



Mask Detection using Convolutional Neural Network

Adhigana Priyatama¹, Husein Abdul Hamid², Yufiz Azhar³

Abstract

Strategi yang efisien untuk menahan pandemi virus covid 2019 (COVID-19) merupakan hal yang wajib untuk meringankan dampak negatif kesehatan masyarakat dan ekonomi global. Dengan tidak adanya obat-obatan yang sangat efektif, banyak tindakan yang digunakan untuk mengelola tingkat infeksi dan menghindari keterbatasan sumber daya rumah sakit. Mengenakan masker adalah salah satu tindakan intervensi nonfarmasi (NPI) yang dapat diimplementasikan secara efektif dengan biaya minimum dan tanpa mengganggu praktek sosial secara drastis. Dalam penelitian ini, peneliti mencoba membangun sebuah sistem pendeteksi penggunaan masker wajah dengan harapan model dengan ketentuan yang dibangun dapat bermanfaat dan dapat dikembangkan lebih baik di masa depan. Data yang digunakan dalam penelitian ini adalah data citra berjumlah 11800 yang didapat dari *website kaggle* dimana data dikumpulkan dari beberapa *dataset* yang tersebar di internet. Dalam proses identifikasinya, digunakan metode *feature extraction* yang kemudian dilakukan tahap klasifikasi menggunakan metode convolutional neural network (CNN). Dari hasil penelitian yang dilakukan, peneliti dapat menemukan metode CNN yang tepat dan mampu menghasilkan pengolahan gambar sehingga mengeluarkan output kemiripan gambar berdasarkan *query* gambar yang diidentifikasi.

1. Introduction

Otak merupakan organ utama yang sangat luar biasa yang terdapat pada tubuh manusia. Dengan otak yang kita punya, kita dapat mengerti segala masukkan yang didapat dari kelima panca indra yang kita punya. Berat dari otak sendiri hanya berkisar setengah kilogram namun dapat menyelesaikan berbagai macam masalah yang bahkan sebuah komputer super tidak dapat menyelesaikannya. Setelah beberapa bulan manusia dilahirkan, manusia telah dapat mengenali wajah orang tuanya dan dapat mulai berbicara. Setahun setelah dilahirkan, manusia telah mempunyai sebuah intuisi pada objek natural, dapat mengikuti berbagai objek dan mengerti maksud dari suatu suara. Ketika mereka sudah masuk pada fase anak – anak, mereka dapat mengerti susunan kata dan memiliki banyak sekali kata dalam kosa katanya.

Membangun sebuah mesin yang memiliki kecerdasan seperti otak manusia tentu tidak mudah, untuk membuat sebuah mesin yang mempunyai kecerdasan buatan kita harus menyelesaikan permasalahan komputasi yang sangat rumit yang sebenarnya terkadang dapat kita selesaikan dengan otak kita sendiri dalam hitungan detik. Untuk melewati permasalahan-permasalahan ini. Kita harus menemukan beberapa cara yang lain untuk memprogram komputer yang telah digunakan dalam dekade ini. Untuk itu telah muncul bidang pengetahuan mengenai kecerdasan buatan yang secara umum disebut deep learning [1].

