

BAB II

LANDASAN TEORI

2.1. Ekspresi

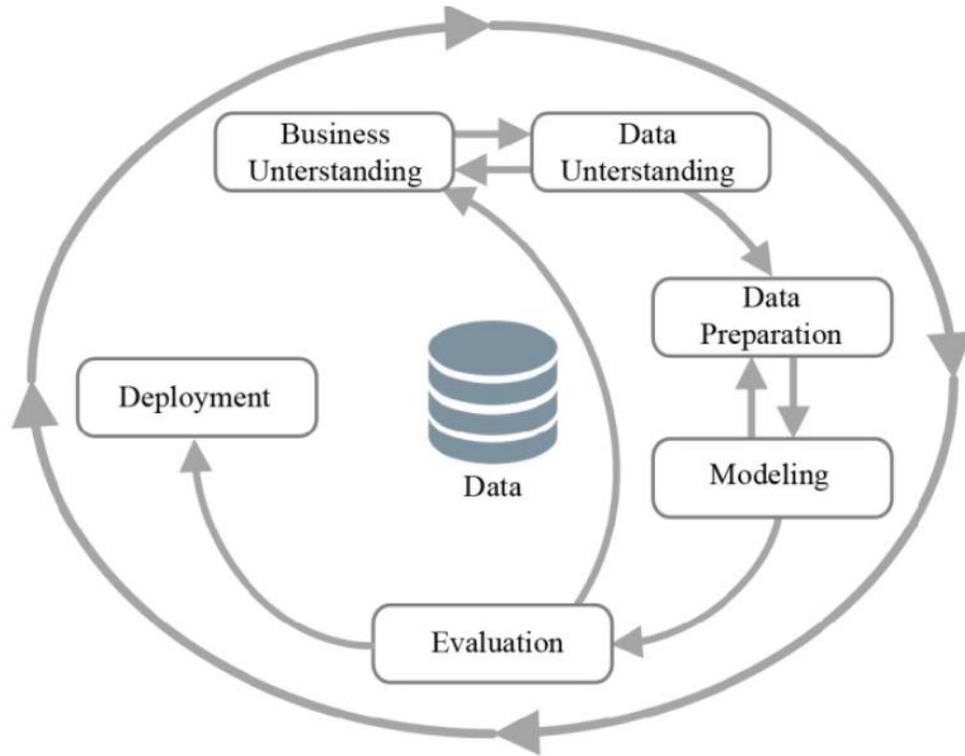
Pada umumnya manusia mengemukakan ekspresi pada wajah mereka sesuai dengan perasaan mereka saat itu dan hal tersebut terjadi pada anak-anak maupun orang dewasa [11]. Ada juga yang mengatakan ekspresi adalah cara manusia untuk mengungkapkan atau proses menyatakan suatu keinginan tertentu dan menghasilkan perubahan bentuk wajah seperti pada bibir, pergerakan alis, perbersaran mata dan lain-lain [12]. Juga dikatakan ekspresi dari wajah dapat diterapkan dalam berbagai bidang seperti kedokteran, pengawasan sistem, *machine learning*, mekanisme robot dan *facial expression* [12]. Tentu manusia sudah sewajarnya mudah mengenali ekspresi dari wajah yang mereka lihat seperti membedakan disaat orang sedang terlihat marah atau senang [12]. Ekspresi manusia dapat menghasilkan emosi positif maupun negatif dan dapat mempengaruhi fisik [13]. Ekspresi seperti takut dan marah sudah diakui sebagai emosi dasar dan diakui banyak kalangan [13]. Ekspresi juga merupakan hal yang penting untuk mengerti tingkah laku dan interaksi manusia [14]. Ekspresi dari seseorang juga mempengaruhi banyak hal dalam perilaku manusia seperti keringat dingin saat cemas, menangis saat sedih, dan pengambilan keputusan pun dapat dipengaruhi [15]. Ada juga dikatakan bahwa

suatu ekspresi tidak hanya mempengaruhi keadaan diri sendiri tapi dapat juga berpengaruh ke orang sekitar [16].

