} p(i) \log_2(p(i)) \quad (6)$$

Semakin tinggi nilai entropi, semakin kompleks citra tersebut. Perlu diketahui, entropi dan energi berkecenderungan berkebalikan. Entropi juga merepresentasikan jumlah informasi yang terkandung di dalam sebaran data.

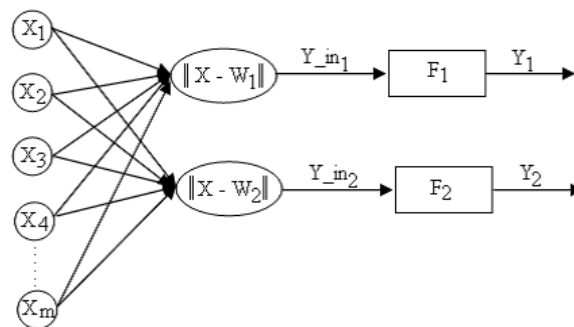
Properti kehalusan biasa disertakan untuk mengukur tingkat kehalusan/kekasaran intensitas pada citra. Definisinya sebagai berikut:

$$R = 1 - \frac{1}{1 + \sigma^2} \quad (7)$$

Dalam rumus yang baru saja ditunjukkan, (adalah simpangan baku. Menurut rumus yang disajikan di atas, nilai R yang rendah menunjukkan bahwa gambar memiliki intensitas kasar. Nilai perlu diketahui, kehalusan hitungan, dan varians perlu dinormalisasi sehingga bahwa mereka termasuk dalam kisaran [0 1] menggunakan metode yang membaginya dengan $(L-1)^2$.

2.2.4 Learning Vector Quantization

Learning Vector Quantization adalah metode adaptif pengklasifikasi yang didasarkan pada prototipe utama yang digunakan untuk memproses data vektor. Algoritma LVQ praktis digunakan dalam masalah terkait dengan kategorisasi data terstruktur, seperti diagnosis medis, analisis data ekspresi gen, mengenali tekstur, dan mengatakan waktu dalam kehidupan nyata. Algoritma LVQ dapat digunakan untuk menyelesaikan masalah yang melibatkan klasifikasi biner (dua kelas) serta klasifikasi multi-kelas. Tujuan utama dari LVQ adalah untuk mencapai tingkat akurasi tertinggi dalam klasifikasi.[14] Mengenai struktur bangunan, kuantisasi vektor pembelajaran dapat diperhatikan di gambar.



Gambar 4. arsitektur jaringan *learning vector quantization*

2.2.5 Klasifikasi

Klasifikasi adalah strategi untuk menemukan model atau kualitas penjelas, untuk mengetahui perbedaan antara metode dan kelas data, dan untuk mencapai kelas akurasi hasil dari sesuatu yang belum diketahui labelnya [15].

2.3 Pengumpulan data

Penyusunan data menggunakan dataset yang dikumpulkan dari internet, yang berjumlah 513 foto rontgen rontgen dada, diperoleh sebagai bagian dari pemeriksaan 170 foto rontgen tersebut. gambar sinar: yang menunjukkan dada sehat dan yang menunjukkan adenokarsinoma di dada. Berikut ini adalah rincian jumlah informasi yang dapat dilihat pada masing-masing citra rontgen individu pada gambar 5 dan tabel 1.



Gambar 5. Jumlah Dataset latih

Tabel 1. Jumlah Dataset latih 150

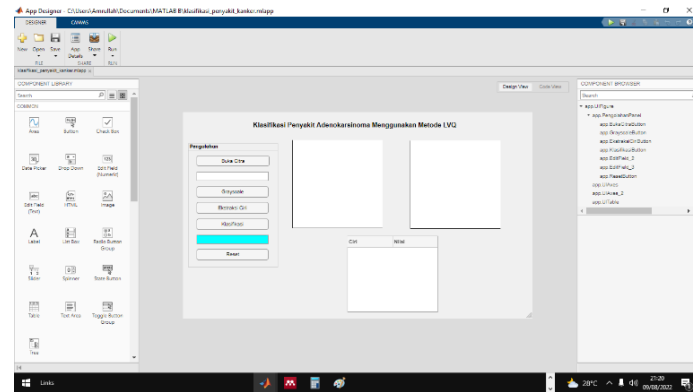
Jenis-jenis Citra	Jumlah data yang di pilih	Jumlah data latih	Jumlah data pengujian
Sehat	50	50	50
Adenokarsinoma	120	120	100
Total	170	170	150

