

Automatic Sign Language Gesture Recognition using Prewitt & Morphological Dilation



Avinash Rai, Kavita Gour

Abstract: Sign languages have their own linguistic structure, grammar and characteristics, and are independent of the rules that govern spoken languages. They are visual languages that rely on hand gestures as well as on bodily and facial expressions. Sign languages in different countries are vastly different from one another, so enabling easy communication is important: not just to break the barrier between hearing and deaf individuals, but also between people who do not sign in the same language. In India, sign language plays an important role in the field of communication among dumb and deaf people. There are different signs associated for communication in every country as per their convenient gestures. Automatic sign language gesture recognition is an approach for recognizing gestures and converts it to its actual meaning and convey either through speech or text as per requirements. Here the system is based on Prewitt Edge Detection that possesses the gestures of sign language and helps to recognize and assign their meanings. The Prewitt is second order derivative that has been used in image processing and computer vision, in the form of edge detection or extraction algorithms where it creates gradient of horizontal and vertical magnitude. System also uses certain pre-processing filtration technique such as morphological dilation for better feature extraction.

Keywords : Sign Language Recognition, Prewitt Edge Detection, Morphological Operation, Dilation, Gesture Recognition.

I. INTRODUCTION

Gestures have a wide range of functions to achieve meaning and meaning in a word. Typically, gestures are regarded as gestures made with the help of hands or using certain facial gestures intended to communicate with others without spoken communication. To create effective and relevant communications, both transmitters and receivers must know the same set of data used to perform the action. Basically, Gesture types are split into two different groups' namely dynamic stimulus and physical touch. A powerful touch is what can be changed with time. Strong gestures are considered to be a permanent mark of other messages. A handshake can be considered a gesture of goodbye or Hello! This is an example of a powerful touch. An ambulance signal

is used under normal action. To know all the information in a message sent using touch you need to understand both ways of moving your body. The process used to extract the details of physical symptoms is called body recognition using predefined information. In cases, when mode used to express your feelings, thoughts and to convey any message is the signs or symbol made by physical behavior which is known as Gesture. It is available in two different forms i.e. hand gesture and body gesture and there are two types of gesture i.e. static gesture having predefined sign and dynamic gesture are the signs which can change over the time. Today, technologies are so much modified in the field of gesture recognition that it can be used as a basic tool of interaction for humans with computers. There is a huge difference between current and conventional technologies which were depending on hardware. Gesture detection finds out the objective of a user through the signs made through their hands and postures made by the body parts. There are many useful applications which employ the technology of hand gesture recognition.



Fig. 1. Sign Language Gestures [2]

Sign languages (also known as sign languages) are languages that use hand-drawn comprehension to convey understanding. Sign languages are produced through the use of texts in conjunction with non-literary works. Sign languages are natural languages that are fully integrated with their grammar and lexicon. Linguists view spoken and signed communication as a kind of natural language,

Revised Manuscript Received on October 20, 2020.

Dr. Avinash Rai*, Dept. of Electronics and Communication, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India. Email: avinashrai@rgtu.net

Ms. Kavita Gour, Dept. of Electronics and Communication, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India. Email: kavilagour96@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

meaning that both arose through the normal aging process and evolved over time without careful planning. Sign language should not be confused with body language, the type of communication that does not respond. Although signaling is primarily used by the deaf and hard of hearing, it is also used by hearing people, such as those who cannot speak physically, those who have speech problems because of a disability or a condition (poor communication and contrast), or those who have deaf family members, such as deaf adult children. It is not known how many sign languages are currently available worldwide. Each country usually has its own native sign language, and some have more than one language[1].

