International Journal of Advanced Science and Technology
Vol. 29, No. 7s, (2020), pp. 2037-2047

`

"MENMOZHI" Voice Perception Wheel Chair for Disabled People

Banumathi.A.C[1], Dr. E.Chandra[2],Dr. D. Sumathi[3]
[1] *Research Scholar, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamilnadu, India*
[2] *Professor & Head, Department of Computer Science, Bharathiar University, Coimbatore- 46, India*
[3] *Associate Professor, Dept. of CSE., Kuppam Engineering College, Kuppam, AP., India.*

Abstract

Speech Recognition has become inevitable for this day modern era Devices. The Physically disabled people are assisted with a Wheel chair which operates with voice perception. Our proposed design perceives the command to be given in Tamil grammar–based recognition unit called "Menmozhi" which means soft spoken in Tamil Language. The user can give various tamil commands like "Munnadisel", "PinnadiSel". There are two different types of command given to the wheel chair, firstly single Tamil Phoneme words and to make it move to short distance inside the room and another is to give continuous Tamil Phoneme words to interact with outside world (Internet). We have achieved a successful speech Recognition rate by Menmozhiup to 99.67% for movement commands and 97.12 % for Internet usage.

1 Introduction:

Speech Recognition is the process of making the machine to understand the words spoken by the humans. It has become inevitable task of this technical age. Future is going to be the dominant rule of the machines with speech conversion which reduces the time of the input.

Recent surveys shows that more than one billion people in the world live with some form of disability, of which nearly 200 million experience considerable difficulties in functioning. In the years ahead, disability will be an even greater concern because its prevalence is on the rise[1]. This is due to ageing populations and the higher risk of disability in older people as well as the global increase in chronic health conditions such as diabetes, cardiovascular disease, cancer and mental health disorders.The disabled people will have difficulties in performing regular activities. With the help of the voice controlled wheel chair the strain of the disabled people is very much reduced. It helps them to have an independent mobilization without depending on others help.

Our proposed project consists of a wheelchair with voice recognition unit, sensors to get away the hurdles and to avoid collision against the walls, furniture, or people and Raspberry pi controlled motors with the Ethernetconnection. Our speech recognition unit is called as "Menmozhi" which consists of Two Phases. Our Project's core process is the speech Recognition. There are two types of Speech Recognition such as Speaker Dependent and Speaker Independent. During the first phase it converts single phoneme Tamil words in to signals for the movement of the wheel chair. The input is given to the device using a MCVECGNUSB of professional 18 inchuni-directional, noise cancelling microphone. The best voice features are selected using the wavelet transformation and the input voice is trained and tested by Adaptive Neuro fuzzy inference system (ANFIS). When the data has to be enhanced for continuous Tamil words and for more difficult words it can be passed to the Deep Belief Network to create its own architecture.

2 The Related works

Various Studies have concluded that the movement, independent mobility includes powered wheel chair, walkers, crutches etc., for the disabled people. It helps them to be independent and does not rely on others for their day to day needs[1].This project uses arduino kit Microcontroller circuit and DC motors to create the movement of wheel chair and Ultrasonic Sensors to detect the hurdles in between wheelchair and the

ISSN: 2005-4238 IJAST

way of direction[2]. The goal of this smart wheelchair project is to enhance an ordinary powered wheelchair using sensors to perceive the wheelchair's surroundings, a speech interface to interpret commands. Intelligent wheelchair will play an important role in the future welfare society[3]. This paper describes the design of a smart, motorized, voice operated wheelchair using embedded system here in this project the microcontroller used is ATMEGA32 for receiver and ATMEGA162 for transmitter. This concept was taken in this paper to reduce the human efforts in driving[4].

