

Kajian Literatur Multimedia Retrieval : Machine Learning Untuk Pengenalan Wajah

Mahdianta Pandia^{1*}, Poltak Sihombing², Purwanto Simamora³, Roberto Kaban⁴

^{1,2,3,4} Fakultas Ilmu Komputer dan Teknologi Informasi, Teknik Komputer, Universitas Sumatera Utara, Medan, Indonesia
Gedung C Fasilkom-TI, Universitas Sumatera Utara, Jl. Alumni No.3, Padang Bulan, Kec. Medan Baru, Kota Medan,
Sumatera Utara, Indonesia, 20155

Email: ¹ mahdianta03@gmail.com, ² poltak@usu.ac.id, ³ purwantosimamora@gmail.com, ⁴ roberto.kaban@yahoo.com

Abstrak– Multimedia retrieval adalah pencarian dan temu kembali mengenai informasi yang terkandung pada konten multimedia. Konten multimedia terdiri dari gambar, teks, video, suara atau kombinasi dari keempatnya. Konten multimedia khususnya foto diupload lebih dari lebih dari 1,2 triliun foto dan video digital diambil setiap tahunnya. Delapan puluh lima (85) % dari konten multimedia tersebut diambil menggunakan *smartphone* dan langsung di upload ke media sosial. Penumpukan konten multimedia akan terus bertambah setiap tahunnya, sehingga membutuhkan waktu untuk menelusurinya kembali di media penyimpanan yang digunakan. Multimedia retrieval dapat mengklasifikasikan konten tersebut berdasarkan pemilik wajah dari konten tersebut. Pengenalan wajah dapat dilakukan dengan baik menggunakan artificial intelligence. Perkembangan artificial intelligence juga terus berkembang hingga munculnya teknologi machine learning. Saat ini banyak dilakukan penelitian mengenai multimedia retrieval menggunakan machine learning yang didukung dengan algoritma AI yang lain, seperti deep learning. Pada kajian literatur ini akan melakukan kajian tentang multimedia retrieval, machine learning dan algoritma yang digunakan dalam pengenalan wajah sehingga diperoleh kesimpulan Tingkat keberhasilan metode multimedia retrieval dan machine learning untuk mengenali wajah.

Kata Kunci: Multimedia Retrieval, Machine Learning, Deep Learning, Supervised Learning

Abstract– *Multimedia retrieval is the search and retrieval of information contained in multimedia content. Multimedia content consists of images, text, video, sound or a combination of the four. Multimedia content, especially photos, is uploaded from more than 1.2 trillion digital photos and videos taken every year. Eighty-five (85)% of the multimedia content was taken using a smartphone and directly uploaded to social media. The accumulation of multimedia content will continue to increase every year, so it takes time to search for it again on the storage media used. Multimedia retrieval can classify the content based on the owner's face of the content. Facial recognition can be done well using artificial intelligence. The development of artificial intelligence also continues to grow until the emergence of machine learning technology. Currently, there is a lot of research being carried out on multimedia retrieval using machine learning which is supported by other AI algorithms, such as deep learning. In this literature review, we will conduct a study on multimedia retrieval, machine learning and algorithms used in face recognition so that conclusions can be drawn on the level of success of multimedia retrieval and machine learning methods for recognizing faces.*

Keywords: Multimedia Retrieval, Machine Learning, Deep Learning, Supervised Learning

1. PENDAHULUAN

Perkembangan teknologi kamera *smartphone* yang sangat canggih membuat media sosial banyak dimanfaatkan oleh pengguna sebagai media penyimpanan konten multimedia. Konten multimedia tersebut telah membentuk big data video dan foto [1]. Big data foto dan video, tersebut terdiri dari 45% wajah [2], sehingga perlu dilakukan penelitian yang dalam dengan skala besar [3].

Big data adalah istilah yang diberikan pada kumpulan data yang berukuran sangat besar dan kompleks, sehingga tidak memungkinkan untuk diproses menggunakan perangkat pengelola pangkalan data konvensional [4]. *Multimedia retrieval* merupakan ilmu yang mempelajari tentang temu kembali (*retrieval*) informasi yang terkandung dalam multimedia. Informasi tersebut berasal dari komponen multimedia seperti audio, video, gambar, teks atau kombinasinya.