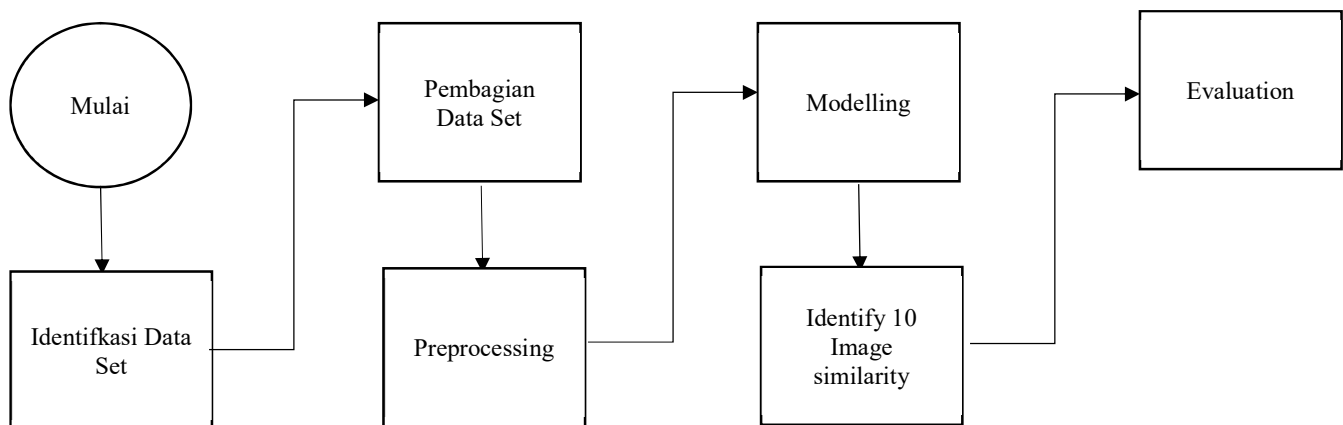
Pada dewasa ini Artificial Intelligence telah berkembang dengan sangat pesat. AI telah digunakan di banyak wilayah penelitian. Pada wilayah penalaran komputer *Content-Based Image Retrieval* (CBIR) telah dikembangkan pada berbagai skema multi-level dari fitur level rendah hingga fitur dengan level yang tinggi. *Convolutional Neural Network* (CNN) telah berhasil digunakan sebagai fitur penjelas yang efisien dan menghasilkan hasil yang akurat. Secara umum, fitur yang didapat dari metode *deep learning* dilatih dengan menirukan persepsi manusia melewati operasi yang bervariasi seperti konvolusi dan penyatuan. *Deep learning* telah menjadi sebuah fitur penjelas yang lebih baik daripada fitur – fitur dengan level yang rendah. Meskipun sekarang model CNN telah menjadi seni dalam penalaran komputer, hal ini tidak menjamin fitur yang didapat dari level tertinggi selalu mendapatkan performa terbaik [2].

Eksperimen yang telah dilakukan pada dataset yang sama berjudul Face Mask Detection Using Convolutional Neural Network yang dilakukan oleh Dr.Beena Ullala Mata, Bhavya, Ashitha dari B.M.S College of engineering yang menggunakan metode CNN. Hasil yang didapat telah mencapai level sistem terbaik berdasarkan fitur lokal dengan pencocokan spasial dan perluasan kueri sehingga memory yang hilang menjadi lebih sedikit [3].

Dataset yang digunakan dalam penelitian yang kami gunakan adalah Face Mask Detection yang didapat dari kaggle. Dataset ini merupakan data citra berjumlah 11800 yang didapat dari website kaggle yang datanya sendiri dikumpulkan dari beberapa dataset yang tersebar di internet. Dataset ini dibagi menjadi 2 kelas yang berisi gambar wajah yang menggunakan masker dengan baik, wajah yang tidak menggunakan masker.

Sasaran dari dilakukannya penelitian ini adalah untuk mengimplementasikan metode convolutional neural network pada sistem pengenalan wajah yang menggunakan masker berbasis pengumpulan gambar. Pada tugas pengumpulan gambar, didukung oleh klasifikasi dengan metode CNN maka pada fitur yang telah diekstrak yang dalam hal ini gambar yang telah dikumpulkan nantinya akan dilakukan klasifikasi.

2. Research Method



Gambar 1. Alur Penelitian

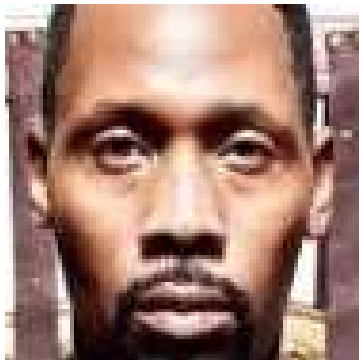
Pada saat ini perkembangan computer vision sudah sangat cepat, banyaknya sistem yang mampu membedakan gambar atau data agar dapat diolah dan menjadi sebuah hasil evaluasi di berbagai aspek bidang seperti Kesehatan, bahan pakan dan lain sebagainya. Salah satu metode yang sering digunakan adalah penggunaan CNN. Convolutional Neural Network sudah sangat terkenal dalam beberapa penggunaan sistem yang digunakan oleh para ilmuwan. Dinamakan Convolutional Neural Network dikarenakan sistem ini sama

seperti jaringan saraf manusia yang saling terhubung satu sama lain untuk menggerakkan tubuh manusia. Di Dalam metode CNN ini terdapat fungsi operasional matematika yang mempelajari keterkaitan hubungan satu dengan lainnya dengan cara perkalian matrik umum di tiap lapisannya.