3 Proposed Project.

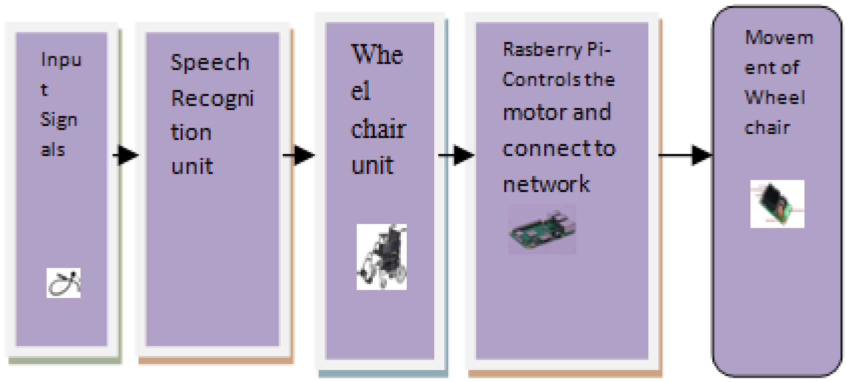


Fig 1: Wheel chair unit

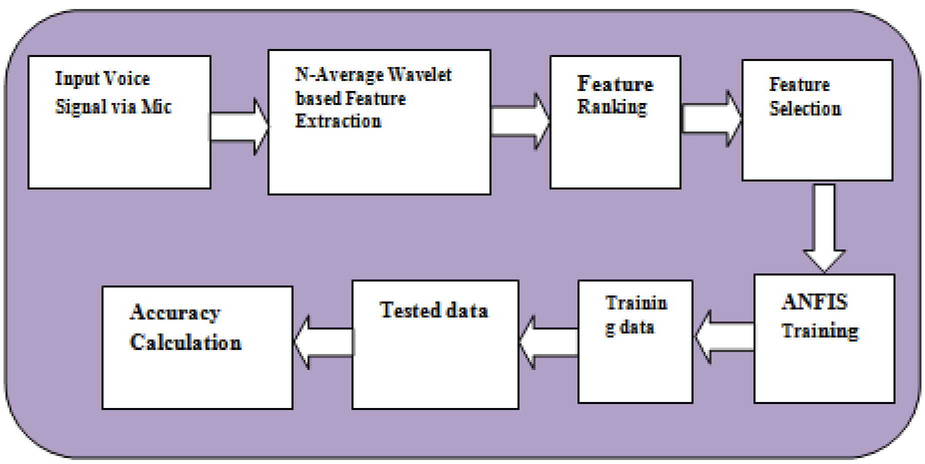


Fig 2: Proposed methodology of voice recognition unit using N-Average Algorithm for Isolated words.

ISSN: 2005-4238 IJAST

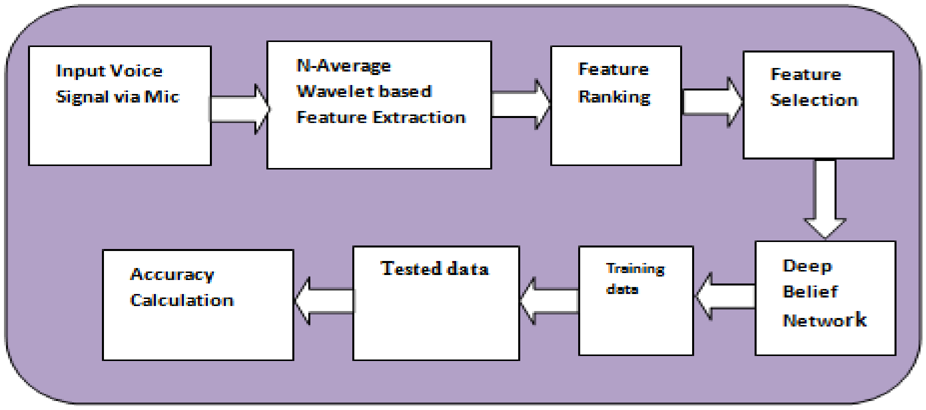


Fig 3: Proposed methodology of voice recognition unit using Deep Belief Network for Continuous words.

Our proposed research develops a voice operated wheelchair which uses the voice commands of single phoneme tamil words as the communicative media. The different commands in Tamil Language to the wheel chair are 1) Munnesel 2) Pinnesel 3) Siridhumunne 4) SiridhuPinne5) ValathupuramThirumbu6) EddhuPuramThirumbu 7) SuvarUllathu8)MuzuvathumThirumbu9)ValliyilThadai10) Samayal Arai Sel11) varaverpu Arai sel 12) vagamagasel13) MedhuvagaSel14) Meylasel15) Kelayvaravum16) SaindhuUttkaravedum17)Aaspathirikuallaippu. The commands to the Wheel Chair are divided into two types as indoor and outside commands. The commands 1 to 9 are movement commands inside a residence using a room locator. The remaining commands 10 to 17 are the commands used for outside environment.