Pada penelitian [3] [5] [1] telah berhasil melakukan pengenalan wajah dengan menggunakan metode berbeda tetapi masih dalam ruang lingkup *machine learning*. Penelitian tersebut menggunakan metode *Features and Deep Learning* [1], *Supervised Deep Hashing For Scalable Faceimage Retrieval* [3] dan menggunakan *artificial intelligence* dengan *machine learning* [5].

Penelitian tersebut berhasil mengenali wajah dengan baik dengan dataset foto digital atau video yang diekstrak ke bentuk frame. Perkembangan saat ini akan jauh lebih menantang dan kompleks untuk dibidang *retrieval multimedia* yang bersumber dari bigdata multimedia. Tantangan bukan hanya sekedar berhasil tetapi mampu mengoptimalkan waktu dalam pencarian dan melakukan *retrieval* dengan cepat, tepat dan dapat memprediksi query pencarian sehingga diperoleh hasil yang akurat.

Penelitian ini difokuskan untuk melakukan kajian *machine learning* dalam pengenalan wajah (*face recognition*). Penelitian terdiri dari tiga bagian yang berbeda. Bagian pertama memaparkan gambaran singkat tentang *multimedia retrieval*, *machine learning beserta komponennya* dan tantangan kedepannya. Bagian kedua menjelaskan proses melakukan

temu kembali dengan menggunakan metode machine learning. Bagian ketiga menyajikan penelitian terdahulu yang melakukan proses pengenalan wajah menggunakan machine learning.

2. METODOLOGI PENELITIAN

2.1. Tahapan Penelitian

Penelitian ini menggunakan metode tinjauan literatur dengan menyusun tahapan pertama, melakukan identifikasi dengan membaca sumber utama dan sumber tambahan, berupa jurnal, buku, hasil penelitian. Tahap kedua melakukan kajian analisis dan membandingkan dengan beberapa penelitian sebelumnya.

Tujuan penelitian ini untuk mengamati dan membandingkan hasil penelitian dari peneliti sebelumnya mengenai mengenali wajah menggunakan machine learning. Sumber literatur di peroleh secara online dari ACM Digital Library, ResearchGate, IEEE, MDPI. Sumber utama penelitian dari MDPI, khususnya penggunaan machine learning dalam mengenali wajah.

2.2. Artificial Intelligence

Artificial intelligent (AI) dalam bahasa Indonesia kecerdasan buatan, mempunyai arti : “intelligence” adalah bahasa Latin “intelligo” yang memiliki arti “saya paham”. AI adalah simulasi dari kecerdasan yang dimiliki oleh manusia yang dimodelkan di dalam mesin dan diprogram agar bisa berpikir seperti halnya manusia. Beberapa invensi awal dalam pemecahan masalah termasuk pekerjaan dasar dalam pembelajaran, representasi pengetahuan, dan kesimpulan sebagai program demonstrasi dalam pengertian bahasa, terjemahan, pembuktian teori, memori asosiatif, dan sistem berbasis pengetahuan [6].

2.3. Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan. Kecerdasan buatan merupakan salah satu bidang keilmuan yang mempelajari tentang komputer sains yang bertujuan untuk membuat komputer bisa memiliki kepintaran, kecerdasan untuk membantu manusia dalam menyelesaikan tugas-tugasnya. Machine learning secara khusus akan menggunakan metode statistika untuk mempelajari data sehingga komputer dapat mempelajari pola pada data tanpa perlu diprogram secara eksplisit [7].

Machine learning memiliki cabang yang menggunakan algoritma jaringan saraf tiruan untuk dapat belajar beradaptasi terhadap sejumlah data yang besar. Algoritma jaringan syaraf tiruan pada *deep learning* terinspirasi dari struktur otak manusia. Algoritma ini memungkinkan mesin untuk melihat pola dari data yang tidak terstruktur atau data yang fiturnya tidak dapat ditentukan secara langsung. Contohnya, data gambar, teks, audio, dan video. Belajar yang dimaksud tidaklah sama dengan proses manusia belajar, mesin mampu belajar apabila ia mampu meng-upgrade parameter, dimana parameter tersebut kurang-lebih merepresentasikan pengetahuan mesin [6].