Jaringan saraf ini menggunakan metode konvolusi dalam mengekstraksi nilai atau volume di setiap lapisannya. Ada 3 lapisan utama yang wajib ada dalam penggunaan CNN yaitu *convolutional layer*, *pooling layer* dan *fully connected layer*. Dalam penggunaannya CNN sebagai algoritma pembelajaran mesin yang dapat mengambil gambar input, menetapkan suatu bobot atau nilai yang dapat dipelajari ke berbagai objek dalam gambar dan dapat membedakan satu dari yang lain. Di sini dataset gambar yang digunakan mengalami banyak lapisan konvolusi dan penyatuan yang menghasilkan gambar yang rata dan kemudian lapisan yang diratakan ini terhubung sepenuhnya untuk mendapatkan *softmax*. Fitur dipelajari atau diidentifikasi selama lapisan konvolusi dan pada tahap akhirnya gambar akan dideteksi menjadi satu berdasarkan kemiripan melalui *feature extraction* atau gambar yang sudah diproses sebagai pembandingnya. Tahapan sistematis yang dilakukan pada proses penelitian kali ini diharapkan mampu untuk menyelesaikan permasalahan dalam mendeteksi dan melihat kemiripan suatu gambar berdasarkan *query* gambar yang di masukkan dengan harapan agar perkembangan dalam mendeteksi suatu objek dapat berjalan dengan cepat dan tepat dalam waktu yang singkat.

A. Dataset

Dataset yang digunakan pada penelitian kali didapatkan dari website Kaggle.com dengan nama “Face Mask Detection ~12K Images Dataset”. Pada data tersebut memiliki 2 kelas utama yaitu “WithMask” dan “WithoutMask” dengan ukuran gambar sebesar 328.92MB. Dikarenakan pada data yang didapatkan pada Kaggle sudah dibagi menjadi data latih data validasi dan data tes maka peneliti tidak perlu untuk membagi data tersebut. Untuk data latih memiliki jumlah data sebesar 10.000, untuk data validasi memiliki data sebesar 800, dan untuk data tes memiliki jumlah data sebesar 992.



Gambar 3. WithoutMask

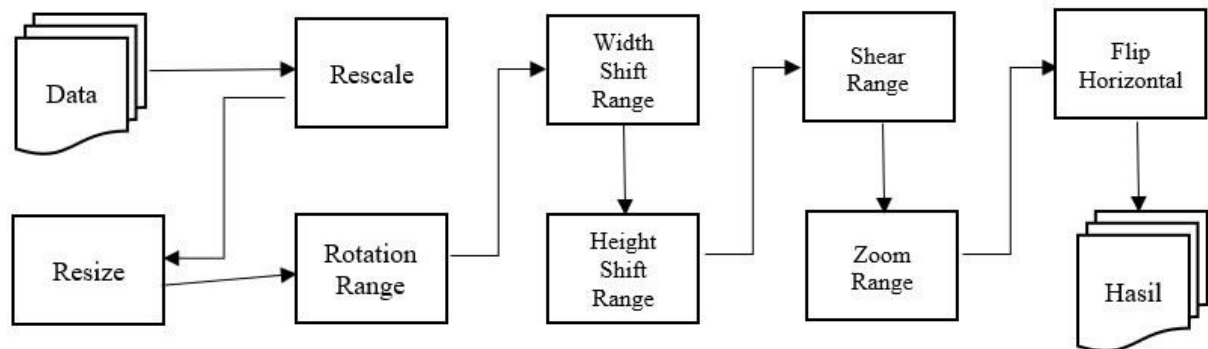


Gambar 2. WithMask

B. Preprocessing

Untuk Data dengan jumlah 10.000 yang terbagi di masing masing data latih, data validasi dan data testing akan mengalami perubahan dengan cara menggunakan tahap preprocessing yang didapatkan dari library keras dengan tujuan data akan disalin ke data augmentasi dan memanipulasi beberapa gambar dengan ketentuan seperti mengubah ukuran gambar, mengubah arah gambar ke samping kanan memperbesar gambar dan lain sebagainya seperti pada contoh gambar dibawah. Dengan cara mengubah ini maka data yang dilatih

