

TUGAS AKHIR

**KLASIFIKASI MONKEYPOX DAN NON-MONKEYPOX
MENGUNAKAN FITUR EKSTRAKSI LBP
DAN ALGORITMA RANDOM FOREST**



Disusun oleh :

Gracivo Elsion Victory 1822250058

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN REKAYASA
UNIVERSITAS MULTI DATA PALEMBANG
PALEMBANG
2023**

**Fakultas Ilmu Komputer Dan Rekayasa
Universitas Multi Data Palembang**

Program Studi Informatika
Tugas Akhir Sarjana Komputer
Semester Gasal Tahun 2022/2023

**KLASIFIKASI MONKEYPOX DAN NON-MONKEYPOX
MENGUNAKAN FITUR EKSTRAKSI LBP
DAN ALGORITMA RANDOM FOREST**

Gracivo Elsion Victory 1822250058

Abstrak

Monkeypox adalah penyakit yang disebabkan oleh virus *monkeypox*. Secara klinis, *monkeypox* sangat mirip dengan penyakit *non-monkeypox* seperti cacar air dan campak. Sebagian orang sulit membedakan antara *monkeypox* dan *non-monkeypox*. LBP merupakan metode ekstraksi fitur yang sederhana dan efisien dalam merepresentasikan ciri tekstur dan *Random Forest* merupakan metode *machine learning* yang dalam proses seleksi fitur dapat mengambil fitur terbaik untuk meningkatkan performa model klasifikasi. Penelitian ini menggunakan *public dataset monkeypox* dan *non-monkeypox* dan memiliki gambar berjumlah 3.192 dan berukuran 224 x 224 pixel. Output LBP menghasilkan *feature vector* dengan ukuran 1 x 59 sebagai input klasifikasi metode *random forest* dengan nilai *n_estimator* yaitu 100, 500 dan 1000. Hasil pengujian menggunakan fitur ekstraksi LBP dan metode *Random Forest* terhadap kelas *monkeypox* dan *non-monkeypox* dengan proporsi dataset 80:20 mendapatkan *n_estimator* terbaik yaitu 500 dengan nilai *accuracy* tertinggi sebesar 85%.

Kata kunci: *Monkeypox, Non-Monkeypox, Local Binary Pattern, Random Forest*



BAB 1

PENDAHULUAN

Pada bab pendahuluan berisi penjelasan terkait latar belakang, rumusan masalah, ruang lingkup, tujuan dan manfaat, dan sistematika penulisan yang berisi garis besar setiap bab.

1.1 Latar Belakang

Monkeypox adalah penyakit yang disebabkan oleh virus *monkeypox*. *Monkeypox* ditemukan pada tahun 1958 di Denmark ketika terjadinya dua kasus seperti cacar pada koloni kera yang dipelihara untuk penelitian, sehingga cacar ini dinamakan *monkeypox*. *Monkeypox* terdeteksi pada manusia pertama kali di Republik Demokratik Kongo (Zaire/DRC) pada tahun 1970 hingga Mei 2022, *monkeypox* menjadi perhatian kesehatan oleh dunia (Sinto dkk., 2022).

Monkeypox menular ketika terjadinya kontak langsung dengan orang atau hewan yang terinfeksi, kemudian bisa juga melalui benda yang terkontaminasi oleh virus tersebut. Secara klinis, *monkeypox* sangat mirip dengan penyakit non-*monkeypox* atau smallpox (seperti cacar air dan campak), yakni dari penampakan ruam seperti bintik merah yang menyebar ke seluruh tubuh secara bertahap. Hal inilah yang menyebabkan sebagian orang sulit membedakan antara penyakit *monkeypox* dan non-*monkeypox* (Sinto dkk., 2022).

Penelitian terhadap *monkeypox* dan *non-monkeypox* pernah dilakukan oleh Ali dkk pada tahun 2022 menggunakan algoritma CNN dengan arsitektur VGG 16 menghasilkan akurasi sebesar 81,48%, Resnet50 menghasilkan akurasi sebesar 82,96%, InceptionV3 menghasilkan akurasi sebesar 74,07%, dan Ensemble menghasilkan akurasi sebesar 79,26% (Ali dkk., 2022).

Saat ini, teknologi selalu berkembang untuk mempermudah pekerjaan manusia, salah satunya adalah Image Classification yang menjadi bagian dari teknologi Computer Vision. Image Classification menduplikasikan kemampuan manusia dalam memahami informasi citra digital, dengan tujuan agar komputer dapat mengenali objek pada citra selayaknya manusia (Nurhikmat, 2018).

