

# Analisa Manfaat Dan Penerimaan Terhadap Implementasi Bahasa Isyarat Indonesia Pada Latar Belakang Komplek Menggunakan Kinect Dan Jaringan Syaraf Tiruan

## (Studi Kasus SLB Karya Mulia 1)

Nehemia Sugianto<sup>1</sup>, Febriliyan Samopa<sup>2</sup>

**Abstrak**— Manusia berinteraksi satu sama lain melalui komunikasi dalam bentuk bahasa. Komunikasi dapat terjadi baik secara verbal maupun non verbal. Komunikasi verbal adalah komunikasi yang menggunakan suara, sedangkan komunikasi non verbal adalah komunikasi yang menggunakan simbol-simbol. Dalam kehidupan sehari-hari, manusia berkomunikasi secara verbal tetapi tidak semua manusia dapat berkomunikasi secara verbal. Para penderita tuna rungu dan tuna wicara menggunakan bahasa isyarat dalam berkomunikasi.

Para penderita tuna rungu dan tuna wicara mengalami kesulitan dalam berkomunikasi dengan orang normal karena perbedaan metode komunikasi. Bagi para tuna rungu dan tuna wicara, bahasa isyarat tersebut adalah umum bagi mereka tetapi asing bagi orang normal. Hal ini dapat mengganggu keharmonisan sosial antara penderita tuna rungu dan tuna wicara dengan orang normal. Untuk itu dibutuhkan perantara alternatif yang dapat menjadi penerjemah antara para penderita tuna rungu dan tuna wicara dengan orang normal.

Dilatarbelakangi oleh permasalahan di atas, maka dilakukan penelitian untuk mengembangkan sebuah aplikasi yang dapat mengenali bahasa isyarat pada latar belakang komplek menggunakan kinect dan jaringan syaraf tiruan serta analisa manfaat dan penerimaan terhadap implementasi di SLB Karya Mulia 1. Proses pengenalan bahasa isyarat ini terdiri dari tiga tahap yaitu tahap masukan, tahap proses dan tahap keluaran. Tahap masukan adalah tahap pengambilan bahasa isyarat yang disajikan oleh pemberi bahasa isyarat menggunakan Kinect Sensor untuk mendapatkan kumpulan depth image dan kumpulan skeleton image. Tahap proses adalah tahap inti dari sistem pengenalan bahasa isyarat. Pada tahap ini, dilakukan pengolahan citra digital untuk mendapatkan fitur-fitur yang akan digunakan sebagai data masukan pada proses klasifikasi (classification) untuk mengenali bahasa isyarat tersebut. Tahap keluaran adalah tahap dimana sistem memberikan hasil bahasa isyarat yang dikenali pada tahap

sebelumnya dan menampilkannya dalam bentuk tulisan, gambar atau suara. Pengembangan aplikasi ini menggunakan bahasa C# dan EmguCV untuk pemrosesan citra digital.

Berdasarkan hasil uji coba, jaringan syaraf tiruan tersebut dapat mengenali bahasa isyarat dengan tingkat akurasi sebesar 85%. Sebagian besar bahasa isyarat dapat dikenali dengan baik tetapi ada beberapa bahasa isyarat yang belum dapat dikenali dengan maksimal dikarenakan kemiripan bentuk tangan pada bahasa isyarat tersebut. Berdasarkan hasil implementasi pada SLB Karya Mulia 1, aplikasi ini memiliki potensi manfaat yang dapat membantu siswa, orang tua dan guru dalam proses belajar mengajar namun perlu digabungkan dengan modul pembelajaran serta mempertimbangkan penggunaan sensor selain Kinect dikarenakan tingkat kompleksitas untuk installasi perangkat keras dan mobilitas perangkat.

**Kata Kunci:** bahasa isyarat, sibi, pengolahan citra digital, jaringan syaraf tiruan, kinect, EmguCV.

**Abstract**— Humans interact with each other through communication called language. Communication can take place either verbally or non-verbally. Verbal communication uses voice while non-verbal communication uses symbols. In daily life, humans communicate verbally but not all people can communicate verbally. The deaf and mute use sign language to communicate. Sign language is non-verbal communication because does not use sound but using gesture (sign, direction, movement of hands), lips, and facial expression to convey intent and mind of a speaker.

The deaf and mute have difficulty to communicate with normal people due to difference of communication method. For the deaf and mute, sign language is common but not common for normal people. This problem can disrupt social harmony between them. Therefore, they need an alternative interpreter between them.

Based on the problem, this research is held to develop an application to recognize sign language in complex background using kinect and artificial neural network and analyze the usefulness and acceptance of the implementation at SLB Karya Mulia 1. The recognition process consists of three main stages i.e. input stage, process stage, output stage. Input stage is capturing the hand image of the signer by using Kinect sensor to retrieve sequence of depth images and sequence of skeleton iamges.

Process stage is main stage of sign language recognition system which consists of some processes i.e. digital image processing to retrieve features which will be used as input in classification process to recognize the sign language. Output

<sup>1</sup> Dosen, Jurusan Teknik Informatika Fakultas Industri Kreatif Universitas Ciputra, Jln. UC Town, Surabaya 60291 INDONESIA (telp: 031-555 5555; fax: 031-876 54321; e-mail: nsugianto@ciputra.ac.id)

<sup>2</sup> Dosen, Jurusan Sistem Informasi Fakultas Teknologi Informasi Institut Teknologi Sepuluh November, Jln. Cokroaminoto 12A Surabaya 60264 INDONESIA (telp: 031-561 3922; fax: 031-568 2887; e-mail: iyan@is.its.ac.id)

stage is producing the meaning of the sign as result and display the result in text, image or sound. This application is developed using C# language and EmguCV library for digital image processing.

Based on testing result, the neural network is able to recognize the sign languages with accuracy level of 85%. Most of sign languages are able to be recognized well but there are some sign languages are not able to be recognized well because of the similarity of hand shape between sign languages. Based on implementation at SLB Karya Mulia 1, this application has potential benefit to help learning process between students, teachers and students' parents but there are some improvement needs such as combining learning module in this application and consideration of using other sensors beside Kinect sensor because of complexity level in device installation and mobility.

**Keywords:** sign language, sibi, image processing, neural network, kinect, EmguCV.

## I. PENDAHULUAN

Manusia berinteraksi satu sama lain melalui komunikasi dalam bentuk bahasa. Komunikasi dapat terjadi baik secara verbal maupun non verbal. Komunikasi verbal adalah komunikasi yang menggunakan suara, sedangkan komunikasi non verbal adalah komunikasi yang menggunakan simbol-simbol. Dalam kehidupan sehari-hari, manusia berkomunikasi secara verbal tetapi tidak semua manusia dapat berkomunikasi secara verbal.

Para penderita tuna rungu dan tuna wicara menggunakan bahasa isyarat dalam berkomunikasi. Bahasa isyarat adalah komunikasi non verbal karena merupakan bahasa yang tidak menggunakan suara tetapi menggunakan bentuk dan arah tangan, pergerakan tangan, bibir, badan serta ekspresi wajah untuk menyampaikan maksud dan pikiran dari seorang penutur. Belum ada bahasa isyarat internasional karena bahasa isyarat di tiap negara belum tentu sama. Para penderita tuna rungu dan tuna wicara di Indonesia berkomunikasi menggunakan bahasa isyarat yang mengacu pada dua sistem yaitu BISINDO (Berkenalan Dengan Sistem Isyarat Indonesia) dan SIBI (Sistem Isyarat Bahasa Indonesia). BISINDO dikembangkan oleh orang tuna rungu sendiri melalui GERKATIN (Gerakan Kesejahteraan Tuna Rungu Indonesia). SIBI dikembangkan oleh orang normal, bukan penderita tuna rungu. SIBI sama dengan bahasa isyarat yang digunakan di Amerika yaitu ASL (*American Sign Language*).

Para penderita tuna rungu dan tuna wicara mengalami kesulitan dalam berkomunikasi dengan orang normal karena perbedaan metode komunikasi. Bagi para tuna rungu dan tuna wicara, bahasa isyarat tersebut adalah umum bagi mereka tetapi asing bagi orang normal. Beberapa orang normal yang tinggal dekat dengan penderita tuna rungu dan tuna wicara mungkin tidak asing akan bahasa isyarat mereka, tetapi asing bagi orang normal pada umumnya. Hal ini dapat mengganggu keharmonisan sosial antara penderita tuna rungu dan tuna wicara dengan orang normal. Untuk itu dibutuhkan

perantara yang dapat menjadi penerjemah antara para penderita tuna rungu dan tuna wicara dengan orang normal agar dapat tercipta komunikasi yang lebih baik. Perantara ini akan menterjemahkan bahasa isyarat menjadi tulisan atau suara.

Dalam mengenali bahasa isyarat, dibutuhkan sejumlah fitur-fitur dari anggota tubuh manusia yang dapat membedakan bahasa isyarat yang satu dengan bahasa isyarat yang lain. Anggota tubuh yang dimaksud meliputi penampil (tangan atau bagian tangan yang digunakan untuk membentuk isyarat), posisi (kedudukan tangan atau kedua tangan terhadap pengisyarat pada waktu berisyarat), tempat (bagian badan yang menjadi tempat awal isyarat dibentuk atau arah akhir isyarat), arah (gerak penampil ketika isyarat dibuat), frekuensi (jumlah gerak yang dilakukan pada waktu isyarat dibentuk), mimik muka, gerak tubuh, kecepatan gerak, dan kelenturan gerak.

Dilatarbelakangi oleh permasalahan di atas, maka dilakukan penelitian untuk mengenali bahasa isyarat pada latar belakang kompleks menggunakan kinect dan jaringan syaraf tiruan serta melakukan analisa manfaat dan penerimaan terhadap implementasi di SLB Karya Mulia 1. Salah satu yang menjadi fokus dalam penelitian ini adalah menghindari penggunaan alat tambahan khusus seperti sarung tangan robotik (robotic glove) atau sarung tangan yang diberi warna khusus dalam mengambil bahasa isyarat.

Sesuai dengan tujuan penelitian yang ingin dicapai dan dengan pertimbangan terbatasnya waktu penelitian, maka penelitian ini terbatas pada aspek-aspek sebagai berikut :

- Bahasa isyarat yang dapat dikenali adalah huruf (huruf A hingga huruf Z) dan angka (angka 0 hingga angka 9)
- Hanya mengenali seorang pemberi isyarat
- Penyampaian bahasa isyarat menggunakan satu tangan yaitu tangan kanan
- Penyampaian bahasa isyarat dilakukan dalam kondisi tubuh tegak (bukan miring) dengan pencahayaan normal (tidak kekurangan atau kelebihan cahaya) dimana posisi tangan di depan tubuh pemberi isyarat dan menghadap sejajar ke kamera
- Tidak terbatas pada warna kulit pemberi isyarat
- Hasil keluaran dari sistem ini berupa tulisan atau suara dari bahasa isyarat yang dikenali
- Bersifat monolog yaitu menterjemahkan bahasa isyarat ke bahasa normal, bukan sebaliknya.

## II. KAJIAN PUSTAKA DAN DASAR TEORI

Bagian ini berisi kajian pustaka berupa penelitian-penelitian terdahulu dari beberapa literatur dan dasar teori yang mendukung agar memberikan landasan dan kerangka berpikir yang kuat dan relevan dengan penelitian ini sehingga dapat mencapai tujuan penelitian yang telah ditetapkan sebelumnya

### A. Bahasa Isyarat

Bahasa isyarat adalah komunikasi non verbal karena merupakan bahasa yang tidak menggunakan suara tetapi menggunakan bentuk dan arah tangan, pergerakan tangan, bibir, badan serta ekspresi wajah untuk menyampaikan maksud dan pikiran dari seorang penutur.