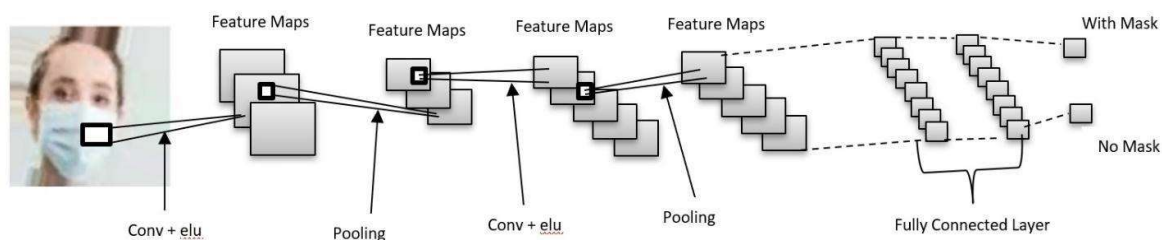
dengan data yang diuji akan mengalami keseimbangan dalam melihat struktur gambar binary yang akan memudahkan data yang akan dilatih.



Gambar 3. Tahap Preprocessing

C. Modelling

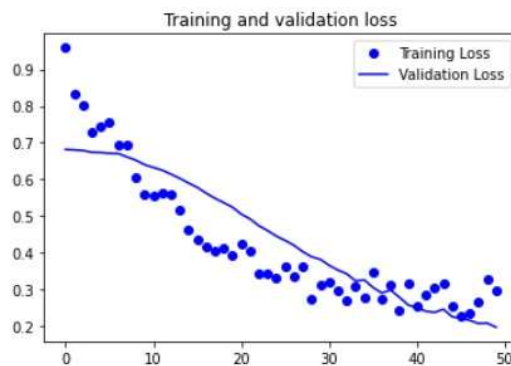
Data yang sudah masuk kedalam proses generator, selanjutnya akan masuk ke dalam proses *modelling*. Proses ini dimaksudkan untuk data yang telah masuk ke dalam direktori latih (*train*) akan di uji dengan direktori uji (*validation*) melalui beberapa proses. Proses yang di maksud adalah pembentukan tiap lapisan atau layer dengan input gambar yang kita inginkan, dengan beberapa feature extraction seperti convolutional, pooling atau juga dapat menggunakan metode pretrained training yang disediakan oleh keras serta fully connected layer. Setiap lapisan yang kita buat harus memiliki keseimbangan antara satu layer dengan layer lainnya agar hasil model yang keluar memiliki akurasi maksimal dan ketepatannya model yang dapat diuji dengan gambar aslinya. Terdapat fungsi aktivasi yang akan kita gunakan dalam setiap layer yang kita bangun yaitu fungsi aktivasi “elu”, fungsi aktivasi ini memiliki keunggulan diantaranya setiap data yang dipelajari di masing masing layer lebih maksimal dan dapat meminimalisir lompatan data (bouncing) saat model sedang dibangun. Selanjutnya terdapat lapisan yang dinormalisasikan dengan cara menambahkan *Batch Normalization* pada FCL agar seluruh data yang sudah dipelajari dapat ditempatkan dengan fitur ekstraksi yang sudah dipelajari. Lalu akan membentuk lapisan output dengan jumlah output yang diinginkan yaitu WithMask dan NoMask.



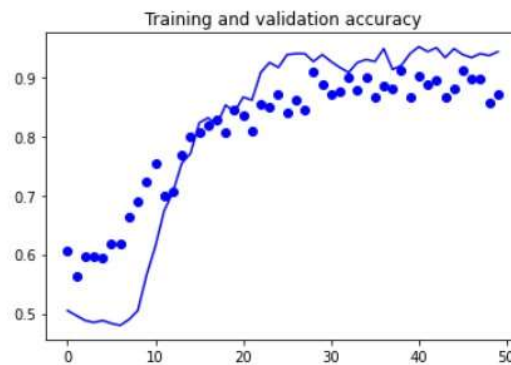
3. Results and Discussion

Selanjutnya data yang sudah kita proses pada arsitektur model akan lanjut ketahap pemrosesan model dengan bantuan template keras yang disediakan oleh tensorflow yaitu penggunaan optimizer SGD yang telah kita atur dengan learning rate sebesar 0.001 dengan output loss binary. Kita definisikan setiap model untuk mempelajari sebanyak 100 dengan mempelajari data di train generator dengan data uji pada validation generator. Output model yang dihasilkan mendapatkan akurasi sebesar 87%.

