

## BAB II

### LANDASAN TEORI

#### 2.1. Tinjauan Pustaka

Pada tahapan ini penulis akan melakukan tinjauan pustka yang telah dilakukan dalam penerapan Deep Learning menggunakan metode *Convolution Neural Network* sehingga, dalam penelitian ini diperlukan tinjauan pustaka sebagai alat dalam penerapan metode ini, agar dapat menghindari pembuatan ulang, mengidentifikasi kesenjangan, mengetahui metode yang sudah diterapkan, mengetahui penelitian yang sama dibidang ini, serta melanjutkan untuk penelitian sebelumnya.

No	Nama Peneliti	Tahun	Judul dan Hasil
1	Farica dan Adhi	2018	Pengenalan Tulisan Tangan Offline Dengan Algoritma Generalized Hough Transform dan Backpropagation. Hasil nilai akurasi dari algoritma generalized Hough transform dan backpropagation sebesar 80%.
2	Teny Handhayani	2017	Pengenalan pola tulisan tangan dengan algoritma support vector mesin (SVM) dengan kernel linear mendapatkan hasil akurasi sebesar 93.6%
3	Riska Yulianti	2019	Pengenalan Pola Tulisan Tangan Suku Kata Aksara Sasak Menggunakan Metode Moment Invariant dan Support

			Vector Machine. Di uji dengan fitur berjumlah 112 fitur mendapatkan akurasi 92.52%.
4	Siwi Prihatiningsih	2019	Analisis performa untuk pengenalan pola tulisan tangan Angka berdasarkan jumlah iterasi dengan algoritma CNN. 98,67% pada validasi, dan 98,99% pada testing
5	Nahila Khunafa Qudsi	2020	Identifikasi Citra Tulisan Tangan Digital Menggunakan Convolutional Neural Network (CNN). citra koresponden yaitu sebesar 98.21 %
6	Christopher Albert Lorentius	2019	Klasifikasi Aksara Tulisan Jawa dengan Menggunakan Metode Cnn = pada epoch ke 75, dengan akurasi mencapai 97,52%.
	Chandra Kusuma Dewa	2018	Convolutional Neural Networks for Handwritten Javanese Character Recognition. CNN model for the handwritten Javanese character dataset cannot reach 90% in all folds
8	Herman	2018	PENGENALAN ANGKA tulisan tangan menggunakan jaringan syaraf tiruan. Hasil pengujian pada data uji

			didapatkan tingkat akurasi sebesar 11,67% dari total data uji
9	Nisa Amalia	2020	pengenalan aksara sunda menggunakan jaringan saraf tiruan backpropagation dan deteksi tepi canny. dilakukan menggunakan data latih terhadap 70 citra menghasilkan nilai akurasi rata-rata sebesar 90%
10	Eliza Hara	2016	Penggunaan Deteksi Tepi (Canny) pada Sistem Pengenalan Tulisan Tangan Aksara Lampung Berbasis Jaringan Syaraf Tiruan. Sistem aplikasi pengenalan aksara Lampung berhasil mengenali per-karakter aksara Lampung sebesar 78% dari 10 (sepuluh) kali pengujian.
11	TANTI OKTAVIANI	2018	transformasi fourier untuk pengenalan tulisan tangan aksara lampung. menghasilkan akurasi sebesar 72,11%.
12	Adhika dan Rinaldi		Pengenalan Aksara Lampung dengan Jaringan Syaraf Tiruan.. Akurasi terhadap data latih sebesar 100% dan

			mendapatkan akurasi untuk data uji yaitu 90.8%.
--	--	--	---

Tabel 2. 1 Tinjauan Pustaka

## 2.2. Teori Yang Digunakan Pada Sistem

### 2.2.1. Pengolahan Citra Digital

Digital Image atau Pengolahan Citra (Digital Image Processing) adalah suatu disiplin ilmu yang mempelajari teknik-teknik untuk mengolah suatu citra. Citra yang dimaksud dalam penelitian ini adalah citra diam yang berasal dari sensor penglihatan berupa kamera. Secara matematis citra berarti fungsi kontinu menggunakan intensitas cahaya pada bidang 2 dimensi. sehingga dapat diproses oleh komputer pribadi digital sebuah gambar dapat direpresentasikan secara digital menggunakan nilai-nilai diskrit. (RD. Kusumanto, 2011)