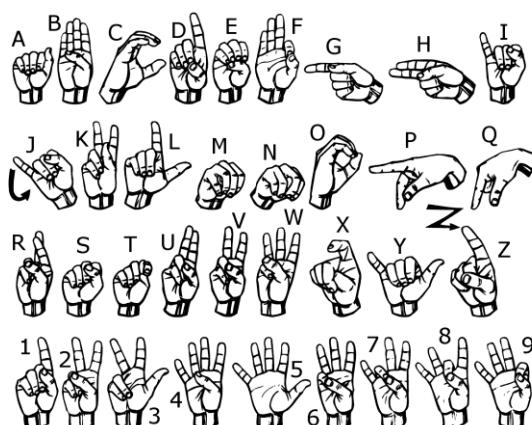
Belum ada bahasa isyarat internasional karena bahasa isyarat di tiap negara belum tentu sama. Ada beberapa bahasa isyarat yang dipakai di suatu negara tetapi tidak ditemukan di negara lain. Bahasa isyarat biasanya berkembang sesuai dengan lingkungan dan budaya setempat. Beberapa bahasa isyarat yang ada adalah *American Sign Language (ASL)*, *French Sign Language (LSF)*, *German Sign Language (DGS)*, dan *Arabic Sign Language (ArSL)*.

Para penderita tuna rungu dan tuna wicara di Indonesia berkomunikasi menggunakan bahasa isyarat yang mengacu pada dua sistem yaitu BISINDO (Berkenalan Dengan Sistem Isyarat Indonesia) dan SIBI (Sistem Isyarat Bahasa Indonesia). BISINDO dikembangkan oleh orang tuna rungu sendiri melalui GERKATIN (Gerakan Kesejahteraan Tuna Rungu Indonesia). SIBI dikembangkan oleh orang normal, bukan penderita tuna rungu. SIBI sama dengan bahasa isyarat yang digunakan di Amerika yaitu *American Sign Language (ASL)*.

#### 1) Sistem Bahasa Isyarat Indonesia (SIBI)

Sistem Isyarat Bahasa Indonesia (SIBI) merupakan salah satu media yang membantu komunikasi sesama kaum tuna rungu di dalam masyarakat yang lebih luas (Departemen Pendidikan Nasional, 2002). Wujudnya adalah tatanan yang sistematis tentang seperangkat isyarat jari, tangan, dan berbagai gerak yang melambangkan kosa kata bahasa Indonesia.

Pada Sistem Bahasa Isyarat Indonesia (SIBI), terdapat 26 isyarat huruf (dimana 24 huruf merupakan isyarat statis dan 2 huruf merupakan isyarat non statis yaitu huruf J dan huruf Z) dan 10 isyarat angka (dimana 10 angka merupakan isyarat statis). Isyarat statis adalah isyarat yang tidak melibatkan pergerakan tangan. Isyarat non statis adalah isyarat yang melibatkan pergerakan tangan.



Gambar 1. Bahasa isyarat huruf dan angka pada sistem bahasa isyarat Indonesia

#### 2) Perkembangan Penelitian Bahasa Isyarat

Sistem pengenalan bahasa isyarat pada umumnya dibagi menjadi dua kategori yaitu bahasa isyarat statis dan bahasa isyarat dinamis (Maraqa, Al-Zboun, Dhyabat, & Abu Zitar, 2012). Bahasa isyarat statis adalah bahasa isyarat yang disajikan dalam bentuk sebuah citra digital saja. Bahasa isyarat dinamis adalah bahasa isyarat yang disajikan dalam bentuk beberapa citra digital dimana terjadi pergerakan anggota tubuh di dalamnya.

Dalam mengenali bahasa isyarat, dibutuhkan sejumlah fitur-fitur dari anggota tubuh manusia yang dapat membedakan bahasa isyarat yang satu dengan bahasa isyarat yang lain. Anggota tubuh yang dimaksud meliputi penampil (tangan atau bagian tangan yang digunakan untuk membentuk isyarat), posisi (kedudukan tangan atau kedua tangan terhadap pengisyarat pada waktu berisyarat), tempat (bagian badan yang menjadi tempat awal isyarat dibentuk atau arah akhir isyarat), arah (gerak penampil ketika isyarat dibuat), frekuensi (jumlah gerak yang dilakukan pada waktu isyarat dibentuk), mimik muka, gerak tubuh, kecepatan gerak, dan kelenturan gerak.

Penelitian tentang pengenalan bahasa isyarat telah menjadi salah satu topik penelitian yang diminati sejak dahulu dan masih menjadi sebuah tantangan yang menarik untuk diteliti lebih dalam lagi hingga sekarang karena pengaruh kecepatan gerakan tangan, tingginya tingkat kebebasan gerakan (*high degrees of freedom*) dan tumpang tindihnya antar telapak tangan atau bagian tubuh yang lain (seperti wajah). Berdasarkan jenis perangkat yang digunakan untuk mengambil fitur-fitur dari anggota tubuh manusia, sistem pengenalan bahasa isyarat yang dilakukan dalam penelitian-penelitian terdahulu dapat dibagi menjadi tiga jenis yaitu sistem pengenalan bahasa isyarat yang menggunakan perangkat keras khusus (seperti sarung tangan robotik atau *data glove with flex sensor*), sistem pengenalan bahasa isyarat yang menggunakan *web camera* atau perangkat sejenisnya dan sistem pengenalan bahasa isyarat yang menggunakan *motion sensing input device* (seperti *Kinect Sensor*, *PlayStation Eye*, *Asus Xtion PRO LIVE*, *Wii Remote*, *Leap Motion* atau perangkat sejenisnya). Tabel perbandingan antara ketiga jenis penelitian terdahulu tentang bahasa isyarat dapat dilihat pada Tabel I.

TABEL I  
TABEL PERBANDINGAN ANTARA BERBAGAI JENIS PENELITIAN-PENELITIAN TERDAHULU TENTANG BAHASA ISYARAT

Parameter Perbandingan	Sistem Pengenalan Bahasa Isyarat Yang Menggunakan Perangkat Keras Khusus	Sistem Pengenalan Bahasa Isyarat Yang Menggunakan Web Camera Atau Perangkat Sejenisnya	Sistem Pengenalan Bahasa Isyarat Yang Menggunakan Motion Sensing Input Device	Data Masukan (Input Data)	Glove Data	RGB Data (Color Data)	Leap Frame Data [7]
Peneliti	[1] Iqbal dkk (2012) [2] Murakami dkk (1991)	[3] Mapari dkk (2012) [4] Maraqqa dkk (2012) [5] Mekala dkk (2011) [6] Lungociu (2011)	[7] Nowicki dkk (2014) [8] Gunawan dkk (2013) [9] Chai dkk (2013)		<ul style="list-style-type: none"> <li>• Thumb finger bending</li> <li>• Index finger bending</li> <li>• Middle finger bending</li> <li>• Ring finger bending</li> <li>• Baby finger bending</li> <li>• 3D hand movement (x,y,z)</li> </ul>		<ul style="list-style-type: none"> <li>• Hands</li> <li>• Fingers</li> <li>• Pointables</li> <li>• Additional Information (gestures recognition, timestamp, rotation, translation, scaling data)</li> </ul>
Jenis Bahasa Isyarat Yang Dikenali	Bahasa isyarat statis [2] Bahasa isyarat dinamis [1,2]	Bahasa isyarat statis [3,4,5,6] Bahasa isyarat dinamis [4,5]	Bahasa isyarat statis [7] Bahasa isyarat dinamis [7,8,9]				Depth Data dan Skeleton Data [8,9]
Sistem Bahasa Isyarat Yang Dikenali	Bhs Isyarat Indonesia [1] Bhs Isyarat Jepang [2]	Bhs Isyarat Amerika [3,5,6] Bhs Isyarat Arab [4]	Bhs Isyarat Amerika [7] Bhs Isyarat Indonesia [8] Bhs Isyarat Cina [9]		Proses Pengolahan Citra Digital (Digital Image Processing)	<ul style="list-style-type: none"> <li>• Main Computation :</li> <li>• Pre-Processing [3,4,5,6]</li> <li>• Hand Detection [3,4,5,6]</li> <li>• Hand Segmentation [3,4,5,6]</li> <li>• Feature Extraction [3,4,5,6]</li> <li>• Additional Computation :</li> <li>• Fingertips Detection [4,5]</li> <li>• Fingertips Labelling [4,5]</li> </ul>	
Perangkat Masukan Utama (Input Device)	Sarung tangan robotik (Data glove with Flex Sensor) [1,2]	Web camera atau perangkat sejenisnya [3,4,5,6]	Leap Motion [7] Kinect Sensor [8,9]		Proses Klasifikasi (Classification) / Pencocokan Pola (Template Matching)	<ul style="list-style-type: none"> <li>• Dynamic Time Warping (DTW) [1]</li> <li>• Recurrent Neural Network [4]</li> <li>• Combinational Neural Network [5]</li> <li>• Backpropagation Neural Network [6]</li> <li>• Support Vector Machine (SVM) [3]</li> <li>• Recurrent Neural Network [4]</li> <li>• Combinational Neural Network [5]</li> <li>• Backpropagation Neural Network [6]</li> <li>• K-Means [7]</li> <li>• Hidden Markov Models (HMMs) [7]</li> <li>• Dynamic Time Warping (DTW) [8]</li> <li>• Euclidean Distance [9]</li> </ul>	
Perangkat Masukan Tambahan (Additional Input Device)	Accelerometer Sensor [1,2]	Colored Glove (sarung tangan yang ujung tiap jarinya diberi warna yang berbeda-beda) [4]	(tidak ada)		Kelebihan Dari Sistem	<ul style="list-style-type: none"> <li>• Tingkat akurasi yang tinggi pada kualitas data masukan</li> <li>• Biaya komputasi yang murah</li> <li>• Mampu mendeteksi pergerakan jari tangan secara akurat dimana berguna bagi</li> <li>• Biaya perangkat masukan murah</li> <li>• Membutuhkan waktu instalasi perangkat masukan yang cepat</li> <li>• Mudah untuk diimplementasikan dalam kehidupan sehari-hari karena perangkat masukan itu kecil, ringan dan mudah dibawa</li> </ul>	<ul style="list-style-type: none"> <li>• Biaya komputasi yang relatif murah dibandingkan yang menggunakan web camera atau perangkat sejenisnya karena perangkat masukan tersebut sudah mampu melakuk</li> </ul>
Batasan Sistem (Constraints)	<ul style="list-style-type: none"> <li>• Tidak terbatas pada batasan latar belakang pemberi isyarat</li> <li>• Latar belakang pemberi isyarat harus berwarna tertentu (seperti hitam atau putih) untuk memudahkan komputasi [3,4,5]</li> </ul>						

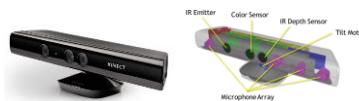
	<ul style="list-style-type: none"> <li>• bahasa isyarat yang melibatkan jari tangan</li> <li>• Tidak terbatas pada latar belakang pemberi isyarat apapun</li> <li>• Faktor cahaya tidak berpengaruh pada kualitas data masukan</li> </ul>	<ul style="list-style-type: none"> <li>• n deteksi seperti hand tracking, finger tracking, skeleton tracking, face tracking dan kamera infra merah</li> <li>• Dengan Leap Motion, mampu mendeteksi pergerakan jari tangan dengan akurat sehingga biaya komputasi menjadi murah dimana berguna bagi bahasa isyarat yang melibatkan jari tangan [7]</li> <li>• Tidak terbatas pada latar belakang apapun</li> <li>• Faktor cahaya tidak berpengaruh pada kualitas data masukan</li> <li>• Relatif mudah untuk diimplementasikan dalam kehidupan sehari-hari karena perangkat masukan itu relatif kecil, ringan dan mudah dibawa</li> </ul>	<p>Kekurangan Dari Sistem</p> <ul style="list-style-type: none"> <li>• Biaya perangkat masukan mahal</li> <li>• Membutuhkan waktu installasi perangkat masukan yang lama karena melibatkan kabel-kabel yang harus dipasang di beberapa anggota tubuh. Susah untuk diimplementasikan dalam kehidupan sehari-hari karena perangkat masukan ini ribet, berat dan susah dibawa</li> </ul>	<ul style="list-style-type: none"> <li>• Latar belakang pemberi isyarat berpengaruh pada kualitas pengenalan bahasa isyarat</li> <li>• Faktor cahaya berpengaruh besar pada kualitas data masukan</li> <li>• Biaya komputasi yang mahal</li> <li>• Membutuhkan penerapan beberapa algoritma tambahan untuk mendekripsi pergerakan jari tangan apabila ingin mengenali bahasa isyarat yang melibatkan jari tangan sehingga biaya komputasi menjadi lebih mahal</li> <li>• Membutuhkan waktu installasi perangkat masukan yang lebih lama dibandingkan web camera atau perangkat sejenisnya (tetapi masih relatif terjangkau)</li> <li>• Jika menggunakan akan Kinect Sensor versi 1, membutuhkan penerapan beberapa algoritma untuk mendekripsi pergerakan jari tangan apabila ingin mengenali bahasa isyarat yang melibatkan jari tangan sehingga biaya komputasi menjadi lebih mahal</li> </ul>
--	---	--	---	--