Model yang sudah terbangun dan dilatih akan disimpan dalam format h5, model ini akan digunakan sebagai uji gambar yang akan kita bangun menggunakan flask sebagai template html yang akan di kombinasikan dengan model yang sudah kita buat. Dengan memanfaatkan metode ini maka program secara langsung akan mampu mendefinisikan gambar yang di input dengan user dan membandingkan kesamaan dan kemiripan suatu gambar dan mengeluarkan hasil prediksi suatu gambar dengan baik. Gambar di bawah ini adalah grafik yang dihasilkan untuk training dan validation loss serta untuk training dan validation accuracy serta classification report :



Gambar 5. Grafik yang diperoleh untuk Training dan Validation Loss



Gambar 6. Grafik yang diperoleh untuk Training dan Validation Accuracy

	precision	recall	f1-score
WithMask	0.96	0.78	0.86
WithoutMask	0.81	0.97	0.88
accuracy			0.87
macro avg	0.89	0.87	0.87
weighted avg	0.89	0.87	0.87

Gambar 7. Classification Report

4. Conclusion

Karena teknologi yang berkembang dengan pesat disertai tren baru yang banyak muncul, ketersediaan dari data wajah yang baru untuk mendeteksi penggunaan masker akan sangat berkontribusi pada perawatan kesehatan masyarakat. dengan bantuan penalaran komputer dan deep learning, kami telah membuat sebuah model. model ini diuji pada gambar yang berbeda. Dengan pengembangan deteksi penggunaan masker, kita dapat mendeteksi jika orang tersebut menggunakan masker wajah dan bisa menyaring mereka ketika akan masuk ke lingkungan masyarakat. implementasi model kami menunjukkan akurasi sekitar 87% yang harapannya dapat ditingkatkan di masa depan menggunakan teknik yang lebih baik dan algoritma yang lebih canggih.

References

- [1] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, "A fast face detection method via convolutional neural network," *Neurocomputing*, vol. 395, no. Bo Li, pp. 128–137, 2020, doi: 10.1016/j.neucom.2018.02.110.
- [2] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Meas. J. Int. Meas. Confed.*, vol. 167, p. 108288, 2021, doi: 10.1016/j.measurement.2020.108288.
- [3] L. Qi, J. Huo, L. Wang, Y. Shi, and Y. Gao, "A mask based deep ranking neural network for person retrieval," *Proc. - IEEE Int. Conf. Multimed. Expo*, vol. 2019-July, pp. 496–501, 2019, doi: 10.1109/ICME.2019.00092.
- [4] T. Mehta and C. Bhensdadia, "Adaptive near duplicate image retrieval using SURF and CNN features," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 5, pp. 104–115, 2019, doi: 10.22266/ijies2019.1031.11.
- [5] Ş. Öztürk, A. Alhudhaif, and K. Polat, "Attention-based end-to-end CNN Framework for Content-Based X-Ray Image Retrieval," *Turkish J. Electr. Eng. Comput. Sci.*, pp. 2680–2693, 2021, doi: 10.3906/elk-2105-242.