### 2.2.2. Jenis Citra Digital

Pengolahan citra digital Secara umum citra digital dapat dibagi. 3, gambar berwarna, gambar hitam putih dan gambar biner:

#### 1. Gambar RGB (merah, hijau, biru).

RGB ialah suatu model warna yang terdiri atas 3 buah warna: merah (Red), hijau (Green), serta biru (Blue), yg ditambahkan menggunakan bermacam-macam cara buat membentuk warna. Kegunaan primer model warna RGB ialah untuk menampilkan citra / gambar pada perangkat elektro, seperti televisi dan computer.

#### 2. Hitam dan putih.

Gambar digital dan hitam putih (skala abu-abu) Setiap piksel memiliki transisi warna dari putih ke hitam. Rentang tadi berarti bahwa setiap piksel bisa diwakili oleh 8 bit, atau 1 byte. Rentang warna di black and white sangat cocok dipergunakan untuk pengolahan arsip gambar. salah satu bentuk manfaatnya digunakan dalam kedokteran (X-ray). Black and white sebenarnya adalah hasil rata-rata dari color image.

#### 3. Binary Image.

Setiap piksel hanya terdiri dari warna hitam atau putih, karena hanya ada 2 warna untuk setiap piksel, maka hanya perlu 1 bit per piksel (0 dan 1) atau Bila di 8 bit ( 0 dan 255), sehingga sangat efisien untuk hal penyimpanan. Gambar yang direpresentasikan memakai biner sangat cocok untuk teks (dicetak atau tulisan tangan), sidik jari (finger print), atau gambar arsitektur.

### **2.2.3. Computer Vision**

Computer vision merupakan sebuah kemampuan sebuah computer yang ke desain agar bisa melihat sebuah object sehingga mampu menampilkan objek digital dan dapat mengoleksi data secara visual personal komputer bisa melakukan beberapa pekerjaan yang tidak mampu dilakukan oleh manusia (Purno & Wibowo, 2016)

### **2.2.4. Artificial intelligence**

Kecerdasan uatan adalah salah satu idang komputasi yang didefinisikan seagai kecerdasan yang diciptakan untuk suatu sistem sehingga sistem terseut mampu erpikir seperti manusia. Kecerdasan uatan erkaitan dengan agaimana komputer melakukan tindakan logis yang dapat diklasifikasikan seagai cerdas. (Prasetyo & Retnowati, 2015)

