Real Time Interface for Assistive E-Learning

Indhumathi R, Geetha A

Abstract*: Assistive system is any hardware or software designed to enable independence for disabled and older people. Technologies are growing up day to day. The disabled and elder people find very difficult to adopt new technologies. It is important to develop effectively available systems to accomplish their incorporation inside the new advances. Many researches are ongoing in assistive technology. This proposed work is concentrated on Assistive human computer interface. Nowadays technology has been improved in many areas particularly in learning. Everyone likes E-Learning i.e. learning by using electronic media. Basically two types of materials are used in E-Learning they are learning by video and learning by text reading. Our aim is used to create an application for effective Assistive E-Learning. The head gestures are used instead of traditional input devices. This work recognizes head gestures in real time to create a head interface between user and computers. It helps to automatically operate the applications such as click, scroll up and scroll down. The Haar cascaded classifier with rotation invariant approach is used to detect head gestures. The detected gestures are recognized as head-left, head-right, head-origin and head-down using Random Forest and Decision Tree. The proposed system has been tested in real time with adobe reader and performance analysis is carried out. The system is found to outperform other existing methods.

Index Terms:Assistive Technology, Head Gestures Detection, Haar cascaded classifier, Random Forest, Decision Tree, E-Learning.

I. INTRODUCTION

Assistive Technology as any tool, equipment, system, or service is designed to help develop, maintain or improve a person with a disability to function in all aspects of his or her life. Many researches are ongoing in assistive technology. Some research areas are

- Alternative Human Computer Interfaces for people with motor disorders
- AT centre and service delivery issues
- ICT-Based learning technologies for disabled people
- Power Mobility
- Acceptance and use of E-health Technology by older adults and professionals
- The development and implementation of "Remote Care"
- Using the cloud to enhance AT

In this research Human Computer Interaction with ICT learning technology based interface implemented for disabled people. In E-Learning two types of materials are to be used. One is video another one is text. Generally operating materials is very essential. Learners have to use their hand often to operate the E-Learning material. It disturbs the learning lot. This proposed work the materials automatically operated by our head gestures. The learner can easily operate the materials. It is most helpful for physically challenged people. The capacity to estimate the head pose of another person is a common human ability that presents a unique challenge for computer vision systems. A number of head pose techniques are used in existing work. The Haar cascaded classifier detects only frontal face of the face. In this proposed work rotation invariant algorithm is used to detect left, right, origin and down. Random forest classifier is used to recognize the head gestures.

A. Decision Tree

Decision tree constructs classification models as a tree struct ure.It breaks down a data set into smaller and smaller subset s and at the same time incrementally develops an associated decision tree.The end result is a tree with nodes of decision and node of leaf. There are two or more branchesin a decisio n node. A leaf node is a classification or decision. The conce pt of a decision tree is a system based on rules.Given the tar gets and features of the training data set, some set of rules w ill be developed for the decision tree algorithm.It is possible to use the same set rules to predict on the test dataset. In decision tree algorithm calculating these nodes and forming the rules will happen using the information gain and gini index calculations.

B. Random Forest

Random forest algorithm is a supervised classification algorithm. As the name recommend, this algorithm makes the forest with various trees. As a rule, the more trees in the forest the more powerful the forest resembles. Similarly in the Random forest classifier, the higher the quantity of trees in the forest gives the high accuracy results. The following are the benefits of Random forest classification contrasted and other classification algorithms.

- The over fitting issue will never come
- The same arbitrary woodland calculation can be utilized for both classification and regression task.
- The random forest algorithm can be accustomed to recognizing the most vital features out of the accessible features from the training dataset.

Revised Manuscript Received on June 07, 2019.

