Use Of Haar Features And Multi-Layer Perception To Support Connectivism.

BHARADWAJ VISHANTH THIYAGARAJAN
B.E IN COMPUTER ENGINEERING.
P.I.C.T (University of Pune)
PUNE, INDIA.
Email: vishanthbharadwaj@gmail.com

***Abstract*: MOOCs have refined the way of teaching where the electronic devices can provide knowledge on the go, which says that knowledge is omnipresent. The connectivism is a technique where the knowledge is contained in the electronic media, which can be delivered as per the requirement. The use of Face detection with Haar based algorithms help to track the current status of the user and perform the teaching accordingly. The Multi-Layer Perception classifier is used to learn about the user attributes and provides the best teaching method to the user. These methods aim to deliver the knowledge completely and in the most effective form in which the user can comprehend. The major parts of the paper include Prediction, Face-Tracking, Face-Detection, Connectivism and "Teach and Learn". The use of multi-perception is for a dual purpose, which includes the training of the previous data and also includes the storage of the real-time data for future prediction purpose. This includes 4 teaching techniques for the students who are physically normal and 2 for the students who are visually impaired or deaf. Hence this proposed project could be a personal assistant for the students by tracking, teaching and also learning from the real-time.**

***Index Terms*—Object Detection, face detection, Machine Learning, MOOCs, Multi Layer-Perception, Haar wavelets.**

INTRODUCTION:

In this paper an approach is introduced for combining machine learning concepts with Haar features to support the connectivism and provide an intelligent teaching technique for students.

The field of computer vision has a huge area of application. This includes the field of ADAS, OCR, OBR, Medical applications, Satellite systems etc. Thus object detection can also be applied in the field of MOOCs to track the activities of the student and improve the education imparted. The Human-human interaction had a basic method of see and teach which had minimal knowledge loss. Use of computer vision can also adapt the basics of human-human interaction and provide an environment where there is a minimal knowledge loss.

The machine learning techniques have been used for a long time for prediction and accuracy check purposes. Here the Multi Layer-perception is a neural-net, which can be used to predict the students' quality and provide him/her an appropriate teaching method. Hence, this improves the accuracy and saves time for the students.

Combining the above algorithms with the connectivism, which was introduced to provide efficient education through electronic media, a novel teaching project is introduced.

METHODOLOGY OF THE PROPOSED WORK:

This paper contains a 4-stage mechanism. Teach and Learn, Face tracking, Face Detection, Connectivism.

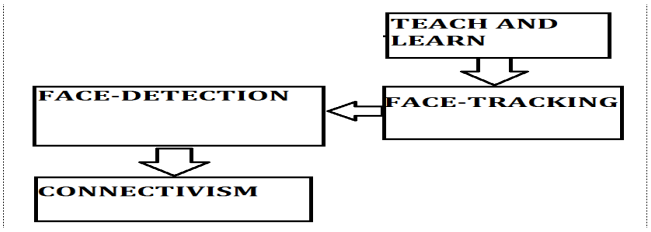


Fig.1. Flow-Chart of the algorithms used.

1.) TEACH AND LEARN:
   a. Multi-Perception algorithm.
2.) FACE-TRACKING:
   a. Mean Shift
   b. Histogram Back-Projection.
   c. CAM shift.
3.) FACE-DETECTION:
   a. HAAR-Features
   b. Integral Image.
   c. ADA-Boost for images.
   d. CAM-shift.
4.) CONNECTIVISM.

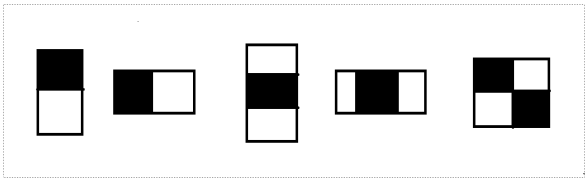
HAAR FEATURES.

The Haar features described by the Viola-Jones algorithm prove that the detected object are classified and converted into white and black pixels. The way in which pixels are converted and organized is different for every object. Hence this makes the object tracking simpler. One of the important features comparable to the haar features is Convolution Kernel.

There are 5 basic types of HAAR features shown below:

978-1-5090-0076-0/15 $31.00 © 2015 IEEE
DOI 10.1109/CICN.2015.144

IEEE computer society



Type I Type II Type III Type IV Type V.

