Assalamualaikum Wr Wb, selamat pagi.

Yang terhormat bapak Aji Seto Arfianto selaku pembimbing saya

Yang terhormat ibu Ratih Ayuninghemi selaku penguji 1

Yang terhormat bapak Mukhamad Angga Gumilang selaku penguji 2

Nama saya rizkika zakka palindungan, saya disini ingin mempresentasikan konsep dan progress skirpsi saya.

Saya mengambil tema kecerdasan buatan dengan pengolahan citra / disebut dgn computer vision.

Judul yang saya angkat adalah SISTEM PENERJEMAH BAHASA ISYARAT MENGGUNAKAN METODE DEEP LEARNING MODEL CONVOLUTIONAL NEURAL NETWORK (CNN).

1. Pertama-tama saya akan menjelaskan latar belakang kenapa saya memilih topik dan judul ini.

* Adanya peningkatan populasi yang mengalami gangguan pendengaran. Menurut penelitian (WHO), terdapat lebih dari 5% populasi dunia atau 466 juta orang mengalami gangguan pendengaran. Diperkirakan pada tahun 2050 lebih dari 900 juta orang atau satu dari setiap sepuluh orang akan mengalami gangguan pendengaran. (usia, bawaan sejak lahir, kecelakan, kondisi lingkungan).
* Selanjutnya Komunitas yang mengalami gangguan pendengaran ini memiliki cara berkomunikasi sendiri tanpa menggunakan bahasa lisan yaitu dengan bahasa isyarat. Sedangkan pada umumnya bahasa isyarat sulit dipahami oleh masyarakat dan membuat komunitas tersebut merasa terasingkan di lingkungannya.
* Untuk mengatasi keterbatasan dalam hal berkomunikasi, teknologi computer vision bisa menjadi solusi. Teknologi ini mampu menjembatani komunikasi antar manusia melalui mesin. (Penerjemah / Edukasi). Teknologi ini memanfaatkan data dari kamera atau sensor yang kemudian diolah menggunakan model deep learning. Sehingga mesin mampu secara akurat mengidentifikasi dan mengklasifikasikan objek kemudian bereaksi terhadap apa yang mesin “lihat”.
* Karena data yg saya pakai berupa data gambar maka saya memilih konsep Deep Learning. Dan salah satu metode Deep Learning yang saat ini memiliki hasil paling signifikan dalam pengenalan citra adalah Convolutional Neural Network (CNN). Hal tersebut dikarenakan CNN berusaha meniru sistem pengenalan citra pada visual cortex manusia sehingga memiliki kemampuan mengolah informasi citra.

karena saya menemukan AI open source yang dikembangkan google (nama frameworknya yaitu mediapipe) yg mana berhasil mendeteksi landmark / anatomy tangan manusia dari hanya inputan 1 kamera di arah depan. Maka saya merubah konsep objek detection dan jumlah kamera dari proposal awal yaitu :

* dari yang menggunakan 2 kamera dengan arah samping dan depan sekarang hanya 1 kamera depan (selfie)
* dari yang menggunakan Teknik objek detection dengan cara : penghilangan bagian wajah dan leher dengan menggunakan bantuan Haar Cascade Classifier, kemudian melakukan skin mask atau deteksi warna kulit untuk memfokuskan pada objek tangan. Sekarang sudah tidak perlu karena dengan bantuan mediapipe saya bisa mendapatkan Region of interest dan informasi landmark / anatomy tangan manusia.

Maka dari itu akhirnya ada perubahan juga di rumusan dan Batasan masalah serta tujuan dan manfaat.