Fitur tekstur adalah salah satu fitur yang digunakan untuk menentukan sebuah objek. Fitur tekstur dapat diekstraksi dengan berbagai macam metode, salah satunya adalah Local Binary Pattern (LBP). LBP merupakan salah satu fitur yang sederhana dan efisien dalam merepresentasikan ciri tekstur. Salah satu kelebihan LBP adalah tidak terpengaruh pada pencahayaan yang tidak merata pada citra, karena LBP mendeskripsikan tekstur secara lokal (Ojala dkk., 2000)

Beberapa penelitian yang menggunakan fitur tekstur LBP sudah diimplementasikan untuk klasifikasi citra seperti kulit, wajah, sidik jari, makanan, dan beberapa objek lainnya. Penerapan fitur LBP untuk mendeteksi limfoblas pada citra sel darah mendapatkan akurasi sebesar 94,32% (Sri Indrawanti & Prakarsa Mandyartha, 2018). Selain itu, penelitian tentang identifikasi penyakit daun tanaman apel mendapatkan akurasi sebesar 91,41% (Putri dkk., 2021)

Oleh karena itu LBP baik digunakan sebagai fitur yang digunakan untuk melakukan ekstraksi fitur terhadap citra monkeypox yaitu ditunjukkan oleh penampakan ruam atau perubahan tekstur pada Seperti pada kulit, oleh karena itu Fitur ekstraksi LBP digunakan dalam penelitian ini.

Salah satu metode machine learning yang sering digunakan untuk klasifikasi adalah Random Forest. *Random forests* adalah suatu metode klasifikasi yang terdiri dari gabungan pohon klasifikasi (CART) yang saling independen. Prediksi klasifikasi diperoleh melalui proses *voting* (jumlah terbanyak) dari pohon-pohon klasifikasi yang terbentuk. *Random forests* merupakan pengembangan dari metode *ensemble* yang pertama kali dikembangkan oleh Breiman (2001) dan digunakan untuk meningkatkan ketepatan klasifikasi.

Metode *Ensemble (Ensemble Method)* merupakan salah satu teknik machine learning dimana menggabungkan beberapa model dasar sehingga akan menghasilkan satu model prediksi yang lebih optimal. Penyebab kesalahan dalam proses training adalah varians, noise, dan bias. Maka dari itu, ensemble method akan membantu meminimalkan faktor penyebab tersebut dimana ensemble method dirancang untuk memperkuat keakuratan dari algoritma Machine Learning. Keuntungan dalam menggunakan *ensemble method* yakni menghasilkan akurasi yang lebih tinggi dan cenderung lebih stabil.

Random Forest termasuk salah satu algoritma yang menggunakan *ensemble method* (Lutins, 2017). Algoritma *Random Forest* mampu memberikan *Accuracy* yang lebih besar dibandingkan dengan algoritma pohon yang lainnya seperti *decision tree*.

Namun demikian, beberapa penelitian menghimbau agar dalam model klasifikasi dilakukan optimasi untuk membuat performa yang dihasilkan dapat lebih baik. Pembentukan parameter untuk penentuan kelas data memiliki ruang pencarian yang besar dan *overfitting* sehingga kelas data yang terbentuk kurang optimal (Elyan & Gaber, 2017). Salah satu cara pengoptimalan dalam menentukan kelas data yaitu dengan menggunakan metode *Bagging (Bootstrap aggregating)* dalam mencegah *overfitting* dan mengurangi *varians* data (Arfiani & Rustam, 2019).

Bagging merupakan suatu teknik *ensemble* dan biasanya teknik ini digunakan dalam klasifikasi untuk memisahkan data latih ke dalam beberapa data latih baru dengan *random sampling* sehingga menghasilkan model berbasis data latih baru (Wahono, 2013). Algoritma *Bagging* dapat digunakan untuk meningkatkan kecepatan klasifikasi dan lebih efisien serta mampu menurunkan kebutuhan memori (Roshan & Asadi, 2020). *Bagging* berfungsi untuk memperbaiki stabilitas dan meningkatkan keakuratan klasifikasi yang dihasilkan oleh metode tunggal (Hasibuan *et al.*, 2019).