S.No	Commands in Tamil	In English	Reaction
1	முன்னேசெல்	Munnesel	Moves forward 30cm
2	பின்னேசெல்	Pinnesel	Moves Backward 30cm
3	நிற்கவும்	Nirkavum	Stop
4	சிறிதுமுன்னே	Sirithumunne	Move forward 15cm
5	சிறிதுபின்னே	SirithuPinne	Move Backward 15cm
6	வலதுபுறம்திரும்பு	ValathupuramThirumbu	Move right Turn 90°
7	இடதுபுறம்திரும்பு	EdadhupuramThirumbu	Move Left Turn 90°
8	முழுவதும்திரும்பு	MuzuvathumThirumbu	Turn 180°
9	மெதுவாகசெல்	MedhuvagaSel	Moves forward 10cm
10	வேகமாகசெல்	vagamagasel	Moves forward 10cm

Table 1: Different single phoneme voice commands and its reaction

4 Wheel Chair construction

The components used in the wheel Chair Project

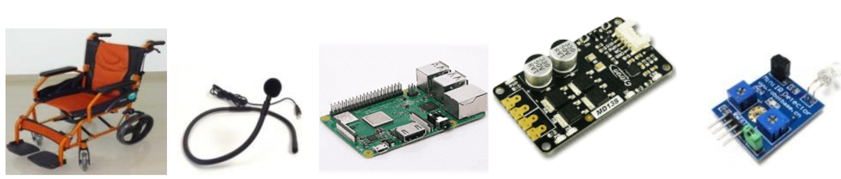


Fig 4 Wheel Chair, MicroPhone, RasberryPi, motor driver module,IR obstacle resistance sensor

ISSN: 2005-4238 IJAST

`

Our proposed project consists of a voice operated cost effective Wheel chair. It consists of cost effective items like Raspberry pi, Microphone to give the instruction, A Battery sensors and motors to move the wheels of the wheel chair and voice recognition unit to convert the given signals to the movement of the chair.

Our Voice recognition unit is called as "Menmozhi". It consists of the following unit which Extracts Feature from the given speech signals by the Application of N-Wavelet algorithm to find the best featured data from the given set of input data. Application of Adaptive Neuro Fuzzy Inference System(ANFIS). This ANFIS helps us to train the input signals and compare the result with the tested input signals and reduces the error to the maximum of 1.49875 e 00050.

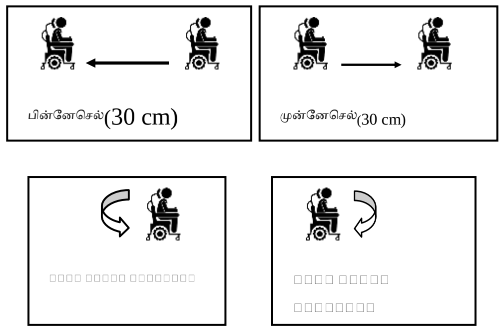


Fig 5. Different movement of wheel chair

4.1Voice Recognition Unit:

The input data is given through the microphone of sensitivity 2.0mV/Pa with frequency response of 100-10000Hz. It is used to convert the voice signal in to electrical signal. Output of the MIC is given to the voice recognition Module.

It is connected to a Raspberry Pi 3 Model B+. The processing of this unit is done with the Matlab2015R software. It consists of two units. One for the recognition of single phoneme Tamil Words done by the Adaptive Neuro Fuzzy Inference System(ANFIS) algorithm and another for the continuous Tamil Phoneme word using Deep Belief Network(DBN). The single tamil words are trained with the help of the ANFIS and tested for the accuracy and efficiency. The continuous tamil words are checked for efficiency using the Deep Belief Networks.

The motor drivers of capacity 12VDC are connected to the wheels and are connected to the Raspberry Pi. One Driver is sufficient. The speed of the motor is very important. If the rpm output of the motor is too high it will cause the fast movement of the wheel chair. At the same time low rotation causes the slow movement of the Wheel chair. we can maintain the speed of motor by giving the proper commands from the Processor Raspberry pi. A powerful battery of 12V battery is used for the motors to provide a sufficient power to the motors with a facility of charging the battery.