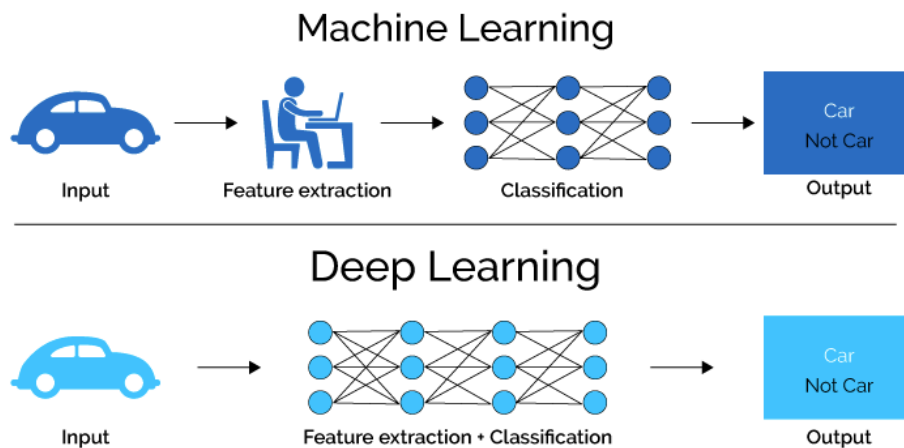
### **2.2.5. Mechine Learning**

Pemelajaran mesin atau mechine learning adalah pemelajaran yang sangat memantu dalam pemecahan masalah memuat segalanya leih mudah. Di industri rumah sakit atau kesehatan machine learning dapat dengan mudah dilakukan misalnya agi seorang dokter untuk dapat mendiagnosis penyakit jantung dengan cepat dalam waktu singkat. Dengan erita yang erkemang tentang pemelajaran mesin itu menjadi mesin yang mampu elajar sendiri tanpa perlu dikendalikan oleh pengguna individu. (Telaumbanua et al., 2020)

### **2.2.6. Deep Learning**

Pembelajaran Mendalam atau sering disebut Pembelajaran Terorganisir Mendalam atau Pembelajaran Progresif adalah bagian dari AI yang terdiri dari menangani pertimbangan tingkat tak terbantahkan yang menunjukkan metode informasi menggunakan sekelompok kapasitas perubahan tidak langsung yang

tersebar di lapisan dan kedalaman. Strategi dan sistem penyelesaian dalam pembelajaran mendalam dapat digunakan baik untuk pembelajaran terkelola, pembelajaran tanpa bantuan dan pembelajaran semi-teratur dalam aplikasi yang berbeda seperti pengenalan gambar, pengenalan suara, pengelompokan teks, dan sebagainya Pembelajaran Mendalam dipandang sebagai mendalam dengan alasan bahwa konstruksi dan jumlah organisasi saraf dalam perhitungan dapat mencapai banyak lapisan. Pembelajaran Mendalam adalah semacam perhitungan organisasi saraf palsu yang menggunakan metadata sebagai informasi dan siklus itu menggunakan berbagai lapisan rahasia dari perubahan informasi tidak langsung untuk memastikan nilai hasil. Perhitungan dalam Pembelajaran Mendalam memiliki unsur yang luar biasa, tepatnya unsur yang dapat memisahkan secara alami. Ini menyiratkan bahwa perhitungannya secara konsekuen dapat menangkap elemen-elemen penting tergantung pada situasi dalam menangani suatu masalah. Perhitungan semacam ini sangat penting dalam kesadaran buatan karena dapat mengurangi masalah pemrograman saat menentukan sorotan cepat. Selanjutnya, perhitungan ini dapat digunakan untuk menangani hal-hal yang memerlukan pengawasan (diarahkan), solo (tanpa bantuan), dan semi-teratur (semi-managed). (Lorentius et al., 2020)