## B. Kinect

*Kinect* (disebut juga sebagai *Kinect Sensor*) adalah sebuah perangkat keras yang terdiri dari kamera-kamera, sebuah *microphone array* dan sebuah *accelerometer* yang mampu melakukan pemrosesan komputasi pada *RGB data*

(*color data*), *depth data* dan *skeleton data*. Pada Kinect Sensor versi 1, terdiri dari empat bagian utama yaitu :

- Sebuah *RGB camera* yang mampu menangkap gambar dan menghasilkan sebuah citra berwarna (*RGB data*) dalam tiga *channel* dengan resolusi maksimal hingga 1280 x 960 *pixel*
- Sebuah *infrared (IR) emitter* dan *infrared (IR) depth sensor*. *Infrared (IR) emitter* memancarkan cahaya infra merah. *Infrared (IR) depth sensor* menangkap pantulan dari cahaya infra merah tersebut. Pantulan dari cahaya infra merah tersebut dikonversi menjadi informasi kedalaman (*depth data*) yang berisi jarak antara obyek dengan sensor.
- Sebuah *multi-array microphone* yang terdiri dari empat mikrofon untuk menangkap suara yang memungkinkan untuk merekam suara, mengetahui lokasi sumber gelombang suara dan arah sumber gelombang suara.
- Sebuah *3-axis accelerometer* yang dikonfigurasi untuk jangkauan 2G dimana G merupakan akselerasi terhadap gravitasi bumi. *Accelerometer* ini dapat digunakan untuk menentukan orientasi *kinect* saat ini.



Gambar 2. Kinect sensor versi 1

### 1) Color Data

*Kinect Sensor* memproduksi sebuah citra bertipe *color data* secara terus menerus pada interval tertentu dengan menggunakan *RGB camera* (Miles, 2012). Isi citra tersebut adalah kumpulan *pixel* dimana tiap *pixel* memiliki empat nilai warna dasar yang menunjukkan warna pada *pixel* tersebut. Empat nilai warna dasar tersebut adalah nilai merah (*red*), hijau (*green*), biru (*blue*) dan tingkat transparansi (*alpha*). Batas nilai setiap warna dasar adalah 0 hingga 255 dimana semakin tinggi nilai tersebut maka semakin kuat warna tersebut.

### 2) Depth Data

*Kinect Sensor* memproduksi sebuah citra bertipe *depth data* secara terus menerus pada interval tertentu dengan menggunakan kamera infra merah (Miles, 2012). Isi citra tersebut sama halnya dengan citra bertipe *color data*, hanya saja setiap *pixel* pada citra tersebut berisi sebuah nilai yang menunjukkan jarak dari kamera menuju sebuah titik obyek pada citra tersebut. Hal ini dapat terjadi karena *Kinect Sensor* memancarkan sebuah jaring yang berisi titik-titik infra merah ke obyek-obyek yang ada di depannya menggunakan kamera infra merah dan menangkap kembali hasil pantulan infra merah tersebut sehingga didapatkan sebuah jarak. Jarak yang diberikan adalah dalam milimeter dan dinyatakan dalam bentuk 13 bit. Dengan kata lain, *Kinect Sensor* dapat mencari kedalaman obyek yang ada di depannya hingga sekitar 8

meter ( $2^{13} = 8192$ , dalam milimeter). Namun, dalam kenyataannya, *Kinect Sensor* hanya mampu mendapatkan kedalaman obyek-obyek yang ada di depannya selama berada dalam jangkauan 0,8 meter hingga 4 meter. Hal ini disebabkan karena kamera infra merah yang digunakan tidak dapat fokus menangkap titik-titik infra merah pada obyek yang terlalu dekat atau terlalu jauh, sama halnya dengan mata kita yang tidak dapat menangkap jelas obyek yang terlalu dekat atau terlalu jauh dari mata kita.

### 3) Skeleton Data

Manusia terdiri dari 206 tulang (*bones*) yang saling terhubung bersama untuk menjadi sebuah kerangka (*skeleton*) yang dapat menyokong tubuh manusia (*body*) (Miles, 2012). Tulang-tulang tersebut memiliki sendi-sendi (*joints*) di antara mereka sehingga mereka dapat bergerak (misalnya sendi kaki, sendi siku dan sebagainya). *Kinect Sensor* menyederhanakan tubuh manusia menjadi 19 tulang yang saling terhubung dengan 20 sendi.

Dengan menggunakan *depth data* yang dihasilkan melalui kamera infra merah, *Kinect Sensor* mampu melakukan pelacakan tubuh (*body tracking*) pada para pemain yang berada di depan sensor. *Kinect Sensor* versi 1 mampu melacak tubuh pemain hingga enam pemain tetapi hanya mampu melacak tubuh secara lengkap pada dua pemain saja. Keempat sisa pemain hanya dapat dilacak secara terbatas saja karena membutuhkan biaya komputasi yang mahal. Untuk mendapatkan hasil pelacakan tubuh yang maksimal, disarankan untuk menghindari penggunaan pakaian yang bisa mengaburkan bentuk tubuh manusia (seperti penggunaan rok, jaket, topi longgar) karena salah satu parameter yang digunakan oleh *Kinect Sensor* untuk mendeteksi tubuh manusia adalah berdasarkan bentuk tubuh manusia.

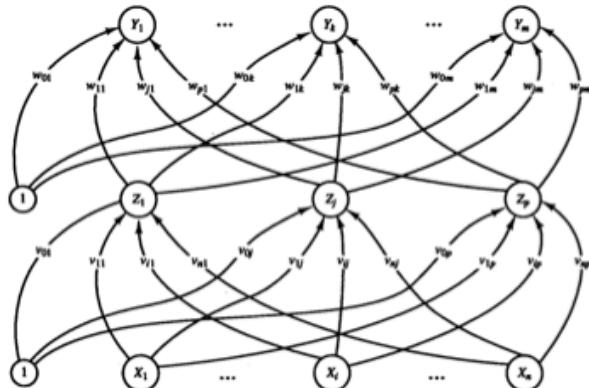
## C. Jaringan Syaraf Tiruan

Jaringan syaraf tiruan adalah sebuah sistem yang bersifat adaptif yang dapat belajar untuk melakukan sebuah fungsi (pemetaan data masukan dan data keluaran) berdasarkan data. Adaptif berarti parameter-parameter yang dimiliki sistem dapat berubah (disebut fase pelatihan). Setelah fase pelatihan selesai dimana parameter-parameter yang dimiliki sistem sudah ditentukan, maka sistem dapat digunakan untuk menyelesaikan permasalahan yang ada (disebut fase uji coba).

### 1) Jaringan Syaraf Tiruan Dengan Banyak Lapisan

Jaringan syaraf tiruan ini terdiri dari sebuah lapisan masukan (*input layer*), satu atau lebih lapisan tersembunyi (*hidden layer*) dan sebuah lapisan keluaran (*output layer*). Lapisan masukan terdiri dari  $n$  neuron. Lapisan tersembunyi terdiri dari  $p$  neuron. Lapisan keluaran terdiri dari  $m$  neuron. Penggunaan lebih dari satu lapisan tersembunyi memang terkadang menguntungkan dalam beberapa kasus, tetapi menggunakan satu lapisan

tersembunyi saja sudah cukup. Setiap neuron pada lapisan tersembunyi dan lapisan keluaran mempunyai sebuah bias. Bias ini sama seperti nilai bobot pada koneksi penghubung dimana nilainya selalu 1.



Gambar 3. Jaringan syaraf tiruan dengan banyak lapisan

Simbol dan istilah yang sering digunakan dalam jaringan syaraf tiruan dengan banyak lapisan dapat dilihat pada Tabel II.

TABEL II  
SIMBOL DAN ISTILAH PADA JARINGAN SYARAF TIRUAN  
DENGAN BANYAK LAPISAN

Simbol	Penjelasan
$\mathbf{x}$	Vektor masukan (digunakan pada fase pelatihan) $\mathbf{x} = (x_1, \dots, x_n)$
$\mathbf{t}$	Vektor keluaran yang merupakan target yang diharapkan (digunakan pada fase pelatihan) $\mathbf{t} = (t_1, \dots, t_m)$
$\delta_k$	Nilai koreksi kesalahan (error) untuk penyesuaian nilai bobot $w_{jk}$ pada neuron $Y_k$ pada lapisan keluaran. Informasi ini yang dirambatkan terbalik ke neuron-neuron yang berhubungan pada lapisan tersembunyi
$\delta_j$	Nilai koreksi kesalahan (error) untuk penyesuaian nilai bobot $v_{ij}$ dirambatkan terbalik dari lapisan keluaran menuju neuron-neuron yang berhubungan pada lapisan tersembunyi
$\alpha$	Laju pembelajaran (learning rate)
$X_i$	Neuron ke $-i$ pada lapisan masukan
$Y_k$	Neuron ke $-k$ pada lapisan keluaran
$Z_j$	Neuron ke $-j$ pada lapisan tersembunyi
$v_{0j}$	Nilai bias pada neuron ke $-j$ pada lapisan tersembunyi
$w_{0k}$	Nilai bias pada neuron ke $-k$ pada lapisan keluaran

Fungsi aktivasi yang dapat digunakan pada jaringan syaraf tiruan dengan banyak lapisan harus mempunyai sifat kontinu, dapat dibedakan (*differentiable*), dan tidak

turun secara monoton (*monotonically non-decreasing*). Dari sejumlah fungsi aktivasi yang ada, fungsi sigmoid biner (*binary sigmoid*) dan fungsi sigmoid bipolar (*bipolar sigmoid*). Berdasarkan fungsi aktivasi yang digunakan, sinyal masukan dan sinyal keluaran harus ditransformasi terlebih dahulu sehingga memiliki pola sinyal masukan dan sinyal keluaran yang sama.

## 2) Pelatihan Jaringan Syaraf Tiruan Dengan Perambatan Balik

Jaringan syaraf tiruan yang telah dibuat harus diberikan pelatihan agar dapat menjadi jaringan syaraf tiruan yang memiliki kepintaran dalam mengklasifikasikan data masukan baru. Kepintaran jaringan syaraf tiruan berada pada nilai bobot pada setiap koneksi penghubung antar neuron di semua lapisan. Jaringan syaraf tiruan yang telah dilatih akan menghasilkan perubahan nilai bobot pada setiap koneksi penghubung antar neuron di semua lapisan.

Beberapa parameter yang mempengaruhi tingkat keberhasilan tahap pelatihan pada jaringan syaraf tiruan adalah inisialisasi bobot, laju pelatihan, momentum, jumlah lapisan tersembunyi, lama pelatihan dan jumlah data pelatihan.

Algoritma perambatan balik (*backpropagation*) adalah salah satu algoritma yang dapat digunakan untuk melatih sebuah jaringan syaraf tiruan banyak lapisan dimana algoritma ini akan membandingkan nilai keluaran yang dihasilkan oleh jaringan syaraf tiruan (*actual output*) dengan nilai keluaran yang seharusnya (*expected output*) untuk menghasilkan *Mean Squared Error (MSE)*. Nilai *MSE* akan digunakan untuk memperbaiki nilai bobot pada setiap koneksi penghubung.

Algoritma ini terdiri dari tiga fase yaitu fase perambatan maju (*feedforward*) terhadap data masukan, fase perambatan terbalik terhadap kesalahan yang dihasilkan, dan fase penyesuaian kembali nilai bobot dan bias. Nilai bobot diisi secara acak pada waktu inisialisasi jaringan. Algoritma pelatihan ini merupakan *supervised training* karena disediakan sejumlah vektor masukan atau pola masukan dimana setiap masukan itu disertai juga dengan vektor keluaran atau pola keluaran.