Machine Learning membutuhkan Data yang valid sebagai bahan belajar (ketika proses training) sebelum digunakan testing untuk hasil output yang optimal. *Machine learning* dibagi menjadi 4 bagian berdasarkan metode pembelajarannya yaitu : (a)*supervised learning*, (b)*unsupervised learning*, (c)*semi-supervised learning* dan (d)*reinforcement learning* [7]. Masing-masing metode ini memiliki cara yang berbeda dalam mengenali, mengolah data sehingga menghasilkan keluaran yang benar. Proses belajar *machine learning* berdasarkan data yang disebut dengan dataset, data set yang telah terkumpul akan dibagi menjadi dua, yaitu data training dan data testing. Data tersebut akan digunakan pada jenis machine learning yang sesuai.

Supervised learning adalah bagian machine learning yang mempelajari sebuah dataset yang telah memiliki label, serta menghasilkan solusi yang diinginkan [7]. Dataset yang telah memiliki label kemudian dipelajari oleh algoritma dan algoritma mempelajari pola dari pasangan data dan label tersebut. Terdapat banyak metode yang ada *regresi linier*, *random forest* dan *support vector machine*.

Unsupervised learning bersifat deskriptif [7], yang akan berguna untuk mengelompokkan atau mengkategorikan data. Algoritma ini tidak mendapatkan training data set, karena algoritma ini bukan bersifat prediktif, sehingga membutuhkan pembelajaran dari data yang telah ada. *Unsupervised learning* bekerja dengan mengelompokkan data berdasarkan kedekatannya disebut dengan *clustering*. Metode ini bekerja dengan menganalisis data yang tidak berlabel untuk menemukan pola tersembunyi dan menentukan korelasinya. Beberapa contoh algoritma yang dapat digunakan dalam unsupervised learning seperti, *K-Means*, *Hierarchical clustering*, *DBSCAN*, dan *Fuzzy C-Means*.

Semi-Supervised Learning merupakan gabungan dari *supervised learning* dan *unsupervised learning*. Dataset yang digunakan ada yang memiliki label ada juga yang tidak memiliki label. Untuk menyelesaikan jenis dataset seperti ini dapat menggunakan algoritma *semi supervised learning*. Algoritma ini berada di tengah-tengah antara algoritma *supervised* dan *unsupervised learning*. Cara kerja algoritma ini adalah dengan menggunakan teknik *unsupervised learning* untuk menemukan dan mempelajari struktur dalam variabel input. Setelah itu, sistem akan menggunakan algoritma *supervised learning* untuk membuat prediksi terbaik dari data tak berlabel dan memasukkan kembali data tersebut ke

algoritma *supervised learning* sebagai data latih, kemudian menggunakan model tersebut untuk membuat prediksi baru dari data input baru. Salah satu contoh dari model semi *supervised learning* adalah *Deep Belief Network* (DBNs).

Reinforcement learning dikenal sebagai model yang belajar menggunakan sistem *reward* dan penalti. Algoritma *reinforcement learning* didefinisikan sebagai metode *machine learning* yang berkaitan dengan cara *software agent* mengambil tindakan di *environmentnya*. *Reinforcement learning* adalah teknik yang mempelajari bagaimana membuat keputusan terbaik, secara berurutan, untuk memaksimalkan ukuran sukses kehidupan nyata.

2.4. Deep Learning

Deep Neural Network (DNN) adalah *artificial neural network* yang memiliki banyak layer. Pada umumnya, deep neural network memiliki lebih dari 3 layers (input layer, $N > 2$ hidden layers, output layer). Karena ada relatif banyak layer, disebutlah deep. Proses pembelajaran pada DNN disebut sebagai *deep learning* [7]. Deep learning adalah bagian dari algoritma machine learning dengan kompleksitas karakteristik yang tinggi [1]. Dataset yang digunakan untuk melakukan pelatihan sudah diberi label [7].

Deep learning adalah bagian dari algoritma mesin learning dengan karakteristik kompleksitas tinggi. Model deep learning dilatih dengan kumpulan data berlabel besar dan arsitektur *neural network* yang mempelajari fitur langsung dari data input tanpa melakukan ekstraksi fitur [1].

