

Random Forest Classifier untuk Deteksi Penderita COVID-19 berbasis Citra CT Scan

Nanik Wuryani¹, Sarifah Agustiani²

^{1,2} Ilmu Komputer, STMIK Nusa Mandiri

¹e-mail: 14002336@nusamandiri.ac.id

²e-mail: 14002398@nusamandiri.ac.id

Diterima	Direvisi	Disetujui
01-01-2020	01-02-2020	01-03-2020

Abstrak - Covid-19 merupakan virus yang menyebar dan meluas sehingga berubah menjadi suatu pandemi dunia. Virus Covid-19 menyerang melalui organ vital manusia yaitu paru-paru, oleh karena itu peneliti lebih berfokus untuk mengidentifikasi Covid-19 pada paru-paru. Penelitian ini dilakukan dengan menggunakan citra *CT Scan* paru-paru dan bertujuan untuk mendeteksi Covid-19 dengan cara mengklasifikasikan citra *CT Scan* paru-paru ke dalam tiga kelas menggunakan algoritma *Random Forest* serta mengkombinasikannya dengan menyertakan beberapa ekstraksi fitur yaitu *Haralick*, *Color Histogram*, dan *Hu-Moments*. Penelitian dimulai dengan memasukkan satu fitur ke dalam percobaan, lalu mengkombinasikan dengan fitur yang lain, kemudian membandingkannya menggunakan klasifikasi oleh algoritma lain seperti *K-Nearest Neighbor (KNN)*, *Decision Tree*, *Linear Discriminant Analysis (LDA)*, *Logistic Regression*, *Support Vector Machine (SVM)*, dan *Naive Bayes*. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* paling cocok jika digunakan untuk klasifikasi citra *CT Scan* paru-paru dengan memasukkan fitur *Haralick* dan *Color Histogram* ke dalam proses yang menghasilkan akurasi sebesar 96,9%, diikuti oleh *KNN* sebesar 96,5%, *Decision Tree* sebesar 95,5%, dan yang paling rendah yaitu *Naive Bayes* sebesar 42,4%.

Kata Kunci: *Random Forest*, klasifikasi, ekstraksi fitur

Abstract - Covid-19 is a virus that spreads and spreads so that it has turned into a global pandemic. The Covid-19 virus attacks through vital human organs, namely the lungs, therefore researchers are more focused on identifying Covid-19 in the lungs. This research was conducted using CT scan images of the lungs and aims to detect Covid-19 by classifying CT Scan images of lungs into three classes using the Random Forest algorithm and combining them by including several feature extractions, namely *Haralick*, *Color Histogram*, and *Hu-Moments*. The research begins by entering one feature into the experiment, then combining it with other features, then comparing it using classifications by other algorithms such as *K-Nearest Neighbor (KNN)*, *Decision Tree*, *Linear Discriminant Analysis (LDA)*, *Logistic Regression*, *Support Vector Machine (SVM)*, and *Naive Bayes*. The results showed that the *Random Forest* algorithm was most suitable when used for the classification of CT Scan images of the lungs by incorporating the *Haralick* and *Color Histogram* features into the process which resulted in an accuracy of 96.9%, followed by *KNN* of 96.5%, *Decision Tree* of 95.5%, and the lowest is *Naive Bayes* at 42.4%.

Keywords: *Random Forest*, classification, feature extractions

PENDAHULUAN

Tahun 2020 diawali dengan merebaknya suatu wabah virus atau pandemi yang sekarang disebut dengan Covid-19. Pandemi ini ditemukan di Wuhan, China pada Desember 2019 (Li et al., 2020) yang kemudian menyebar ke seluruh dunia dengan sangat cepat. Meskipun pandemi ini tidak ganas, namun dengan penyebaran yang sangat cepat itu membuat orang-orang khawatir karena ternyata virus ini memicu penyakit bawaan yang diderita oleh seseorang. Para peneliti sangat gencar dan bersaing untuk menangani

virus ini. Selain dari pihak medis, semua peneliti atau bahkan mahasiswa-mahasiswa di semua Universitas saling berlomba untuk ikut serta dalam menangani pandemi.