- [6] H. Prasetyo and B. A. Putra Akardihya, "Batik image retrieval using convolutional neural network," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 6, pp. 3010–3018, 2019, doi: 10.12928/TELKOMNIKA.v17i6.12701.
- [7] K. R. Kruthika, Rajeswari, and H. D. Maheshappa, "CBIR system using Capsule Networks and 3D CNN for Alzheimer's disease diagnosis," *Informatics Med. Unlocked*, vol. 14, no. November 2018, pp. 59–68, 2019, doi: 10.1016/j.imu.2018.12.001.
- [8] Z. Rian, V. Christanti, and J. Hendryli, "Content-Based Image Retrieval using Convolutional Neural Networks," *Proc. - 2019 IEEE Int. Conf. Signals Syst. ICSigSys 2019*, pp. 1–7, 2019, doi: 10.1109/ICSIGSYS.2019.8811089.
- [9] K. Zhang, T. Li, B. Liu, and Q. Liu, "Co-saliency detection via mask-guided fully convolutional networks with multi-scale label smoothing," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 3090–3099, 2019, doi: 10.1109/CVPR.2019.00321.
- [10] S. K. Sundararajan, B. Sankaragomathi, and D. S. Priya, "Deep Belief CNN Feature Representation Based Content Based Image Retrieval for Medical Images," *J. Med. Syst.*, vol. 43, no. 6, 2019, doi: 10.1007/s10916-019-1305-6.
- [11] K. Kanwal, K. T. Ahmad, R. Khan, A. T. Abbasi, and J. Li, "Deep learning using symmetry, FAST scores, shape-based filtering and spatial mapping integrated with CNN for large scale image retrieval," *Symmetry (Basel)*, vol. 12, no. 4, p. 612, 2020, doi: 10.3390/SYM12040612.
- [12] P. Oberdiek, M. Rottmann, and G. A. Fink, "Detection and retrieval of out-of-distribution objects in semantic segmentation," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2020-June, pp. 1331–1340, 2020, doi: 10.1109/CVPRW50498.2020.00172.
- [13] J. Murray-Bruce and P. L. Dragotti, "Estimating localized sources of diffusion fields using spatiotemporal sensor measurements," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3018–3031, 2015, doi: 10.1109/TSP.2015.2419187.
- [14] L. M. Dang, S. I. Hassan, S. Im, and H. Moon, "Face image manipulation detection based on a convolutional neural network," *Expert Syst. Appl.*, vol. 129, pp. 156–168, 2019, doi: 10.1016/j.eswa.2019.04.005.
- [15] 홍종욱, "3월 1일의 밤은 대한민국의 봄이었다 — 권보드래, □3월 1일의 밤: 폭력의 세기에 꾸는 평화의 꿈□ (돌베개, 2019) —," *Concept Commun.*, vol. null, no. 23, pp. 301–316, 2019, doi: 10.15797/concom.2019..23.009.
- [16] S. Karnila, S. Irianto, and R. Kurniawan, "Face Recognition using Content Based Image Retrieval for Intelligent Security," *Int. J. Adv. Eng. Res. Sci.*, vol. 6, no. 1, pp. 91–98, 2019, doi: 10.22161/ijaers.6.1.13.
- [17] S. K. Vipparthi, S. Murala, A. B. Gonde, and Q. M. J. Wu, "Local directional mask maximum edge patterns for image retrieval and face recognition," *IET Comput. Vis.*, vol. 10, no. 3, pp. 182–192, 2016, doi: 10.1049/iet-cvi.2015.0035.

- [18] Y. Cai, Y. Li, C. Qiu, J. Ma, and X. Gao, "Medical image retrieval based on convolutional neural network and supervised hashing," *IEEE Access*, vol. 7, pp. 51877–51885, 2019, doi: 10.1109/ACCESS.2019.2911630.
- [19] R. Dong, M. Liu, and F. Li, "Multilayer Convolutional Feature Aggregation Algorithm for Image Retrieval," *Math. Probl. Eng.*, vol. 2019, 2019, doi: 10.1155/2019/9794202.
- [20] W. W. Y. Ng, J. Li, X. Tian, H. Wang, S. Kwong, and J. Wallace, "Multi-level supervised hashing with deep features for efficient image retrieval," *Neurocomputing*, vol. 399, pp. 171–182, 2020, doi: 10.1016/j.neucom.2020.02.046.
- [21] N. Chifa, A. Badri, and Y. Ruichek, "Rotation-invariant approach using mask to content-based image retrieval," *ACM Int. Conf. Proceeding Ser.*, vol. Part F148262, pp. 11–14, 2019, doi: 10.1145/3323933.3324066.
- [22] J. Lei, Y. Song, B. Peng, Z. Ma, L. Shao, and Y. Z. Song, "Semi-Heterogeneous Three-Way Joint Embedding Network for Sketch-Based Image Retrieval," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 9, pp. 3226–3237, 2020, doi: 10.1109/TCSVT.2019.2936710.
- [23] Z. Ma *et al.*, "Shoe-Print Image Retrieval with Multi-Part Weighted CNN," *IEEE Access*, vol. 7, no. c, pp. 59728–59736, 2019, doi: 10.1109/ACCESS.2019.2914455.
- [24] H. Y. Lee, H. K. Lee, and Y. H. Ha, "Spatial color descriptor for image retrieval and video segmentation," *IEEE Trans. Multimed.*, vol. 5, no. 3, pp. 358–367, 2003, doi: 10.1109/TMM.2003.814792.
- [25] D. Gao *et al.*, "FashionBERT: Text and Image Matching with Adaptive Loss for Cross-modal Retrieval," *SIGIR 2020 - Proc. 43rd Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 2251–2260, 2020, doi: 10.1145/3397271.3401430.