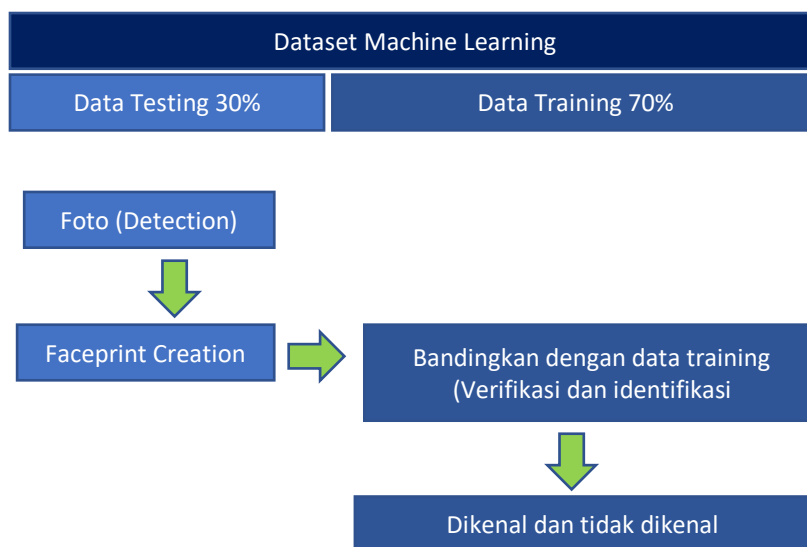
2.5. Pengenalan Wajah

Pengenalan wajah (face recognition) adalah sebuah proses pengenalan wajah dengan menggunakan metode tertentu, dalam perkembangannya penelitian menggunakan machine learning juga sudah digunakan. Algoritma Neural Network dan Deep learning sering digunakan untuk pengenalan pola terutama pengenalan wajah.

Secara umum pengenalan wajah dibagi menjadi 3 tahap, yaitu

1. Detection, pada tahap ini sistem akan menentukan apakah pada sebuah gambar terdapat pola wajah, dengan mengekstrak pola.
2. Faceprint creation, adalah proses pembuatan cetakan wajah dapat dilakukan dengan dua cara yaitu pendekatan geometris (*geometric approach*) dan pendekatan fotometrik (*photometric approach*)
3. Verifikasi atau identifikasi Proses verifikasi dilakukan dengan membandingkan foto yang dideteksi dengan data yang tersimpan di pangkalan data. Proses identifikasi, membandingkan foto yang dideteksi dengan seluruh wajah yang ada di database.

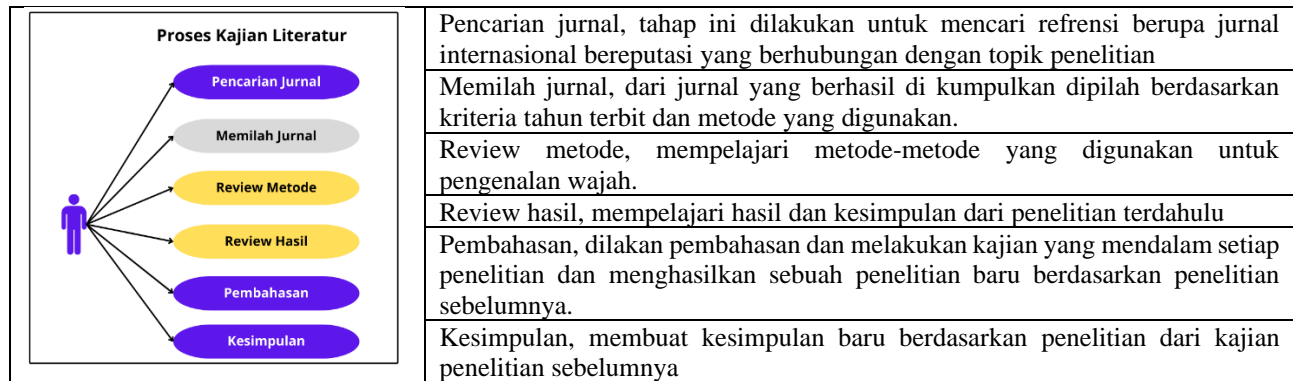
Pengenalalan wajah menggunakan machine learning membutuhkan dataset, semakin banyak data dan semakin baik kondisi data akan menghasilkan out yang lebih baik. Pembagian dataset menjadi data training dan testing adalah 80:20, 70:30 atau 60:40, tergantung ukuran dan jumlah data.



Gambar 1. tahapan pengenalan gambar menggunakan machine learning

3. HASIL DAN PEMBAHASAN

Kajian literatur multimedia retrieval pengenalan wajah menggunakan *machine learning* dilakukan dengan studi pustaka dari penelitian terdahulu dengan membaca sumber utama dan sumber pendukung dari berbagai jurnal internasional dan juga buku yang berhubungan dengan topik penelitian. Tahapan penelitian yang dilakukan ditunjukkan seperti pada gambar 2 dibawah ini.