Terdapat beberapa penelitian yang menggunakan teknik *Bagging* diantaranya yaitu, Prasajo & Haryatmi (2021) melakukan penelitian untuk mengklasifikasi data kanker ovarium menggunakan *Bagging* dengan metode *Random Forest*. Pada penelitian tersebut, tingkat akurasi tertinggi yang diperoleh yaitu sebesar 98.2% menggunakan metode *Random Forest* dan mengalami peningkatan sebesar 2.8% setelah diterapkan teknik *Bagging* pada metode *Random Forest*.

Proses seleksi fitur *random forest* mampu mengambil fitur terbaik untuk meningkatkan performa pada model klasifikasi. Adanya fitur seleksi membuat *random forest* mampu bekerja pada data yang besar dengan parameter yang kompleks secara efektif. (Devella dkk., 2020). Penelitian yang menggunakan metode *random forest* seperti klasifikasi kanker kulit menggunakan *random forest* menghasilkan akurasi sebesar 97,3% (Nandhini dkk., 2019). Kemudian penelitian mengidentifikasi virus COVID-19 dari chest x-ray menghasilkan akurasi sebesar 99.17% ((Thepade dkk., 2020).

Berdasarkan penjelasan penelitian sebelumnya, ekstraksi fitur LBP memiliki tingkat akurasi yang baik untuk mengenali ciri tekstur pada objek, serta metode *random forest* juga memiliki tingkat akurasi yang baik dalam klasifikasi sebuah objek. Oleh karena itu akan dilakukan penelitian klasifikasi *monkeypox* dan *non-monkeypox* menggunakan fitur ekstraksi LBP dan algoritma *random forest*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan, maka perumusan dalam penelitian ini adalah menerapkan fitur ekstraksi LBP dan algoritma *random forest* sebagai metode klasifikasi *monkeypox* dan *non-monkeypox*.

1.3 Ruang Lingkup

Ruang lingkup dari penelitian ini yaitu :

1. Menerapkan LBP sebagai ekstraksi fitur tekstur.
2. Menggunakan algoritma *random forest* sebagai metode klasifikasi.
3. Menggunakan parameter random forest yaitu dengan nilai *n_estimator* yaitu 100, 500, 1000.
4. Objek gambar yang diolah adalah citra *Monkeypox* dan *Non-Monkeypox* (meliputi cacar air dan campak) menggunakan public dataset yang terdapat pada situs kaggle (<https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset>).
5. Objek gambar berupa citra RGB yang berjumlah 3.192 gambar dengan dimensi masing-masing berukuran berukuran 224 x 224 pixel. Dataset dibagi menjadi data latih dan data uji dengan rasio 60:40, 70:30, dan 80:20.
6. Format dari objek gambar yang digunakan adalah *.jpg.
7. Menggunakan bahasa pemrograman Python dan Google Collab sebagai perangkat lunak uji implementasi.

1.4 Tujuan dan Manfaat

Adapun tujuan dari penelitian ini adalah menerapkan fitur ekstraksi LBP dan algoritma *Random Forest* untuk klasifikasi *Monkeypox* dan *Non-Monkeypox*. Berikut manfaat dari penelitian ini adalah :

1. Mengetahui tingkat Precision, Recall, Accuracy dan F1-Score dari hasil klasifikasi *Monkeypox* dan *Non-Monkeypox* menggunakan fitur ekstraksi LBP dan algoritma Random Forest.
2. Mengetahui n_estimator terbaik dari hasil nilai *Precision*, *Recall*, *Accuracy*, dan *F1 Score* berdasarkan keseluruhan kelas.
3. Menambah hasil penelitian tentang klasifikasi *Monkeypox* dan *Non-Monkeypox* yang menggunakan fitur ekstraksi LBP dan algoritma Random Forest.

1.5 Sistematika Penulisan

Sistematika penulisan merupakan gambaran dan penjelasan singkat mengenai bab - bab yang telah disusun oleh penulis dalam laporan skripsi. Penulisan skripsi terdiri dari lima bab, dimana tiap bab terdiri dari sub bab. Susunan garis besar sistematika penulisan dapat dilihat dibawah ini :

BAB 1 PENDAHULUAN

Pada bab pendahuluan ini berisi penjelasan terkait dengan latar belakang, perumusan masalah, ruang lingkup, tujuan dan manfaat, dan sistematika penulisan penelitian.