Pada tahap perambatan maju, setiap neuron di lapisan masukan ( $X_i$ ) menerima sinyal dan meneruskan sinyal tersebut ke semua neuron di lapisan tersembunyi ( $Z_1, \dots, Z_p$ ). Kemudian, setiap neuron di lapisan tersembunyi itu ( $Z_i$ ) menghitung sinyal keluaran atau respon ke setiap neuron di lapisan keluaran ( $Y_1, \dots, Y_m$ ) dengan menggunakan fungsi aktivasi yang telah ditentukan. Kemudian, setiap neuron di lapisan keluaran itu ( $Y_i$ ) menghitung sinyal keluaran atau respon terhadap sinyal masukan yang diberikan ke jaringan tersebut dengan menggunakan fungsi aktivasi yang telah ditentukan.

Pada tahap perambatan balik, setiap neuron pada lapisan keluaran ( $Y_i$ ) membandingkan sinyal keluaran yang dihasilkan itu (nilai aktual) dengan nilai keluaran yang seharusnya (nilai yang diharapkan) untuk menentukan nilai kesalahan (*error delta*). Nilai keluaran yang seharusnya itu didapatkan dari data pelatihan dimana

data pelatihan terdiri dari sejumlah pasangan nilai masukan dan nilai keluaran. Berdasarkan nilai kesalahan tersebut, dapat dihitung nilai faktor  $\delta_k$  ( $k = 1, \dots, m$ ) yang digunakan untuk mendistribusikan terbalik nilai kesalahan pada neuron tersebut di lapisan keluaran ( $Y_k$ ) ke semua neuron di lapisan sebelumnya (lapisan tersembunyi) yang terhubung dengannya. Nantinya, nilai faktor  $\delta_k$  juga digunakan untuk menyesuaikan nilai bobot pada koneksi penghubung dan nilai bias antara lapisan tersembunyi dengan lapisan keluaran. Dengan cara yang sama pula, dapat dihitung nilai faktor  $\delta_j$  ( $j = 1, \dots, p$ ) yang digunakan untuk mendistribusikan terbalik nilai kesalahan pada neuron tersebut di lapisan tersembunyi ( $Z_j$ ) ke semua neuron di lapisan sebelumnya (lapisan masukan) yang terhubung dengannya. Memang tidak begitu penting untuk mendistribusikan terbalik nilai kesalahan ke semua neuron di lapisan masukan, tetapi faktor  $\delta_j$  digunakan untuk menyesuaikan nilai bobot pada koneksi penghubung dan nilai bias antara lapisan tersembunyi dengan lapisan keluaran.

Pada tahap penyesuaian kembali nilai bobot dan bias, semua nilai bobot pada koneksi penghubung dihitung lagi sesuai dengan nilai faktor  $\delta$  yang didapatkan secara serentak. Nilai bobot pada koneksi penghubung dan bias pada setiap neuron di lapisan tersembunyi menuju lapisan keluaran didasarkan pada nilai faktor  $\delta_k$  dan sinyal keluaran yang dihasilkan dari fungsi aktivasi. Nilai bobot pada koneksi penghubung dan bias pada setiap neuron di lapisan masukan menuju lapisan tersembunyi didasarkan pada nilai faktor  $\delta_j$  dan sinyal keluaran yang dihasilkan dari fungsi aktivasi.

### III. METODA PENELITIAN

Bagian ini menjelaskan tahapan-tahapan yang dilakukan dalam penelitian ini dalam urutan yang sistematis, termasuk di dalamnya membahas metoda yang digunakan untuk mengenali bahasa isyarat pada latar belakang kompleks sehingga penelitian dapat mencapai tujuan penelitian dalam waktu yang telah ditetapkan. Tahapan-tahapan pada metodologi penelitian dapat dilihat pada Gambar 4.

#### A. Tahap Persiapan Penelitian

Penelitian ini dimulai dengan mengamati cara komunikasi antara para penderita tuna rungu dan tuna wicara dengan orang normal sehingga mampu mengidentifikasi permasalahan yang terjadi sebagai latar belakang penelitian ini dan dirumuskan dalam sebuah rumusan masalah yang akan dijawab melalui penelitian ini. Dipaparkan juga tujuan dan manfaat serta batasan masalah dalam penelitian ini karena keterbatasan waktu dan sumber daya yang dimiliki peneliti.

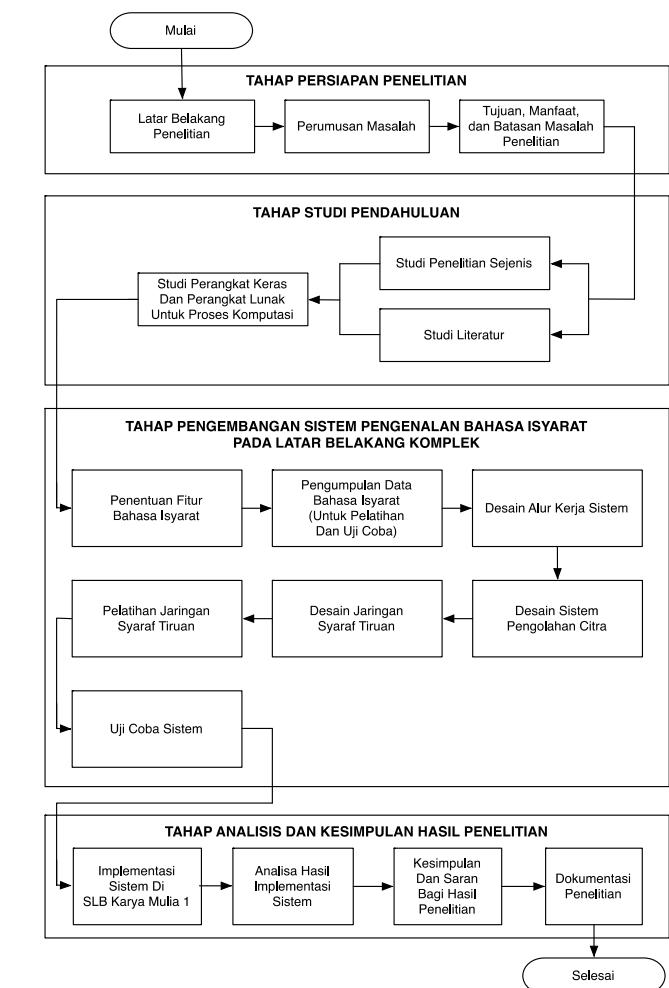
#### B. Tahap Studi Pendahuluan

Untuk menunjang penelitian ini, dilakukan beberapa studi pendahuluan agar mendapatkan fakta-fakta yang dapat memperkuat latar belakang, mempertajam

perumusan masalah, memantapkan tujuan dan manfaat serta memfokuskan ruang lingkup penelitian. Studi pendahuluan ini meliputi studi penelitian sejenis, studi literatur dan studi perangkat lunak untuk proses komputasi.

#### 1) Studi Penelitian Sejenis

Dalam mengenali bahasa isyarat, dibutuhkan sejumlah fitur-fitur dari anggota tubuh manusia yang dapat membedakan bahasa isyarat yang satu dengan bahasa isyarat yang lain. Anggota tubuh yang dimaksud meliputi penampil (tangan atau bagian tangan yang digunakan untuk membentuk isyarat), posisi (kedudukan tangan atau kedua tangan terhadap pengisian pada waktu berisyarat), tempat (bagian badan yang menjadi tempat awal isyarat dibentuk atau arah akhir isyarat), arah (gerak penampil ketika isyarat dibuat), frekuensi (jumlah gerak yang dilakukan pada waktu isyarat dibentuk), mimik muka, gerak tubuh, kecepatan gerak, dan kelenturan gerak.



Gambar 4. Tahapan metodologi penelitian

Penelitian tentang pengenalan bahasa isyarat telah menjadi salah satu topik penelitian yang diminati sejak dahulu dan masih menjadi sebuah tantangan yang menarik untuk diteliti lebih dalam lagi hingga sekarang karena

pengaruh kecepatan gerakan tangan, tingginya tingkat kebebasan gerakan (*high degrees of freedom*) dan tumpang tindihnya antar telapak tangan atau bagian tubuh yang lain (seperti wajah). Berdasarkan jenis perangkat yang digunakan untuk mengambil fitur-fitur dari anggota tubuh manusia, sistem pengenalan bahasa isyarat yang dilakukan dalam penelitian-penelitian terdahulu dapat dibagi menjadi tiga jenis yaitu sistem pengenalan bahasa isyarat yang menggunakan perangkat keras khusus (seperti sarung tangan robotik atau *data glove with flex sensor*), sistem pengenalan bahasa isyarat yang menggunakan *web camera* atau perangkat sejenisnya dan sistem pengenalan bahasa isyarat yang menggunakan *motion sensing input device* (seperti *Kinect Sensor*, *PlayStation Eye*, *Asus Xtion PRO LIVE*, *Wii Remote*, *Leap Motion* atau perangkat sejenisnya).

Sistem pengenalan bahasa isyarat yang menggunakan perangkat keras khusus telah dilakukan sejak lama seperti yang dilakukan oleh Murakami dkk pada tahun 1991 dan Iqbal dkk pada tahun 2012. Penelitian-penelitian tersebut bertujuan untuk mengenali bahasa isyarat statis dan dinamis. Penelitian yang dilakukan oleh Murakami dkk bertujuan untuk mengenali bahasa isyarat Jepang. Penelitian yang dilakukan oleh Iqbal dkk bertujuan untuk mengenali bahasa isyarat Indonesia. Secara garis besar, sistem pengenalan bahasa isyarat ini mengambil data masukan dari sebuah perangkat tertentu yang nantinya akan diolah menjadi fitur-fitur yang mendeskripsikan beberapa anggota tubuh pemberi isyarat. Fitur-fitur tersebut selanjutnya digunakan pada proses klasifikasi sehingga menghasilkan data keluaran yaitu arti bahasa isyarat tersebut. Perangkat masukan utama (*input device*) yang digunakan adalah sarung tangan robotik (*data glove with flex sensor*) dimana data masukan yang didapatkan berupa nilai teukan tiap jari tangan. Perangkat masukan tambahan (*additional input device*) yang digunakan adalah *accelerometer sensor* dimana data masukan yang didapatkan berupa nilai pergerakan tangan dalam koordinat 3D. Proses komputasi yang terjadi pada sistem pengenalan bahasan isyarat ini hanyalah proses klasifikasi atau pencocokan pola. Tidak ada proses pengolahan citra digital. Beberapa metode klasifikasi atau pencocokan pola yang digunakan adalah *Recurrent Neural Network*, *Dynamic Time Warping (DTW)*.

Sistem pengenalan bahasa isyarat yang menggunakan *web camera* atau perangkat sejenisnya juga telah dilakukan sejak lama seperti yang dilakukan oleh Lungociu dkk dan Mekala dkk pada tahun 2011 serta Maraqa dkk dan Mapari dkk pada tahun 2012. Penelitian-penelitian tersebut bertujuan untuk mengenali bahasa isyarat statis dan dinamis. Penelitian yang dilakukan oleh Mapari dkk, Mekala dkk dan Lungocio dkk bertujuan untuk mengenali bahasa isyarat Amerika. Penelitian yang dilakukan oleh Maraka dkk bertujuan untuk mengenali bahasa isyarat Arab. Secara garis besar, sistem pengenalan bahasa isyarat ini mengambil data masukan dari sebuah perangkat tertentu yang nantinya akan diolah menjadi fitur-fitur yang mendeskripsikan beberapa anggota tubuh

pemberi isyarat. Fitur-fitur tersebut selanjutnya digunakan pada proses klasifikasi sehingga menghasilkan data keluaran yaitu arti bahasa isyarat tersebut. Perangkat masukan utama (*input device*) yang digunakan adalah *web camera* atau perangkat sejenisnya dimana data masukan yang didapatkan berupa kumpulan citra bertipe RGB data (*color data*) yang berurutan. Pada penelitian yang dilakukan oleh Maraqa dkk, perangkat masukan tambahan (*additional input device*) yang digunakan adalah *colored glove* (sarung tangan yang ujung tiap jarinya diberi warna yang berbeda-beda) dimana berguna bagi proses pendektsian dan penamaan tiap jari tangan pemberi isyarat sehingga proses komputasi menjadi lebih murah dan lebih akurat dalam mendekripsi tiap jari tangan pemberi isyarat. Proses komputasi yang terjadi pada sistem pengenalan bahasan isyarat ini adalah proses pengolahan citra digital dan proses klasifikasi. Proses pengolahan citra digital berfungsi untuk mengolah data masukan menjadi fitur-fitur yang digunakan sebagai data masukan pada proses klasifikasi. Pada umumnya, proses pengolahan citra digital yang dilakukan meliputi *pre-processing*, *hand detection*, *hand segmentation*, *feature extraction*. Pada penelitian yang dilakukan oleh Maraqa dkk dan Mekala dkk, ditambahkan juga *fingertips detection* dan *fingertips labelling* pada proses pengolahan citra digital. Metode klasifikasi yang digunakan adalah *Support Vector Machine (SVM)*, *Recurrent Neural Network*, *Combinational Neural Network*, *Backpropagation Neural Network*.