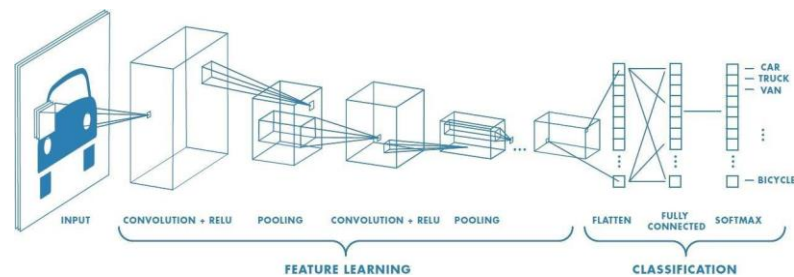


Gambar 2. 1 Mechine Learning & Deep Learning

Sumber : <https://warstek.com/deepmachinelearning>

### 2.2.7. Convolutional Neural Network

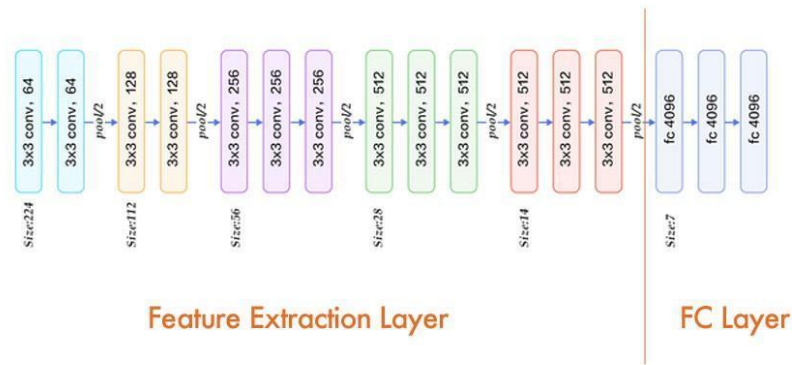
Convolutional Neural Organization (CNN) adalah sejenis organisasi saraf yang biasanya digunakan dalam informasi gambar. CNN dapat dimanfaatkan untuk membedakan dan melihat objek dalam sebuah gambar. Pada umumnya, tidak sepenuhnya berbeda dari organisasi saraf standar. CNN terdiri dari neuron yang memiliki bobot, predisposisi dan kapasitas inisiasi. Siklus yang diselesaikan oleh Convolutional Neural Organization adalah: Convolutional Layer, Non-Linearity Layer (ReLU Layer), Pooling Layer, dan Completely Associated Layer. Rekayasa Convolutional Neural Organization (CNN). (Lorentius et al., 2020) dapat dilihat pada Gambar 2.2



Gambar 2. 2 Arsitektur CNN

### 2.2.8. Arsitektur VGG16

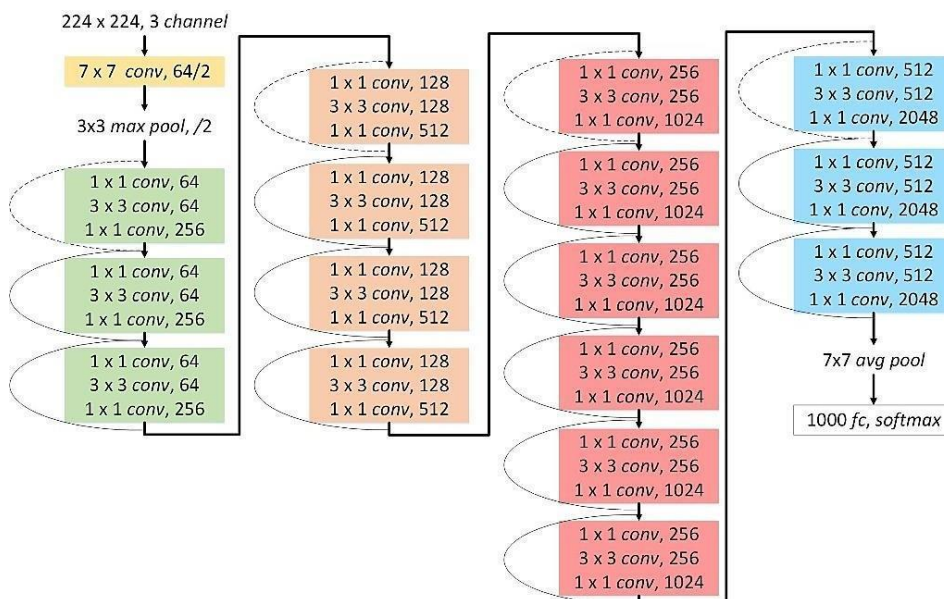
Arsitektur VGG terdiri dari convolutional lock yang terdiri dari rectified linear unit (ReLU) dan maxpolling seagai link yang kemudian erakhir pada full connected layer dengan 1000 layer dalam hal ini fungsi yang digunakan adalah softmax. Fungsi dari ReLU dan maxpolling pada arsitektur ini adalah untuk memperkecil ukuran lok proses konvolusi hanya dengan menggunakan satu langkah dan padding. (Sandhopi et al., 2020) Arsitektur VGG16 bisa dilihat pada gambar 2.3