BAB 2 LANDASAN TEORI

Pada bab landasan teori berisi penjelasan tentang teori-teori pendukung yang berkaitan dengan penelitian dan penelitian terkait untuk mendukung penelitian yang dibuat.

BAB 3 METODE PENELITIAN

Pada bab metode penelitian ini terdiri dari lingkungan pengembangan program serta strategi / metodologi pemecahan masalah.

BAB 4 HASIL DAN PEMBAHASAN

Pada bab hasil dan pembahasan ini terdiri dari penjelasan implementasi dari penggunaan fitur tekstur LBP dan hasil dari metode klasifikasi *Random Forest*.

BAB 5 PENUTUP

Pada bab penutup ini berisi kesimpulan dan saran mengenai hasil dari penelitian yang telah selesai diimplementasi.





DAFTAR PUSTAKA

- Ahsan, M. M., Uddin, M. R., Farjana, M., Sakib, A. N., Momin, K. Al, & Luna, S. A. (2022). *Image Data collection and implementation of deep learning-based model in detecting Monkeypox disease using modified VGG16*. <http://arxiv.org/abs/2206.01862>
- Ali, S. N., Ahmed, M. T., Paul, J., Jahan, T., Sani, S. M. S., Noor, N., & Hasan, T. (2022). *Monkeypox Skin Lesion Detection Using Deep Learning Models: A Feasibility Study*. <http://arxiv.org/abs/2207.03342>
- Arfiani, A., & Rustam, Z. (2019). Ovarian cancer data classification using *Bagging* and random forest. *AIP Conference Proceedings*, 2168(November).
- Breiman, L. (2001). *Random Forests*. <https://doi.org/https://doi.org/10.1023/A:1010933404324>
- Cahaya, F. N., Pebrianto, R., & M, T. A. (2021). Klasifikasi Buah Segar dan Busuk Menggunakan Ekstraksi Fitur Hu-Moment , Haralick dan Histogram. *IJCIT*, 6(1), 57–62.
- Chauhan, D., Walia, R., Singh, C., Deivakani, M., Kumbhkar, M., & Professor, A. (2021). Detection of Maize Disease Using Random Forest Classification Algorithm. In *Turkish Journal of Computer and Mathematics Education* (Vol. 12, Nomor 9).
- 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)*, 7(2), 310–320. <https://doi.org/10.35957/jatisi.v7i2.289>
- Elyan, E., & Gaber, M. M. (2017). A genetic algorithm approach to optimising random forests applied to class engineered data. *Information Sciences*, 384, 220–234.

- Fachruddin, M. I. (2015). Perbandingan Metode Random Forest Classification Dan Support Vector Machine Untuk Deteksi Epilepsi Menggunakan Data Rekaman Electroencephalograph (EEG). *Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember*, 1–83.
- Fadila, I. (2021). *Apa Saja Perbedaan Cacar Air dan Campak?*
<https://hellosehat.com/parenting/kesehatan-anak/infeksi-anak/perbedaan-cacar-air-dan-campak/>
- Fatni, Z. (2021). *Klasifikasi Citra Magnetic Resonance Imaging (MRI) Otak Dalam Mengidentifikasi Tumor Menggunakan Algoritma Random Forest.*
- Gautam, A., & Raman, B. (2020). Skin cancer classification from dermoscopic images using feature extraction methods. *IEEE Region 10 Annual International Conference, Proceedings/TENCON, 2020-November*, 958–963.
<https://doi.org/10.1109/TENCON50793.2020.9293863>
- Hasibuan, M. J. A., Rusgiyono, A., & Safitri, D. (2019). Pemodelan Produk Domestik Regional Bruto (Pdrb) Di Provinsi Jawa Tengah Menggunakan Bootstrap Aggregating Multivariate Adaptive Regression Splines (*Bagging Mars*). *Jurnal Gaussian*, 8(1), 139–148.
- Kaplan, K., Kaya, Y., Kuncan, M., & Ertunç, H. M. (2020). Brain tumor classification using modified local binary patterns (LBP) feature extraction methods. *Medical Hypotheses*, 139. <https://doi.org/10.1016/j.mehy.2020.109696>
- Khasanah, N., Komarudin, R., Afni, N., Maulana, Y. I., & Salim, A. (2021). Skin Cancer Classification Using Random Forest Algorithm. *SISFOTENIKA*, 11(2), 137. <https://doi.org/10.30700/jst.v11i2.1122>
- Khultsum, U., & Subekti, A. (2021). Penerapan Algoritma Random Forest dengan Kombinasi Ekstraksi Fitur Untuk Klasifikasi Penyakit Daun Tomat. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 186.
<https://doi.org/10.30865/mib.v5i1.2624>
- Lutins, E. (2017). Ensemble Methods in Machine Learning: What are They and Why Use Them?
<https://towardsdatascience.com/ensemble-methods-in-machine-learning-what-are-they-and-why-use-them-68ec3f9fef5f>