Sebuah masalah yang tengah dihadapi adalah bagaimana cara yang efektif untuk mendeteksi gejala-gejala atau tanda-tanda dari seseorang yang terkena Covid-19. Salah satu cara yang dapat dilakukan untuk mendeteksi gejala covid-19 adalah dengan menggunakan Citra *CT Scan* dada, yang mana dari Ct scan ini dapat mengidentifikasi bagaimana bentuk atau infeksi yang ditimbulkan oleh Covid-19.

Berdasarkan pengamatan para pakar, infeksi yang ditimbulkan ditunjukkan dengan adanya lesi atau kabut yang terlihat pada citra CT scan. Namun untuk melakukan proses pengidentifikasian secara manual akan memerlukan waktu yang cukup lama, untuk itu perlu adanya sebuah teknologi yang memungkinkan dapat mengidentifikasi gejala Covid-19 pada Citra CT-scan, dalam hal ini dilakukan dengan teknik pengolahan citra.

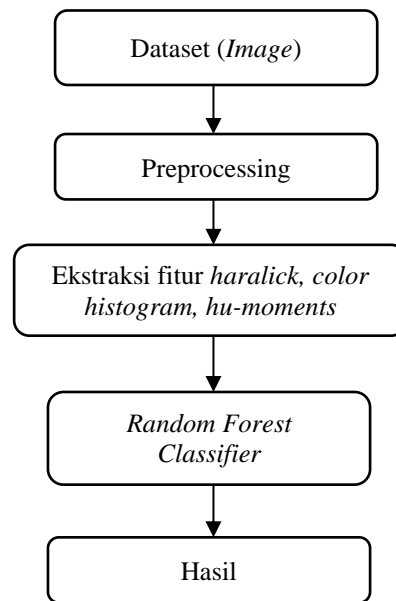
Proses pengidentifikasian citra CT-scan ini dapat diidentifikasi berdasarkan warna, tekstur, dan juga ciri lainnya. Beberapa paper telah menunjukkan penelitian dengan menggunakan berbagai metode, beberapa diantaranya adalah menerapkan deep learning untuk klasifikasi dan segmentasi Covid-19 pada citra CT Scan (Amyar, Modzelewski, & Ruan, 2020). Thepade, dkk (Thepade, Bang, Chaudhari, & Dindorkar, 2020) menggunakan ekstraksi fitur GLCM dan mengkombinasikannya dengan *wavelet transform*, serta menggunakan KNN dan SVM sebagai *classifier* untuk mengidentifikasi Covid-19 pada citra X-Ray paru. Penelitian ini menghasilkan nilai akurasi paling tinggi pada saat menggunakan *KNN Classifier*, yaitu sebesar 92,6%.

Penggunaan ekstraksi fitur juga dapat diterapkan pada berbagai permasalahan lain seperti penggunaan fitur *haralick* pada klasifikasi bentuk lengkung gigi yang dilakukan oleh Shinta (Priyantono, Rachmawan, Budi, & Kirana, 2020). Fitur *haralick* digunakan untuk membantu proses klasifikasi oleh jaringan syaraf tiruan dan algoritma propagasi balik. Metode ini menghasilkan keakuratan sebesar 66%, sehingga masih diperlukan analisis lebih lanjut agar dapat mengklasifikasikan lengkung gigi dengan lebih baik. Beberapa peneliti menambahkan ekstraksi fitur GLCM dan *Color Histogram* pada klasifikasinya terhadap citra makanan menggunakan algoritma *Learning Vector Quantization (LVQ)* meskipun hasil yang didapatkan masih rendah yaitu sebesar 53,33% (Simarmata, Sari, & Adinugroho, 2019). Devella, dkk (Devella, Yohannes, & Rahmawati, 2020) memanfaatkan ekstraksi fitur SIFT untuk klasifikasi motif songket Palembang. Ekstraksi SIFT digunakan untuk mendapatkan *keypoint* yang dapat mempengaruhi *noise* pada citra. Klasifikasi dilakukan dengan algoritma *Random Forest* dan menghasilkan akurasi yang sangat baik. Ekstraksi fitur terbukti sangat penting, sehingga dimanfaatkan pada *preprocessing* citra. Seperti yang dilakukan oleh Hernando, dkk (Hernando, Widodo, & Dewi, n.d.) dimana mereka memanfaatkan fitur warna dan fitur tekstur pada klasifikasi citra drone untuk mengetahui jenis-jenis penggunaan lahan.