4.2Working of the Wheel Chair

Here we use a voice recognition module to control performance of the Wheel chair for the movement using motor drivers. Voice recognition module is trained by giving the 10 commands. The commands are converted into hex code with the Raspberry pi processor. When command is given the program in the corresponding processor sends the Hex code to the motors for the movement and thereby controls the movement or rotation of the motor. This is the basic working principle of the 'voice controlled wheel

ISSN: 2005-4238 IJAST

`

chair'. The battery and a battery charger unit is supported for power supply. The wheels of the Wheel chair can be moved avoiding all the obstacles with the help of the Ultrasonic sensors. The Raspberry pi processor connected to the voice unit and sensors gives the instruction to move and to avoid obstacles, stairs, walls etc., Raspberry pi acts as a core processor for connectinguser. it to the networking and to utilize the internet facility for the user. Room locator software is also installed to the Raspberry processor to give the correct route, and direction of the House or Campus. The environment is learned and the direction of the rooms is instructed to the Wheel chair for the Navigation. The batteries connected to motors can be recharged and used to move the motors. A Display unit can also be connected to see the instruction given to the Wheel Chair.

5 Speech Recognition unit

Our Speech Recognition unit is a Tamil grammar–based recognition unit called "Menmozhi" which means soft spoken in Tamil Language. It consists of two phases the one for single Phoneme words and another for continuous phoneme words. The single Phoneme input is given to the wheel chair through the microphone. It also consists of a Display unit to give the input through the buttons available. Several families of wavelets that have proven to be especially useful are included in this voice toolbox. Wavelet families are used for the feature extraction and feature selection process.There are fifty three types of wavelets available such as, Haar, Daubechies, Biorthogonal, Coiflets, Symlets, Morlet, Mexican Hat, Meyer, Other Real Wavelets, Complex Wavelets. Apart from the fifty three wavelets for our proposal we have selected haar, db1, sym3, coif1, bior3.1, rbio3.1 wavelets for the better feature values. From the 10 voice signal, Energy feature is extracted for all 6 wavelets. The feature values are normalized then ranked. The normalized feature values are sent to the Adaptive Neuro Fuzzy Interference System (ANFIS) algorithms and then the performance is analyzed[4]. A separate set of data of about 250 voices has been recorded for the Deep Belief Network and the Architectures are layered and the performance is tested.

5.1FEATURE SELECTION USING N-AVERAGE WAVELET

Step1: Acquiring energy feature values from each audio signal.

Step2: Normalizing the feature values by dividing the individual values by maximum value of each feature.

Let wavelet Features: F1,F2,F3....F53.

n=Number of audio signals considered for training.

$$\frac{F1(1,2\ldots n)}{MaxF1(1,2\ldots n)}$$

Step3:Taking minimum and maximum of individual feature values.

F_{min}=Min($F1(1,2\ldots n)$)

F_{max} =Max($F1(1,2\ldots n)$)

Energy values for 10 Input Voice Signals:

Probably the Max(F1(1,2...n)) will be 1 for all the feature values.

Step4: Taking difference between the $F_{min}$and F_{max}

ISSN: 2005-4238 IJAST

`

$F_{Diff=}F_{min}\sim F_{max}$

Haar	db1	sym3	coif1	bior3.1	rbio3.1	Target
9.7075	49.7075	6.5187	6.5277	5.3788	4.4455	1
9.6648	49.6648	6.5699	6.5729	5.2303	4.2141	2
49.9063	49.9063	6.5643	6.5816	5.4624	4.8700	3
9.6567	49.6567	6.5587	6.5616	5.2573	4.2136	4
9.5881	49.5881	6.5334	6.5360	5.3399	4.1096	5
9.8801	49.8801	16.596	6.6052	5.2923	4.7327	6
49.6616	49.6616	6.5835	6.5833	5.2088	24.104	7
9.8520	49.8520	16.591	6.5979	5.2181	4.8003	8
9.6191	49.6191	6.5574	6.5604	5.3138	3.9744	9
9.5987	49.5987	6.5235	6.5274	5.3216	4.1629	10

Table2: Feature values before Normalization

Haar	db1	sym3	coif1	bior3.1	rbio3.1	Target
.99601	.99601	.99534	.99532	.99671	.98293	1
.99516	.99516	.99843	.99805	.99088	.97362	2
1	1	0.99809	.99857	1	1	3
.99499	.99499	.99775	.99737	.99194	.97360	4
.99362	.99362	0.99623	.99583	.99518	.96942	5
.99947	.99947	1	1	0.99332	.99447	6
.99509	.99509	.99924	.99868	.99003	.96919	7
.99891	.99891	.99969	.99955	.99040	.99719	8
.99424	.99424	.99767	0.9973	.99416	.96398	9
.99383	.99383	.99563	.99531	.99447	.97156	10