Indhumathi R, Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Tamilnadu, India,

Geetha A, Associate Professor, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Tamilnadu, India,

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

II. RELATED WORKS

Gerard Biau [1] discussed Random Forest model. Anwar saeed, Ayoub AI Hamadi and Ahmed Ghonein [2] proposed a frame based approach to estimate head pose. The Viola and Jones Haar –like face detector is used to estimate the head pose. SVM classifier is used which gives performance of 75% accuracy. Rushikesh T.Bankar and Suresh S.Salankar [3] developed a system to recognize head gestures in real time. The gestures are recognized from the video sequences. Alvise Memo Pietro Zanuttigh [4] presented the combined usage of a head-mounted display and a multi-modal sensor setup. Reliable gesture recognition is obtained through a real-time algorithm exploiting novel feature descriptors arranged in a multi-dimensional structure fed to an SVM classifier. Shubhada P Deshmukh, Manasi S.Patwardhan and Anjali R.Mahajan [5] proposed a head gesture recognition system that is performed by using SVM classifier. This work provides required time efficiency and good accuracy of 97% for facial expression and 98% accuracy for head gestures. Rushikesh T.Bankar and Suresh S.Salankar [6] proposed system based on a sequence of data acquisition that is image capturing, pre processing means filtering, feature extraction that is rectangular features and a parallel stage with a cascade of classifiers design and classification. Va Zen-ping Bian, Junhui hou, et al [7] used non linear function mapping purpose.The Model used FM-Facial Position and Expression Mouse system. Vrushali Y Kulkarni, Pradeep K Sinha [8] proposed effective learning and classification of Random Forest model. Eesha Goel, Er. Abhilasha [9] discussed nature and working of Random Forest classifier.

III. HEAD GESTURES DETECTION

A. Flow Diagram of the Proposed Work

Initially, the proposed system has to be designed to detect automatically the user's face and track it through time to recognize head gestures in real time. Subsequently, a new information fusion procedure has to be proposed to acquire data from computer vision algorithms and use its results to carry out a robust recognition process. Finally, an intelligent interface has to be developed to replace the conventional input devices. The proposed work is developed to interact with a computer using vision-based interface without any hardware. The adobe reader operations are carried out using head gestures. The overall flow diagram of the proposed work is shown in Figure 1.

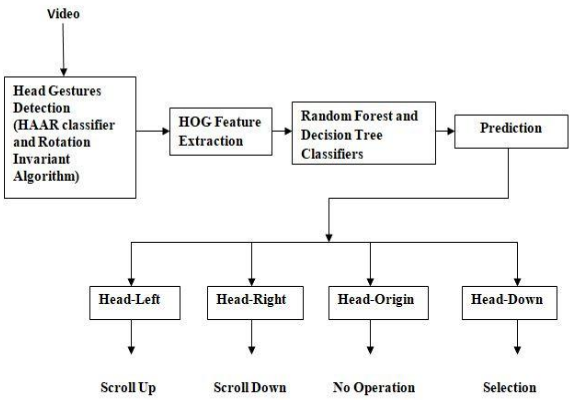


Fig1: Flow diagram of head interface with Text Reader

B. Haar Classifier and Rotation Invariant Algorithm

Head gestures are the movements made by the head movement.[7].The rotation invariant and Haar cascaded approaches are used to find head gestures such as head left, head right, head origin and head down as shown in Fig 2.

Fig2a: Orgin

Fig 2b: Right

Fig 2c: Left

Fig2d: Down

Fig 2:Head Gestures

Haar Classifier

Haar-like highlights [10], [11] are sure features in a digital image which are utilized in object detection. They are named so because of their closeness with Haar wavelets. The Haar-like features are utilized progressively object recognition.Viola et built up the calculation out of the blue.[10], [11] and was later reached out by Lien hart et al [12]. The Haar course calculation [13] is an object recognition calculation used to find faces where a course work is prepared from a great deal of positive and negative pictures. It is utilized to recognize faces. The classifier is arranged with a base size of 150 x 150 pixels and a scaling component of 1.1, The greatest face in the picture is picked as an area of intrigue. For the discovery of the faces Haar features are the principle part of the Haar cascade classifier. Each element results in a solitary esteem which is determined by subtracting the aggregate of pixels under white square shape from the total of pixels under dark rectangle shape as shown in Fig 3.