Manusia memiliki berbagai macam ekspresi yang dapat dikenali menjadi 7 tipe utama yang berbeda, yaitu *happiness, sadness, anger, fear, disgust, surprise* dan *neutral* [17]. Ekman juga mengatakan 7 tipe tersebut yaitu *happiness, sadness, anger, fear, disgust, surprise* dan *neutral* merupakan tipe ekspresi utama [18]. Menurut Goleman manusia sudah memiliki ekspresi dasar atau utama yaitu *fear, anger, joy, surprise, disgust, sadness, contempt* semenjak dari lahir serta ekspresi *neutral*[19]. Schindler juga memakai 7 ekspresi utama untuk melakukan penelitian karena sudah terdapat banyak *dataset* yang dapat dipakai untuk mempelajari emosi seseorang [1].

2.2. CRISP-DM

Cross Industry Standard Process for Data Mining atau CRISP-DM merupakan metode yang menyusun projek *Data Mining* dalam fase-fase yang saling terkait, mengubah proses menjadi dalam bentuk iteratif dan interaktif [20].



Gambar 2.1. CRISP-DM

Sumber [21]

Berikut 6 tahapan pada langkah-langkah di metode CRISP-DM seperti pada gambar 2.1:

1. Business Understanding

Business Understanding merupakan tahap untuk menentukan tujuan pada masalah *Data Mining* dengan mengidentifikasi penyebab masalah tersebut dan efek yang akan terjadi jika masalah tersebut berkelanjutan [21].

2. Data Understanding

Data Understanding merupakan tahap pengumpulan data dan sekaligus mengeksplorasi, mendeskripsikan serta memeriksa kualitas data yang digunakan [22]. Tahap ini juga membuat hipotesa mengenai informasi yang belum terlihat berdasarkan pengalaman dan asumsi yang sesuai [21].

3. *Data Preparation*

Data Preparation merupakan tahap untuk pengambilan data yang relevan pada proyek dan mempersiapkannya untuk proses *Data Mining*. Hal tersebut termasuk *preprocessing* data seperti *reduction filtering* dan sekaligus *feature generation* sesuai dengan tujuan pada proyek *Data Mining* [21].

4. *Modelling*

Modelling merupakan tahap konstruksi model *Data Mining* dengan menentukan parameter yang sesuai untuk algoritma yang dipilih dan untuk menjalankan proses *Data Mining* dari data pada tahap *data preparation* [21].

5. *Evaluation*

Evaluation merupakan tahap dimana model yang sudah di *training* di *test* dengan data nyatanya sekaligus melihat hasil dari model tersebut sesuai dengan tujuan bisnis yang ditentukan atau tidak [21].

6. *Deployment*

Deployment merupakan tahap terakhir dimana model yang dibuat akan diimplementasikan dengan *software* atau merupakan hasil dari laporan. Tahap ini juga menjelaskan tahap-tahap untuk melakukan *deployment* untuk model yang dibuat [21].