Sistem pengenalan bahasa isyarat yang menggunakan *motion sensing input device* baru dilakukan beberapa tahun terakhir seperti yang dilakukan oleh Chai dkk dan Gunawan dkk pada tahun 2013 serta Nowicki dkk pada tahun 2014 seiring dengan munculnya perangkat tersebut. Penelitian-penelitian tersebut bertujuan untuk mengenali bahasa isyarat statis dan dinamis. Penelitian yang dilakukan oleh Nowicki dkk bertujuan untuk mengenali bahasa isyarat Amerika. Penelitian yang dilakukan oleh Gunawan dkk bertujuan untuk mengenali bahasa isyarat Indonesia. Penelitian yang dilakukan oleh Chai dkk bertujuan untuk mengenali bahasa isyarat Cina. Secara garis besar, sistem pengenalan bahasa isyarat ini mengambil data masukan dari sebuah perangkat tertentu yang nantinya akan diolah menjadi fitur-fitur yang mendeskripsikan beberapa anggota tubuh pemberi isyarat. Fitur-fitur tersebut selanjutnya digunakan pada proses klasifikasi atau pencocokan pola sehingga menghasilkan data keluaran yaitu arti bahasa isyarat tersebut. Perangkat masukan utama (*input device*) yang digunakan adalah *Leap Motion* (Nowicki, Pilarczyk, Wasikowski, & Zjawin, 2014) atau *Kinect Sensor* (Gunawan & Salim, 2013, Chai et al., 2013) dimana data masukan yang didapatkan berupa kumpulan leap frame data yang berurutan (jika menggunakan *Leap Motion*) atau kumpulan *depth data* dan *skeleton data* yang berurutan (jika menggunakan *Kinect Sensor*). Proses komputasi yang terjadi pada sistem pengenalan bahasan isyarat ini hanyalah proses klasifikasi atau pencocokan pola. Tidak ada proses pengolahan citra

digital. Metode klasifikasi atau pencocokan pola yang digunakan adalah *Support Vector Machine (SVM)*, *K-Means* dan *Hidden Markov Models*, *Dynamic Time Warping (DTW)*, *Euclidean Distance*.

## 2) Studi Literatur

Studi literatur bertujuan untuk memberikan dasar yang kuat dan dapat dipertanggungjawabkan secara ilmiah dalam merancang sistem pengenalan bahasa isyarat sehingga dapat menjawab rumusan permasalahan yang telah ditentukan.

## 3) Studi Perangkat Keras Dan Perangkat Lunak Untuk Proses Komputasi

Studi perangkat keras dan perangkat lunak untuk proses komputasi bertujuan untuk membantu peneliti dalam menentukan perangkat keras dan perangkat lunak apa saja yang dibutuhkan untuk membangun sistem pengenalan bahasa isyarat dengan mempertimbangkan tingkat akurasi, kecepatan komputasi, harga dan beberapa aspek pertimbangan lain. Salah satu yang menjadi fokus dalam penelitian ini adalah menghindari penggunaan alat tambahan khusus seperti sarung tangan robotik (robotic glove) atau sarung tangan yang diberi warna khusus dalam mengambil bahasa isyarat.

Perangkat keras yang dimaksud adalah perangkat keras apa saja yang dibutuhkan untuk melakukan proses pengambilan gambar bahasa isyarat sebagai data masukan (*input device*) serta menampilkan arti bahasa isyarat sebagai data keluaran (*output device*). Perangkat lunak yang dimaksud adalah bahasa pemrograman dan *library* apa saja yang digunakan untuk melakukan proses komputasi (proses pengolahan citra digital dan proses klasifikasi bahasa isyarat). Perangkat keras yang dapat digunakan untuk melakukan proses pengambilan gambar bahasa isyarat sebagai data masukan (*input device*) adalah *motion sensing input device* seperti *Kinect Sensor*, *PlayStation Eye*, *Asus Xtion PRO LIVE*, *Wii Remote*, *Leap Motion* atau perangkat sejenisnya.

TABEL III

TABEL PERBANDINGAN BEBERAPA PERANGKAT KERAS *MOTION SENSING INPUT DEVICE*

Parameter Perbandingan	<i>Kinect Sensor Version 1</i>	<i>Asus Xtion PRO LIVE</i>	<i>Leap Motion</i>
Vendor	Microsoft	Asus	Leap Motion Inc.
<i>RGB Camera</i>	✓	✓	✗
<i>Depth Camera</i>	✓	✓	✓
<i>Microphones</i>	✓	✓	✗
Fitur <i>Hand Tracking</i>	✓	✓	✓
Fitur <i>Finger Tracking</i>	✗	✗	✓
Fitur <i>Skeleton Tracking</i>	✓	✓	✗
Fitur <i>Face Tracking</i>	✓	✓	✗
Fitur <i>Voice Recognition</i>	✓	✓	✗
View Range	57°Horizontal, 43°Vertikal	58°Horizontal, 45°Vertikal, 70°Diagonal	
Harga	\$ 216	\$ 169.99	\$ 79.99
Ukuran Perangkat	Cukup besar	Cukup besar	Sangat Kecil
Persiapan Installasi Perangkat	Lama untuk installasi	Lama untuk installasi	Cepat untuk installasi
<i>Power Adapter</i>	<i>Dedicated power cord + USB</i>	USB	USB
Operating System	Windows	Windows, Android, Linux	Windows, Mac

Perangkat lunak yang dapat digunakan untuk melakukan proses komputasi (proses pengolahan citra digital dan proses klasifikasi bahasa isyarat) adalah *MATLAB* dan *OpenCV*. Dari beberapa pilihan perangkat lunak tersebut, *OpenCV* yang dipilih untuk digunakan pada sistem pengenalan bahasa isyarat ini. Pemrosesan citra digital menggunakan *MATLAB* relatif lebih lambat tetapi lebih cepat dan mudah dalam melakukan pemrograman. Pemrosesan citra digital menggunakan *OpenCV* relatif lebih cepat dan dapat digunakan dalam penerapan aplikasi mandiri tetapi lebih susah dalam melakukan pemrograman. Bahasa pemrograman yang digunakan dalam pengembangan sistem pengenalan bahasa isyarat ini adalah bahasa *C#* pada *Visual Studio 2010 Express*. Dengan kata lain, dibutuhkan *Kinect SDK (Software Development Kit)* untuk mengakses *Kinect Sensor versi 1* dan *EmguCV (OpenCV untuk versi C#)* untuk melakukan proses pengolahan citra digital.

## C. Tahap Pengembangan Sistem Pengenalan Bahasa Isyarat Pada Latar Belakang Kompleks

Tahap ini merupakan tahap inti dari penelitian ini yang meliputi penentuan fitur bahasa isyarat, pengumpulan data bahasa isyarat (untuk pelatihan dan uji coba), desain alur kerja sistem, desain sistem pengolahan citra, desain

jaringan syaraf tiruan, pelatihan jaringan syaraf tiruan, uji coba sistem

### 1) Penentuan Fitur Bahasa Isyarat

Fitur merupakan atribut penting dari sebuah obyek yang dapat mendeskripsikan, membedakan dan mengenali obyek tersebut dibandingkan obyek lain. Fitur-fitur tersebut akan digunakan sebagai data masukan pada proses klasifikasi untuk mengenali bahasa isyarat tersebut. Pemilihan fitur bergantung pada sifat obyek itu sendiri (seperti warna, tekstur, bentuk, ukuran, pergerakan dan sebagainya) dan kondisi lingkungan dari obyek yang akan dikenali (seperti kondisi latar belakang obyek, kehadiran obyek lain yang tidak diinginkan, iluminasi dan sebagainya). Pemilihan fitur yang tepat akan mengurangi kompleksitas komputasi (karena melibatkan faktor konsumsi waktu dan memori) dan menghasilkan tingkat akurasi pengenalan yang baik. Semakin banyak bahasa isyarat yang dapat dikenali oleh sistem pengenalan bahasa isyarat dan perbedaan yang tipis antara bahasa isyarat yang satu dengan bahasa isyarat yang lain, maka dibutuhkan lebih banyak fitur agar dapat mengenali bahasa isyarat dengan tepat.

Ada dua hal penting dalam mengenali bahasa isyarat yang dapat digunakan dalam menentukan fitur-fitur yang akan diekstrak yaitu bentuk tangan (*hand shape*) dan pergerakan tangan (*hand movement*) (Mekala, Gao, Fan, & Davari, 2011). Warna tangan tidak dianggap sebagai fitur dikarenakan warna tangan setiap manusia berbeda dan perbedaan warna tangan tidak mempengaruhi arti sebuah bahasa isyarat. Tekstur dan ukuran tangan juga tidak dianggap sebagai fitur dikarenakan perbedaan tekstur dan ukuran tangan tidak mempengaruhi arti sebuah bahasa isyarat.

Berdasarkan observasi, eksperimen dan studi penelitian-penelitian sebelumnya, maka ada dua fitur yang digunakan dalam mengenali bahasa isyarat yaitu bentuk tangan (*hand shape*) dan gerakan tangan (*hand movement*). Kedua fitur tersebut dapat dilakukan dengan menggunakan *depth image* dan *skeleton image*. Dengan kata lain, tidak melibatkan *RGB image* sama sekali

### 2) Pengumpulan Data Bahasa Isyarat (Untuk Pelatihan Dan Uji Coba)

Penyajian bahasa isyarat dilakukan oleh empat orang pemberi bahasa isyarat yaitu dua orang penderita tuna rungu dan tuna wicara dari SLB Karya Mulia 1 dan dua orang normal (1 orang pria dan 1 orang wanita) dengan umur 20-30 tahun. Sebelum pengambilan bahasa isyarat dilakukan, pemberi bahasa isyarat diberikan pelatihan sederhana dalam melakukan gerakan bahasa isyarat agar kualitas gerakan yang dihasilkan seragam dan sesuai dengan pedoman bahasa isyarat SIBI. Pengambilan bahasa isyarat ini sengaja tidak hanya melibatkan penderita tuna rungu dan tuna wicara saja tetapi melibatkan juga orang normal dengan pertimbangan agar data bahasa isyarat yang didapatkan lebih obyektif (tidak

subyektif ke seorang pemberi bahasa isyarat saja) sehingga kemampuan jaringan syaraf tiruan dalam mengenali bahasa isyarat lebih baik dan mendekati realita (Kouichi, 2012).

Setiap pemberi bahasa isyarat melakukan 10 kali repetisi untuk setiap bahasa isyarat sehingga didapatkan 1.440 kumpulan *depth image* dan 1.440 kumpulan *skeleton image*. Dari total citra digital yang didapatkan, 80% digunakan untuk tahap pelatihan dan 20% digunakan untuk tahap uji coba jaringan syaraf tiruan.

### 3) Desain Alur Kerja Sistem

Secara garis besar, sistem pengenalan bahasa isyarat terdiri dari tiga bagian yaitu tahap masukan, tahap proses, dan tahap keluaran.

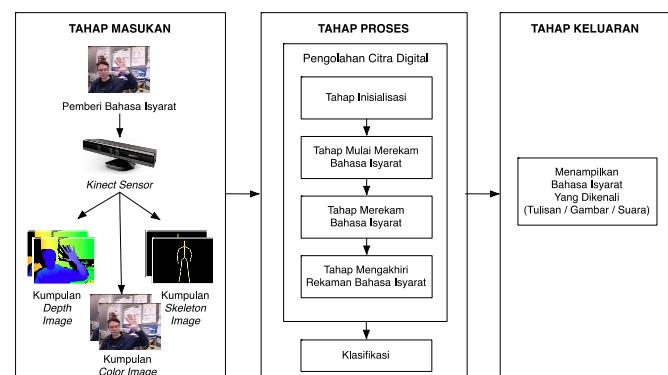
Tahap masukan adalah tahap pengambilan bahasa isyarat yang disajikan oleh pemberi bahasa isyarat menggunakan *Kinect Sensor* untuk mendapatkan kumpulan *depth image* dan kumpulan *skeleton image* untuk diproses pada tahap selanjutnya.