1. Rumusan dan Batasan Masalah
2. Tujuan dan Manfaat
3. Bahasa Abjad Bisindo
4. Pengembangan Software
5. Membuat Dataset A-J (Dataset)
   1. Membuat struktur folder

Disini saya membuat folder dataset dengan nama label (index 0 - 9 yang mereferensikan A-J). Setiap folder terdapat 1016 data gambar dengan ukuran 180x120 px, jadi total keseluruhan data yaitu 10160 data gambar

* 1. Deteksi Tangan / get Landmark

Saya menggunakan framework mediapipe untuk mendapatkan ROI dengan menggunakan fitur Hand Tracking yg hasilnya berupa landmark. Nantinya landmark akan digunakan untuk membuat binding box ROI dan Menggambar rangka tangan

* 1. Simpan gambar sesuai folder (atau struktur model)

1. Membuat Model (Training)
   1. import dataset A-J

memasukkan semua data image kedalam satu array secara berurutan

memasukkan semua data label kedalam satu array secara berurutan

* 1. Split dan shuffle Data

Data array dan label yang berpasangan tadi akan di acak dan di bagi kedalam 3 bagian (testing, training validation)

* 1. Distribution Plot

Mengecek apakah persebaran data sudah balance atau tidak (berikut adalah contoh dari persebaran data training)

* 1. Preprocessing and Reshaping data

Mengubah ukuran menjadi 32x32 px

Konversi warna dari BGR ke Gray

Membuat efek kontras dan menjadikan hasil nilai menjadi 0 / 1 (black white)

Merubah channel dan ukuran sesuai kebutuhan keras tensorflow (32, 32, 1)

* 1. Image Augmentation

Setiap data akan dilakukan proses augmentasi berupa merubah ukuran lebar, tinggi, zoom, pemotongan, rotasi

* 1. One Hot Encode

Metode ini merepresentasikan data bertipe kategori sebagai vektor biner yang bernilai integer, 0 dan 1, dimana semua elemen akan bernilai 0 kecuali satu elemen yang bernilai 1, yaitu elemen yang memiliki nilai kategori tersebut.

* 1. Mendesign dan Proses Training Model

Saya mencontoh CNN LENET dari tutorial text and number detection

* 1. Hasil dan Simpan Model

1. Testing Model (Klasifikasi dengan Webcam)
   1. Deteksi Objek
   2. Preprocessing and Reshaping data
   3. Memulai Prediksi dan Menunjukkan Hasil

# CONVOLUTION

* Cara kerja CNN mirip seperti kita mengidentifikasi gambar, (ilustrasi ada di pojok kanan atas) :
* Ketika kita melihat gambar baru, kita memindai/melihat gambar mungkin dari kiri ke kanan dan atas ke bawah untuk memahami fitur-fitur berbeda dari gambar.
* Langkah selanjutnya adalah kita menggabungkan berbagai fitur lokal yang sudah dipindai untuk mengklasifikasikan gambar.
  + Inilah cara kerja CNN, Sehingga CNN mambu mengidentifikasi ketika gambar meskipun gambar tersebut diputar, ukurannya dirubah atau dilihat dalam pencahayaan yang berbeda sehingga sebuah objek akan tetap dikenali sebagai objek yang sama.
* konsep utama convolutional adalah
  + sliding window. Proses mentransformasi suatu window menjadi suatu nilai numerik (filter).
  + Filter. Proses merepresentasikan aspek lokal paling informatif
  + Pooling adalah mengekstrak informasi paling informatif(semacam meringkas). Semua vektor yang dihasilkan pada tahap sebelumnya (sliding window) dikombinasikan (pooled) menjadi satu vektor c (baru) .
  + Normalisasi (mengubah nilai negatif menjadi nol)
* Setelah melewati berbagai operasi convolution dan pooling, kita akan memiliki satu vektor yang kemudian dilewatkan pada multilayer perceptron (fully connected) untuk melakukan sesuatu misal klasifikasi gambar