Gambar 2. Hahap penelitian dan penjelasannya.

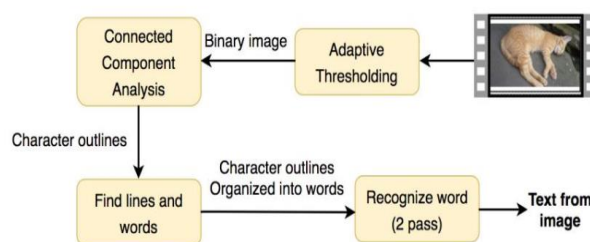
Berikut ini tabel kajian literatur penelitian terdahulu yang berhubungan dengan topik penelitian

Tabel 1. Penelitian terdahulu dan metode yang digunakan

Tahun	Metode	Judul Artikel	Penulis
2021	AI Machine Learning	AI-Based Semantic Multimedia Indexing and Retrieval for Social Media on Smartphones	Stefan Wagenpfeil dkk
2021	Machine Learning Deep Matric	Verifikasi Wajah Menggunakan Deep Metric Learning pada Data Wajah dengan Disparitas Umur yang Besar	Syauqi Sabili dkk
2022	Big Data Extensive Features Deep Learning	Content-Based Video Big Data Retrieval with Extensive Features and Deep Learning	Thuong-Cang Phan dkk
2022	Large-Scale Image Retrieval	A New Comparative Study of Dimensionality Reduction Methods in Large-Scale Image Retrieval	Mohammed Amin Belarbi dkk
2023	Machine Learning Principal Component Analysis	Tracking Unauthorized Access Using Machine Learning and PCA for Face Recognition Developments	Vasile-Daniel Păvăloaia dkk

Perjalanan penelitian pengenalan wajah menggunakan machine learning dan juga menggunakan algoritma yang sesuai dengan machine learning terus berkembang. Semua metode memberikan hasil yang baik, pada kajian literatur ini algoritma yang sering digunakan adalah *deep learning*.

Optical Character Recognition (OCR) adalah teknik untuk mengidentifikasi karakter pada gambar [1]. Library yang digunakan untuk OCR salah satunya adalah Tesseract, library ini bersifat *open source* yang dikembangkan oleh google. Keunggulannya adalah dengan akurasi hingga 97% [1], mendukung pengenalan banyak bahasa, berjalan di banyak platform dapat juga diinterintegrasikan dengan OpenCV. Proses identifikasi akan dilakukan secara berurutan sesuai dengan beberapa tahapan, seperti terlihat pada gambar 2 [1].



Gambar 2. Algoritma OCR menggunakan library Tesseract

Content-Based Image Retrieval (CBIR) System adalah sistem temukembali konten berbasis gambar. Kerangka kerja CBIR adalah sistem komputer digunakan untuk memvisualisasikan, menemukan kembali dan mengambil gambar digital dari database yang besar [8]. Secara tradisional cara satu-satunya untuk melakukan retrieval image adalah dengan menggunakan kata kunci atau hanya dengan browsing. Selain itu, kumpulan data citra digital telah membuka pendekatan pencarian berbasis konten [8]. Fokus CBIR adalah ekstraksi fitur yang efisien dengan menggunakan transformasi wavelet daripada struktur indeks untuk mendukung retrieval yang cepat. Pada CBIR terus perkembangan sehingga muncul sistem QBIC yang paling terkenal saat ini. QBIC dikembangkan oleh IBM untuk kueri CBIR [8].

Hasil dari penelitian [1] menunjukkan bahwa metode yang digunakan dengan *deep learning* terdistribusi mencapai akurasi 96% dan waktu pemrosesan dipersingkat 50% dibandingkan dengan metode lain tanpa *Spark*. Hasil penelitian [8] hasil percobaan mereka dengan menggunakan CBIR memperoleh kecepatan pencarian 30% hingga 40% dan kecepatan hingga 60% hingga 70% dari memori komputer. Hasil penelitian [9] perangkat lunak yang dibangun untuk mencapai FR terbukti efisien, andal dan berguna untuk adaptasi di masa depan.