Pada paper ini, kami mengusulkan metode klasifikasi menggunakan algoritma *Random Forest* dengan menambahkan beberapa ekstraksi fitur seperti *Haralick*, *Color Histogram*, dan *Hu-moments*. Hasil eksperimen menunjukkan akurasi yang sangat baik jika dibandingkan dengan klasifikasi tanpa

menggunakan ekstraksi fitur, atau jika dibandingkan dengan menggunakan algoritma *classifier* lain. Metode ini dapat digunakan untuk jenis citra yang sama dan masalah yang serupa, sehingga penelitian selanjutnya dapat dilakukan dengan metode ini ataupun mengembangkannya menjadi lebih baik. Selanjutnya pada paper ini akan dipaparkan metode yang digunakan pada penelitian, berikut pembahasan dan kesimpulan.

METODOLOGI PENELITIAN



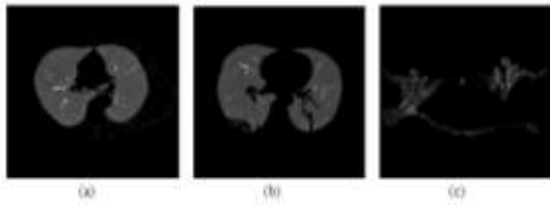
Sumber: Hasil Penelitian (2021)

Gambar 2. Alur Metode Penelitian

Tahapan penelitian dilakukan seperti yang terlihat pada Gambar 2, yaitu :

1. Dataset (Image)

Dataset yang digunakan diperoleh dari Kaggle berupa citra CT scan Paru sebanyak 19.685 citra dengan tiga kelas di dalamnya, yaitu kelas Positif Covid (pCT) sebanyak 4.001 citra, Negatif Covid (nCT) sebanyak 9.979 citra, dan kelas tanpa informasi (NiCT) sebanyak 5.705 citra (Ning et al., 2020). Citra pada dataset ini memiliki ukuran yang sama yaitu 512 x 512 piksel. Masing-masing kelas memiliki ciri citra yang berbeda-beda, namun pada kelas NiCT terlihat sangat berbeda karena citra yang didapat adalah citra CT scan yang tidak sempurna.



Sumber: W.Ning (2020)

Gambar 1. Citra dataset tiap kelas: (a) citra positif Covid-19, (b) citra negatif Covid, (c) citra tanpa informasi

2. Preprocessing

Tahapan awal penelitian adalah dengan melakukan *preprocessing*, dimana tahap ini akan membantu mempercepat ataupun mempermudah proses eksperimen. *Preprocessing* dilakukan dengan menyeragamkan ukuran citra menjadi 500x500 piksel. Selain untuk menyeragamkan, pada tahap ini juga dapat dilakukan *resize* untuk memperkecil ukuran citra agar eksperimen berjalan lebih cepat.