Table 3: Feature values After Normalization(0-1 Scale)

Step5:Arranging the FDifffor all the 53 features in largest to smallest order and taking the first five maximum features

6.2 ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

ANFIS was introduced by Jang [5]. ANFIS is used for modeling, controlling, and parameter estimation intimation in complex systems [6]. ANFIS is a combination of artificial neural network (ANN) and fuzzy inference system (FIS). Combining the ANN and fuzzy-set theory can provide advantages and overcome the disadvantages in both techniques. The ANFIS model can be trained without relying solely on expert knowledge sufficient for a fuzzy logic model. The ANFIS model has the advantage of having both numerical and linguistic knowledge. ANFIS also uses the ANN's ability to classify data and identify patterns. Compared to the ANN, the ANFIS model is more transparent to the user and causes less memorization errors. Consequently, several advantages of the ANFIS exist, including its adaptation capability, nonlinear ability, and rapid learning capacity [7].

ISSN: 2005-4238 IJAST

`

ANFIS constructs a fuzzy inference system (FIS) whose membership function parameters are derived from the training examples. The most commonly used fuzzy inference systems are Mamdani and Sugeno. The main difference between Mamdani and Sugeno is that the output membership functions of the Sugeno system are either linear or constant [5]. However, the output membership functions of the Mamdani system can be triangular, Gaussian, etc. In this study, the Sugeno-type fuzzy inference system was used, because the Sugeno-type system is more computationally efficient than the Mamdani type. The Mamdani type is more reliant on expert knowledge. However, the Sugeno type is trained by real data.

In order to explain the ANFIS architecture, we assumed that there are two inputs: x and y. Two fuzzy if–then rules for a first-order Sugeno fuzzy model can be expressed as follows:

Rule 1: If x is A_1 and y is B_1, then $f_1 = p_1 x + q_1 y + r_1$,

Rule 2: If x is A_2 and y is B_2, then $f_2 = p_2 x + q_2 y + r_2$,

where Ai and Bi are the fuzzy sets, fi is the output, and pi , qi , and ri are the design parameters that are determined during the training process. The ANFIS architecture used to implement the two rules is shown in Figure 3.

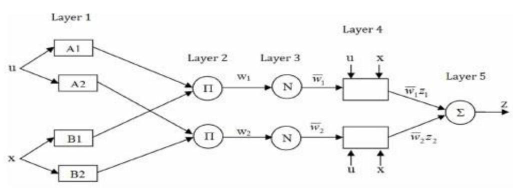
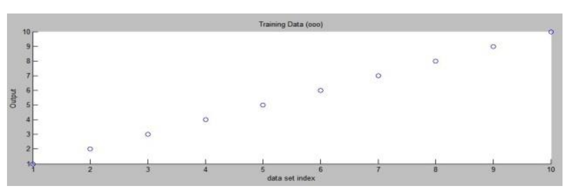


Figure 6. The architecture of the adaptive neuro-fuzzy inference system (ANFIS) model with two inputs, one output, and two rules

The selected data is loaded for training and the Fuzzy Inference system is generated sub clustering algorithm. With 100 epoch of training the error rate comes down to a minimum of Epoch 100:error= 2.9187e-006 .



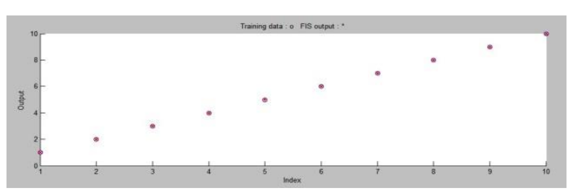


Fig5: Training data Loaded Fig:6 Trained Data

5.3 Deep Belief Network

In order to recognize the continuous Tamil words in our project we use the Deep Belief Network in the second phase of the Voice Recognition unit. To overcome the limitation of earlier neural networks,

ISSN: 2005-4238 IJAST

`

professor Geoffrey Hinton introduces Deep Belief Networks. Deep Belief Networks (DBN) consists of two different types of neural networks – Belief Networks and Restricted Boltzmann Machines. In contrast to perceptron and backpropagation neural networks, DBN is unsupervised learning algorithm. It can develop their Network Layers to the unspoken words by adjusting the weights between the inputs so that the network can generate similar to the observed data. Deep Belief Network a DBN normally is the stack of many layers of RBM model. Hinton and Salakhutdinov (2006) showed that RBMs can be stacked and trained in a greedy manner to form so-called Deep Belief Networks (DBN). DBNs are graphical models which learn to extract a deep hierarchical representation of the training data.