Gambar 2. 3 Arsitektur dari VGG16

### 2.2.9. Arsitektur Resnet50

ResNet: ResNet adalah jaringan saraf tiruan yang diangun di atas struktur sel piramidal di korteks sereral. Jaringan ekerja dengan menggunakan koneksi yang dilewati atau pintasan untuk melewati anyak lapisan. Pada lapisan pertama ukuran kernel adalah  $7 \times 7$  dengan 6 filter kemudian ditransfer dengan ukuran kernel maksimum  $3 \times 3$ . Kelompok lapisan pertama terdiri dari tiga lok diikuti oleh empat enam dan tiga lok. Di antara lok-lok ini ada skip koneksi antar lapisan. Arsitektur Resnet50 dapat di lihat pada Gambar 2.4



Gambar 2. 4 Arsitektur ResNet-50

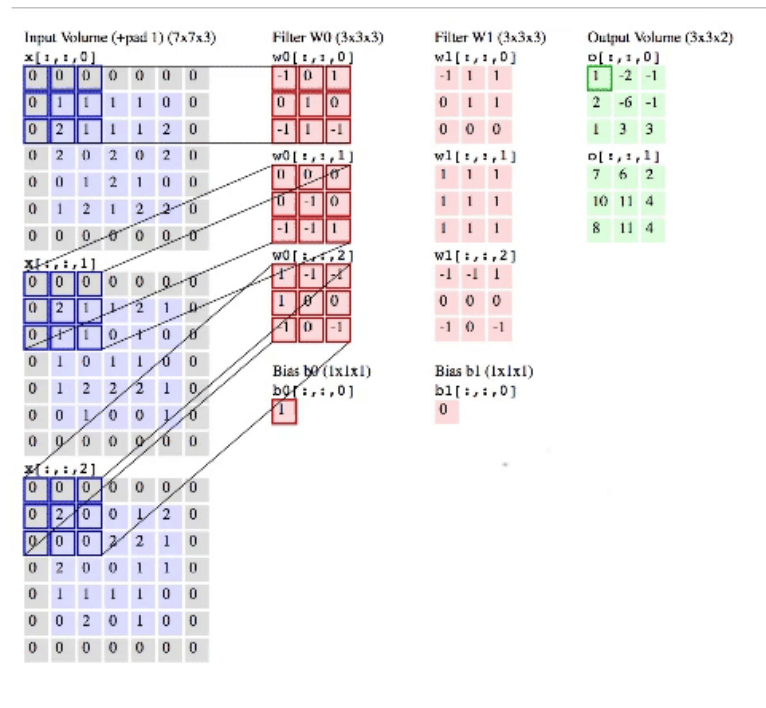


### 2.2.10. Feature Extraction Layer

Feature Extraction ialah proses untuk melakukan encoding, dari sebuah citra atau image yang akan menjadi feature berupa angka, yang akan menjelaskan gambar tersebut, pada CNN terdapat dua bagian feature extraction yang ada di Convolutional Layer dan Pooling Layer.

### 2.2.11. Convolutional Layer

Convolutional Layer adalah layer yang terdiri dari berbagai neuron yang telah tersusun sehingga membentuk suatu filter dengan tinggi dan Panjang pixel. Contoh nya Layer pertama pada feature extraction ialah layer convolutional dengan ukuran layer  $3 \times 3 \times 3$ . Panjang layer 3, tinggi layer 3, dan jumlah channel atau ketebalan layer berjumlah 3 tergantung jumlah channel pada gambar input. Cara kerjanya ialah filter berukuran  $3 \times 3$  tadi akan di geser ke seluruh bagian gambar input, setiap penggeseran yang dilakukan bergantung pada jumlah stride, jika stride berjumlah 2 maka filter akan bergeser sebanyak 2 kolom, kemudian akan menghasilkan sebuah output yang bisa disebut activation map, untuk proses Convolutional Layer dapat dilihat pada gambar dibawah:



Gambar 2. 5 Contoh Proses Convolutional Layer

### 2.2.12. Convolution 1D 2D 3D

Terdapat 3 jenis convolution yang dimana dapat digunakan tergantung dataset yang akan diolah, sebagai contoh :

