

BAB II KAJIAN PUSTAKA

2.1 Penelitian Relevan

Penelitian terdahulu yang terkait tentang pengenalan Bahasa isyarat yang ada dalam penelitian ini diantaranya ialah :

Jurnal dengan judul Implementasi Penerjemah Bahasa Isyarat Pada Bahasa Isyarat Indonesia (BISINDO) Dengan *Metode Principal Component Analysis* (PCA) oleh (Borman Priyopradono, 2018). Tunarungu dan tunawicara melakukan komunikasi dengan menggunakan bahasa isyarat, bahasa isyarat yang sering digunakan di Indonesia adalah metode BISINDO (Bahasa Isyarat Indonesia) yang digalakkan oleh GERKATIN (Gerakan Kesejahteraan Tunarungu Indonesia). Namun penyandang tunarungu dan tunawicara memiliki kesulitan dalam melakukan komunikasi dengan orang normal walaupun telah ada buku panduan karena buku tersebut hanya menjelaskan cara penggunaan bahasa isyarat saja.

Dengan adanya permasalahan diatas maka dibutuhkan sebuah aplikasi penerjemah bahasa isyarat Indonesia (BISINDO) menggunakan Pengolahan Citra Digital untuk manipulasi sebuah gambar digital dengan skala dua dimensi. Dan menggunakan metode *principal component analysis* (PCA) untuk *ekstraksi* pada citra agar data tangan dapat diterjemahkan. Hasil dari penelitian tersebut adalah dengan melakukan pengujian data 26 huruf, tetapi yang dapat dikenali hanya 24 huruf, karena dua huruf tidak terbaca oleh sistem. dari 10 orang memberikan 10 isyarat tangan rata-rata keberhasilan 90,6%.

Penelitian tentang Pengenalan Isyarat Tangan Berupa Angka Pada Video Percakapan Bahasa Isyarat Indonesia (Bisindo) Menggunakan Metode *Ekstraksi* Warna Kulit Dan *Deep Learning* Oleh (Rofiqoh Zahro, 2019). Penelitian ini menggunakan metode *ekstraksi* warna kulit untuk segmentasi pengenalan organ tubuh dan memisahkan bagian gambar yang memiliki warna kulit manusia dengan menggunakan tipe HSV (*Hue Saturation Value*).

Sedangkan Metode *Deep Learning* disini digunakan untuk mengklasifikasikan data yaitu data training yang akan membaca sebuah bahasa isyarat tangan. Namun metode yang telah dilakukan sebelumnya menghasilkan nilai akurasi 91,6 % untuk data uji dengan menggunakan satu tangan dan 62,5 % untuk data uji menggunakan dua tangan. Pada saat pendeteksian objek wajah, sistem masih menghitung nilai dari area wajah tersebut serta objek isyarat hanya huruf vokal saja.

Selain itu adapun penelitian tentang Pengenalan Abjad Pada Video Bahasa Isyarat Indonesia (Bisindo) Menggunakan Metode *Ekstraksi* Warna Dan *Deep Learning* Dengan Penghapusan Wajah oleh (Evia Riska Fitriana, 2020). Pada penelitian tersebut, pengenalan abjad pada video bahasa isyarat indonesia dengan menggunakan metode *ekstraksi* warna kulit dapat memisahkan warna kulit dengan *background* dan memisahkan isyarat *gesture* tangan, sedangkan *Haar Cascade* digunakan untuk mendeteksi wajah pada gambar serta menghapus wajah dan memblok kotak wajah menjadi hitam, dan metode *Deep Learning* digunakan untuk membaca bahasa isyarat dengan benar pada gambar yang jelas dengan nilai akurasi 92 % .

Namun hasil dari penelitian diatas kurang efektif karena metode yang digunakan hanya dapat membaca satu gambar saja (statis) dan belum bisa membaca gerakan sehingga informasi yang diperoleh kurang banyak dan akurasinya kecil. Berdasarkan pada penelitian sebelumnya maka terdapat perbedaan pada penelitian yang akan dilakukan saat ini yaitu penelitian ini menggunakan metode *Deep Gated Recurrent Unit* dimana metode ini berfungsi untuk melihat video percakapan bahasa isyarat indonesia (BISINDO) secara keseluruhan atau beruntut sehingga bisa menghasilkan informasi yang utuh dengan akurat dan tepat.