3. Feature Extraction

Feature extraction adalah sebuah teknik pengambilan ciri pada suatu bentuk citra agar tahapan eksperimen berikutnya memiliki hasil yang maksimal. *Feature extraction* dapat dilakukan dengan mengambil beberapa ciri unik seperti fitur warna, fitur tekstur, dan lainnya. Dengan kata lain, ekstraksi fitur adalah cara untuk mengenali pola dari suatu citra. Penggunaan ekstraksi fitur akan meningkatkan akurasi hasil eksperimen. Beberapa *feature extraction* yang digunakan pada penelitian ini adalah:

a. Haralick

Fitur *haralick* merupakan fitur deskriptor tekstur yang biasa digunakan dalam analisis citra. Fitur ini dihitung dengan mengurangi level keabuan pada citra yang biasa disebut dengan proses kuantisasi. Fitur *haralick* harus digunakan dengan melakukan proses kuantisasi yang sama pada citra yang digunakan (Htun, Han, & Hlaing, n.d.). *Haralick* memberikan informasi mengenai intensitas dalam piksel tertentu dengan posisi yang terhubung dengan piksel tetangga. Homogenitas bayangan (F1) yang merupakan kemiripan antar piksel, dijelaskan pada persamaan (1) dimana $p(k,l)$ adalah posisi elemen dalam matrik.

$$F1 = \sum_{k=0}^N \sum_{l=0}^N \frac{dy}{dx}$$

Tekstur fitur *haralick* memiliki 14 fitur yang dirumuskan ke dalam persamaan-persamaan untuk menghitung intensitas piksel yang berdekatan (Perumal, Narayanan, & Rajasekar, 2020).

b. Color Histogram

Histogram dapat terdiri dari 48 bin warna yang mendefinisikan rentang nilai piksel. Penelitian dilakukan dengan menggunakan 8 bin warna, namun karena citra yang digunakan adalah citra abu, maka nilai histogram tidak akan memberikan pengaruh yang signifikan terhadap hasil klasifikasi citra (Hussain, Rao, & Praveen, 2013).

c. Hu-moments

Fitur *Hu-moments* digunakan agar nilai invariant citra tidak terpengaruh dari perubahan skala, rotasi, dan refleksi. Ekstraksi fitur *moment* dapat dilakukan dengan moment geometris pusat pada citra. Perhitungannya dapat dihitung menggunakan persamaan berikut (Žunić, Hirota, & Rosin, 2010):

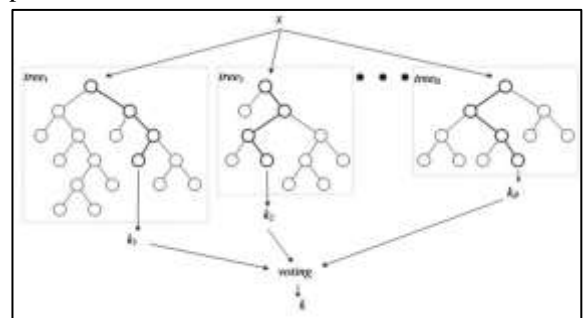
$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_c)^p (y - y_c)^q f(x, y) dx dy$$

dimana $x_c = \frac{m_{10}}{m_{00}}$, $y_c = \frac{m_{01}}{m_{00}}$ adalah koordinat pusat.

Penggunaan ekstraksi fitur dimaksudkan agar dapat memperjelas perbedaan-perbedaan yang ada pada citra baik secara tekstur, ketajaman warna, dan lain-lain.

4. Random Forest Classifier

Random Forest didefinisikan sebagai prinsip umum suatu ansambel acak dari suatu pohon keputusan (Breiman, 2001). Bentuk umum dari model klasifikasi *Random Forest* dapat dilihat pada Gambar 3.



Sumber: Breiman (2001)

Gambar 3. Arsitektur Algoritma *Random Forest*

Konstruksi *Random Forest* dapat dilakukan dengan tahapan sebagai berikut:

- a. Menggambar sampel bootstrap *n-tree* dari data.
- b. Menumbuhkan pohon untuk setiap kumpulan data *bootstrap*. Di setiap simpul pohon, pilih variabel entri secara acak untuk dipisahkan, lalu tumbuhkan pohon sehingga setiap *node* terminal memiliki tidak kurang dari kasus ukuran *node*.
- c. Informasi agregat dari pohon *n-tree* untuk prediksi data baru seperti *voting* mayoritas untuk klasifikasi.
- d. Hitung tingkat kesalahan *out-of-bag* (OOB) dengan menggunakan data bukan dalam sampel *bootstrap*.