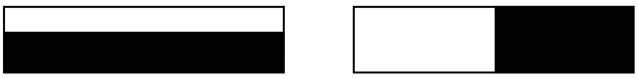


Fig 3a: Edge Features

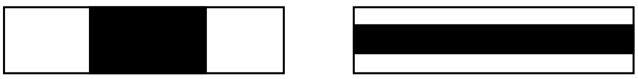


Fig3b: Line Features

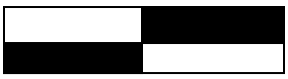


Fig 3c: Four Rectangle Features

Fig 3: HARR Features

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Rotation Invariant Algorithm

The Rotation Invariant Algorithm is given below:
1) Haar classifier checks for the nearness of frontal faces in the input picture.
2) If frontal face is found it detects the face.
3) If frontal face is not found, a rotational matrix is formed with $\boldsymbol{R}$ with $\boldsymbol{\theta} = \pm\mathbf{30^o}$, to transform the image. The angle of $\boldsymbol{\theta}$ = $\pm$ $\mathbf{30^o}$ is chosen because the Haar classifier in this case has been developed with inherent rotational features up to $\pm\mathbf{30^o}$.
4) The Haar classifier looks for the face in the transformed picture.
5) If for any of the set values of $\boldsymbol{\theta}$, the Haar classifier does not detect the face, the algorithm concludes that no face is found.

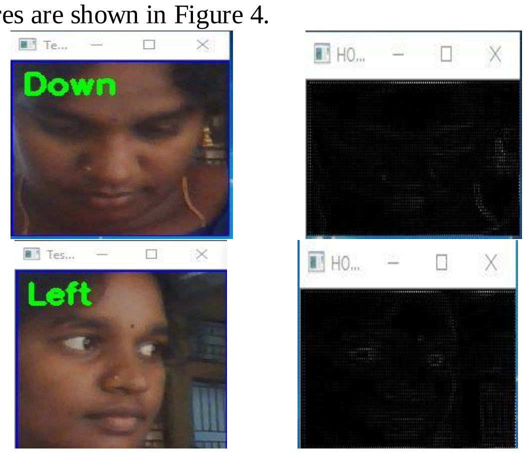
C. Hog Features

HOG descriptors are principally utilized in computer vision and AI for item recognition. We can likewise utilize HOG descriptors to measure and speak to shape just as surface, be that as it may. In other words, HOG has five phases

1. Normalizing the picture before to depiction.
2. Gradients computing in both the x and y headings.
3. Acquiring weighted votes in spatial and orientation cells.
4. Differentiation normalizing overlapping spatial cells.
5. Gathering all Histograms of Oriented gradients to shape the final feature vector.

The orientations, pixels per cell and cells per block are the most important parameters for the HOG descriptor. These three parameters (along with the size of the input image) effectively control the dimensionality of the resulting feature vector. In this proposed work Orientations is 9, Pixels-per-cell is 10X10 and Cells-per-block is 2X2 used for Hog feature vector generation.

The HOG descriptor restores a genuine esteemed component vector.HOG is implemented in both OpenCV and scikit-image. The OpenCV implementation is less flexible than the scikit-image implementation, and thus we will primarily use the scikit-image implementation. There are 6156 features derived. Among these features 100 higher gradient features are selected to display the hog. The Hog features are shown in Figure 4.



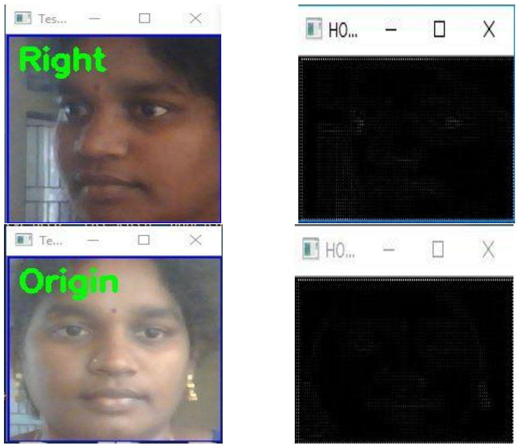


Figure 4: HOG Images

D. Training with Random Forest

The pseudo code for random forest algorithm can split into two stage i.Random forest creation pseudo code. ii. Pseudo code to perform prediction from the created random forest classifier.