# NEURAL NETWORK

* Artificial Neural Networks (Jaringan Syaraf Tiruan) juga disebut sebagai ANN (JST)
* **Algoritma Machine Learning ini** terinspirasi pada otak manusia, bagaimana manusia berpikir, belajar dan mengambil keputusan.
* Atau **juga bisa diibaratkan seperti** kumpulan fungsi yang saling berelasi satu dengan yang lainnya (output pada neuron sebelumnya akan menjadi input pada neuron selanjutnya).
* Dilihat di layar PPT pojok kanan atas , Jadi **ANN ini diibaratkan seperti** fungsi yang menerima sinyal input selanjutnya mengubah state menggunakan fungsi aktivasi (relu,sigmoid,tanh) untuk menghasilkan output.
* Dilihat gambar pojok kiri bawah. Sifat neuron seperti berikut :
  + Terdapat nilai **inputan** yang diterima.
  + Disetiap relasi node terdapat **bobot(Weights)** untuk menentukan aktif tidaknya neuron
  + **fungsi aktivasi** akan memprosesinput dan bobot (weight) menghasilkan output.
* Dilihat gambar pojok kanan bawah, ANN memiliki :
  + **Input layer**.
  + Satu atau lebih **hidden layer**, di setiap hidden layer akan memiliki unit **bias**
  + **Output layer**

# CONVOLUTION

* Definisi dan apa itu **CNN**
* Adalah singkatan dari Convolutional Neural Network yang merupakan ANN khusus untuk memproses data yang memiliki bentuk input seperti gambar 2D seperti matriks.
* CNN biasanya digunakan untuk deteksi dan klasifikasi gambar.
* Tujuan :

Kita ingin memahami (memproses) objek pada gambar pada berbagai macam posisi yang mungkin (**invarian**).

* Cara kerja CNN mirip seperti kita mengidentifikasi gambar, (ilustrasi ada di pojok kanan atas) :
* Ketika kita melihat gambar baru, kita memindai/melihat gambar mungkin dari kiri ke kanan dan atas ke bawah untuk memahami fitur-fitur berbeda dari gambar.
* Langkah selanjutnya adalah kita menggabungkan berbagai fitur lokal yang kami pindai untuk mengklasifikasikan gambar.
* Inilah cara kerja CNN, Sehingga CNN mambu mengidentifikasi ketika gambar diputar, ukurannya berbeda atau dilihat dalam pencahayaan yang berbeda, sebuah objek akan dikenali sebagai objek yang sama.
* Untuk ilustrasi konsep **sliding window dan filter** ada di gambar yang berada ditengah.
* Warna biru merepresentasikan satu window,
* kemudian kotak ungu merepresentasikan aspek lokal paling informatif (disebut filter) yang dikenali oleh window.
* Dengan kata lain, kita mentransformasi suatu window menjadi suatu nilai numerik (filter).
* Window ini kemudian digeser-geser sebanyak T kali. Keseluruhan operasi ini disebut sebagai convolution.
* Pergeseran ini disebut stride yaitu seberapa banyak data yang digeser untuk window baru.
* Ide utama **pooling** (ilustrasi ada di gambar kanan bawah)
  + Adalah mengekstrak informasi paling informatif(semacam meringkas).
  + Semua vektor yang dihasilkan pada tahap sebelumnya (sliding window) dikombinasikan (pooled) menjadi satu vektor c (baru) .
  + Ada beberapa teknik pooling, diantaranya: max pooling, average pooling, dan K-max pooling
* Setelah melewati berbagai operasi convolution dan pooling, kita akan memiliki satu vektor yang kemudian dilewatkan pada multilayer perceptron (fully connected) untuk melakukan sesuatu (tergantung permasalahan), misal klasifikasi gambar, klasifikasi sentimen