Restricted Boltzmann Machines (RBM) A (re-stricted) Boltzmann Machine (RBM) (Salakhut-dinov et al,. 2007) is a parameterized generative model representing a joint probability distribution.By Giving some training data, learning of an RBM means adjusting the RBM parameters to maximize thelikelihood of the training data under the model. Re-stricted Boltzmann machines consist of two layerscontaining visible and hidden neurons.The energy function E(v; h) of an RBM is:

$E(v; h) = b^0v \;\; c^0h \;\; h^0W v;$

where W represents the weights connecting hid-den and visible units and b, c are the offsets of the visible and hidden layers respectively. The joint probability distribution is then given by the exponential family $P(v; h) = {}_{Z}{}^{1}e^{E(v;h)}$; where Z is a normalisation factor. The likelihood of some data

X R^n is thus L (X) := $Õ_{v2X}å_h$ P(v; h) and b, c, and W are chosen to maximize this likelihood (or equivalently minimize the negative log likelihood):

$arg_{b;c;W}$ min log L (X) =å log åP(v; h):v2X h

The hidden neurons extract relevant features from the observations, and these features can serve as input to another RBM. By stacking RBMs in this way, we can learn a higher level representation.Practical training strategies in practice of the DBN training often consists of two steps: greedy layer-wise pre-training and fine tuning. Layer-wise pre-training involves training the model parameters.In order to train the DBN we have tried to use a larger set of data since we realized that more data was needed to train the DBN. This larger set has been obtained from different age group of speakers like Male, Female, Children, Adolescent, Aged, Middle-aged etc. Nearly 250 such samples has been recorded by using an Android Cell Phone.

The DBN structure has been Configured by specifying the number of hidden layers and the number of units per data input given. We used one hidden layer to give three layers in total. We used two different configuration of training units, set empirically (and semi-arbitrarily) from data samples – DBN 250 with 250 units in each of the three layers and DBN 500, with 500 units per layer.

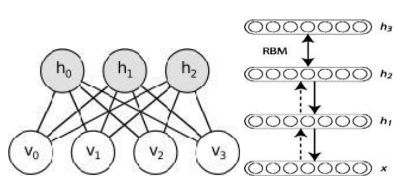


Fig 6 . RBM machine

where W represents the weights connecting hidden and visible units and b, c are the offsets of the visible and hidden layers respectively. The joint probability distribution is then given by the exponential family $P(v; h) = {}_{Z}{}^{1}e^{E(v;h)}$;; where Z is a normalization factor.

ISSN: 2005-4238 IJAST

`

$p(h_j = 1/h) = \sigma(b_j + \sum v_i w_{ij})$ ..(2)

where i= 2, . . . , n, and ai is the bias of the ith visible unit.

	Methods	DBN 250			DBN500		
Input signals	Positives	Trained	Tested	Accuracy	Trained	Tested	Accuracy
Female	7688	0.9016	0.9273	0.9143	0.9032	0.9282	0.9155
Male	4616	0.7001	0.7621	0.7298	0.6945	0.7821	0.7357
Children	4396	0.4771	0.9627	0.6380	0.7138	0.7318	0.7227
Adult	4347	0.8797	0.8042	0.8403	0.8476	0.8548	0.8512
Middle Aged	2518	0.6397	0.3737	0.4718	0.6098	0.6330	0.6212
Aged	2108	0.7108	0.4255	0.5323	0.7085	0.4900	0.5794
Adolescent	1467	0.6806	0.3758	0.4842	0.6813	0.3599	0.4710
Toddler	1304	0.8102	0.7891	0.7995	0.8890	0.7063	0.7872

Table 4. Input voice for various specimen features

As mentioned above , the DBN as a miraculous deep model is a stack of RBMs [11]. Figure 2 illustrates the architecture of the DBN with k hidden layers and it's layer-wise pre-training process.

$$\mathbf{A}_k(\mathbf{x}) = s(\mathbf{b}_k + \mathbf{W}_k s(\quad + \mathbf{W}_2 s(\mathbf{b}_1 + \mathbf{W}_1 \mathbf{x}))), \qquad (3)$$

The activation of the kth hidden layer with respect to input sample x can be computed aswhere Wu and bu (u = 1, 2, . . . , k) are, respectively, the weighting matrices and hidden bias vectors of the jth RBM. Furthermore, s is the logistic sigmoid function s(x) = 1/(1 + e x).