Random Forest pseudo code:

- Randomly select "k" features from all out "m" features. Where k << m
- Among the "k" features, ascertain the node "d" utilizing the best part point.
- Split the node into child nodes utilizing the best split.
- Repeat 1 to 3 stages until "l" number of nodes has been come to.
- Build forest by rehashing stages 1 to 4 for "n" number occasions to make "n" number of trees.

Random forest prediction pseudo code:

To perform prediction utilizing the trained random forest algorithm utilizes the beneath pseudo code.

1. Takes the test features and utilize the tenets of each haphazardly made decision tree to foresee the result and stores the predicted result (target)
2. Calculate the votes in favor of each predicted target.
3. Consider the high voted predicted focus as the last prediction from the random forest algorithm.

E. Training with Decision Tree

Decision Tree calculation has a place with the group of supervised learning calculations. Decision Tree is to make a training model which can use to anticipate class or estimation of target factors by taking in choice tenets surmised from earlier data (training data). The decision tree calculation endeavors to take care of the issue, by utilizing tree portrayal. Each interior node of the tree relates to a trait, and each leaf node compares to a class label.

Decision Tree calculation pseudo code:

1. Place the best quality of the dataset at the foundation of the tree.
2. Split the training set into subsets. Subsets ought to be made so that every subset contains data with a similar value for a characteristic.
3. Repeat stage 1 and stage 2 on every subset until you discover leaf nodes in every one of the parts of the tree.

IV.IMPLEMENTATION

A. Environment Setup

Machine learning environment has been set by installing python 3.5 versions along with anaconda library. OpenCV (version 3.3.0) library is successfully linked with python 3.5 interpreter. Using conda install command required libraries such as numpy, PIL, Scikitlearn, scipy, sklearn, pyautogui, glob, are imported.

B. Dataset Used

Real time dataset is used for this proposed work. Dataset are taken from various kinds of learners. The Haar cascaded classifier and rotation invariant algorithm used to detect head pose. The Haar cascaded detects frontal faces. The Rotation Invariant algorithm added with Haar to detect all sides of head.

C. Training Phase

The head gestures are captured from web camera and labeled as four types such as head-left, head-right, head-origin and head-down. About 1200 training samples such as head-left 300, head-right 300, head-origin 300, head-down 300 are used for training. Every frame is resized as 150 x 150 and all images are converted as gray scale for better performance.

D. Testing Phase

About 200 samples such as head-left 50, head-right 50, head-origin 50, head-down 50 are tested. The video frames are resized as 150 x 150 and converted into gray image. The Random forest model classifies the frames as head-left, head-right, head-origin and head down.

V. INTERFACE WITH E-LEARNING

Pyautogui module in python is used for creating an interface with the computer. This library consists of functions for keyboard and mouse event handling operations. The tested head gestures are interfaced with adobe reader to perform operations. The scroll up, scroll down and selection operations of pdf (portable document file) reader are performed by head left, right and down. The interface is works well in real time. The output screen shots are shown below.

A.PDF Reader:

The figure 5 shows the screen shots of head gestures interfaced with pdf reader.