II. RELATED WORK

Bhumika Gupta et al. [3] proposed architecture for sign language recognition which is based on K-nearest correlation as a tool for extracting features. System has been tested to recognize alphabets by their gestures using both hands. But the Classification is generally used for classifying either regular or irregular objects from the images to better prediction. Hand gestures/sign languages are features that belong to the training based data models where supervised learning is generally used for better recognition. Segmentation is not a method that can implement sign language recognition system because it requires visual analysis what kind of gesture it is. G.Anantha Rao et al. [4] proposed architecture which is based on CNN. The system is mainly depends over the training samples along with machine iterations or propagation. But the weakness of CNNs network is it has been trained within the amount of knowledge that you provide but does not work effectively if an unknown source is found. If you provide them with less, then expect the CNNs perform poorly. CNNs may have many parameters and datasets that can be run over-fitting problem because CNN needs massive amount of knowledge to quench the thirst. So, if you give much amount of data, then CNNs will become more strong and willing to offer you better performance, but when you give less data then CNNs becomes extremely weak. Large dataset may increases the time and space complexity that may requires much time to execute and respond poorly. Snehal Madhukar Daware et al. [5] proposed architecture which is based on morphological operations. Morphological operation is a part of image enhancement tool in the field of image processing. It enhances the images either by eliminating the unnecessary pixels which is known as erosion but where it loses that is called dilation. But by using these operations or algorithms a complete gesture cannot be portrayed well because it can only have active pixels and rest will be eroded, it means that it segments an image in black and white pixels (either 0 or 1) that does not dignify the hand gesture. Muthu Mariappan H et al. [6] proposed a system which is based on certain preprocessing tools and classification techniques such as morphological operations and fuzzy classification resp. Clustering is the process of grouping the similar data items within the clusters and detaining the dissimilar items. In fuzzy classification, the data items may belong to more than one cluster. From various fuzzy clustering methods, fuzzy c-means clustering (FCM) algorithm is generally used that

can be useful for supervised learning as well as unsupervised one as per the requirements. If clustering has been done after erosion then the useful data may get erode or missing the information that directly leads towards precision. K. Revanth et al. [7] proposed a system which is based on Support Vector Machine for classification. System uses skin color masking technique for extracting region of interest and eliminating the background on the basis of skin color. For skin segmentation; here the system uses OpenCV. On perceptual tasks, they are massively outclassed by deep neural networks. On structured data, they are outperformed by gradient boosted trees. Sruthi C. et al. [8] proposed a system which is based on deep learning methodology i.e. CNN. This work solves the problem of Indian signature static alphabet recognition through a philosophy-based approach. An in-depth learning technique can be used to create a model called an evolutionary neural network (CNN) signet that can detect signals, support data-monitored learning, and often divide the whole process into CNN training and model testing. The concentrated layer consists of a set of matrices, multiplied by the output of the previous layer in the convection process, some of which are characteristic (such as edges, color grade or pattern) or complex (e.g. shape, nose, normal) or mouth. So, those matrices are called filters or kernels. Sign language recognition is a complex issue because different sign languages have similar gestures. A strong learning module is required to achieve the goal with a high level of accuracy. Salma Hayani et al. [9] proposed a system that uses Convolutional Neural Network (CNN) for recognizing sign language gestures. The system recognizes Arabic sign languages of different categories of letters and numbers. Many icons may have similar gestures that can seriously confuse the system to get the right symbol. If the machine is not trained with enough patterns and the system is inefficient for complex gestures, template matching may give erroneous results and re-propagation with CNN may not work effectively, thus reducing the optimal gesture recognition rate.

III. PROPOSED WORK

System is intended to recognize sign language gestures with their appearance that extracted from edge detection technique. System analyzes the gesture and portrayed it horizontally as well as vertically and compare with the stored templates and return result where max keypoints become similar. Here the system is based on Prewitt Edge Detection and morphological dilation.

A. Prewitt Operator

Prewitt operator is used in image processing, especially in edge detection algorithms. Technically, it is a differential operator that estimates the gradient of the image intensity function. At each point in the image, the result of the particular operator is the corresponding gradient vector or standard of this vector. The private operator relies on wrapping the image with small, separable and numeric value filters in the horizontal and vertical directions and is therefore less costly in



terms of calculations such as Sobel and Kayyali operators. On the other hand, the approximate gradient it produces is relatively raw, especially for high frequency variations in the image. Simply put the operator calculates the gradient of the image intensity at each point, which gives the greatest possible increase in light from dark and the rate of change in that direction. So the result shows how quickly the image becomes “abrupt” or “smooth” at that point, so that part of the image represents an edge, as well as how that edge is oriented. In practice, the magnitude (potential of the edge) calculation is more reliable and easier to understand than the direction calculation.