Fig. 2. Five basic HAAR-feature types.

A 24x24 training window is used for the training purpose, which includes the features, like scale, position and type which can provide upto 1,60,000+ features.

INTEGRAL IMAGE.

In order to calculate the area of the above-mentioned rectangles that is presented in the haar features the use of Integral Image is made. The algorithm claims that, the knowledge of the co-ordinates of the four corners of the rectangle is mandatory which is used to calculate the area of the rectangle.

$$\sum(a)+\sum(L) \qquad (1)$$

a= co-ordinates of the rectangle of the required edge
L=co-ordinates of the edge located to the left of required edge.

4	1	2
6	3	8
5	9	7

→

4	5	7
10	14	24
15	28	45

Fig. 3. Theoretical Example for calculation of the mentioned formula .

Here the L.H.S is the normal image co-ordinates and the R.H.S is the co-ordinates generated using Integral Image.

PRACTICAL EXAMPLE:

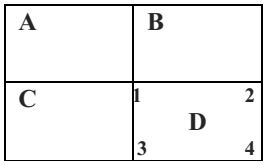


Fig. 4. Integral Image representation in terms of rectangle.

co-ordinates in two-dimensional system. Now the area of Block (D) can be given by the formula:-

Area(D)= ((1+4) – (2+3)). Area of Z can also be represented in the terms of the area of the other blocks. It is given by:-

Area(D) = A + (A+B+C+D) – ((A+B)+(A+C))

I. ADA-BOOST:

Haar-like features create 1,60,000+ features for every object for detection purpose. ADA-BOOST is a machine-learning algorithm used to track the best features for detection. Following the detection of these features a weighted compilation of all these features are made to check whether the given image matches the trained object or not.

Following formula demonstrates the features below:

$$F(x)=\alpha 1 f1(x)+\alpha 2 f2(x)+\alpha 3 f3(x)+\alpha 4 f4(x)... \qquad (2)$$

α = assigned weight.
$f_N(x)$ = weak classifiers (tracked features).
F(x) = compiling features to generate strong classifiers.
Consider the following figure:

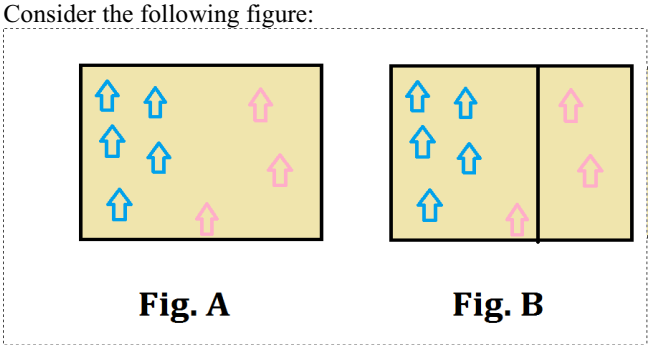


Fig. 5. Weights of positive and negative image.

The blue arrows shown in the figure are the true positive images or the required objects whereas the pink arrows shown are the negative images present in the frame. In figure (A) shown above we separate positive and negative images by drawing a black line leading to figure (B).

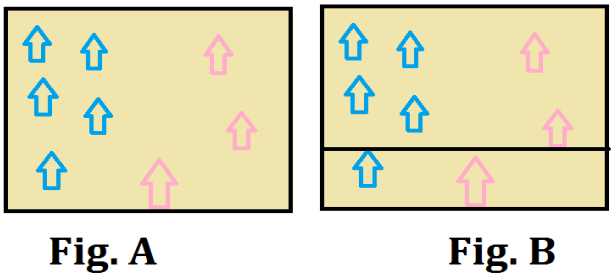


Fig. 6. Weights of positive and negative image.

Now it can be observed that there is a false detection in figure B. Hence, increase the weights of the falsely detected image, which leads to figure (A) above. Now again rearrange and classify the images based on their weights leading to Figure B. There are again two falsely detected images in Figure B. Hence again repeat the procedure where we increase the weights of the false detection Figure (B). Finally we separate the images with higher weights from the images with lower weights. After separation it becomes convenient to identify the positive figure and negative figure.

The use of ADA-BOOST algorithm is to find a rectangular threshold, which has the feature to separate the positive and negative images.