Tahap proses adalah tahap inti dari sistem pengenalan bahasa isyarat. Pada tahap ini, dilakukan pengolahan citra digital (tahap inisialisasi, tahap mulai merekam bahasa isyarat, tahap merekam bahasa isyarat, tahap mengakhiri rekaman bahasa isyarat) yang didapatkan pada tahap masukan untuk mendapatkan fitur-fitur yang akan digunakan sebagai data masukan pada proses klasifikasi (*classification*) untuk mengenali bahasa isyarat tersebut.

Tahap keluaran adalah tahap dimana sistem memberikan hasil bahasa isyarat yang dikenali pada tahap sebelumnya dan menampilkannya dalam bentuk tulisan, gambar atau suara.

### 4) Desain Sistem Pengolahan Citra

Secara garis besar, sistem pengolahan citra terdiri dari lima bagian yaitu tahap inisialisasi, tahap mulai merekam bahasa isyarat, tahap merekam bahasa isyarat, tahap mengakhiri rekaman bahasa isyarat. Desain sistem pengolahan citra dapat dilihat pada Gambar 6.



Gambar 5. Alur kerja sistem pengenalan bahasa isyarat

Tahap inisialisasi adalah tahap awal dalam pengolahan citra. Tahap ini berfungsi untuk mendeteksi apakah pemberi isyarat dalam keadaan siap untuk memberi

isyarat atau belum. Jika pemberi isyarat dalam keadaan siap, maka akan dilanjutkan ke tahap selanjutnya yaitu tahap mulai merekam bahasa isyarat. Untuk mendeteksi kesiapan pemberi isyarat, dilakukan beberapa proses pada yaitu :

- Deteksi jumlah pemberi syarat  
Proses ini bertujuan untuk mendeteksi jumlah pemberi isyarat. Jumlah pemberi isyarat yang diijinkan adalah seorang pemberi isyarat. Pendekripsi jumlah pemberi isyarat dapat dilakukan dengan menggunakan *skeleton image*.
- Deteksi tangan kanan  
Proses ini bertujuan untuk mendeteksi keberadaan tangan kanan pemberi isyarat. Pendekripsi tangan kanan dapat dilakukan dengan menggunakan *skeleton image*.
- Deteksi jarak tangan kanan ke kamera  
Proses ini bertujuan untuk mendeteksi apakah pemberi isyarat berada dalam jarak yang cukup ke kamera apa tidak. Dikatakan berada dalam jarak yang cukup jika pemberi isyarat berada antara 1,8 meter hingga 2 meter dari depan kamera (dimana didapatkan dari penelitian-penelitian sebelumnya). Pendekripsi jarak tangan kanan ke kamera dapat dilakukan dengan menggunakan *skeleton image*.
- Deteksi posisi tangan kanan di depan tubuh  
Proses ini bertujuan untuk mendeteksi apakah tangan kanan berada di depan tubuh apa tidak. Alasan yang mendasari mengapa tangan kanan harus berada di depan tubuh adalah asumsi bahwa tangan kanan adalah bagian tubuh yang paling dekat dengan kamera dimana hal ini dibutuhkan untuk memudahkan proses ekstraksi tangan kanan pada tahap selanjutnya. Pendekripsi posisi tangan kanan dapat dilakukan dengan menggunakan *skeleton image*.
- Deteksi apakah tangan kanan berhenti bergerak  
Proses ini bertujuan untuk mendeteksi apakah tangan kanan berhenti bergerak sehingga dapat dilanjutkan pada tahap selanjutnya yaitu tahap mulai merekam bahasa isyarat. Pendekripsi apakah tangan kanan berhenti bergerak dapat dilakukan dengan menggunakan ambang batas (*threshold*).

Tahap mulai merekam bahasa isyarat merupakan langkah awal untuk merekam bahasa isyarat dimana berguna bagi proses ekstraksi fitur gerakan tangan. Tahap ini berfungsi untuk mendeteksi posisi awal tangan kanan kemudian disimpan dalam sebuah penampung. Dengan kata lain, tahap ini menyimpan *frame* pertama dari kumpulan *skeleton image*. Pendekripsi posisi awal tangan kanan dapat dilakukan dengan menggunakan *skeleton image*.

Tahap merekam bahasa isyarat merupakan langkah untuk menyimpan arah gerakan tangan dalam kurun waktu tertentu (hingga gerakan isyarat berhenti) ke dalam sebuah larik penampung arah gerakan tangan dimana berguna bagi proses ekstraksi fitur gerakan tangan. Apabila

kapasitas larik penampung arah gerakan tangan sudah penuh, maka arah gerakan tangan di urutan awal akan dihilangkan dan dilakukan pergeseran mundur sehingga dapat menyimpan arah gerakan tangan yang terbaru. Jika pemberi isyarat tidak lagi bergerak, maka akan dilanjutkan ke tahap selanjutnya yaitu tahap mengakhiri merekam bahasa isyarat. Untuk merekam bahasa isyarat, dilakukan beberapa proses pada tahap ini yaitu :

- Deteksi posisi tangan kanan saat ini  
Proses ini bertujuan untuk mendeteksi posisi tangan kanan saat ini kemudian disimpan dalam sebuah penampung. Dengan kata lain, tahap ini menyimpan frame kedua hingga akhir dari kumpulan *skeleton image*. Pendekripsi posisi tangan kanan saat ini dapat dilakukan dengan menggunakan *skeleton image*.
- Hitung arah gerakan tangan kanan saat ini  
Proses ini bertujuan untuk menghitung arah gerakan tangan kanan saat ini. Arah gerakan tangan kanan saat ini dapat didapatkan dengan membandingkan posisi tangan kanan saat ini dengan posisi tangan kanan pada frame sebelumnya. Arah gerakan tersebut dikonversi sesuai 20 arah gerakan tangan kemudian disimpan dalam sebuah larik penampung arah gerakan.
- Deteksi apakah tangan kanan berhenti bergerak  
Proses ini bertujuan untuk mendeteksi apakah tangan kanan berhenti bergerak sehingga dapat dilanjutkan pada tahap selanjutnya yaitu tahap mengakhiri rekaman bahasa isyarat. Pendekripsi apakah tangan kanan berhenti bergerak dapat dilakukan dengan menggunakan ambang batas (*threshold*).

Tahap mengakhiri rekaman bahasa isyarat merupakan tahap akhir dalam pengolahan citra. Tahap ini berfungsi untuk mengakhiri rekaman bahasa isyarat dimana berguna bagi proses ekstraksi fitur gerakan tangan dan fitur bentuk tangan. Fitur-fitur yang didapatkan pada tahap ini akan digunakan sebagai data masukan pada jaringan syaraf tiruan di proses selanjutnya yaitu klasifikasi (classification). Untuk mengakhiri rekaman bahasa isyarat, dilakukan beberapa proses pada tahap ini yaitu :

- Deteksi kategori bahasa isyarat  
Bertujuan untuk mendeteksi apakah bahasa isyarat ini adalah bahasa isyarat statis atau bahasa isyarat dinamis.
- Normalisasi urutan arah gerakan tangan kanan  
Proses ini bertujuan untuk menormalisasi kumpulan arah gerakan tangan yang disimpan dalam larik penampung menjadi 6 arah gerakan tangan.
- Segmentasi tangan kanan  
Proses ini bertujuan untuk mengekstraksi (segmentasi) tangan kanan dan lengah bawah dari latar belakang kompleks. Posisi tangan kanan diasumsikan adalah obyek yang paling dekat dengan kamera sehingga proses ekstraksi akan

lebih akurat jika dilakukan pada depth image dibandingkan dengan cara mengambil tangan dan lengan bawah berdasarkan warna kulit karena diversifikasi warna kulit manusia. Proses ekstrasi dilakukan dengan cara mengambil semua titik pada gambar yang tidak melebihi nilai ambang batas yaitu

$$F = \{(p|z(p) < z_0 + z_D\} \quad (1)$$

$F$  merupakan gambar tangan dan lengan bawah (*foreground*) yang akan diekstrak.  $(p|z(p))$  merupakan sebuah titik pada *depth image* pada koordinat  $z$  dengan nilai kedalaman sebesar  $z(p)$ .  $z_0$  adalah titik kedalaman terkecil atau yang paling dekat dengan kamera dari semua titik pada gambar tersebut dan  $z_D$  adalah 0.2 meter yang merupakan jarak antara tangan dengan lengan bawah. Segmentasi tangan kanan dapat dilakukan dengan menggunakan *depth image*.

- Normalisasi ukuran hasil segmentasi tangan kanan  
Proses ini bertujuan untuk menormalisasi ukuran hasil segmentasi tangan kanan ke ukuran yang sudah ditetapkan yaitu 256 piksel x 256 piksel (dimana didapatkan dari penelitian-penelitian sebelumnya). Normalisasi ukuran dilakukan secara proposisional sehingga tidak mengubah bentuk asli tangan kanan. Hasil normalisasi ukuran hasil segmentasi tangan adalah gambar tangan yang berada di posisi tengah (baik secara horizontal dan vertikal). Normalisasi ukuran hasil segmentasi tangan kanan ini dapat mengurangi kompleksitas komputasi dalam mengolah citra (waktu dan memori) pada proses selanjutnya.
- Penghalusan hasil segmentasi tangan kanan  
Proses ini bertujuan untuk menghaluskan hasil segmentasi tangan kanan agar mengurangi tingkat gangguan (*noise*) yang ada pada hasil segmentasi. Algoritma yang digunakan adalah algoritma *Gaussian Blur* dengan nilai sigma sebesar 2 (dimana didapatkan dari hasil eksperimen).
- Deteksi tepi tangan kanan  
Proses ini bertujuan untuk mendeteksi garis tepi (kontur) tangan kanan yang nantinya akan digunakan untuk menganalisa bentuk tangan kanan di proses selanjutnya. Algoritma yang digunakan adalah algoritma *Canny*. Hasil dari proses ini adalah sebuah vektor 2 dimensi yang berisi koordinat-koordinat  $x$  dan  $y$  yang mendefinisikan garis tepi tangan kanan.
- Hitung shape signature  
Proses ini bertujuan untuk menghitung *shape signature* pada garis tepi (kontur) tangan kanan yang nantinya akan digunakan pada proses selanjutnya. *Shape signature* ( $s$ ) adalah sebuah vektor 1 dimensi yang berisi karakter-karakter dari garis tepi, dalam hal ini adalah garis tepi tangan.

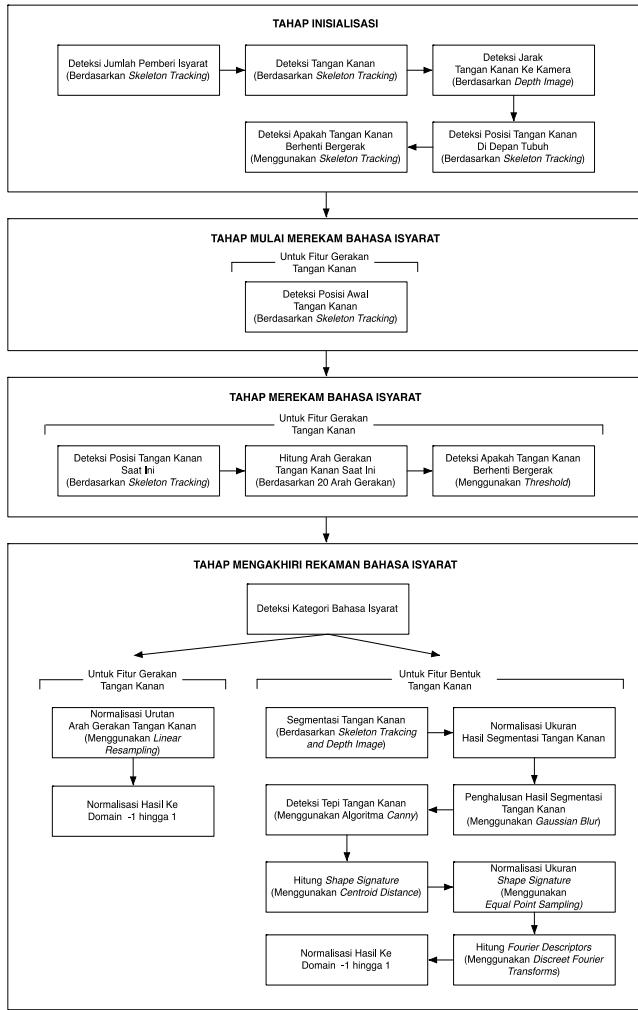
Ada empat tipe dari *shape signature* seperti yang dijelaskan di bab 2, yaitu *complex coordinates*, *centroid distance*, *curvature signature*, *cumulative angular function* (Lungociu, 2011). Proses ini menggunakan tipe *centroid distance* karena tipe ini dapat stabil terhadap masalah translasi dan penskalaan. Hasil dari proses ini adalah sebuah vektor 1 dimensi yang berisi sejumlah *shape signature*,  $s(t)$  dimana  $t = 0, 1, \dots, N$ .