In order to obtain better feature representation, the DBN utilizes deep architecture and adopts the layer-wise pre-training to optimize the inter-layer weighting matrix [11]. The training algorithm of the DBN will be given in the next section in detail.

ISSN: 2005-4238 IJAST

`

6 Future Scope:

Speech Recognition is used in many areas nowadays. For Physically disabled people Voice command is used to operate the Wheel Chair and also it is helpful access the World Wide Web, Sending Mail. The commands are given to the device by means of Tamil words. The given words are processed for Feature Extraction, feature selection and it is send to the voice recognition module. According to the instruction given by the trained data the wheels of the Wheel chair moves with the help of the Raspberry pi and the motors attached to it. In Future the Wheel chair can be charged by using the Solar Panel. The weight of the Wheel chair can be enhanced by increasing the load bearing capacity of the wheel chair. In future the wheel chair can be converted into crutches.

Conclusion:

The design of the Wheel Chair has been developed taking into consideration of many factors like, reliability, cost economical, power consumption and possibility to adapt to the current technology. Speech Recognition is applied in two ways. Single Tamil words are given for giving instruction to the Wheel Chair for basic movements like front, Back, Left and Right using the Neuro Fuzzy Inference System and the Whereas the continuous tamil words are given to the Internet Services using the Deep Belief Network. A good efficient result is achieved using the above algorithms. The Overall performance of the two different algorithms is explained with its pros and cons. With the help of the ANFIS and DBN algorithms we have tried to bring the accuracy and efficiency of the given Tamil voices to the Wheel chair. These techniques when applied to the latest machine will defiantly help us to produce more powerful systems in Future.

Reference

1. http://www.who.int/disabilities/world_report/2011/report.pdf
2. F.Daerden and D. Lefber, “The Concepts and design of pleated pneumatic artificaila muscles”, International journal of fuzzy Practices”, IJFP,vol.2, No.3, 2011.PP 41-45.
3. Mohammad Ilyas Malik et al, International Journal of Computer Science and Mobile Computing, Vol.6 Issue.6, June- 2017, pg. 411-419
4. Suryawanshi et al., International Journal of Advanced Research in Computer Science and Software Engg. 3(5)
5. Nikhil Meshram.1 Shubham Khobragade2 Snehal Ganvir3,” Voice Operated Wheelchair”, (IJSRD/Vol. 6/Issue 01/2018/208)
6. Shanthi Therese,Chelpa Lingam, International Journal of Scientific Engineering and Technology, June 2013.,Review of Feature Extraction Techniques in Automatic Speech Recognition.
7. Speech Recognition Technique: A Review Sanjib Das Department of Computer Science, SukantaMahavidyalaya, (University of North Bengal), India, International Journal of Engineering Research and Applications (IJERA) May/Jun 2012.
8. Bengio Y, “Deep learning of representations: looking forward,” in: Statistical Language and Speech Processing, pp. 1--37, Springer, 2013.
9. Li Deng, “A Tutorial Survey of Architectures, Algorithms, and Applications for Deep Learning” to appear in APSIPA Transactions on Signal and Information Processing, Cambridge University Press, 2014.
10. Wen, Xiaohu& Feng, Qi &Haijiao, yu& Wu, Jun & Si, Jianhua & Chang, Zongqiang& Xi, Haiyang. (2015). Wavelet and adaptive neuro-fuzzy inference system conjunction model for groundwater level predicting in a coastal aquifer. Neural Computing and Applications. 26. 10.1007/s00521-014-1794-7.
11. Jang, J.-S.R. Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm.

ISSN: 2005-4238 IJAST

In Proceedings of the AAAI'91 Proceedings of the Ninth National Conference on Artificial Intelligence, Anaheim, CA, USA, 14–19 July 1991.

12. Amid, S.; Gundoshmian, T.M. Prediction of output energies for broiler production using linear regression, ANN (MLP, RBF), and ANFIS models. Environ. Prog. Sustain. Energy 2017, 36, 577–585
13. Banumathi., et.al., "Speech Recognition of Continuous Tamil phoneme using DBN", International Journal of Innovative Research in Computer and Communication Engineering ,Vol. 4, Issue 7, July 2016.

ISSN: 2005-4238 IJAST