2.4 Desain

Persiapan ini Untuk melaksanakan skema yang direncanakan untuk penelitian, perlu terlebih dahulu menyelesaikan pra-pemrosesan gambar rontgen dada skala abu-abu. Setelah itu citra akan diekstraksi tekstur karakteristiknya, kemudian proses klasifikasi akan dilanjutkan dengan menggunakan Learning Vector Quantization. Berikut adalah gambar dari sistem yang dipelajari, yang dapat dilihat dibawah ini pada Gambar 6:

2.5 Implementasi

Setelah proses desain selesai, langkah selanjutnya adalah melakukan eksekusi desain yang telah dibuat. Grayscale merupakan langkah awal dalam proses ekstraksi, dan setelah itu dilakukan ekstraksi ciri tekstur. Hasil ekstraksi ciri tersebut disajikan sebagai skor, yang kemudian digunakan untuk klasifikasi melalui Learning Vector Quantization. Tahapan pertama dilakukan Grayscale untuk mendapatkan citra grayscale.



Gambar 6. Tampilan App design interface pada matlab

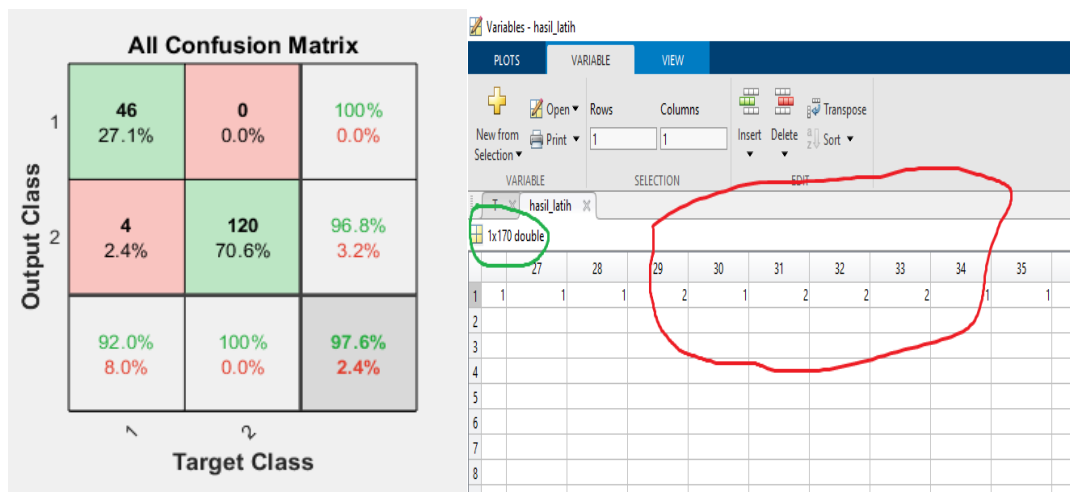
2.6 Pengujian

Setelah dilakukan proses implementasi dilanjutkan dengan proses pengujian pada citra latih dan citra uji untuk mendapatkan hasil dan mengetahui akurasi yang didapatkan. Pengujian dilakukan dengan menggunakan metode Learning Vector Quantization sebagai metode untuk klasifikasi dataset. Selanjutnya, hasil klasifikasi menggunakan LVQ dihitung untuk mendapatkan tingkat keberhasilan dari metode yang digunakan, dengan menggunakan metode Confusion Matrix yang dimana akan menghitung nilai accuracy, precision dan recall

3. HASIL DAN PEMBAHASAN

3.1 Tahap pelatihan

Tahap pelatihan pada metode yang digunakan dilakukan dengan menggunakan image train. Praktek komposisi gambar yang digunakan terdiri dari 120 foto adenokarsinoma rontgen dada dan 50 foto rontgen dada sehat. Klasifikasi dibagi menjadi dua kelompok yang berbeda. Kelas 1 lima puluh foto rontgen dada yang sehat. Menerima kelas dua untuk gambar kanker pada rontgen dada. Hasil dari pelatihan Langkah ditampilkan pada Gambar 7.