Fig. 2. Input Hand Gesture



Fig. 3. Prewitt Affected Gesture

Prewitt works with horizontal and vertical edges separately and then compute the magnitude of them. The kernels are shown below-

$$H_x = \begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix} * I$$

$$V_y = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * I$$

Here * denotes the convolution operator. Horizontal kernel goes towards left direction and vertical kernel goes towards upward direction. H_x is the horizontal kernel mask and V_y is the vertical kernel mask.

$$G = \sqrt{H_x^2 + V_y^2}$$

G denotes the gradient magnitude that combined the result of horizontal and vertical kernels. G is the resultant variable for computing the complete edges from the input image and further proceed the matrix for better predicting the sign with high precision.

B. Histogram Equalization

A histogram of an image is a graphical representation of intensities. In simple terms, it has been assumed that at each intensity value; it represents the number of pixels. Histogram equalization is an adjustment of entropy or contract for image enhancement for better recognition or information extraction. It is a part of pre-processing data for better segmentation and feature extraction that symbolize the hand gesture with better prediction.

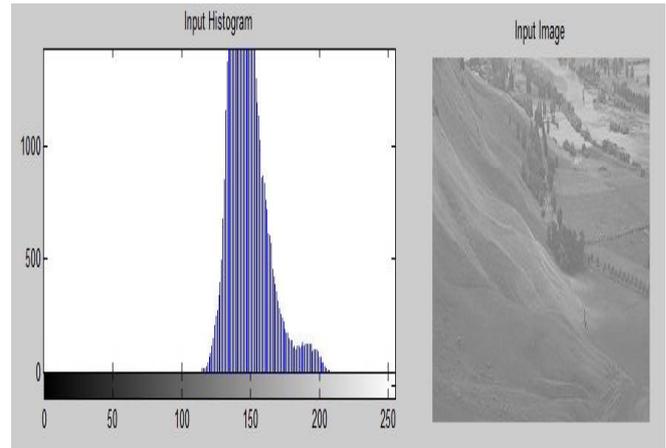


Fig. 4. Histogram Equalization

C. Morphological Dilation

Morphological dilation is a process of filling holes that may erode during edge detection. It is a process of enumerating the active pixels from binary image and converting the odd one into active one. Dilation enhances the broken edges and form the proper gesture that helps to identify the actual sign.



Fig. 5. Dilation

Every sign has a unique gesture and system trained with those samples that have been manually captured with real time images. The color histogram graph of an image represents the number of pixels for each color components.

Automatic Sign Language Gesture Recognition using Prewitt & Morphological Dilation

The histogram equations cannot be applied separately for red, green, and blue components because it causes dramatic changes in the image's color balance. However, if the image firstly converted to another color mode, such as the HSL / HSV color space, then the algorithm can be applied to the luminance and the value channel as a result of changes over the color and saturation of the image.

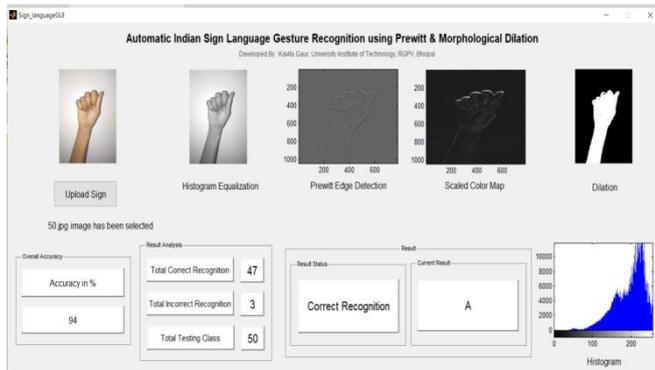


Fig. 6. Proposed Work GUI

As per the flow chart, first of all an input image will be acquired for preprocessing and once the histogram equalization has been applied system will apply prewitt edge detection for edge extraction and then scaled color mapping will be shown for better visibility then dilation will be applied for filling holes and then a template will be created and input image will be compared with the template and if extracted points are greater than the threshold value the respected sign will be returned otherwise no sign will be found.