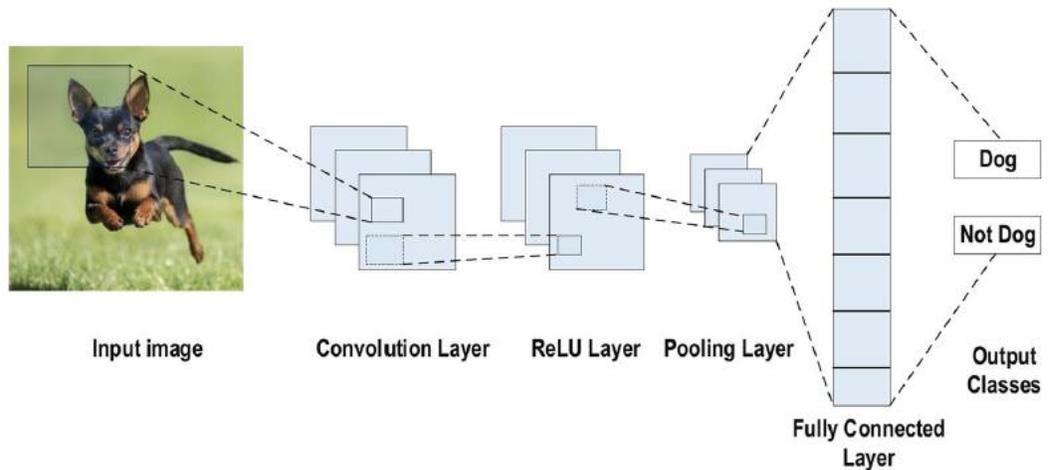
2.3. *Deep Learning*

Deep learning adalah sub bidang dari *machine learning* yang mencoba untuk mempelajari abstraksi *level* tinggi dalam data dengan memanfaatkan arsitektur hirarkis. Ini merupakan pendekatan yang terlihat dan sudah secara luas di terapkan dalam *Traditional Artificial Intelligence Domains*, seperti *Semantic Parsing Transfer Learning*, *Natural Language Processing*, *Computer Vision*, dan banyak lagi. Terdapat 3 alasan utama untuk meledaknya *Deep Learning* pada hari ini yaitu Peningkatan pesat kemampuan *Chip Processing* (Contoh : *GPU units*), Sangat menurunnya harga dari *computing hardware*, dan kemajuan pesat dari algoritma *machine learning*. Berbagai pendekatan *deep learning* sudah secara terus menerus dibahas dan didiskusikan pada beberapa tahun ini. Salah satu nya Schmidhuber yang berpendapat akan pentingnya inspirasi dan kontribusi teknikal dalam format *historical timeline*, sedangkan Bengio memeriksa rintangan - rintangan dari penelitian *deep learning* dan mengusulkan beberapa arah penelitian kedepannya. *Deep network* terlihat sukses untuk tugas *computer vision* karena mereka dapat menyaring fitur yang sesuai selagi melakukan diskriminasi secara bersamaan [23].

2.4. *Convolutional Neural Network*

Pada ruang lingkup dari DL, CNN merupakan algoritma yang paling terkenal dan sering dipakai. Keuntungan utama yang dimiliki CNN dibanding pendahulunya yaitu dapat secara otomatis mengenali fitur yang relevan tanpa perlu campur tangan manusia [23].

CNN sudah sangat sering di diterapkan pada bidang yang berbeda, termasuk *Computer Vision*, *Speech Proccesing*, *Face Recognition*, dan sebagainya. Struktur dari CNNs terinspirasi dari *neuron* yang terdapat pada otak manusia dan binatang, serupa dengan *convolutional neural network*. Secara spesifik, dalam pikiran seekor kucing, sel-sel kompleks yang pola-nya berurutan membentuk *visual cortex* kucing tersebut; Pola sekuensial ini dapat disimulasi dengan CNN. Goodfellow et al. mengidentifikasi tiga manfaat utama dari CNN: representasi yang setara, interaksi yang minim; dan parameter yang dapat diakses bersamaan. Tidak seperti jaringan konvensional yang sepenuhnya terhubung, pembagian bobot dan koneksi lokal pada CNN digunakan untuk memanfaatkan secara sepenuhnya struktur data berdimensi 2D seperti *signal* dari gambar. Teknik ini memanfaatkan jumlah parameter yang sangat sedikit, dimana menjadikan proses *training* menjadi simpel dan mempercepat keseluruhan jaringan [23]. Berikut bentuk arsitektur CNN pada umumnya di gambar 2.2.

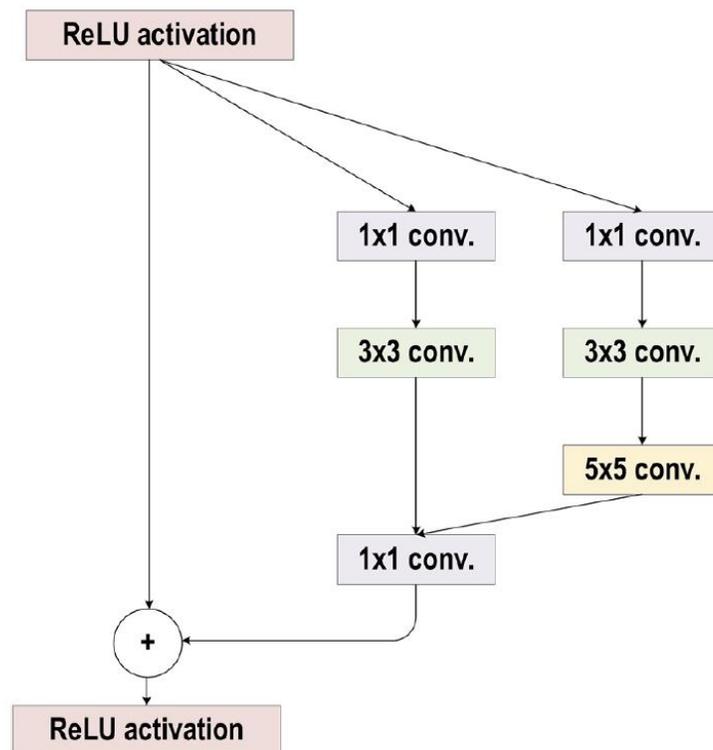


Gambar 2.2. Contoh Arsitektur CNN

Sumber [23]