2.2 Landasan Teori

Dalam sebuah penelitian terdapat landasan dalam penalaran teori yang digunakan, adapun landasan teori pada penelitian ini sebagai berikut:

2.2.1 Tunarungu

Tunarungu adalah seseorang yang memiliki kekurangan atau hilangnya kemampuan mendengar baik sebagian atau seluruhnya yang diakibatkan oleh tidak berfungsinya indra pendengaran (telinga) baik secara sebagian atau keseluruhan, sehingga orang tersebut tidak bisa menggunakan indra pendengarannya dalam kehidupan sehari-hari. Hal tersebut dapat berdampak dalam kehidupan secara kompleks terutama pada kemampuan berbahasa sebagai alat komunikasi yang sangat penting.

Gangguan mendengar yang dialami oleh orang tunarungu dapat menyebabkan terhambatnya perkembangan bahasa, karena perkembangan tersebut, sangat penting untuk berkomunikasi dengan orang lain. Berkomunikasi dengan orang lain membutuhkan bahasa dengan artikulasi atau ucapan yang jelas sehingga pesan yang akan disampaikan dapat tersampaikan dengan baik dan mempunyai satu makna, sehingga tidak ada salah tafsir makna yang dikomunikasikan (Winarsih, Tunarungu, 2017).

2.2.2 Bahasa Isyarat Indonesia (BISINDO)

Bahasa isyarat adalah bahasa yang *universal* karena bahasa isyarat dipakai oleh penyandang tunarungu ketika mereka berbicara atau melakukan komunikasi dengan orang normal atau dengan sesamanya (Riskha, 2020). Bahasa Isyarat Indonesia pertama kali dicetuskan oleh Dimiyati Hakim selaku ketua DPP PERTRIA adalah seorang tunarungu, Dimiyati menciptakan bahasa isyarat Indonesia ini, dikarenakan banyak terjadinya pertikaian dan polemik terkait penerapan dan penggunaan bahasa isyarat di Indonesia (Zahro, 2019).

Penggunaan bahasa isyarat dalam pergaulan komunitas tunarungu harus memiliki 3 syarat utama diantaranya ialah kecepatan, keringkasan, dan kepastian, dalam hal ini BISINDO bisa memenuhi semuanya karena BISINDO merupakan bahasa isyarat konseptual yang berkaitan dengan budaya dan karakter komunitas tunarungu Indonesia. Dengan BISINDO penyandang tunarungu dapat mengutarakan pemikiran dan pendapatnya secara bebas serta dapat

mengekspresikan dirinya bagaikan insan manusia masyarakat Negeri Indonesia yang bermartabat serta sesuai dengan falsafah hidup. Bahasa Isyarat Indonesia ini, dibesarkan dan disebarluaskan lewat sebuah organisasi yaitu GERKATIN (Gerakan buat Kesejahteraan Tunarungu Indonesia).



Gambar 2.1. Bahasa isyarat Tangan

2.2.3 Video

Menurut Kamus Besar Bahasa Indonesia, video merupakan rekaman gambar hidup atau dengan kata lain video merupakan tayangan gambar bergerak yang disertai dengan suara. Video sebenarnya berasal dari bahasa Latin, *video-vidivisum* yang artinya melihat (mempunyai daya penglihatan). Menurut Azhar Arsyad (Arsyadi, 20017) video adalah sebuah gambar - gambar dalam *frame*, di mana *frame* demi *frame* diproyeksikan melalui lensa proyektor secara mekanis sehingga pada layar terlihat gambar hidup. Video dapat menyajikan informasi,

memaparkan proses, menjelaskan konsep-konsep yang rumit, mengajarkan keterampilan, menyingkat atau memperpanjang waktu, dan mempengaruhi sikap.

Menurut Daryanto (Arsyadi, 20017) video adalah rekaman digital dari sebuah gambar yang dijadikan satu sehingga dapat bergerak. Berdasarkan pengertian menurut para ahli di atas, dapat disimpulkan bahwa video adalah gabungan dari beberapa *frame* atau gambar yang dijadikan satu sehingga dapat bergerak dan dapat kita lihat.

2.2.4 Pengolahan Citra Digital

Pengolahan citra merupakan pemrosesan sebuah citra menjadi citra yang kualitasnya lebih baik, tujuan dari pengolahan citra agar mudah dibaca oleh manusia atau mesin (komputer). Beberapa contoh operasi pengolahan citra adalah pada pengubahan kontras citra, penghilangan derau dengan operasi penyaringan (*filtering*), penghasilan tepi objek, penajaman (*sharpening*), pemberian warna semu (*pseudocoloring*), dan sebagainya menurut Leksono, dkk, 2011 dalam skripsi (Zahro, 2019).