ADA BOOST ALGORITHM:

For n rounds we consider:

1.) Weighted error is evaluated for every feature and the best feature is chosen.
2.) Examples taken are to be re-weighted:
Incorrect examples: - Weight is Higher.
Correct examples: - Weight is Lower.

The outcome of the above mentioned method is good classifiers combining the features of all the weak classifiers, which is able to detect the desired object.

II. CASCADING CLASSIFIERS:

There is a limitation with the above-mentioned approach. The algorithm scans the same image repeatedly with different window size. This is gauging and energy consumption also occurs. Consider there are 'm' positive objects in an image. The number of negative images is >m. The algorithm must be effective to scan the faces with high speed and efficiency and must quickly discard the non-faces. Hence this condition can be satisfied by using the cascading the classifiers, which has higher speed than the linear combination of the classifiers which are quite time consuming. Hence cascading the classifier is a method of combining the strong classifiers. This ensures that the features are scanned in stage-by-stage manner. Every stage has a duty to check whether the given image contains a face or not. If the given image contains a face and satisfies classifier 1, it will be passed to the next stage of classifiers else the image window will be discarded.

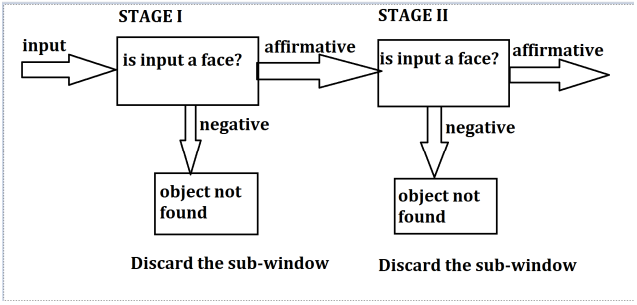


Fig. 7. Schematic flow of object detection.

The following are the 3 rules for designing a classifier:

1. Summation of number of stages.
2. Summation of number of features.
3. Threshold for every single strong classifier.

III. FACE TRACKING ALGORITHM:

A. *Mean Shift Algorithm:*

This real-time tracking algorithm is used to scan the given image for the requires object (In this scenarios, face is the required object).

Following are the rules for the algorithm:

1. Image to be converted into HSV format.
2. Select variable H (hue) of the image.
3. Choose the size of the search window.
4. Choose the initial location of the window.
5. Find the center of mass of the object.
6. Search the center window around center of mass
7. If convergence occurs jump to step 8, else jump to step 5.
8. Compute the parameters.

B. *Histogram Back projection:*

Following steps are included for this algorithm:

1. A window to be chosen for tracking the face.
2. Extract the hue values from the given window.
3. Formation of histogram using the hue values.
4. The histogram is as a reference as and when an image is captured in that frame where their values of pixels are converted to hues.

C. *CAM Shift:*

This algorithm can adapt to the screen size for the object, which is tracked. The size of the window is adjusted for every single iteration and is calculated on the moment of the previous window.

Consider S as the size and M as the Moment.

$$S = 2 * \sqrt{M_{00} / 256} \quad (4)$$

Width of the window = S

Length of the Window = 1.2*S

M_{00}= Moment of order-zero.

CONNECTIVISM – TOOL FOR LEARNING:

The human-human interaction is the age old and an efficient methodology to impart education. In this digital age, Connectivism is an ideology, which can be used to expand the utility and efficiency of human-computer interaction. Connectivism defines the knowledge, which is present in the non-human electronic appliances. Prior to connectivism there existed 3 other methodologies.

They are:

1. Behaviorism.
2. Constructivism.
3. Cognitivism.

These above 3 methodologies are not highly efficient in delivering the knowledge from electronic appliances to humans. Here it is also observed that the previous methods are a bit slow in the use of resources and are inefficient. Hence to overcome these hurdles Downes and Siemens introduces a methodology known as connectivism, which is used to spread the knowledge by utilizing all the resources.

CONNECTIVISM PRINCIPLES:

1. Expanded range of opinions for delivering knowledge.
2. Information services and their special nodes connected by learning as a tool.
3. Increased focus on acquiring knowledge than currently present knowledge.
4. Information can dwell in non-human appliances.
5. Efficient knowledge sharing can be achieved using proper maintenance of connectivism and cultivating them.
6. Maintenance of connections established between ideas concepts and fields.
7. Accuracy and updated information is the currency in Connectivism.
8. Learning process is fueled up by decision-making.