# NEURAL NETWORK

* **Pertama, pengantar singkat :** Artificial Neural Networks (Jaringan Syaraf Tiruan) juga disebut sebagai ANN (JST)
* **Algoritma Machine Learning ini** terinspirasi pada otak manusia, bagaimana manusia berpikir, belajar dan mengambil keputusan, analogy sama seperti :
  1. ketika manusia menyentuh sesuatu zat yang panas bagaimana neuron dalam tubuh kita mengirimkan sinyal ke otak (mengirim input ke jaringan saraf).
  2. Otak kemudian menghasilkan (output) impuls untuk menarik diri dari area panas.
  3. Berdasarkan pengalaman, ketika kita mendapatkan sinyal yang sama (menyentuh permukaan panas) kita dapat memprediksi output/decision yang tepat (dari proses learning).
* Dilihat di layar PPT pojok kanan atas , Jadi **ANN ini diibaratkan seperti** fungsi yang menerima sinyal input dengan mengubah state menggunakan fungsi aktivasi untuk menghasilkan output.
* Atau **juga bisa diibaratkan seperti** kumpulan fungsi yang saling berelasi satu dengan yang lainnya (output pada neuron sebelumnya akan menjadi input pada neuron selanjutnya).
* Dilihat gambar pojok kiri bawah. Sifat neuron seperti berikut :
  + Output akan berubah berdasarkan **input** yang diterima.
  + kekuatan sinyal (aktif tidaknya neuron) disetiap neuron berdasarkan pada nilai **bobot(Weights)** .
  + dan hasil komputasi input dan bobot (weight) akan di proses didalam **fungsi aktivasi** menghasilkan output.
* Dilihat gambar pojok kanan bawah, ANN memiliki :
  + Lapisan **input**.
  + Satu atau lebih **hidden layer**, setiap lapisan tersembunyi akan memiliki unit **bias**
  + **Output layer**
  + **Bobot(Weights)** yang terkait dengan setiap koneksi
  + **Fungsi aktivasi** yang mengubah sinyal **input** dari sebuah node ke sinyal **output**
* Untuk apa bobot ini, apa fungsi aktivasi dan apa persamaan kompleks ini?
  + Bobot adalah bagaimana neural networks belajar.
  + Terjadi penyesuaian bobot untuk menentukan kekuatan sinyal, sehingga bobot membantu kita menghasilkan output yang berbeda.
* Untuk fungsi aktivasi, untuk apa mereka digunakan?
  + Fungsi aktivasi membantu memutuskan apakah kita perlu menembakkan neuron atau tidak (sigmoid = 0-1).
  + Fungsi aktivasi adalah mekanisme dimana neuron memproses dan mengirimkan informasi melalui jaringan saraf.
* Tahapan :
  1. Secara **acak** menginisialisasi **bobot** ke nilai mendekati nol tetapi tidak sama dengan nol.
  2. Menerapkan **forward propagation** dari kiri ke kanan dengan **mengkalikan bobot** ke nilai **input** dan kemudian menggunakan sebagai fungsi aktivasi (relu,sigmoid,tanh).
  3. Melakukan prediksi output dan bandingkan prediksi output dengan nilai output yang diinginkan. (**cross entropy function** adalah cost function non negatif dan kisarannya antara 0 dan 1)
  4. Selanjutnya melakukan **backpropagation** dari kanan ke kiri dan menyesuaikan nilai bobot pada setiap neuron. Bobot disesuaikan berdasarkan seberapa besar bobot yang bertanggung jawab atas error (learning).
* **backpropagation** merupakan mekanisme feedback yang kadang-kadang kita dapatkan dari orang tua, mentor, teman sebaya kita. Feedback membantu kita menjadi orang yang lebih baik.

**backpropagation** adalah algoritma cepat untuk melakukan learning. Ini memberi tahu kita bagaimana cost function akan berubah ketika kita mengubah bobot dan bias. Sehingga mengubah perilaku jaringan saraf.

* **Epoch** adalah ketika satu kali learning pada semua dataset, satu **forward propagation** dan satu **backpropagation** untuk data training.
* **Learning rate** mengontrol seberapa besar kita harus menyesuaikan bobot sehubungan dengan loss gradient.