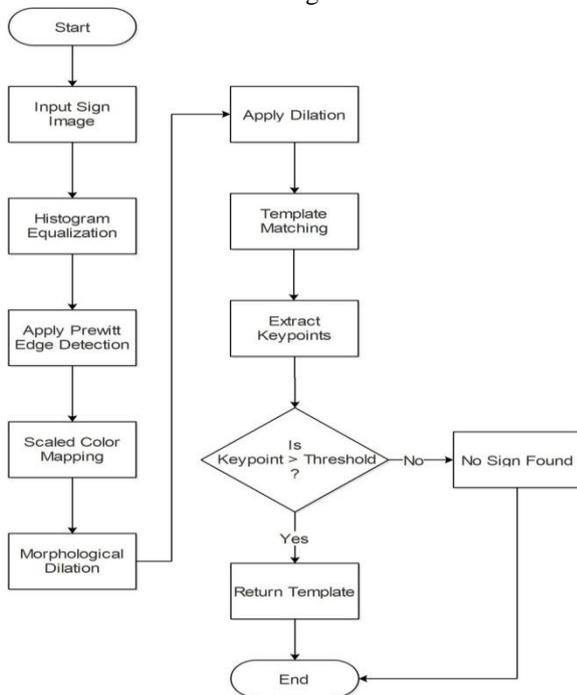


Fig. 7. Flow Chart

IV. RESULT ANALYSIS

The simulation studies involve the various trails with distinct sign language gestures. There are total number of 50 trails where 47 trails recorded as correct recognition and 3 as incorrect that may includes true positive, true negative, false positive and false negative. True positive means that there are

certain trails that positively detected which returns correct recognition and few images that may contain valid sign but system is not able to detect; that entertained in the category of true negative. Similarly as false negative invalid sign detected as any template, whereas false positive means invalid sign rejected positively. So, by observing all these datasets, the perceived accuracy is 94.00 %. CR- Correct Recognition, IR- Incorrect Recognition, TTC – Total Testing Class

$$CR = 47, IR = 3, TTC = 50$$

$$Accuracy = \frac{TTC - IR}{TTC} * 100 \%$$

$$= \frac{50 - 3}{50} * 100 \%$$

$$= 94.00 \%$$

Table- I: Result Analysis

s_no	image_name	recognized_string	result	Score
1	1.jpg	One	CR	1
2	2.jpg	Two	CR	1
3	3.jpg	Zero	CR	1
4	4.jpg	Four	CR	1
5	5.jpg	Five	CR	1
6	6.jpg	Six	CR	1
7	7.jpg	Seven	CR	1
8	8.jpg	Eight	CR	1
9	9.jpg	Nine	CR	1
10	10.jpg	Ten	CR	1
11	11.jpg	Three	CR	1
12	12.jpg	Afternoon	CR	1
13	13.jpg	B	CR	1
14	14.jpg	Bad	CR	1
15	15.jpg	C	CR	1
16	16.jpg	College	CR	1
17	17.jpg	Unable to Recognize	IR	0
18	18.jpg	Unable to Recognize	IR	0
19	19.jpg	F	CR	1
20	20.jpg	Good Bye	CR	1
21	21.jpg	Good	CR	1
22	22.jpg	H	CR	1
23	23.jpg	Unable to Recognize	IR	0
24	24.jpg	J	CR	1
25	25.jpg	K	CR	1
26	26.jpg	L	CR	1
27	27.jpg	M	CR	1
28	28.jpg	Man	CR	1
29	29.jpg	Morning	CR	1
30	30.jpg	N	CR	1
31	31.jpg	Night	CR	1
32	32.jpg	O	CR	1
33	33.jpg	P	CR	1
34	34.jpg	Q	CR	1

35	35.jpg	R	CR	1
36	36.jpg	S	CR	1
37	37.jpg	School	CR	1
38	38.jpg	Sorry	CR	1
39	39.jpg	Strong	CR	1
40	40.jpg	T	CR	1
41	41.jpg	Thank You	CR	1
42	42.jpg	U	CR	1
43	43.jpg	V	CR	1
44	44.jpg	W	CR	1
45	45.jpg	Woman	CR	1
46	46.jpg	Work	CR	1
47	47.jpg	X	CR	1
48	48.jpg	Y	CR	1
49	49.jpg	Z	CR	1
50	50.jpg	A	CR	1

K. Revanth [7]	90.54
S. Hayani [9]	90.02
Proposed	94.00

V. CONCLUSION & FUTURE SCOPE

Automatic sign language recognition using Prewitt and morphological dilation is an approach for recognizing sign language with better precision. System is able to recognize signs effectively with less false alarm rate. Most of the system uses machine learning methods to train the system with various samples for a single hand gesture. But a large dataset can consume the large amount of memory that increases the execution time where it is very important to communicate as earlier as possible with high level of accuracy. Here the system trained with various samples and able to recognize correct sign with relative features. System acquired 94.00 % of accuracy with less error rate or incorrectly recognition. In future; accuracy can be enhanced with various image enhancement technique and feature extraction method that later can be implemented for removing communication barriers.