The decision-making also comprises of two patterns. Some decision-makings must be knowledge driven which implies assuming the environment to be an ideal one. Whereas

some decision though comply with the available knowledge should not be taken but decided based on experience. This is called experience-driven approach. Hence appropriate methods must be used at appropriate situations.

Briefly a word can be put as the connectivism, which is not simply a theory but an ideology, which can be used to cultivate and improvise the spreading of knowledge and also increase the human-computer interaction. The use of the above algorithms of machine learning and computer vision can adapt to the surrounding and perform the teaching in an efficient manner.

TEACH AND LEARN TECHNIQUE:

The "teach and learn" technique is dual-purpose methodology proposed which improves the knowledge of the students and also increases the accuracy of the software by imparting machine learning.

Following steps are involved in the teach and learning technique:

1. Check whether the student is deaf or blind or normal.
2. If student is deaf impart the Teaching technique T1.
3. Impart technique T2 for the blind students.
4. For normal students, start with the 11 questions to scan the attributes of the students. Out of these 11 questions 09 are randomized from the stored 150 psychological questions.
5. Scan the answers given by the students and scan for the teaching method to be imparted which is from the set T= [T1, T3, T4, T5, T6] obtained by initial training.
6. After teaching ask the student for satisfactory results. If affirmative update the database with the data scanned for the student. Else discard the data and leave the control in student's hand.
7. Repeat step 1 to step 6 for every new student.

Let A be the set of attributes taken for scanning the value of students. A= {Marks, Attendance, Concentration-level, Regularity-level, Public speaking confidence-level, Presentation on paper-level, Concept understanding-level, Imagination Skills-level, English understanding level, Self-learning level}

NOTE: All the levels are based on 4 types namely VERY HIGH (>90%), HIGH(75%-90%) , MEDIUM (40-75%), VERY LOW (<40%).

TEACHING METHODS:

T1 = Pure Vocal teaching by scanning concentration.
T2 = Concentration scanning for display and blink screen on deviation.
T3 = Teaching with basic flow diagrams and co-relation using vocal assistance and concentration detection.
T4 = Teaching with videos with pause and play methods Controlled by concentration detection.
T5 = Basic Book Reading.
T6 = Teaching using Question-Answer Format which acts as a flow for the knowledge gain.

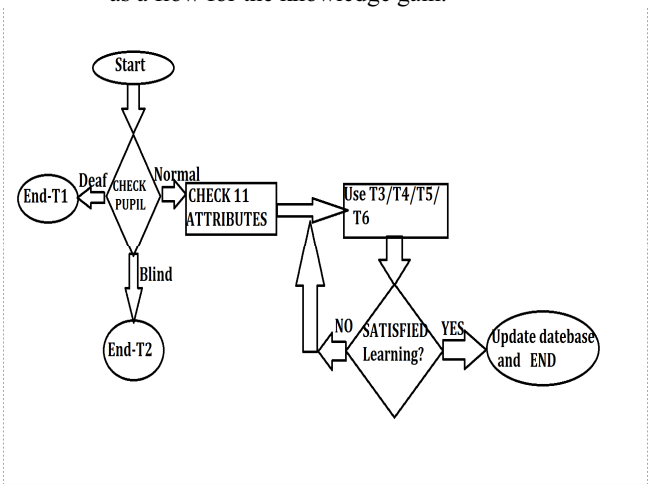


Fig. 8. Flow Of Teach and Learn Technique.

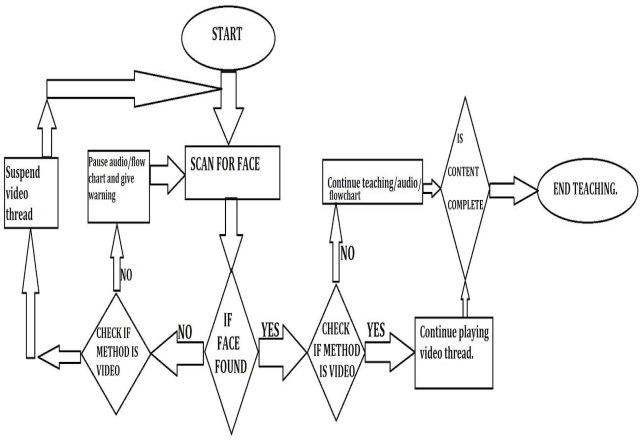
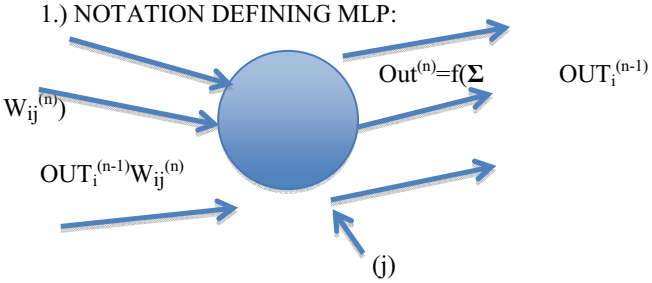