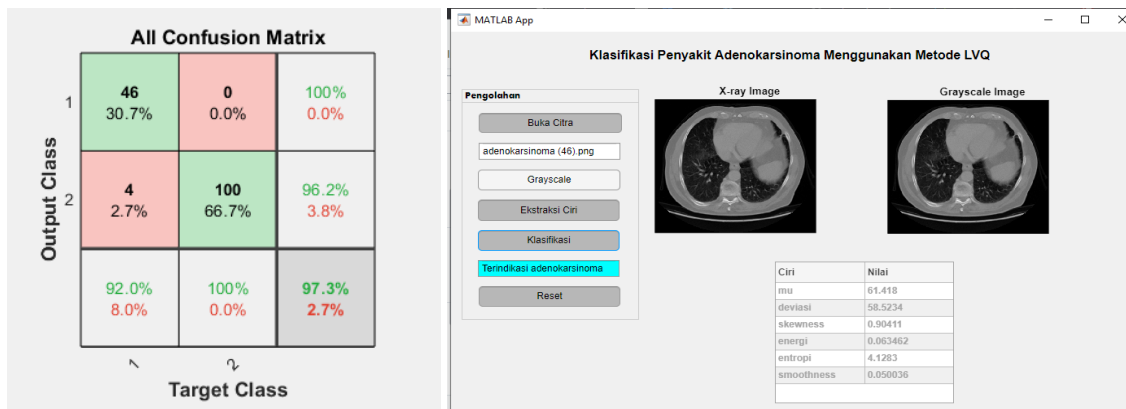


Gambar 7. Confusion Matrix tahap latih dan data salah tahap latih

Hasil tahapan Gambar 7 menyajikan pelatihan yang digunakan untuk mendemonstrasikan metode yang mampu mengenali gambar dengan akurasi 97,6 persen untuk setiap kelas benar dari data gambar yang digunakan. matriks konfusi yang dikembangkan selama tahap pelatihan Seperti yang terlihat pada Gambar 7, sistem mampu mengenali bahwa sebanyak 46 data citra pelatihan (27,1 persen) merupakan citra sehat rongten dada (kelas 1), sedangkan 120 data citra pelatihan (70,6 persen) menunjukkan gambaran rongten dada yang mengandung adenokarsinoma (tingkat 2). Dan 4 gambar yang salah (2,4 persen)

3.2 Tahap Uji

Metode divalidasi melalui pemanfaatan kumpulan data yang bervariasi melalui pelatihan Langkah. Jumlah data yang digunakan untuk foto rongten dada citra sehat adalah lima puluh foto, sedangkan jumlah data yang digunakan untuk foto rongten dada adenokarsinoma adalah seratus foto. Confusion matrix yang dikembangkan selama tahap pengujian dapat ditunjukkan pada Gambar 9. Temuan yang disajikan pada Gambar 9 menunjukkan bahwa, pada tahap ini dalam teknik pengujian, juga mampu mengenali setiap kumpulan data gambar dengan yang sesuai. kelas. Sebanyak 50 data gambar uji berhasil dideteksi sebagai rongten dada gambar Sehat, yang menyumbang 30,7% dari total (kelas 1). Gambar rongten dada yang menunjukkan adenokarsinoma (kelas 2) juga berkinerja baik dan terdeteksi dari 100 gambar uji (66,7 persen).



Gambar 8. Tahap Pengujian Confusion Matrix dan Tampilan Hasil Klasifikasi pada Aplikasi Matlab



Gambar 9. Grafik hasil akurasi

Pie Chart Gambar 9 menunjukkan bahwa hasil yang dicapai akurasi 97,6 persen selama tahap pelatihan warna-warni biru dengan total 170 gambar, dan hasil yang dicapai akurasi 97,3 persen selama tahap pengujian menggunakan 150 gambar tidak berbeda nyata dari satu sama lain. Rata-rata kesalahan yang diperkenalkan ke masing-masing dari Step train dan pengujian adalah empat gambar yang tidak dikenali pada tingkat 2,70 persen.

4. KESIMPULAN

Berdasarkan hasil penelitian menggunakan pendekatan Learning Vector Quantization Optimization Identification Adenocarcinoma from X-Ray Image, dapat ditarik kesimpulan sebagai berikut: Dengan bantuan perangkat lunak Matlab, pendekatan tiruan saraf jaringan yang efektif menggunakan Learning Vector Quantization (LVQ) telah dibuat (2). Akurasi yang diperoleh _ dari pelatihan dengan 170 data latih dan 150 data uji mendapatkan tingkat akurasi 97,333% dengan parameter jumlah neuron hidden layer 10 dan laju pembelajaran 0,1,0.2 maksimal 100 epoch.