- Normalisasi ukuran *shape signature*  
Proses ini bertujuan untuk menormalisasi ukuran *shape signature* yang dihasilkan pada proses sebelumnya menjadi sebuah angka dalam batas yang telah ditentukan. Normalisasi ukuran *shape signature* ini dapat mengurangi kompleksitas komputasi pada proses selanjutnya, memberikan perlakuan yang sama bagi berbagai tipe *shape signature*, dan menghilangkan detail bentuk garis tepi tangan kanan yang tidak signifikan.
- Hitung *Fourier Descriptor*  
Proses ini bertujuan untuk menghitung *fourier descriptors* ( $FD$ ) menggunakan *Discrete Fourier Transforms* berdasarkan *shape signature* yang telah dinormalisasi pada proses sebelumnya.
- Normalisasi hasil ke domain -1 hingga 1  
Proses ini bertujuan untuk melakukan normalisasi hasil ke domain -1 hingga 1. Hasil dari ekstraksi fitur bentuk tangan kanan dengan fitur gerakan tangan kanan memiliki domain yang berbeda. Hasil dari ekstraksi fitur bentuk tangan kanan memiliki domain 0 hingga 1. Hasil dari ekstraksi fitur gerakan tangan kanan memiliki domain 1 hingga 8. Perbedaan domain tersebut akan mempengaruhi hasil pengenalan pola sehingga perlu dinormalisasi ke domain yang sama.

##### 5) Desain Jaringan Syaraf Tiruan

Jaringan syaraf tiruan yang digunakan untuk mengenali bahasa isyarat adalah *backpropagation neural network*. Jaringan syaraf tiruan ini memiliki tiga lapisan yaitu sebuah lapisan masukan, sebuah lapisan tersembunyi dan sebuah lapisan keluaran.

Lapisan masukan merupakan lapisan yang terdiri dari 32 neuron dimana 16 neuron untuk fitur bentuk tangan kanan ( $X_1, X_2, \dots, X_{15}, X_{16}$ ) dan 16 neuron untuk fitur gerakan tangan kanan ( $X_{17}, X_{18}, \dots, X_{31}, X_{32}$ ). Data untuk 32 neuron tersebut merupakan hasil ekstraksi fitur pada proses sebelumnya.



Gambar 6. Desain sistem pengolahan citra

Lapisan tersembunyi merupakan lapisan yang terdiri dari sejumlah neuron ( $Z_1, Z_2, \dots$ ) dimana jumlah neuron tersebut didapatkan dari beberapa rumus di bab 2 dan selanjutnya dilakukan eksperimen untuk mendapatkan jumlah neuron yang akan menghasilkan pengenalan pola yang maksimal. Fungsi aktivasi yang digunakan pada lapisan ini adalah fungsi sigmoid biner. Nilai keluaran yang dihasilkan dari fungsi aktivasi tersebut adalah sebuah nilai desimal berkisar 0 hingga 1.

Lapisan keluaran merupakan lapisan yang terdiri dari 36 neuron dimana 26 neuron untuk huruf A hingga huruf Z ( $Y_1, Y_2, \dots, Y_{25}, Y_{26}$ ) dan 10 neuron untuk angka 0 hingga angka 9 ( $Y_{27}, Y_{28}, \dots, Y_{35}, Y_{36}$ ). Fungsi aktivasi yang digunakan pada lapisan ini adalah fungsi sigmoid biner. Nilai keluaran yang dihasilkan dari fungsi aktivasi tersebut adalah sebuah nilai desimal berkisar 0 hingga 1. Nilai keluaran tersebut selanjutnya diberi ambang batas pada nilai 0,8 agar menghasilkan nilai akhir 0 dan 1. Jika nilai keluaran  $\geq 0,8$  maka nilai akhirnya adalah 1. Jika nilai keluaran  $< 0,8$  maka nilai akhirnya adalah 0. Neuron yang menghasilkan nilai akhir 1 merupakan neuron yang terpilih sebagai hasil dari klasifikasi.

Pada jaringan syaraf tiruan ini, setiap neuron pada

lapisan tersembunyi dan lapisan keluaran mempunyai sebuah bias. Bias ini sama seperti nilai bobot pada koneksi penghubung dimana nilainya selalu 1.

#### 6) Pelatihan Jaringan Syaraf Tiruan

Algoritma pelatihan yang digunakan pada jaringan syaraf tiruan ini adalah algoritma *backpropagation*. Inisialisasi bobot pada koneksi penghubung menggunakan algoritma *Nguyen*. Penentuan laju pelatihan dan momentum akan dilakukan secara *trial-and-error*. Lama pelatihan akan ditentukan berdasarkan jumlah *epoch* dan *MSE*. Jumlah data pelatihan yang digunakan adalah 20% dari total data bahasa isyarat yang didapatkan pada proses sebelumnya. Bobot akhir yang dihasilkan pada pelatihan jaringan syaraf tiruan akan digunakan pada uji coba.

#### 7) Uji Coba Sistem

Uji coba sistem akan dilakukan pada lima orang pemberi bahasa isyarat yaitu seorang penderita tuna rungu dan tuna wicara (pria) dan dua orang normal (2 orang pria dan 2 orang wanita) dengan umur 20-30 tahun. Sebelum uji coba sistem dilakukan, pemberi bahasa isyarat diberikan pelatihan sederhana dalam melakukan gerakan bahasa isyarat agar kualitas gerakan yang dihasilkan seragam dan sesuai dengan pedoman bahasa isyarat SIBI. Uji coba sistem ini sengaja tidak hanya melibatkan penderita tuna rungu dan tuna wicara saja tetapi melibatkan juga orang normal untuk mengetahui seberapa besar tingkat akurasi pengenalan bahasa isyarat pada berbagai macam variasi data masukan.

Perangkat keras yang dipakai selama pengembangan sistem pengenalan bahasa isyarat ini adalah sebuah komputer dengan spesifikasi sebagai berikut *processor* 2,4 GHz Intel Core i5, *cache memory* 512 KB, *main memory* 4 GB, *graphic card* Intel HD Graphics 3000 384 MB, *hard drive* 500 GB, *operating system* Windows 7 64 bit.

Perangkat lunak yang dipakai selama pengembangan sistem pengenalan bahasa isyarat ini adalah *Visual C# 2010 Express* sebagai bahasa pemrograman, *Kinect SDK* untuk mengambil aliran data dari *Kinect Sensor*, dan *EmguCV library* yang digunakan untuk proses pengolahan citra.

## IV. HASIL PENELITIAN DAN PEMBAHASAN

Bab ini menjelaskan proses-proses yang dilakukan dalam penelitian ini untuk menganalisa dan mengambil kesimpulan. Proses-proses tersebut meliputi proses pengumpulan data bahasa isyarat untuk pelatihan dan uji coba jaringan syaraf tiruan, penentuan desain jaringan syaraf tiruan, proses pelatihan jaringan syaraf tiruan dengan beberapa macam struktur jaringan syaraf tiruan, uji coba jaringan syaraf tiruan dengan beberapa macam struktur jaringan syaraf tiruan, analisis hasil uji coba jaringan syaraf tiruan dan analisis manfaat dan penerimaan hasil implementasi yang dilakukan di SLB Karya Mulia 1.

### A. Proses Pengumpulan Data Bahasa Isyarat

Penyajian bahasa isyarat dilakukan oleh empat orang pemberi bahasa isyarat yaitu dua orang penderita tuna rungu dan tuna wicara (pria) dari SLB Karya Mulia 1 dan dua orang normal (1 orang pria dan 1 orang wanita) dengan umur 20-30 tahun. Sebelum pengambilan bahasa isyarat dilakukan, pemberi bahasa isyarat diberikan pelatihan sederhana dalam melakukan gerakan bahasa isyarat agar kualitas gerakan yang dihasilkan seragam dan sesuai dengan pedoman bahasa isyarat SIBI. Pengambilan bahasa isyarat ini sengaja tidak hanya melibatkan penderita tuna rungu dan tuna wicara saja tetapi melibatkan juga orang normal dengan pertimbangan agar data bahasa isyarat yang didapatkan lebih obyektif (tidak subyektif ke seorang pemberi bahasa isyarat saja) sehingga kemampuan jaringan syaraf tiruan dalam mengenali bahasa isyarat lebih baik dan mendekati realita (Kouichi, 2012).

Setiap pemberi bahasa isyarat melakukan 10 kali repetisi untuk setiap bahasa isyarat sehingga didapatkan 1.440 set data dimana setiap set data berisi 32 fitur (16 fitur bentuk tangan kanan dan 16 fitur gerakan tangan kanan). Dari 1.440 set data yang didapatkan, 80% digunakan untuk tahap pelatihan dan 20% digunakan untuk tahap uji coba jaringan syaraf tiruan.

### B. Proses Penentuan Desain Jaringan Syaraf Tiruan

Lapisan masukan merupakan lapisan yang terdiri dari 32 neuron dimana 16 neuron untuk fitur bentuk tangan kanan ( $X_1, X_2, \dots, X_{15}, X_{16}$ ) dan 16 neuron untuk fitur gerakan tangan kanan ( $X_{17}, X_{18}, \dots, X_{31}, X_{32}$ ). Data untuk 32 neuron tersebut merupakan hasil ekstraksi fitur pada proses sebelumnya.

Lapisan tersembunyi merupakan lapisan yang terdiri dari sejumlah neuron ( $Z_1, Z_2, \dots, Z_n$ ) dimana jumlah neuron tersebut didapatkan dari beberapa rumus di bab 2 dan selanjutnya dilakukan eksperimen untuk mendapatkan jumlah neuron yang dapat menghasilkan pengenalan pola yang maksimal. Fungsi aktivasi yang digunakan pada lapisan ini adalah fungsi sigmoid biner. Nilai keluaran yang dihasilkan dari fungsi aktivasi tersebut adalah sebuah nilai desimal berkisar 0 hingga 1.

Lapisan keluaran merupakan lapisan yang terdiri dari 36 neuron dimana 26 neuron untuk huruf A hingga huruf Z ( $Y_1, Y_2, \dots, Y_{25}, Y_{26}$ ) dan 10 neuron untuk angka 0 hingga angka 9 ( $Y_{27}, Y_{28}, \dots, Y_{35}, Y_{36}$ ). Fungsi aktivasi yang digunakan pada lapisan ini adalah fungsi sigmoid biner. Nilai keluaran yang dihasilkan dari fungsi aktivasi tersebut adalah sebuah nilai desimal berkisar 0 hingga 1. Nilai keluaran tersebut selanjutnya diberi ambang batas pada nilai 0,8 agar menghasilkan nilai akhir 0 dan 1. Jika nilai keluaran  $\geq 0,8$  maka nilai akhirnya adalah 1. Jika nilai keluaran  $< 0,8$  maka nilai akhirnya adalah 0. Neuron yang menghasilkan nilai akhir 1 merupakan neuron yang terpilih sebagai hasil dari klasifikasi.

Pada jaringan syaraf tiruan ini, setiap neuron pada lapisan tersembunyi dan lapisan keluaran mempunyai

sebuah bias. Bias ini sama seperti nilai bobot pada koneksi penghubung dimana nilainya selalu 1.