Pengolahan citra (*Image Processing*) merupakan sebuah proses pengolahan piksel-piksel dalam citra digital untuk melakukan tujuan tertentu. Citra yang dimaksud disini adalah berupa foto diam (gambar) dan foto bergerak (video) (Riska, 2020). Darma Putra membagi pengolahan citra digital menjadi tiga kategori, yaitu rendah, menengah, dan tinggi. Kategori rendah hanya melibatkan operasi-operasi sederhana seperti pra pengolahan untuk mengurangi derau, pengaturan kontras, dan pengaturan ketajaman citra. Pengolahan kategori menengah melibatkan operasi-operasi seperti segmentasi dan klasifikasi citra. Sedangkan pengolahan kategori tinggi melibatkan proses pengenalan dan deskripsi citra. Teknik yang dilakukan dalam pengolahan citra ialah dengan mentransformasikan citra menjadi citra yang lain (Putra, 20018). Namun, dalam hal ini hasil keluaran dari citra memiliki kualitas lebih baik daripada citra masukan.



Gambar 2.2. Contoh pengolahan citra pengenalan jarak jauh dengan Biometrika wajah (sumber : Darma Putra)

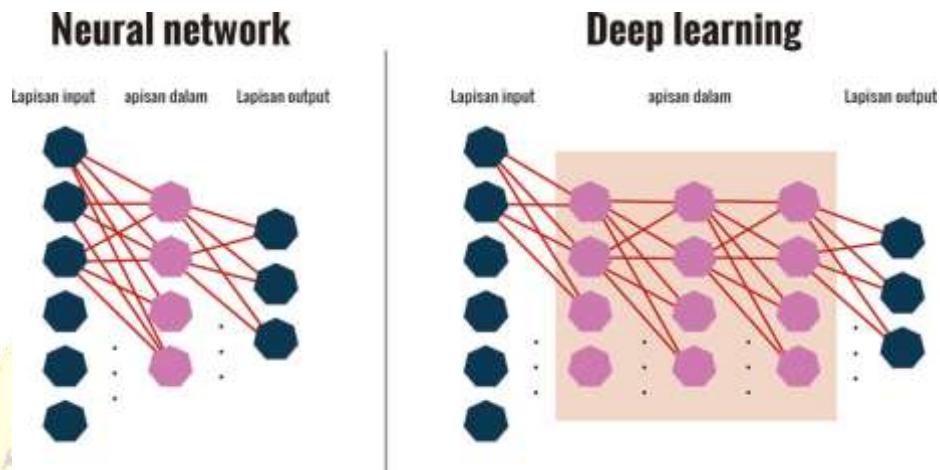
2.2.5 Deep Learning

Deep Learning merupakan bagian dari *machine learning* yang diperoleh dari pengembangan Jaringan Syaraf Tiruan (JST) dimana didalamnya terdapat banyak *layer* (lapisan) yang digunakan untuk permasalahan dataset yang besar (Zahro, 2019). Teknik *deep learning* ini dapat memberikan arsitektur yang sangat kuat dengan menambahkan lebih banyak lapisan maka model pembelajaran tersebut bisa mewakili data citra berlabel lebih baik lagi.

Deep learning berfungsi untuk melatih komputer dengan cara belajar dari contoh agar bisa mengklasifikasikan sebuah gambar, suara, maupun video sehingga komputer dapat melakukan tindakan yang dianggap alami oleh manusia. Aplikasi konsep jaringan syaraf tiruan dapat ditangguhkan pada algoritma *machine learning* yang sudah ada sehingga komputer sekarang bisa belajar dengan kecepatan, akurasi, dan skala yang besar.

Prinsip tersebut terus berkembang sehingga *deep Learning* semakin sering digunakan pada komunitas riset dan industri untuk membantu memecahkan banyak masalah data besar. Dalam menyelesaikan permasalahan tentang prediksi atau klasifikasi, arsitektur pada teknik *deep learning* lebih “*deep*” dibandingkan dengan teknik dalam *machine learning* lainnya. Pada gambar 2.3. dapat disimpulkan

bahwasanya *deep learning* dapat meningkatkan transmisi dan pemrosesan informasi *interlayers*, sehingga memungkinkan untuk meningkatkan ketepatan dan keserbagunaan jumlah fitur serta meningkatkan akurasi prediksi.