1. Convolution 1D digunakan untuk input yang berukuran Panjang x jumlah channel Panjang dapat di presentasikan sebagai waktu atau ukuran lainnya, misalnya sinyal sepanjang 4 detik,
2. Convolution 2D biasanya dipakai untuk input yang memiliki Lebar x Tinggi x Jumlah Channel, Contohnya ialah input gambar yang memiliki tinggi gambar, lebar gambar, dan 3 channel RGB.
3. Convolution 3D digunakan untuk inputan berukuran lebar x tinggi x waktu x jumlah channel, biasanya terdapat dari sebuah video.

### 2.2.13. Stride

Stride ialah parameter atau yang mengatur berapa banyak penggeseran filter, Jika stride bernilai 2 maka filter akan bergeser sebanyak 2 kotak secara Horizontal maupun secara Vertikal, ukuran stride juga akan mempengaruhi proses pengambilan informasi dari gambar, semakin kecil langkah stride contohnya 1 stride maka akan menghasilkan informasi yang banyak, sedangkan semakin besar ukuran stride maka informasi yang didapat akan semakin sedikit karena filter akan melewati nilai-nilai pada sebuah gambar, meskipun menggunakan stride dengan ukuran kecil tidak selalu mendapatkan hasil performa yang baik.

### 2.2.14. Padding

Padding atau bisa disebut zero padding ialah parameter yang mengatur dan menentukan jumlah pixel yang akan diberi nilai 0, yang akan ditambahkan di setiap sudut pada gambar input, Hal ini bertujuan untuk memanipulasi input dari convolutional layer, dengan menggunakan zero padding, kita bisa mendapatkan

ukuran dimensi output yang sama dengan input sehingga dapat memperbanyak feature yang akan di extrac dan dapat meningkatkan performa karena convolutional layer akan berfokus di infomasi yang berada di tengah zero padding tersebut, pada gambar di atas, dimensi dari input adalah 5x5, jika dilakukan convolutional layer maka akan mendapatkan ukuran 2x2 jika menggunakan filter 3x3, namun dengan menambahkan zero padding, maka hasil extraction berukuran 3x3 sehingga lebih banyak infomasi yang didapat. Untuk menghitung dimensi maka bisa menggunakan rumus sebagai berikut.

$$output = \frac{W - N + 2P}{S} + 1$$

W = Panjang/Tinggi Input

N = Panjang/Tinggi Filter

P = Zero Padding

S = Stride

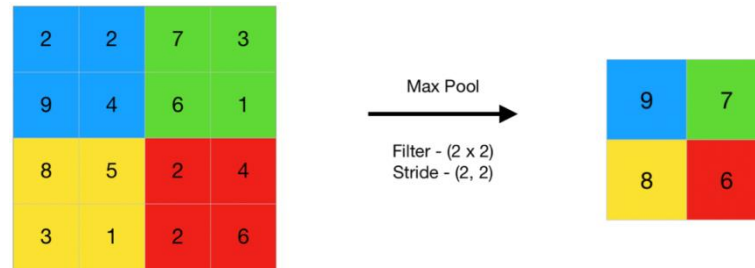
### 2.2.15. Fungsi Aktivasi

Pada tahapan ini, nilai yang dihasilkan convolutional layer akan digunakan fungsi aktivasi. Ada beberapa fungsi aktifasi yang umum digunakan pana CNN, salah satunya adalah reLu, cara kerj dari reLu ialah akan mengubah nilai dari konvolusi menjadi 0 sebagai contoh pada proses convolution mendapatkan nilai -1, -0,2 maka tugas relu mengubah nilai tersebut menjadi 0.