Fig 5a: Head-Left, Scroll-up

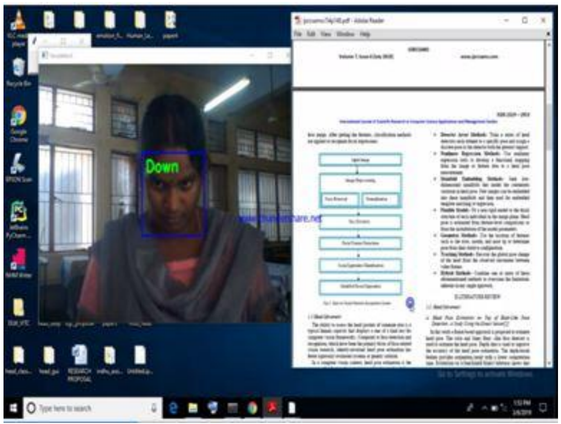


Fig 5b: Head-Down, Left-click

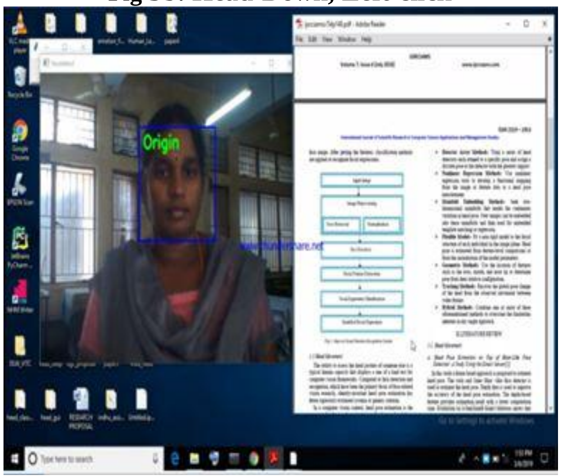
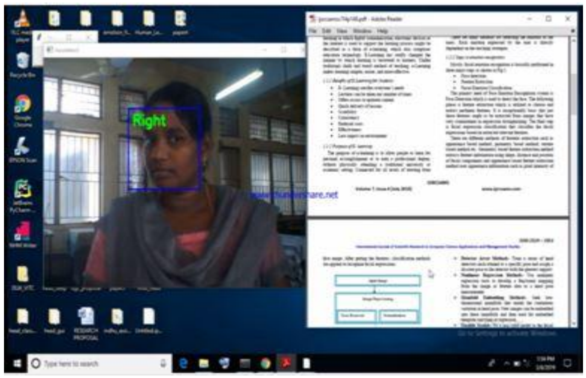


Fig 5c: Head-Origin, No action

Fig 5d: Head-Right, Scroll-down



VI. RESULTS

The performance of the proposed system with Random Forest and Decision Tree models are measured using the following performance metrics.

A. Confusion Matrix

A confusion matrix is a table that is used to describe the performance of a classification model on a set of test data for which the true values are known. The proposed system is tested with 50 samples for each gesture. The result of confusion matrix for Random Forest and Decision Tree are is shown in table 1 and table 2 respectively.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table 1: Confusion Matrix for Random Forest

		PREDICTED CLASS			
		Head-Left	Head-Right	Head-Origin	Head-Down
ACTUAL CLASS	Head-Left	38	0	5	7
	Head-Right	2	44	0	4
	Head-Origin	10	3	27	10
	Head-Down	0	1	0	49

Table 2: Confusion Matrix for Decision Tree

		PREDICTED CLASS			
		Head-Left	Head-Right	Head-Origin	Head-Down
ACTUAL CLASS	Head-Left	26	2	17	5
	Head-Right	6	34	6	4
	Head-Origin	14	5	16	15
	Head-Down	10	3	16	21

B. Classification Report

The Performance metrics are used to find out how effective the model is based on some metric using test datasets. Different performance metrics are used to evaluate different Machine Learning Algorithms. Here the metric for evaluation of machine learning algorithms used are Precision, Recall, F-score and Accuracy. The classification reports for Random Forest and Decision Tree are given in table 3 and table 4 respectively.

Precision:

Precision *P* is the number of correctly classified positive examples divided by the total number of examples that are classified as positive as shown in the Eq 1.

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \qquad (1)$$

Recall:

Recall *R* shown in Eq 2 is the number of correctly classified positive examples divided by the total number of actual positive examples in the test set.

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (2)$$

F-Score:

It is hard to compare two classifiers using two measures. F-score shown in Eq 3 combines precision and recall into one measure.

$$\mathrm{F-score} = \frac{\mathrm{2PR}}{\mathrm{P+R}} \qquad (3)$$

Accuracy:

Accuracy is calculated by the equation shown in Eq 4.based on the confusion matrix the accuracy is calculated. The accuracy shows how effectively the classifier predicts the testing data. In this work the Random Forest produces 90% accuracy where the Decision Tree produces 74%.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \qquad (4)$$

Table 3: Classification Report for Random Forest

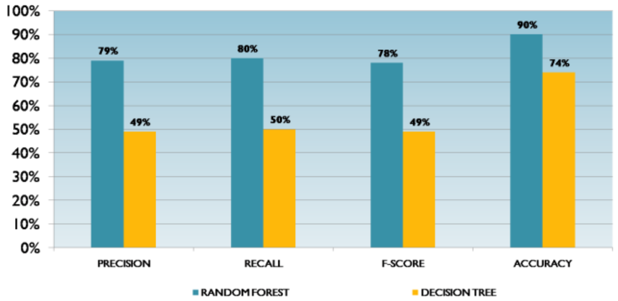
	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Head-Left	76	76	76	88
Head-Right	88	91	89	95
Head-Origin	54	84	66	86
Head-Down	98	70	82	89
Avg/Total	79	80	78	90