DAFTAR PUSTAKA

- [1] H. Sung *et a* Berdasarkan studi yang telah dilakukan, penulis telah memperoleh rekomendasi berikut untuk menawarkan saran kepada peneliti:l., “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
- [2] R. A. Wati, H. Irsyad, M. Ezar, and A. Rivan, “Klasifikasi Pneumonia Menggunakan Metode Support Vector Machine,” *J. Algoritm.*, vol. 1, no. 1, pp. 21–32, 2020.
- [3] S. Tinggi and I. Komputer, “Jaringan Syaraf Tiruan LVQ Berbasis Parameter HSV dalam Penentuan Uang Jaringan Syaraf Tiruan LVQ Berbasis Parameter HSV dalam Penentuan Uang Rupiah Palsu,” vol. 13, no. January 2019, pp. 47–52, 2020.
- [4] A. Izaaz, “Klasifikasi Radang Kandung Kemih dan Nefritis dari Pelvis Ginjal Menggunakan Metode Learning Vector Quantization,” *J. Tek. Elektro dan Komput. TRIAC*, vol. 8, no. 2, pp. 35–38, 2021, doi: 10.21107/triac.v8i2.11893.
- [5] A. F. Herdajanti, Y. D. Setiyaningrum, G. F. Shidik, R. A. Pramunendar, and A. Z. Fanani, “(Hog) and Learning Vector Algorithm Quantization (Lvq) in Classification Carica,” pp. 519–522, 2019.
- [6] F. Tawakal and A. Azkiya, “Diagnosa Penyakit Demam Berdarah Dengue (DBD) menggunakan Metode Learning Vector Quantization (LVQ),” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 4, no. 3, p. 56, 2020, doi: 10.14421/jiska.2020.43-07.
- [7] S. T. Informatika, “Implementasi Deteksi Real Time Klasifikasi Jenis Kendaraan Di Indonesia Menggunakan Metode YOLOV5,” vol. 6, pp. 13971–13982, 2022.
- [8] A. Syaeful, M. I. Fadillah, I. Muftadi, and D. Iskandar, “Klasifikasi Citra Bunga Dahlia Berdasarkan Warna Menggunakan Metode Otsu Thresholding Dan Naïve Bayes,” vol. 6, pp. 575–582, 2022.

- [9] Yuli Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.
- [10] D. Retnoningrum, A. W. Widodo, and M. A. Rahman, “Ekstraksi Ciri Pada Telapak Tangan Dengan Metode Local Binary Pattern (LBP),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, pp. 2611–2618, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [11] L. Amatullah, I. Ein, and M. M. Santoni, “Identifikasi Penyakit Daun Kentang Berdasarkan Fitur Tekstur dan Warna Dengan Menggunakan Metode K-Nearest Neighbor,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, no. April, pp. 783–791, 2021.
- [12] S. P. Adenugraha, V. Arinal, and D. I. Mulyana, “Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode KNN dan PCA Berdasarkan Citra RGB dan HSV,” *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 9, 2022, doi: 10.30865/mib.v6i1.3287.
- [13] R. Mardisa, I. S. Nasution, and K. Siregar, “Klasifikasi Kualitas Fisik Kopi Beras Arabika Menggunakan Pengolahan Citra Dengan Metode K-Nearest Neighbor (K-NN),” *J. Ilm. Mhs. Pertan.*, vol. 7, no. 2, pp. 514–522, 2022.
- [14] T. Villmann, M. Kaden, W. Hermann, and M. Biehl, “Learning vector quantization classifiers for ROC-optimization,” *Comput. Stat.*, vol. 33, no. 3, pp. 1173–1194, 2018, doi: 10.1007/s00180-016-0678-y.
- [15] A. Muharam and I. Azmi, “Jurnal Ilmiah Teknik Informatika dan Komunikasi KLASIFIKASI KANKER PAYUDARA MENGGUNAKAN ALGORITMA K-NN DENGAN OPTIMASI BOBOT (FORWARD) Jurnal Ilmiah Teknik Informatika dan Komunikasi,” vol. 2, no. 2, 2022.