### 2.2.16. Pooling Layer

Pooling layer menerima output dari convolutional layer di mana ukuran data gamar akan dikurangi. Pada prinsipnya pooling layer terdiri dari filter dengan ukuran tertentu dan stridestepnya kemudian erpindah ke seluruh area feature map. Seagian besar dalam arsitektur CNN metode pooling yang digunakan adalah Max pooling. Agregasi maksimum memagi lapisan akumulasi output menjadi beberapa mesh dan kemudian setiap filter offset mendapatkan nilai maksimum dari setiap mesh. Tergantung pada panjang langkah yang dihasilkan adalah seagian kecil dari

ukuran aslinya yang berguna untuk mengurangi ukuran data sehingga mengurangi jumlah parameter. Proses Pooling layer dapat di lihat pada gambar 2.2 (Nurfita & Ariyanto, 2018)



Gambar 2. 6 Max Pooling Layer

Sumber: <https://www.geeksforgeeks.org>

### 2.2.17. Fully Connected Layer

Lapisan yang terhuung penuh menerima masukan dari lapisan kelompok keluaran seagai peta fitur. Fear map masih dalam entuk array multidimensi sehingga akan mementuk kemali feature map dan menghasilkan vektor n-dimensi dimana n adalah jumlah output layer yang harus dipilih program. Misalnya kelas terdiri dari 500 neuron maka softmax akan diterapkan untuk mengemalikan daftar proailitas maksimum untuk masing-masing dari 10 lael kelas menjadi pengklasifikasi akhir jaringan. (Nurfita & Ariyanto, 2018) Fully connected layer merupakan layer terakhir dimana data erupa matriks x-dimensi diuah menjadi matriks linier (1-dimensi) sehingga klasifikasi dapat dilakukan dengan leih mudah. (Lorentius et al., 2020).

### 2.2.18. Dropout

Dropout merupakan upaya untuk menghindari elajar terlalu anyak dan untuk mempercepat proses elajar. Overfitting merupakan kondisi dimana hampir semua data yang telah mengalami proses training memiliki rate yang aik namun terdapat cacat pada proses prediksinya. Dalam sistem kerjanya Dropout menghilangkan sementara neuron erupa lapisan tersemunyi atau terlihat dalam jaringan. (Ilahiyah & Nilogiri, 2018).

### 2.2.19. Supervised Learning & Unsupervised Learning

Supervised learning adalah jenis algoritma pembelajaran mesin yang menggunakan kumpulan data yang diketahui (dataset pelatihan) untuk memuat prediksi. Metode ini pertama-tama memutuskan kumpulan data erlael. (Buslim 2019) Pembelajaran tanpa pengawasan adalah metode yang tidak memerlukan pelatihan. dan tidak ada guru. Guru di sini adalah lael kelas dari data jadi dalam pembelajaran tanpa pengawasan unsupervised learning kita harus mengelompokkan data ke dalam beberapa kelas. (Budi, 2018).

### 2.3. Teori Pendukung Lainnya

#### 2.3.1. Aksara Lampung

Aksara Lampung memiliki 20 aksara dan 12 himpunan sub-karakter. Di karakter utama akan menelurkan berbagai jenis mode membaca. Total kombinasi bacaan aksara Lampung adalah 560 suku kata. (Aryantio & Munir, 2015)



Gambar 2. 7 Aksara Lampung

Sumber: [https://id.wikipedia.org/wiki/Aksara\\_Lampung](https://id.wikipedia.org/wiki/Aksara_Lampung)

#### 2.3.2. Pengenalan Pola

Sebagai aturan, pengakuan desain (design acknowledgment) adalah ilmu untuk mengkarakterisasi atau menggambarkan sesuatu yang bergantung pada estimasi kuantitatif highlight (atribut) atau properti utama dari sebuah artikel. Contoh sebenarnya adalah unsur yang bercirikan dan dapat diinstruksikan serta diberi nama. Sidik jari adalah ilustrasi contoh. Contohnya dapat berupa bermacam-

macam estimasi atau hasil pengecekan dan dapat dikomunikasikan dalam dokumentasi vektor atau kisi (Putra, 2010). Pengakuan desain adalah metode yang terlibat dengan pengumpulan informasi matematika dan simbol termasuk gambar secara alami oleh PC, alasan pengumpulan adalah untuk melihat item dalam gambar. (Widodo, etal., 2006).