Metode penelitian ini didapatkan dari hasil perbandingan beberapa algoritma *classifier* dan ekstraksi fitur.

HASIL DAN PEMBAHASAN

Hasil penelitian pada eksperimen diperoleh dengan membandingkan hasil klasifikasi dari berbagai algoritma *classifier* dan kombinasi fitur yang digunakan. Beberapa kombinasi ekstraksi fitur menghasilkan nilai akurasi yang tinggi dan ada pula yang menghasilkan nilai akurasi yang rendah, sehingga tidak cocok jika diterapkan pada citra yang digunakan. Selain metode yang diusulkan, dilakukan pula beberapa pengujian untuk menentukan metode mana yang menghasilkan akurasi terbaik. Percobaan ini dilakukan dengan menguji beberapa algoritma klasifikasi lainnya diantaranya *Logistic Regression* (LR), *Linear Discriminant Analysis* (LDA), *Decision Tree Classifier* (CART), *Gaussian NB* (NB) dan *SVM* yang dikombinasikan dengan beberapa fitur yang digunakan. Berikut merupakan hasil pengujian yang telah dilakukan dalam penelitian ini:

1. Hasil Klasifikasi Menggunakan Fitur *Haralick*
Pengujian pertama dilakukan dengan menerapkan fitur *haralick* pada beberapa algoritma klasifikasi, penggunaan fitur *haralick* ini memberikan performa akurasi sangat baik pada setiap algoritma yang dapat dilihat pada Tabel 1.

Tabel 1. Hasil Klasifikasi Menggunakan Fitur *Haralick*

No	Algoritma	Akurasi
1	<i>Random Forest</i>	0,957
2	<i>K-Nearest Neighbor</i>	0,956
3	<i>Decision Tree</i>	0,921
4	<i>LDA</i>	0,855

5	<i>SVM</i>	0,827
6	<i>Logistic Regression</i>	0,811
7	<i>Naive Bayes</i>	0,754

Sumber: Hasil Penelitian (2021)

Tabel 1 menjelaskan bahwa penggunaan fitur *haralick* menghasilkan nilai akurasi tertinggi pada algoritma *Random Forest* dengan nilai akurasi sebesar 95.7%.

2. Hasil Klasifikasi Menggunakan *Color Histogram*

Pengujian kedua dilakukan dengan menerapkan fitur *color histogram* yang menunjukkan peningkatan performa akurasi yang lebih baik hampir pada semua algoritma dibandingkan fitur sebelumnya.

Tabel 2. Hasil Klasifikasi Menggunakan Fitur *Color Histogram*

No mor	Algoritma	Akurasi
1	<i>K-Nearest Neighbor</i>	0,961
2	<i>Random Forest</i>	0,960
3	<i>Decision Tree</i>	0,942
4	<i>SVM</i>	0,842
5	<i>Logistic Regression</i>	0,813
6	<i>LDA</i>	0,797
7	<i>Naive Bayes</i>	0,769

Sumber: Hasil Penelitian (2021)

Tabel 2 menunjukkan bahwa penggunaan fitur *color histogram* ini meningkatkan nilai akurasi yang lebih baik pada algoritma *Random Forest* dengan nilai sebesar 96% jika dibandingkan dengan fitur sebelumnya yang memiliki nilai sebesar 95.7%. Meskipun dalam pengujian ini *K-Nearest Neighbor* memiliki nilai akurasi terbaik yaitu 96.1%.