2.5. DenseNet

DenseNet pertama kali muncul pada tahun 2016 oleh *Facebook AI Research*. DenseNet menggunakan konsep konvolusi penggabungan, dan fungsi aktivasi ULT yang sama untuk bekerja. Fitur yang terpenting pada arsitektur ini adalah jaringan blok nya yang padat. Dalam blok padat ini, setiap lapisan mengambil semua peta fitur sebelumnya sebagai masukan, sehingga membantu proses pelatihan dengan mengurangi masalah gradien-hilang. Masalah gradien-menghilang ini muncul di jaringan yang sangat dalam di mana mereka begitu dalam sehingga ketika kita menyebarkan kesalahan ke jaringan, kesalahan ini berkurang di setiap langkah dan akhirnya menjadi 0 [23]. Berikut penggambaran arsitektur CNN pada gambar 2.3.

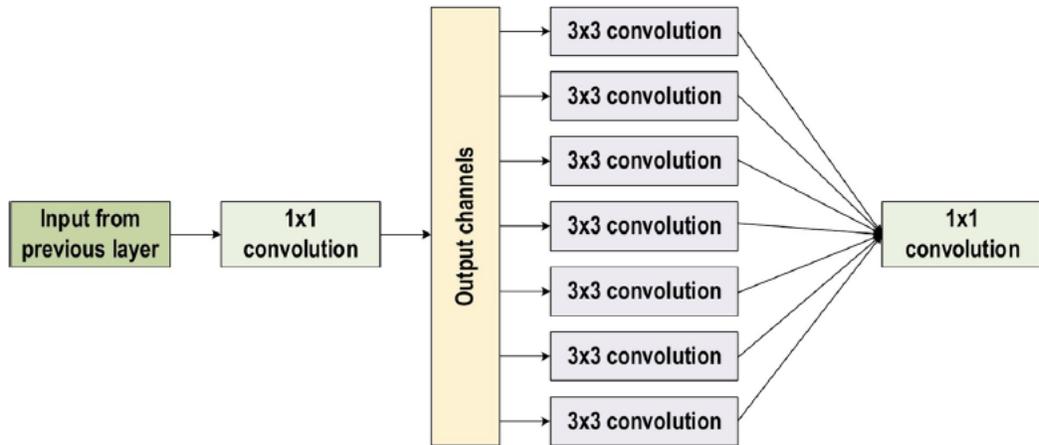


Gambar 2.3. Contoh Arsitektur DenseNet

Sumber [23]

2.6. Xception

Arsitektur *Extreme Inception* merupakan karakteristik utama pada Xception. Tujuan utama Xception sendiri adalah *layer* konvolusi yang dapat dipisahkan secara mendalam. Model Xception menyesuaikan dengan blok awal yang membuat lebih lebar dan dapat menukar satu dimensi (3x3) darj konvolusi 1 x 1 untuk mengurangi kompleksitas komputasi. Jaringan Xception menjadi lebih efektif secara komputasi melalui penggunaan saluran *decoupling* dan korespondensi spasial [23]. Berikut arsitektur tersebut digambarkan pada gambar 2.4.



Gambar 2.4. Contoh Arsitektur Xception

Sumber [23]

2.7. Data Augmentation

Data augmentation merupakan teknik untuk meningkatkan keragaman data *training* dengan menerapkan transformasi gambar secara acak. Data augmentasi ini dilakukan untuk mengekstrak lebih banyak informasi dibanding dengan data aslinya. Berikut teknik yang digunakan untuk melakukan *data augmentation* yaitu *cropping*, *rotation*, *flipping*, *shift*, *shear* dan lainnya. Data augmentasi ini juga digunakan untuk mengurangi *overfitting* pada model [24].