### **2.3.3. Tulisan Tangan**

Tulisan tangan adalah fitur iometrik karena setiap orang memiliki gaya tulisan tangan yang unik. Keunikan ini dapat digunakan sebagai identifikasi iometrik dapat digunakan sebagai karakteristik iometrik yang berguna untuk eragai keperluan seperti sistem keamanan pada perangkat teknis nomor dan melacak identitas seseorang. Fitur iometrik yang digunakan dalam sistem keamanan memiliki tingkat keamanan yang tinggi karena bersifat unik bagi setiap orang. (Handhayani, 2017)

### **2.3.4. Python**

Python merupakan ahasa pemrograman lintas platform yang banyak digunakan pada aplikasi teknologi saat ini dan yang akan datang khususnya di idang IoT. (Kadarina dan Inu Fajar 2019)

Bahasa Pemrograman Python merupakan ahasa pemrograman populer yang memiliki keunggulan sebagai erikut:

1. Kemudahan penggunaan untuk mengemangkan perangkat lunak perangkat keras Internet of Things aplikasi aplikasi we dan video game.
2. Selain mudah diaca sehingga kodenya mudah dipahami ahasa pemrograman ini memiliki lirary yang sangat esar dan sangat kaya.
3. Ini adalah ahasa dukungan yang sangat aik untuk ekosistem Internet of Things.

### **2.3.5. Keras**

Keras adalah API jaringan saraf tingkat tinggi yang ditulis dalam ahasa pemrograman Python dan mampu erjalan di TensorFlow CNTK dan Theano. Keras dikemangkan dengan tujuan untuk mempercepat penelitian. Keras juga merupakan perpustakaan open source. Keras erfokus pada kemampuan yang ramah pengguna modular dan dapat diperluas. (Lorentius et al., 2020)

### **2.3.6. Cropping**

Cropping adalah proses memperkecil ukuran citra dengan cara mengkrop citra ke koordinat tertentu pada area citra. Untuk proses cropping memotong eberapa objek di area gambar untuk mendapatkan bagian tertentu dari gambar dengan ukuran tertentu. (Saifullah et al., 2016)

### **2.3.7. Ekstraksi Fitur**

Ekstraksi objek adalah proses pengindeksan dataase gambar beserta isinya. Secara matematis setiap ekstraksi objek adalah pengkodean vektor n-dimensi yang disebut vektor fitur. Komponen vektor fitur dihitung dengan analisis gambar dan teknik pemrosesan dan digunakan untuk membandingkan satu gambar dengan gambar lainnya. Ekstraksi ciri diklasifikasikan menjadi 3 jenis yaitu level rendah level sedang dan level tinggi. Level rendah adalah ekstraksi fitur berdasarkan konten visual seperti warna dan tekstur level sedang adalah ekstraksi berdasarkan wilayah citra yang ditentukan oleh segmentasi sedangkan level tinggi adalah ekstraksi fitur skor berdasarkan informasi semantik yang terkandung dalam citra. (Sugiartha et al., 2016)

### **2.3.7. Grayscale**

Citra grayscale merupakan citra grayscale dengan nilai intensitas tertinggi 255 dari putih menjadi hitam dengan nilai intensitas terendah. Skala ini digunakan dalam skala au-au. Citra digital dengan skala au-au 8-bit memiliki  $2^8 = 256$  kemungkinan warna yaitu 0 (minimum) hingga 255 (maksimum). (Wahyudi et al., 2015)

### **2.3.8. Image Segmentation**

Segmentasi citra adalah proses dimana komputer membagi dan memotong citra menjadi beberapa bagian/wilayah. Dalam penelitian ini, segmentasi citra akan digunakan untuk memisahkan dokumen dengan aksara Lampung menjadi bagian-bagian aksara Lampung. (Lorentius et al., 2020)