3. Hasil Klasifikasi Menggunakan Fitur *Haralick* dan *Hu-Moments*

Pengujian selanjutnya dilakukan dengan mengkombinasikan dua fitur sekaligus yaitu *haralick* dan *hu-Moments*. Dalam pengujian ini tingkat akurasi yang dihasilkan mengalami penurunan, berbeda dari pengujian sebelumnya, hal ini dapat dilihat pada Tabel 3 yang menunjukkan bahwa akurasi tertinggi dihasilkan oleh algoritma *Random Forest* dengan nilai sebesar 72.2%, dengan demikian, kombinasi dua fitur ini tidak menghasilkan performa yang cukup baik, terutama pada algoritma *Naive Bayes* yang memiliki nilai akurasi terendah yaitu sebesar 20.5%.

Tabel 3. Hasil Klasifikasi Menggunakan Fitur *Haralick* dan *Hu-moments*

Nomor	Algoritma	Akurasi
1	Random Forest	0,722
2	Decision Tree	0,711
3	K-Nearest Neighbor	0,684
4	LDA	0,541
5	Logistic Regression	0,507
6	SVM	0,507
7	Naive Bayes	0,205

Sumber: Hasil Penelitian (2021)

4. Hasil Klasifikasi Menggunakan Fitur *Haralick dan Color Histogram*

Seperti pengajian sebelumnya, pengujian ini dilakukan dengan mengkombinasikan dua fitur yaitu *haralick* dan *color histogram*. Hasil akurasi yang dihasilkan dalam pengujian ini diperlihatkan pada Tabel 4.

Tabel 4. Hasil Klasifikasi Menggunakan Fitur *Haralick dan Color Histogram*

Nomor	Algoritma	Akurasi
1	Random Forest	0,969
2	K-Nearest Neighbor	0,967
3	Decision Tree	0,955
4	LDA	0,865
5	Logistic Regression	0,853
6	SVM	0,849
7	Naive Bayes	0,771

Sumber: Hasil Penelitian (2021)

Berdasarkan Tabel 4 dapat dilihat bahwa penggabungan dua fitur ini menghasilkan performa akurasi yang sangat baik dibandingkan pengujian sebelumnya, hal ini ditunjukkan dari hasil akurasi pada setiap algoritma yang mengalami peningkatan, dengan akurasi terbaik sebesar 96.9% yang diperoleh dari algoritma *Random Forest*.

5. Hasil Klasifikasi Menggunakan Fitur *Haralick, Hu-Moments, dan Color Histogram*

Pengujian terakhir dilakukan dengan menggabungkan semua fitur yang digunakan dalam penelitian ini yaitu *haralick*, *color histogram* dan *hu-Moments*. Penggabungan tiga fitur ini menghasilkan performa yang sama baiknya meskipun tidak lebih tinggi dari penggabungan dua fitur *haralick* dan *color histogram*.

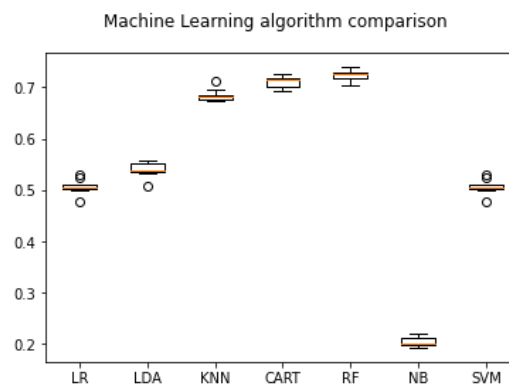
Tabel 5. Hasil Klasifikasi Menggunakan Fitur *Haralick, Color Histogram, dan Hu-moments*

Nomor	Algoritma	Akurasi
1	Random Forest	0,968
2	K-Nearest Neighbor	0,965
3	Decision Tree	0,955
4	LDA	0,864
5	Logistic Regression	0,850

6	SVM	0,842
7	Naive Bayes	0,424

Sumber: Hasil Penelitian (2021)

Berdasarkan Tabel 5 di atas, *Random Forest* masih memiliki nilai akurasi terbaik sebesar 96.8% bahkan hampir di semua pengujian nilai akurasi tertinggi diperoleh dari algoritma *Random Forest*. Hal ini menunjukkan bahwa *Random Forest* merupakan algoritma yang sangat baik untuk melakukan klasifikasi Covid19. Komparasi algoritma dapat dilihat lebih jelas pada boxplot seperti pada Gambar 3.