Hasil dari penelitian [5] berhasil membangun aplikasi berbasis AI pada *smartphone* untuk melakukan multimedia retrieval pengenalan wajah. Hasil penelitian ini memberikan dasar yang kuat untuk penelitian, optimasi dan teknologi lebih lanjut, terkhusus pada ponsel cerdas dapat diperluas pengindeksan dan pengambilan multimedia saat ini dan di masa mendatang.

Hasil dari penelitian [10], menawarkan keempat model penelitian yaitu (InceptionV3, ResNet50V2, VGG19, dan Xception), pada saat pengujian diperoleh nilai akurasi terbaik dari proses testing, yaitu pada model Xception.

4. KESIMPULAN

Dalam penelitian kajian literatur ini, dimulai dengan mendefinisikan konsep multimedia image retrieval system secara mendasar dan juga teknologi *machine learning* yang dapat digunakan untuk melakukan retrieval khususnya pengenalan wajah. Laporan penelitian merupakan pembahasan dari beberapa penelitian sebelumnya beserta metode yang digunakan, dari hasil pengamatan kami berhasil melakukan akurasi mencapai 96% dengan metode *deep learning* dan proses pencarian dipersingkat menjadi 50% dibandingkan dengan metode tanpa *Spark*. *Multimedia retrieval* menggunakan machine learning berhasil melakukan pengenalan wajah dan juga menjadi metode yang cocok dikembangkan dimasa depan.

REFERENCES

- [1] A.-C. P. H.-P. C. T.-N. T. Thuong-Cang Phan, "Content-Based Video Big Data Retrieval with Extensive Features and Deep Learning," *Applied Sciencs*, Vol. 12, No. 1, pp. 1-26, 2022.
- [2] P. B. S. N. Neeraj Kumar, "FaceTracer: A Search Engine for Large Collections of Images with Faces," *European Conference on Computer Vision*, pp. 340-353, 2008.
- [3] Z. L. X. Z. Jinhui Tang, "Supervised Deep Hashing For Scalable Faceimage Retrieval," *Pattern Recognition*, vol. 75, pp. 25-32, 2018.
- [4] W. S. E. D. T. D. Nurul Fajriyah, "Implementasi Teknologi Big Data Di Era Digital," *Jurnal Informatika*, vol. 1, pp. 1-7, 2022.
- [5] F. E. P. M. K. M. H. Stefan Wagenpfeil, "AI-Based Semantic Multimedia Indexing and Retrieval for Social Media on Smartphones," <https://www.mdpi.com/>, vol. 12, pp. 1-30, 2021.
- [6] N. L. A. W. S. T. Kirana Rukmayuninda Ririh, "Studi Komparasi Dan Analisis Swot Pada Implementasi Kecerdasan Buatan (Artificial Intelligence) Di Indonesia," *Jurnal Teknik Industri*, vol. 15, pp. 122-133, 2020.
- [7] J. Putra, *Pengenalan konsep pembelajaran mesin dan deep learning* (Tokyo. Jepang, 2019), Tokyo. Jepang, 2019.
- [8] S. M. G. B. S. A. M. A. C. Mohammed Amin Belarbi, "A New Comparative Study of Dimensionality Reduction Methods in Large-Scale Image Retrieval," *Big data and cognitive computing*, vol. 6, pp. 1-22, 2022.
- [9] G. H. Vasile-Daniel Păvăloaia, "Tracking Unauthorized Access Using Machine Learning and PCA for Face Recognition Developments," *Information*, vol. 14, pp. 1-19, 2023.
- [10] R. F. R. d. E. M. Y. Syauqi Sabili, "Verifikasi Wajah Menggunakan Deep Metric Learning pada Data Wajah dengan Disparitas," *JURNAL TEKNIK ITS*, vol. 10, pp. 432-437, 2021.
- [11] A. R. U. S. J. A. N. A. N. I. R. Afshan Latif, *Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review*, Hindawi, 2019.
- [12] C. R. C. T. Antonio Maria Rinaldi, *A Knowledge-Driven Multimedia Retrieval System Based on Semantics and Deep Features*, MDI, 2020.

- [13] M. R. K. K. S. T. M. M. S. P. Ishan Ratn Pandey, Face Recognition Using Machine Learning, International Research Journal of Engineering and Technology, 2019.
- [14] M. I. I. R. T. K. M. A. Fahima Tabassum, "Human face recognition with combination of DWT and machine learning," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, 2022.
- [15] H. AliElmahmudi, "Deep face recognition using imperfect facial data," *Future Generation Computer Systems*, 2019.