Proses pelatihan jaringan syaraf tiruan ini bertujuan untuk mengetahui struktur jaringan syaraf tiruan yang mana yang memiliki nilai bobot terbaik dari setiap struktur jaringan syaraf tiruan yang ada dalam mengenali bahasa isyarat termasuk jika diberikan data masukan baru. Berdasarkan penjelasan di bab 2 tentang penentuan jumlah neuron pada lapisan tersembunyi, digunakan sejumlah macam jaringan syaraf tiruan dengan struktur sebagai berikut :

**TABEL IV**  
MACAM-MACAM STRUKTUR JARINGAN SYARAF TIRUAN YANG  
DIGUNAKAN PADA PROSES PELATIHAN JARINGAN SYARAF TIRUAN

Struktur Jaringan	Penjelasan
32-33-36	32 neuron pada lapisan masukan, 33 neuron pada lapisan tersembunyi, 33 neuron pada lapisan keluaran
32-46-36	32 neuron pada lapisan masukan, 46 neuron pada lapisan tersembunyi, 36 neuron pada lapisan keluaran
32-60-36	32 neuron pada lapisan masukan, 60 neuron pada lapisan tersembunyi, 36 neuron pada lapisan keluaran

### C. Proses Pelatihan Jaringan Syaraf Tiruan

Proses pelatihan jaringan syaraf tiruan ditentukan oleh beberapa parameter sebagai berikut :

- Parameter inisialisasi bobot  
Proses pelatihan jaringan syaraf tiruan ini dimulai dengan inisialisasi bobot secara acak
- Parameter *Mean Square Error (MSE)*  
Nilai *Mean Square Error (MSE)* yang digunakan untuk proses pelatihan ini ada empat skenario yaitu “0,1”, “0,01”, “0,001”, “0,0001”. Apabila pada sebuah iterasi (epoch) pelatihan, jaringan syaraf tiruan tersebut telah mencapai nilai *MSE*, maka pelatihan jaringan syaraf tiruan akan diberhentikan.
- Parameter laju pelatihan (*learning rate*)  
Nilai laju pelatihan (*learning rate*) yang digunakan untuk proses pelatihan ini adalah 0,2. Apabila nilai ini terlalu besar maka proses pelatihan jaringan syaraf tiruan semakin cepat tetapi pelatihan menjadi tidak stabil sehingga berpotensi untuk mencapai kondisi *local minima*.
- Parameter momentum  
Nilai momentum yang digunakan untuk proses pelatihan ini adalah 0,5. Penggunaan momentum biasanya digunakan agar kondisi konvergensi berlangsung terlalu lama dan mencegah terjadinya *local minima*.
- Parameter lama pelatihan (*epoch*)  
Jumlah *epoch* maksimal yang digunakan untuk proses pelatihan ini adalah 1.000 *epoch*. Jumlah *epoch* ini disesuaikan pula dengan hasil pelatihan jika terjadi kondisi tertentu seperti belum tercapainya nilai *MSE (Mean Square Error)* atau

untuk mendeteksi apakah ada atau tidaknya *local minima*. Penyesuaian ini berupa penambahan jumlah *epoch* dalam proses pelatihan. Jika setelah penambahan jumlah *epoch* ternyata tidak mengakibatkan penurunan nilai *MSE* (*Mean Square Error*), maka ada kemungkinan terjadinya outlier pada data pelatihan.

#### D. Proses Uji Coba Jaringan Syaraf Tiruan

Uji coba sistem bertujuan untuk mengetahui tingkat pengenalan bahasa isyarat dari jaringan syaraf tiruan yang telah dibentuk apabila diberikan data masukan baru (selain data yang digunakan untuk tahap pelatihan). Uji coba sistem dilakukan dengan menggunakan data bahasa isyarat yang didapatkan pada tahap pengumpulan data (20% dari total data bahasa isyarat). Data bahasa isyarat akan dimasukkan pada sistem pengenalan bahasa isyarat sebagai data masukan dan akan menghasilkan data keluaran. Data keluaran tersebut selanjutnya akan dibandingkan dengan data keluaran yang seharusnya sehingga dapat diketahui kemampuan jaringan syaraf tiruan yang telah dibentuk dalam mengenali bahasa isyarat.

TABEL V  
HASIL UJI COBA JARINGAN SYARAF TIRUAN

Bahasa Isyarat	Tingkat Keseksian	Bahasa Isyarat	Tingkat Keseksian	Bahasa Isyarat	Tingkat Keseksian
A	24/40	M	23/40	Y	40/40
B	39/40	N	21/40	Z	39/40
C	39/40	O	38/40	0	38/40
D	27/40	P	38/40	1	18/40
E	28/40	Q	39/40	2	37/40
F	40/40	R	40/40	3	39/40
G	39/40	S	22/40	4	40/40
H	38/40	T	23/40	5	38/40
I	40/40	U	37/40	6	19/40
J	39/40	V	38/40	7	40/40
K	38/40	W	21/40	8	38/40
L	40/40	X	40/40	9	39/40
Total Tingkat Kesalahan = 15/1.440 = 15%					
Total Tingkat Akurasi = 100% - 15% = 85%					

#### E. Analisis Hasil Uji Coba Jaringan Syaraf Tiruan

Berdasarkan uji coba yang telah dilakukan, jaringan syaraf tiruan dengan 32 neuron pada lapisan masukan, 33 neuron pada lapisan tersembunyi, 33 neuron pada lapisan keluaran mampu mengenali bahasa isyarat dengan tingkat akurasi sebesar 85%.

Bahasa isyarat dinamis (huruf J dan huruf Z) dapat dikennenli dengan sempurna dikarenakan model pergerakan tangan untuk kedua huruf tersebut tidak mirip. Bahasa isyarat dinamis tidak berbenturan dengan bahasa isyarat statis dikarenakan perbedaan fitur yang digunakan dalam mengenali bahasa isyarat

Sebagian besar, bahasa isyarat statis dapat dikenali dengan cukup baik dikarenakan model pergerakan tangan antara huruf tersebut tidak mirip. Namun ada beberapa huruf dan angka yang belum bisa dikenali dengan maksimal dikarenakan kemiripan bentuk tangan pada bahasa isyarat tersebut. Huruf dan angka tersebut adalah A, D, E, M, N, S, T, W, 1, 6. Bentuk tangan untuk huruf A mirip dengan huruf E. Huruf M mirip dengan huruf N. Huruf S mirip dengan huruf T. Huruf D mirip dengan angka 1. Huruf W mirip dengan angka 6.

Sistem pengenalan bahasa isyarat ini mampu mengenali bahasa isyarat tanpa terbatas pada skala, rotasi, pencahayaan dan warna kulit pada data masukan yang diberikan oleh pemberi isyarat.

#### F. Analisis Manfaat Dan Penerimaan Hasil Implementasi

Setelah dilakukan uji coba jaringan syaraf tiruan, dilakukan analisa manfaat dan penerimaan hasil implementasi pada SLB Karya Mulia 1. Dilakukan wawancara terhadap seorang guru SLB dan dua orang tua dari siswa-siswi SLB Karya Mulia 1 perihal manfaat dan penerimaan aplikasi pengenalan bahasa isyarat ini. Wawancara dilakukan setelah guru dan orang tua siswa-siswi melakukan percobaan beberapa kali pada aplikasi pengenalan bahasa isyarat.

Berdasarkan hasil wawancara, didapatkan bahwa aplikasi pengenalan bahasa isyarat ini memiliki potensi manfaat yang dapat membantu siswa, orang tua dan guru dalam proses belajar namun perlu dilakukan beberapa pengembangan lanjutan seperti memasukkan unsur pembelajaran pada aplikasi. Unsur pembelajaran bisa pembelajaran formal dengan mengambil materi belajar dasar dalam mengenal bahasa isyarat yang dirupakan dalam permainan seperti mengenal angka, mengenal huruf atau belajar hitung menghitung sederhana.

Selain itu, diperlukan pertimbangan penggunaan sensor selain Kinect karena dinggap cukup susah untuk diimplementasikan dalam kegiatan proses belajar mengajar dikarenakan tingkat kompleksitas untuk installasi perangkat keras dan tingkat mobilitas perangkat.

#### V. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pembahasan pada bab sebelumnya, dapat ditarik kesimpulan sebagai berikut :

- Sistem pengenalan bahasa isyarat ini mampu mengenali bahasa isyarat tanpa terbatas pada skala, rotasi, pencahayaan dan warna kulit pada data masukan yang diberikan oleh pemberi isyarat.
- Proses segmentasi tangan dapat dilakukan dengan baik dikarena menggunakan depth data.
- Proses ekstraksi fitur bentuk tangan menggunakan Fourier Descriptor dapat menghasilkan fitur bentuk tangan dengan maksimal. Proses ekstraksi fitur pergerakan tangan menggunakan skeleton data dapat menghasilkan fitur pergerakan tangan dengan

- maksimal.
- Jaringan syaraf tiruan mampu mengenali bahasa isyarat statis dan dinamis dengan tingkat akurasi sebesar 85%.
  - Terdapat beberapa huruf yang tidak dapat dikenali dengan maksimal dikarenakan kemiripan bentuk tangan pada bahasa isyarat tersebut.
  - Penggunaan sensor Kinect dianggap cukup susah untuk diimplementasikan dalam kehidupan sehari-hari dikarenakan tingkat kompleksitas untuk installasi perangkat keras
- Berdasarkan hasil penelitian dan pembahasan pada bab sebelumnya, terdapat beberapa saran yang dapat dipertimbangkan sebagai masukan untuk penelitian berikutnya :
- Menambahkan fitur lain untuk menyelesaikan permasalahan huruf statis yang mirip dengan huruf lain agar tingkat akurasi pengenalan bahasa isyarat meningkat.
  - Mempertimbangkan penggunaan sensor selain Kinect agar lebih rendah tingkat kompleksitas penggunaannya agar dapat digunakan dalam kehidupan sehari-hari
  - Perlu dikaitkan dengan metode pembelajaran bahasa isyarat agar proses pembelajaran menjadi lebih menyenangkan

#### DAFTAR PUSTAKA

- [1] Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., & Chen, X. (2013). *Sign Language Recognition and Translation with Kinect*. Beijing: Institute of Computing Technology (CAS).
- [2] Departemen Pendidikan Nasional. (2002). *Kamus Sistem Isyarat Bahasa Indonesia*. Jakarta: Direktorat Pendidikan Luar Biasa Proyek Pengembangan Sistem Dan Standarisasi Pengelolaan Pendidikan Luar Biasa.
- [3] Gunawan, A., & Salim, A. (2013). *Pembelajaran Bahasa Isyarat Dengan Kinect Dan Metode Dynamic Time Warping*. Jakarta: Mathematics and Statistics Department, School of Computer Science, Binus University.
- [4] Iqbal, M., & Supriyati, E. (2012). *Ekstraksi Ciri Pada Pengenalan Sistem Isyarat Bahasa Indonesia Berbasis Sensor Flex Dan Accelerometer*. Seminar Nasional Embedded System, 1-8.
- [5] Lungociu, C. (2011). *Real Time Sign Language Recognition Using Artificial Neural Networks*. INFORMATICA, 56(4), 75-84.
- [6] Mapari, R., & Kharat, G. (2012). *Hand Gesture Recognition Using Neural Network*. International Journal of Computer Science and Network (IJCSN), 1(6), 56-60.
- [7] Maraqa, M., Al-Zboun, F., Dhyabat, M., & Abu Zitar, R. (2012). *Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks*. Journal of Intelligent Learning Systems and Applications (JILSA), 4(1), 41-52.
- [8] Mekala, P., Gao, Y., Fan, J., & Davari, A. (2011). *Real Time Sign Language Recognition based on Neural Network Architecture*. System Theory - SSST (IEEE 43rd Southeastern Symposium), 197-201.
- [9] Mekala, P., Zafer Erdogan, S., & Fan, J. (2010). *Automatic Object Recognition Using Combinational Neural Networks in Surveillance Networks*. 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010), 387-391.
- [10] Miles, R. (2012). *Start Here! Learn The Kinect API (1st ed.)*. USA: O'Reilly Media, Inc.
- [11] Murakami, K., & Taguchi, H. (1991). *Gesture Recognition Using Recurrent Neural Networks*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 237-242.
- [12] Nowicki, M., Pilarczyk, O., Wasikowski, J., & Zjawin, K. (2014). *Gesture Recognition Library For Leap Motion Controller*. Poland: Institute of Computing Science, Faculty of Computing, Poznan University of Technology.