Sumber: Hasil Penelitian (2021)

Gambar 3. Boxplot komparasi algoritma pada klasifikasi Covid-19

Boxplot di atas memperlihatkan nilai akurasi pada masing-masing algoritma, dimana nilai paling tinggi adalah yang mendekati nilai 1, sedangkan nilai paling rendah adalah yang mendekati nol.

6. Akurasi Tiap *Fold*

Penelitian yang dilakukan menggunakan iterasi sebanyak 10 kali, dimana iterasi ini merupakan operasi untuk memaksimalkan model yang telah kita bentuk. Sebanyak 10 *Cross Validation* ini memiliki masing-masing nilai, akan ditampilkan pada Tabel 6.

Tabel 6. 10- *Cross Validation Value*

Fold ke-	Akurasi
1	0,962
2	0,970
3	0,973
4	0,961
5	0,972
6	0,969
7	0,968
8	0,966
9	0,970
10	0,974

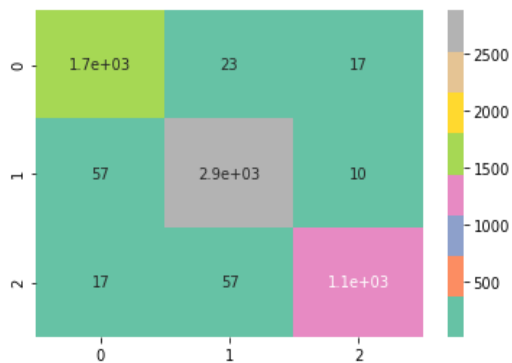
Sumber: Hasil Penelitian (2021)

Setiap iterasi ditampilkan akurasi yang dihasilkan, dan menunjukkan bahwa iterasi

meningkat seiring bertambahnya iterasi yang dilakukan.

7. Confusion Matrix

Hasil penelitian juga disajikan dengan *confusion matrix*, dimana kita dapat melihat proyeksi dari hasil klasifikasi pada tiap kelas.



Sumber: Hasil Penelitian (2021)

Gambar 4. Confusion Matrix Hasil Klasifikasi Algoritma Random Forest

Berdasarkan *confussion matrix* yang ditampilkan pada Gambar 4, terlihat bahwa data yang tepat klasifikasinya berada pada bidang diagonal, sedangkan data yang kurang tepat klasifikasinya menyebar di sekitar diagonal. *Confussion matrix* diatas menunjukkan bahwa hanya sedikit citra yang memiliki akurasi kurang tepat yaitu sebanyak 181 citra.

KESIMPULAN

Berdasarkan hasil penelitian, deteksi Covid-19 dapat dilakukan dengan mengklasifikasikan citra CT paru-paru dengan memanfaatkan kombinasi beberapa ekstraksi fitur dan menggabungnya dengan algoritma *classifier*. Beberapa algoritma memiliki hasil yang cukup baik bahkan sangat baik, tetapi ada pula algoritma dengan hasil yang sangat rendah karena masing-masing algoritma menghasilkan akurasi yang berbeda pada setiap pengujian terutama pada kombinasi fitur ekstraksi yang digunakan. Pada penelitian ini, *Random Forest* memiliki nilai akurasi yang paling tinggi jika dibandingkan dengan metode-metode lain. Tabel 4 menunjukkan hasil klasifikasi dengan dua ekstraksi fitur yaitu *Haralick* dan *Color Histogram* yang menghasilkan akurasi sebesar 0,969 atau 96,9% pada algoritma *Random Forest*. Beberapa algoritma juga menghasilkan akurasi yang cukup baik, namun ada satu algoritma dengan hasil akurasi yang sangat rendah yaitu *Naive Bayes* dengan akurasi sebesar 0,205 atau 20,5% yang diperoleh dari hasil penggabungan dua fitur yaitu *haralick* dan *hu-*