Fig. 9. Flow Chart for concentration detection.

MULTI-LAYER PERCEPTION(MLP):

Single-Layer perceptions cannot deal with the problems of non-linearity like the XOR problems. Hence the introduction to Multi-Layer Perception was made where N inputs can have N patterns to generate N outputs.

1.) NOTATION DEFINING MLP:



$OUT_i^{(n-1)}$ $W_{ij}^{(n)}$ is the activation received by each unit j(circle) in the layer N and sends $OUT_i^{(n)}$ to the next layer of units. The convention methodology of input layer is layer 0, but when consideration is a N layer network, it has N layers of weights and N non-input layers for the processing. The following diagram explains a two layer

MLP:

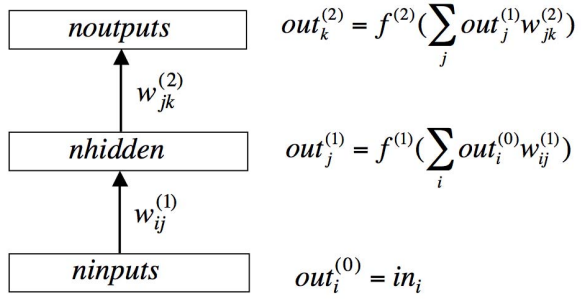


Fig. 10. Flow of a 2-layer MLP.

Thus it becomes clear about the addition of more layers into the network can also be opted even if 2 layers are sufficient. Hence for different layers, different activation functions can be used ($f^{(n)}(x)$) or for different units within a layer.

STEPS FOR TRAINING A MLP:

1.) Select the number of inputs, outputs and training patterns. (in_i^p ,out_i^p, i=[1,n-inputs], j=[1,n-outputs], p=[1, n-patterns]).
2.) Select n-inputs to setup the network along with n-outputs an N-1 hidden layers of n-hiddenlayers to train. Activate connections between each layer(n)to previous layer (n-1) using weight connections $w_{ij}^{(n)}$.
3.) Random Initial Weights to be generated.
4.) Appropriate error function $E(w_{jk}^{(n)})$ and learning rate (η).
5.) Weight update equation is applied to each weight $w_{jk}^{(n)}$ for each training pattern p.
6.) Repeatation of step 5 is done till Error function is insignificant.

RESULTS:

No.	Attribute1 Numeric	Attr2 Numeric	Attr3 Nominal	Attr4 Nominal	Attr5 Nominal	Attr6 Nominal	Attr7 Nominal	Attr8 Nominal	Attr9 Nominal	Attr10 Nominal	Attr11 Nominal	Answer Nominal
1	63.0	200.0 H	H	VH	M	M	M	M	M	M	M	B
2	73.0	184.0 H	VN	H	VH	H	H	VH	H	H	B	
3	83.0	187.0 VH	VH	VHH	H	H	H	VH	VH	VH	A	
4	88.0	190.0 H	H	VH	M	VH	H	M	VH	VH	C	
5	60.0	180.0 H	M	H	M	M	M	M	M	M	C	
6	90.0	181.0 H	VH	VH	M	VH	VH	M	VH	VH	C	
7	97.0	200.0 VH	VH	VH	VH	VH	VH	VH	VH	VH	A	
8	90.0	200.0 VH	M	VH	H	VH	VH	VH	VH	VH	A	
9	46.0	200.0 VH	L	H	L	L	L	VH	L	L	E	
10	71.0	150.0 M	M	VH	M	H	H	M	H	H	C	
11	99.0	175.0 VH	VH	VH	H	VH	VH	VH	VH	VH	E	
12	100.0	182.0 VH	VH	VH	VH	VH	VH	VH	VH	VH	A	
13	75.0	193.0 H	M	H	M	H	H	M	H	H	B	
14	55.0	188.0 M	L	H	M	M	L	M	L	L	E	
15	76.0	177.0 H	H	H	M	H	H	M	H	H	C	
16	42.0	147.0 L	L	M	L	L	L	M	L	L	E	
17	97.0	200.0 VH	VH	VH	H	VH	VH	VH	VH	VH	A	
18	75.0	185.0 H	H	VH	H	H	H	H	H	H	C	
19	82.0	185.0 VH	VH	VH	H	H	H	VH	VH	VH	B	
20	71.0	190.0 H	VH	VH	H	H	H	VH	H	H	A	
21	83.0	180.0 VH	VH	VH	VH	H	H	VH	VH	VH	A	
22	70.0	200.0 VH	H	VH	VH	H	H	VH	H	H	B	
23	88.0	200.0 VH	H	V	H	VH	H	VH	VH	VH	A	
24	77.0	200.0 VH	H	VH	H	H	H	VH	H	H	C	