- Nandhini, S., Sofiyan, M. A., Kumar, S., & Afridi, A. (2019). International Journal of Management and Humanities (IJMH) Skin Cancer Classification using Random Forest. *International Journal of Management and Humanities (IJMH)*, 4(3). <https://doi.org/10.35940/ijmh.C0434.114319>
- Naresh, Y. G., & Nagendraswamy, H. S. (2016). Classification of medicinal plants: An approach using modified LBP with symbolic representation. *Neurocomputing*, 173, 1789–1797. <https://doi.org/10.1016/j.neucom.2015.08.090>
- Nurhikmat, T. (2018). *IMPLEMENTASI DEEP LEARNING UNTUK IMAGE CLASSIFICATION MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN) PADA CITRA WAYANG GOLEK*. Universitas Islam Indonesia.
- Pankaja, K., & Suma, V. (2020). Plant Leaf Recognition and Classification Based on the Whale Optimization Algorithm (WOA) and Random Forest (RF). *Journal of The Institution of Engineers (India): Series B*, 101(5), 597–607. <https://doi.org/10.1007/s40031-020-00470-9>
- Pietikäinen, M., Hadid, A., Zhao, G., & Ahonen, T. (2011). *Computer Vision Using Local Binary Patterns*. Springer. https://doi.org/10.1007/978-0-85729-748-8_14
- Prasetyo, E. (2015). *Data Mining: Konsep dan Aplikasi menggunakan MATLAB*.
- Prasojo, B., & Haryatmi, E. (2021). Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest. *Jurnal Nasional Teknologi dan Sistem Informasi*, 7(2), 79–89.
- Priyanto. (2017). *Pengolahan Citra Digital Teori dan Aplikasi Nyata*. Informatika Bandung.
- Putri, F. D., Ramadhani, K. N., & Yunanto, P. E. (2021). *Identifikasi Penyakit pada Daun Tanaman Apel Menggunakan Local Binary Pattern (LBP) dan Color Histogram*.

- Roshan, S. E., & Asadi, S. (2020). Improvement of *Bagging* performance for classification of imbalanced datasets using evolutionary multi-objective optimization. *Engineering Applications of Artificial Intelligence*, 87(May 2019), 103319.
- S, A. A. N. F., & Erliyani, E. (2022). *CACAR MONYET DI INDONESIA DAN DUNIA: AKANKAH JADI PANDEMI JUGA?* egsaugm. <https://egsa.geo.ugm.ac.id/2022/09/07/cacar-monyet-di-indonesia-dan-dunia-akankah-jadi-pandemi-juga/>
- Sinto, R., Shakinah, S., Pitawati, N. L. P., & Sitompul, P. A. (2022). *PENYAKIT MONKEYPOX*. <http://infeksiemerging.kemkes.go.id>
- Solichin, A. (2017). *Mengukur Kinerja Algoritma Klasifikasi dengan Confusion Matrix*.
- Sri Indrawanti, A., & Prakarsa Mandyartha, E. (2018). Deteksi Limfoblas pada Citra Sel Darah Menggunakan Fitur Geometri dan Local Binary Pattern. In *JNTETI* (Vol. 7, Nomor 4).
- 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. *Electronic Letters on Computer Vision and Image Analysis*, 19(3), 85–97. <https://doi.org/10.5565/REV/ELCVIA.1277>
- Wahono, R. S. (2013). *Combining Particle Swarm Optimization based Feature Selection and Bagging Technique for So ware Defect PRediction*.
- Yusuf, B., Qalbi, M., Dwitawati, I., Ellyadi, M., Studi Pendidikan Teknologi Informasi, P., & Tarbiyah dan Keguruan, F. (n.d.). Implementasi Algoritma Naive Bayes Dan Random Forest Dalam Memprediksi Prestasi Akademik Mahasiswa Universitas Islam Negeri Ar-Raniry Banda Aceh. In *Jurnal Pendidikan Teknologi Informasi* (Vol. 4, Nomor 1).