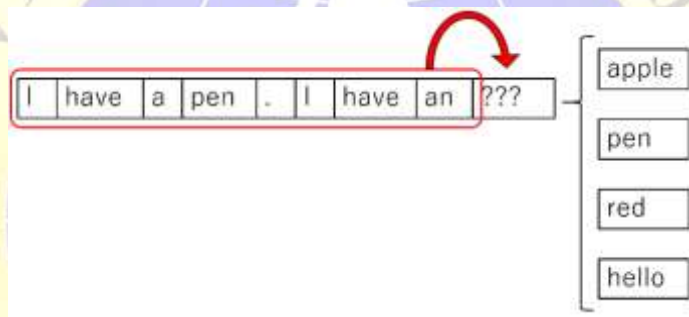
Gambar 2.3. Alur *deep learning*

2.2.6 Recurrent Neural Network

Recurrent Neural Network (RNN) adalah sebuah jaringan saraf berulang, mengapa dikatakan demikian karena nilai *neuron* pada *hidden layer* sebelumnya akan diproses kembali sebagai data input. *Neuron* yang ada pada *hidden layer* akan disimpan ke dalam *context layer* lalu akan terus dilakukan *update* data hingga kondisi RNN terpenuhi (Rafika Puspa Wardana, 2020). Selain itu RNN juga dapat memproses sejumlah data yang bersambung (*sequential data*).

Seperti halnya manusia, *Recurrent Neural Network* menyimpan informasi dari masa lalu dengan melakukan *looping* dalam arsitekturnya. RNN ini dibentuk dari *neuron* buatan yang ada pada *hidden layer* dengan satu atau lebih umpan balik yang berulang. Menurut (Salahenejad, 2018) RNN memiliki tiga lapisan yaitu *layer input*, layer tersembunyi yang berulang, dan *layer output*. *Layer input* memiliki unit *input*, terkoneksi penuh ke unit tersembunyi yang ada di *layer* tersembunyi. Unit tersembunyi tersebut terhubung satu sama lain secara berulang.

Layer tersembunyi bisa didefinisikan sebagai “memori” atau ruang keadaan yang berdimensi tinggi dengan dinamika *non-linier* untuk mengingat dan memproses informasi masa lalu. Keadaan tersembunyi akan merangkum semua informasi unik yang diperlukan sebagai keadaan terakhir dari jaringan, melalui serangkaian langkah waktu. Informasi itu lalu terintegrasi, sehingga mampu menentukan perilaku jaringan di masa depan dan melakukan prediksi yang akurat di *layer output*. *Recurrent neural network* merupakan metode yang kompleks dan dinamis karena nilai *output* yang dihasilkan tidak hanya dipengaruhi oleh nilai *input* saja melainkan juga dari nilai *output* sebelumnya. Arsitektur RNN dapat dilihat pada ilustrasi Gambar 2.4.



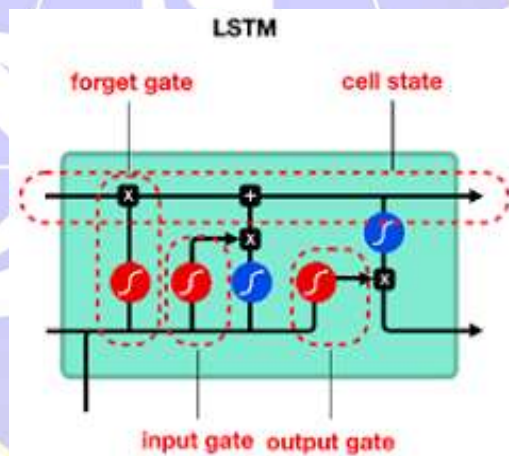
Gambar 2.4. Simulasi RNN

2.2.7 Long Short Term Memory

LSTM pertama kali dibuat oleh Sepp Hochreiter dan Jürgen Schmidhuber pada tahun 1997. LSTM juga disempurnakan oleh banyak orang, seperti Felix Gers, Fred Cummins, Santiago Fernandez, Justin Bayer, Daan Wierstra, Julian Togelius, Faustian Gomez, Matteo Gagliolo, dan Alex Graves (Zaman, 2019). *Long Short Term Memory* (LSTM) merupakan salah satu jenis dari *Recurrent Neural Network* yang dibuat untuk menghindari terjadinya ketergantungan jangka panjang pada *Recurrent Neural Network*. Selain itu fungsi dari LSTM juga ialah dapat menangani penghafalan dan mengingat kembali untuk jangka panjang, khususnya data yang sangat besar. Dan prinsip dari LSTM sendiri yaitu dapat menggunakan unit

memorinya untuk mengingat informasi yang jaraknya jauh dan melacak berbagai atribut teks yang sedang diproses (Habibie, 2018).