Table 1. Data for Attributes in set A mentioned above

Here in Figure 13 attributes are in the order mentioned in set A.

Abbreviations: VH= Very High, H=High, M=Medium, L=Low, Attr=Attributes.

Answer Section: A = t1. B=t3. C=t4. D=t5. E=t6. Here t_i is the teaching method described above.

```
Correctly Classified Instances          33               73.3333 %
Incorrectly Classified Instances        12               26.6667 %
Kappa statistic                          0.6507
Mean absolute error                      0.1292
Root mean squared error                  0.3097
Relative absolute error                 40.6544 %
Root relative squared error             75.8668 %
Total Number of Instances               45

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0         0.025     0          0         0          0.595    B
                 1         0.086     0.769      1         0.87       1        A
                 0.75      0.081     0.667      0.75      0.706      0.909    C
                 0.857     0.129     0.75       0.857     0.8        0.919    E
                 0.625     0.027     0.833      0.625     0.714      0.949    D
Weighted Avg.    0.733     0.081     0.671      0.733     0.695      0.905

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
  0  3  2  0  0 |  a = B
  0 10  0  0  0 |  b = A
  1  0  6  1  0 |  c = C
  0  0  1 12  1 |  d = E
  0  0  0  3  5 |  e = D
```

Fig. 11. Accuracy for the trained data from the students attributes.

The data trained from the students has an accuracy of 73.33%, which increases as the software is deployed in the real-time scenarios.

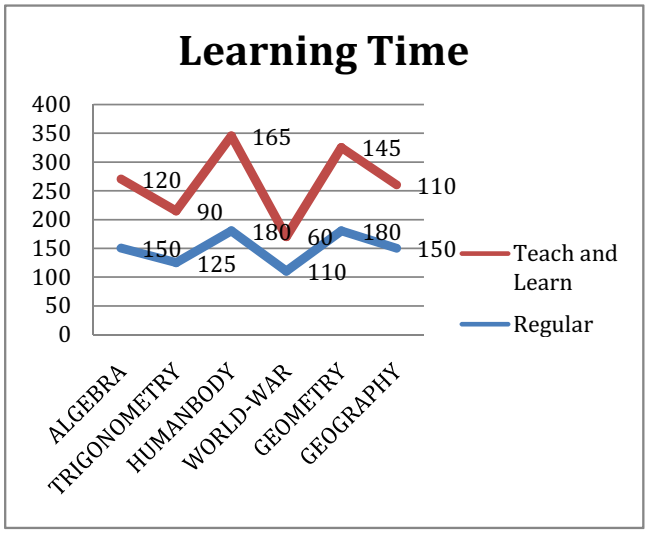


Fig. 12. Comparison graph for learning time for the subjects mentioned

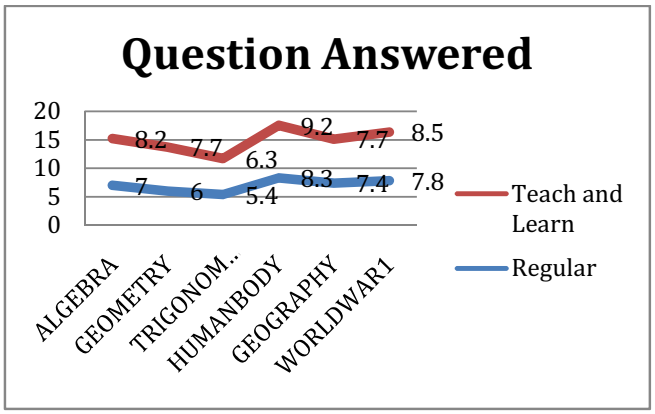


Fig. 13. Comparison graph for average questions answered for the subjects mentioned.