moments. Dengan demikian dapat disimpulkan bahwa penggunaan ekstraksi fitur sangat mempengaruhi hasil klasifikasi terutama pada saat melakukan kombinasi dari beberapa fitur, namun hal ini juga harus disesuaikan dengan tipe citra yang dipakai. Penggunaan metode ini membutuhkan waktu yang cukup lama untuk memproses citra dalam jumlah banyak. Penelitian selanjutnya dapat dilakukan dengan menggunakan *deep learning* sehingga waktu yang digunakan lebih singkat, serta dapat memaksimalkan hasil klasifikasi yang dilakukan.

REFERENSI

- Amyar, A., Modzelewski, R., & Ruan, S. (2020). Multi-task Deep Learning Based CT Imaging Analysis For COVID-19: Classification and Segmentation. *Computers in Biology and Medicine*.
<https://doi.org/10.1101/2020.04.16.20064709>
- Breiman, L. (2001). Random forests. *Machine Learning*.
<https://doi.org/10.1023/A:1010933404324>
- Devella, S., Yohannes, Y., & Rahmawati, F. N. (2020). Implementasi Random Forest Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*.
<https://doi.org/10.35957/jatisi.v7i2.289>
- Hernando, D., Widodo, A. W., & Dewi, C. (n.d.). Pemanfaatan Fitur Warna dan Fitur Tekstur untuk Klasifikasi Jenis Penggunaan Lahan pada Citra Drone. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN, 2548, 964X*.
- Htun, M. L., Han, T. T., & Hlaing, N. N. (n.d.). Content Based Image Retrieval Using Color and Texture Features. *Image, 1, 2*.
- Hussain, C. A., Rao, D. V., & Praveen, T. (2013). Color histogram based image retrieval. *Int J Adv Engg Tech/IV/III/July-Sept, 63, 66*.
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... Feng, Z. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *New England Journal of Medicine*.
<https://doi.org/10.1056/nejmoa2001316>
- Ning, W., Lei, S., Yang, J., Cao, Y., Jiang, P., Yang, Q., ... others. (2020). *iCTCF: an integrative resource of chest computed tomography images and clinical features of patients with COVID-19 pneumonia*.
- Perumal, V., Narayanan, V., & Rajasekar, S. J. S. (2020). Detection of COVID-19 using CXR and CT images using Transfer Learning and

- Haralick features. *Applied Intelligence*.
<https://doi.org/10.1007/s10489-020-01831-z>
- Priyantono, M. B., Rachmawan, A. A., Budi, L. A. P., & Kirana, K. C. (2020). Sistem Prediksi Gejala Virus Korona dengan Metode Forward Chaining. *JTERA (Jurnal Teknologi Rekayasa)*.
<https://doi.org/10.31544/jtera.v5.i1.2019.111-118>
- Simarmata, S. Y. E., Sari, Y. A., & Adinugroho, S. (2019). Klasifikasi Citra Makanan Menggunakan Algoritme Learning Vector Quantization Berdasarkan Ekstraksi Fitur Color Histogram dan Gray Level Co-occurrence Matrix. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*.
- Thepade, S. D., Bang, S. V., Chaudhari, P. R., & Dindorkar, M. R. (2020). Covid19 Identification from Chest X-ray Images using Machine Learning Classifiers with GLCM Features. *ELCVIA Electronic Letters on Computer Vision and Image Analysis*.
<https://doi.org/10.5565/rev/elcvia.1277>
- Žunić, J., Hirota, K., & Rosin, P. L. (2010). A Hu moment invariant as a shape circularity measure. *Pattern Recognition*, 43(1), 47–57.
<https://doi.org/https://doi.org/10.1016/j.patcog.2009.06.017>