REFERENCES

1. <https://www.pinterest.com.au/pin/491103534343877366/>
2. <https://www.semanticscholar.org/paper/An-Automated-System-for-Indian-Sign-Language-Kaur-Gill/ebcceb337be93d44e3d6635ad5964e1e8bf aeb2c>
3. B. Gupta, P. Shukla and A. Mittal, "K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion," 2016 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, 2016, pp. 1-5, doi: 10.1109/ICCCI.2016.7479951.
4. G. A. Rao, K. Syamala, P. V. V. Kishore and A. S. C. S. Sastry, "Deep convolutional neural networks for sign language recognition," 2018 Conference on Signal Processing And Communication Engineering Systems (SPACES), Vijayawada, 2018, pp. 194-197, doi: 10.1109/SPACES.2018.8316344.
5. Daware, Snehal & Kowdiki, Manisha. (2018). Morphological Based Dynamic Hand Gesture Recognition for Indian Sign Language. 343-346. 10.1109/ICIRCA.2018.8597417.
6. H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862125.
7. K. Revanth and N. S. M. Raja, "Comprehensive SVM based Indian Sign Language Recognition," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2019, pp. 1-4, doi: 10.1109/ICSCAN.2019.8878787.
8. S. C.J. and L. A., "Signet: A Deep Learning based Indian Sign Language Recognition System," 2019 International Conference on Communication and Signal Processing (ICCS), Chennai, India, 2019, pp. 0596-0600, doi: 10.1109/ICCS.2019.8698006.
9. S. Hayani, M. Benaddy, O. El Meslouhi and M. Kardouchi, "Arab Sign language Recognition with Convolutional Neural Networks," 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), Agadir, Morocco, 2019, pp. 1-4, doi: 10.1109/ICCSRE.2019.8807586.
10. A. Subhash Chand, A. S. Jalal and R. Kumar Tripathi, "A survey on manual and non-manual sign language recognition for isolated and continuous sign," Applied Pattern Recognition, 2016.
11. <https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm>
12. H. Cooper, B. Holt and R. Bowden, "Sign language recognition," Springer, 2011.

```

[34] '34.jpg' 'Q' 'CR' [1]
[35] '35.jpg' 'R' 'CR' [1]
[36] '36.jpg' 'S' 'CR' [1]
[37] '37.jpg' 'School' 'CR' [1]
[38] '38.jpg' 'Sorry' 'CR' [1]
[39] '39.jpg' 'Strong' 'CR' [1]
[40] '40.jpg' 'T' 'CR' [1]
[41] '41.jpg' 'Thank You' 'CR' [1]
[42] '42.jpg' 'U' 'CR' [1]
[43] '43.jpg' 'V' 'CR' [1]
[44] '44.jpg' 'W' 'CR' [1]
[45] '45.jpg' 'Woman' 'CR' [1]
[46] '46.jpg' 'Work' 'CR' [1]
[47] '47.jpg' 'X' 'CR' [1]
[48] '48.jpg' 'Y' 'CR' [1]
[49] '49.jpg' 'Z' 'CR' [1]
[50] '50.jpg' 'A' 'CR' [1]

Total Correct Recognition = [47]
Total Incorrect Recognition = [3]
Total Testing Class = [50]
Accuracy = [94]
    
```