LSTM juga memiliki kemampuan untuk menambah dan menghapus informasi dari *cell state*. Kemampuan ini disebut dengan *gates*. *Gates* disini berfungsi sebagai pengatur apakah informasi yang diperoleh akan diteruskan atau diberhentikan. *Gates* terdiri dari *sigmoid layer* dan *pointwise multiplication operation*. LSTM memiliki 3 jenis *gates* diantaranya ialah *forget gate*, *input gate*, dan *output gate*. *Forget gate* adalah *gate* yang memutuskan informasi mana yang akan dihapus dari *cell*. *Input gate* adalah *gate* yang memutuskan nilai dari *input* untuk di diperbarui pada *state* memori. *Output gate* adalah *gate* yang memutuskan apa yang akan dihasilkan *output* sesuai dengan *input* dan memori pada *cell*. Ilustrasi LSTM terlihat pada Gambar 2.5.



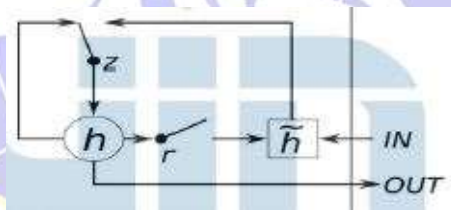
Gambar 2.5. Ilustri LSTM

2.2.8 Gated Recurrent Unit

GRU adalah arsitektur yang diciptakan oleh Kyunghun Cho pada tahun 2014, *Gated Recurrent Unit* (GRU) merupakan sebuah metode yang di peroleh dari perkembangan metode *Recurrent Neural Network* yang mana metode ini dapat melakukan prediksi pada data time-series (deret waktu). Selain itu GRU juga memberikan *konvergensi* yang lebih cepat dan hasilnya bisa di bandingkan dengan LSTM. Berdasarkan hasil penelitian R. Fu, 2016 dalam jurnal (saputra, 2020)

menunjukkan bahwa metode GRU memiliki nilai kesalahan (error) yang lebih kecil dibandingkan metode LSTM, *Gated Recurrent Unit* (GRU) sering digunakan dalam bidang industri dan akademik, GRU disini dapat kita umpamakan seperti manusia karena ketika manusia ingin membuat keputusan sekarang, maka ia tidak perlu melihat atau menggunakan semua informasi masa lalu untuk membuat keputusan tersebut.

Tujuan dari GRU adalah untuk membuat setiap *recurrent unit* agar bisa menangkap setiap hubungan (*dependensi*) dalam skala waktu yang berbeda-beda yang dilakukan secara adaptif. GRU mempunyai sebuah komponen pengatur alur informasi yang disebut dengan *gate* dan GRU disini mempunyai 2 *gate* yaitu, *reset gate* dan *update gate* (Rafika Puspa Wardana, 2020). *Reset gate* berfungsi untuk menentukan bagaimana cara menggabungkan inputan baru dengan informasi masa lalu, dan *update gate* berfungsi untuk menentukan berapa banyak informasi masa lalu yang harus tetap disimpan. Struktur dari GRU bisa dilihat pada Gambar 2.6. Pada struktur GRU di bawah ini, r melambangkan *reset gated*, z melambangkan *update gate*, sedangkan h dan \tilde{h} melambangkan *activation* dan *candidate activation*.



Sumber : Skripsi Rafika

Gambar 2.6. Struktur GRU

2.2.9 Open CV Python

OpenCV (*Open Source Computer Vision*) merupakan sebuah *library* yang menyediakan infrastruktur umum untuk aplikasi *Computer Vision*. OpenCV berada di bawah lisensi BSD (*Berkeley Software Distribution*) sehingga dapat memudahkan pengguna untuk memodifikasi code dan resmi dipakai secara komersial. OpenCV dibuat menggunakan bahasa pemrograman C++, namun OpenCV juga menyediakan *interface* untuk bahasa pemrograman lain seperti *Python Java*, dan *Matlab* (Zahro, 2019).

Secara teori OpenCV cara kerjanya mengikuti sistem visual manusia yaitu dengan melihat objek melalui "penglihatan/mata" dan citra pada objek tersebut diteruskan ke otak untuk diproses sehingga dapat mengerti objek apa yang tampak pada pandangan mata manusia. OpenCV merupakan salah satu cabang dari *Artificial intelligence* (kecerdasan buatan) yang digunakan untuk pengembangan atau analisis isi suatu gambar atau video dengan menggunakan kamera untuk mengambil gambar lalu diolah oleh komputer. OpenCV di desain untuk efisiensi dalam komputasi dan difokuskan pada aplikasi *real-time*. OpenCV juga dapat digunakan pada komputer berbasis *windows, linux, OSX, BSD* dan *IOS*.