Table 4: Classification Report for Decision Tree

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Head-Left	52	46	49	73
Head-Right	68	77	72	87
Head-Origin	32	28	30	63
Head-Down	42	47	44	73
Avg/Total	49	50	48	74

C. Comparison of Models

The comparison of the performances of Random Forest and Decision Tree are shown in Figure 6. In this proposed work the Random Forest performance is higher than the Decision



Tree. The Decision Tree produced 74% accuracy and the Random Forest produced 90% accuracy.

Fig 6: Comparison of Random Forest and Decision Tree

VII. CONCLUSION AND FUTURE WORK

In this proposed work head gestures are detected and recognized as head-left, head-right, head-origin and head-down by Random Forest and Decision Tree classifiers. The detected gestures are interfaced with adobe reader. The interface worked successfully. The proposed work can be extended to more applications and other human body gestures.

REFERENCES

1. Gerard Biau," Analysis of a Random Forests Model", Journal of Machine Learning Research, 1063-1095, vol 13,2012.
2. Anwar saeed, Ayoub AI Hamadi and Ahmed Ghonein," Head pose estimation on top of Haar-like face Detection:A study using the kinect sensor", ISSN1424-8220,Sensor,vol 15, pp 20945-20966, 2015
3. Rushikesh T.Bankar and Suresh S.Salankar," Head Gesture Recognition System Using Adaboost Algorithm with Obstacle Detection", 7th International Conference on Emerging Trends in Engineering & Technology, IEEE, 2015.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

4. Alvise Memo Pietro Zanuttigh," Head-Mounted Gesture Controlled Interface for Human-Computer Interaction", Multimed Tools Appl,Springer Science + Business Media Newyork, 2018.
5. Shubhada P Deshmukh, Manasi S.Patwardhan and Anjali R.Mahajan," Feedback Based Real Time Facial and Head Gestures Recognition for E-Learning System", Association for Computing Machinery.CoDS-COMAD'18,Gao,India , 2018.
6. Rushikesh T.Bankar and Suresh S.Salankar," Head Gesture Recognition System using Adaboost Algorithm with Obstacle Detection", 7th International Conference on Emerging Trends in Engineering & Technology, IEEE, 2015.
7. Bian,Junhui Hou, Lap-Pui Chau,and Nadia Magnenat-Thalmann ,"Facial Position and Expression Based HumanComputer Interface for Persons with Tetraplegia" , Institute of Electrical and Electronics Engineers IEEE,2015.
8. Vrushali Y Kulkarni, Pradeep K Sinha," Effective Learning and Classification using Random Forest Algorithm", International Journal of Engineering and Innovative Technology (IJEIT), ISSN: 2277-3754, vol 3, Issue 11, May 2014.
9. Eesha Goel, Er. Abhilasha," Random Forest: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 128X, Vol 7, Issue 1, January 2017.
10. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Computer Vision and Pattern Recognition, Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1,IEEE, 2001.
11. P. Jones, P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," in University of Rochester. Charles Rich. Citeseer,2001.
12. R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," in Image Processing International Conference on, vol. 1,IEEE, 2002.
13. Ijaz Khan,Hadi Abdullah and Mohd Shamian Bin Zanial,"Efficient Eyes and Mouth Detection Algorithm using Combination of Viola Jones and Skin color Pixel Detection",International Journal of Engineering and Applied sciences,2012

AUTHORS PROFILE

R.Indhumathi M.Sc., B.Ed., M.Phil, Research scholar in Department of Computer and Information Science at Annamalai University. She has published 3 international journals. Her area of research is image processing and computer vision.

Dr.A.Geetha M.E., PhD, Associate Professor in Department of Computer Science and Engineering at Annamalai University. She has published 15 National and International journals. Her specialization is image processing and computer vision.