Fig. 8. Console Output

Graph 1 Result Analysis

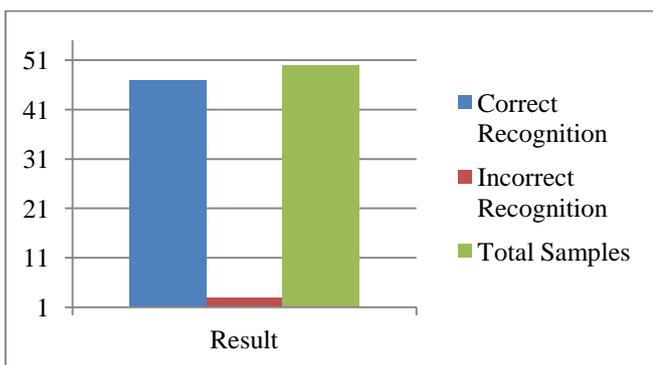


Table- II: Result Comparison

	Accuracy %
B. Gupta [3]	89.50
G. Anantha Rao [4]	88.50
H. Muthu Mariappan [6]	75.00



13. S. Ong and S. Ranganath, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005.
14. M. Al-Rousan and M. Hussain, "Automatic recognition of Arabic sign language finger spelling," International Journal of Computers and Their Applications, 2001.
15. K. Assaleh and M. Al-Rousan, "Recognition of Arabic sign language alphabet using polynomial classifiers," EURASIP Journal on Applied Signal Processing, 2005.
16. O. Al-Jarrah and A. Halawani, "Recognition of gestures in Arabic sign language using neuro-fuzzy systems," Artificial Intelligence, 2001.
17. T. Shanableh, K. Assaleh and M. Al-Rousan, "Spatio-temporal feature-extraction techniques for isolated gesture recognition in Arabic sign language," IEEE Transactions on Systems, Man, and Cybernetics Part B (Cybernetics), 2007.
18. M. AL-Rousan, K. Assaleh and A. Tala'a, "Video-based signer-independent Arabic sign language recognition using hidden Markov models," ELSEVIER, 2009.
19. M. Maraqa, F. Al-Zboun, M. Dhyabat and R. Abu Zitar, "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks," Intelligent Learning Systems and Applications, 2012
20. R. Alzohairi, R. Alghonaim, W. Alshehri, S. Aloqeely, M. Alzaidan and O. Bchir, "Image based Arabic Sign Language Recognition System," International Journal of Advanced Computer Science and Applications, 2018.
21. Y. LeCun, L. Bottoux, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," IEEE, November 1998.
22. A. Krizhevsky, I. Sutskever and G. Hinton, "Imagenet classification with deep convolutional neural networks," Proceedings of the 25th international conference on neural information processing systems, 2012.
23. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint, 2014.
24. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinov, C. Hill and A. Arbor, "Going Deeper with Convolutions," IEEE Xplore, 2015.
25. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2015.
26. D. Coomans and D. Massart, "Alternative k-nearest neighbour rules in supervised pattern recognition: Part 1. k-Nearestneighbour classification by using alternative voting rules," AnalyticaChimicaActa 136, 15-27, 1982.
27. C. Williams and M. Seeger, "Using the Nyström method to speed up kernel machines," 2001.
28. Y-W. Chang, C-J. Hsieh, K-W. Chang, M. Ringgaard and C-J. Lin, "Training and testing low-degree polynomial data mappings via linear SVM," Journal of Machine Learning Research, april 2010.
29. J-P. Vert, K. Tsuda and B. Schölkopf, "A primer on kernel methods," Kernelmethods in computationalbiology, 47, 35-70, 2004. Volume(issue), paging if given. Available: [http://www.\(URL\)](http://www.(URL))

AUTHORS PROFILE



Dr. Avinash Rai has done B.E. (Electronics), M.E. (VLSI Design), Ph.D. (Wireless Sensor Network) and is currently working as Assistant Professor in Department of Electronics & Communication Engineering at University Institute of Technology, Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal (M.P). Dr. Avinash Rai does researches in Wireless Sensor Network, Telecommunication Engineering and Electronics Engineering. His most recent publication is "Enhancing Energy of Sensor Node to Increase Efficiency in Wireless Sensor Network"



Kavita Gour has done B.E. (Electronics & Communication) and is currently an M.E. scholar (Digital Communication) in the Department of Electronics & Communication Engineering at University Institute of Technology, Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal (M.P). Kavita Gour has undertaken research work presented in this paper under the guidance of Dr. Avinash Rai which is required in the partial fulfillment for the award of M.E. in digital communication.