2.8. Penelitian Terdahulu

Tabel 2.1. Penelitian Terdahulu

No.	Nama Jurnal	Jurnal	Tahun	Penulis	Hasil Penelitian
1	2017 Seventh International Conference on Information	<i>An Efficient Real-Time Emotion Detection Using</i>	2017	Binh T. Nguyen, Minh H. Trinh, Tan V. Phan, Hien D. Nguyen	Penelitian ini menggunakan model <i>machine learning</i> dengan model

	<i>Science and Technology (ICIST)</i> . 251-255.	<i>Camera and Facial Landmarks</i>			SVM mendapatkan akurasi terbaik sebesar 70.65% dengan 3 jenis emosi.
2	2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS). 423-427.	<i>Facial Expression Recognition Using Facial Landmarks and Random Forest Classifier</i>	2018	M. I. N. P. Munasinghe	Penelitian ini menggunakan pendekatan <i>machine learning</i> dengan model <i>random forest</i> yang mendapatkan akurasi sebesar 72% dengan 3 jenis emosi.
3	<i>IEEE Transactions on Affective Computing</i> . 1-1.	<i>Deep Facial Expression Recognition: A Survey</i>	2020	Shan Li and Weihong Deng	Penelitian ini mengulas <i>dataset</i> yang digunakan untuk pengenalan ekspresi pada wajah. Salah satunya adalah RAF-DB.
4	ULTIMA Computing, Vol. XI, No. 1, 34-38.	Mengidentifikasi Mood Mahasiswa Berdasarkan Ekspresi Wajah dengan Menggunakan <i>Discrete Wavelet Transform</i> dan Fuzzy K-Nearest Neighbor	2019	Nur Inzani Reski Amalia, Jayanti Yusmah Sari	Penelitian ini meneliti 3 jumlah ekspresi dengan menggunakan metode klasifikasi fuzzy k-nearest neighbour.

Artikel yang berjudul “*An Efficient Real-Time Emotion Detection Using Camera and Facial Landmarks*” yang ditulis oleh Binh T. Nguyen, Minh H. Trinh, Tan V. Phan, Hien D. Nguyen [4] menggunakan model *machine learning* untuk pengenalan ekspresi pada wajah dengan metode nya menggunakan SVM. Selain itu juga model yang dibuat mencapai akurasi 70.65% dengan 3 jumlah klasifikasi ekspresi.

Artikel yang kedua berjudul “*Facial Expression Recognition Using Facial Landmarks and Random Forest Classifier*” dan ditulis oleh M. I. N. P. Munasinghe [8]. Penelitian ini berfokus pada pembuatan model *machine learning* untuk pengenalan ekspresi pada wajah menggunakan algoritma *random forest*. Model yang dibuat mendapatkan hasil akurasi sebesar 72% dengan jumlah ekspresi yang diklasifikasi adalah 3.

Artikel yang ketiga berjudul “*Deep Facial Expression Recognition: A Survey*” yang ditulis oleh Shan Li and Weihong Deng [25]. Penelitian ini membahas *dataset* yang digunakan untuk melakukan training pada model pengenalan ekspresi. Salah satu *dataset* yang dibahas adalah RAF-DB.

Artikel yang keempat berjudul “Mengidentifikasi *Mood* Mahasiswa Berdasarkan Ekspresi Wajah dengan Menggunakan *Discrete Wavelet Transform* dan *Fuzzy K-Nearest Neighbor*” [26]. Penelitian ini berfokus dalam pembuatan program untuk mengidentifikasi *mood* mahasiswa berdasarkan ekspresi wajah menggunakan algoritma *Fuzzy K-Nearest Neighbor* dan hanya mengklasifikasi 3 jenis emosi yaitu marah, senang, dan sedih.

Berdasarkan penelitian terdahulu yang ada pada tabel 2.1, banyak penelitian yang meneliti pembuatan pengenalan ekspresi menggunakan model *machine learning* selain itu juga jenis ekspresi yang diklasifikasi hanya 3 jenis ekspresi. Sedangkan pada penelitian ini, model yang ingin dikembangkan menggunakan pendekatan *deep learning* dan mengklasifikasi 7 jenis ekspresi. *Dataset* yang digunakan juga merupakan *dataset* pada penelitian sebelumnya yaitu RAF-DB.