DISCUSSION.

Here the Table 1 shows the data for the attributes. These attributes were scanned from the students based on their 9th grade marks in mathematics and their attendance. Apart from them 150 questions were chosen from Internet, which were used to decide the 9 attributes other than scores and attendance. 200 students were divided in a group of 4 and each module from t1, t4, t5, t6, t7 were given to them with the subjects shown in the graphs and 50 students were asked to study using their regular style. Later the learning time was recorded which is shown in the graph Figure 12. After the students completed the studies an examination was conducted based on these subjects where 20 objective questions were asked and an average was taken out from the correctly answered questions. From this the graph figure 13 was calculated. The graph shown in the figures are cumulative graph with individual values described on every point on the figure to increase the readability of the graph. For the blind students we can track their response time for the attribute questions asked. If the student does not respond within the speculated time then the software considers him blind and imparts teaching t1. Teaching t1 is also found efficient to in some cases where the student is not blind. For deaf students a vocal voice is given and options are asked to find the spoken word. If the student does not answer it is assumed that student is deaf. Message spoken is "PLEASE TICK THE FOLLOWING WORD" and word is spoken. The software after completing the teaching will update the data of the present student in the database, which in turn increases the accuracy using the multi-perception algorithm. Hence more the software is put to use more intelligence is acquired. This is the advantage of TEACH and LEARN technique.

CONCLUSION:

Here the use of multi-perception algorithm is made as the first phase and then followed with the flavor of Haar features for face detection, which breaths intelligence and also provides an extended version of human-computer interaction. The 6 teaching methods deployed in this project suffice the needs of the deaf students and blind students plus support the enhancement of the learning by the normal students. The 11 attributes are used to analyze the type of learning method to be imparted for the student for which the students were exposed to these different teaching methods and an appropriate method was chosen for each student. Connectivism is a technique used to provide efficient communication for the human-computer interface. The basic principles of connectivism expand the range of information and knowledge delivery. The Teach and Learn method is used as a dual-purpose system for the improvement of software and the student. Here after the student completes the learning the satisfactory report is taken which if affirmative, updating of the database with the students' attributes is done to improve the accuracy. If not the control is given to students where students can themselves choose the teaching technique. Thus the data is discarded. This helps in improving the accuracy and improves the performance of the software. Accuracy results are 73.3%, which can speed up the students' learning speed and accuracy by 1.3-1.5 times as per the results observed. For teaching a specific person their attributes can be hardcoded and Eigen-Face algorithm can be used for face recognition in place of the Haar based algorithms.

Hence we can say that the technology proposed can improve the students' learning capabilities significantly.

ACKNOWLEDGEMENT:

I would like to thank Mathematician T.Ananthi for helping me to test the software on the students and obtaining other necessary data from them for this research.

REFERENCES:

[1] Bharadwaj Vishanth Thiyagarajan, "Facial Detection and facial feature points' detection with the help of KLT algorithm", IJARCSMS, vol. 2,issue 8, august 2014.

[2] Bharadwaj Vishanth Thiyagarajan, "Revelution in Teaching/Moocs by the use of real-time face detection", IEEE ICECCT Conference, pp. 503-510, March, 2015.

[3] M. H Yang, "Detecting faces images, A survey," IEEE Transations Pattern Analysis and Machine Inteligence vol. 24, no. 1, pp. 34–58, 2002.

[4] Frances Bell, "Connectivism: Its Place in Theory-Informed Research and Innovation in Technology-Enabled Learning", IRRODL journal, March-2014.

[5] P. Viola and M. Jones, "Robust Real-time Object Detection," 2nd international workshop on statistical and computational theories of vision - modeling, learning, computing, and sampling, pp. 1-25, 2001.

[6] Smith M.K. (1999), "Learning Theory: The encyclopedia of informal education." retrieved from www.infred.org.

[7] Siemens, G(2005), "Connectivism: A learning theory for digital age", Retrieved from www.elearnspace.org.

[8] M.W Garrdner and S.R Dorling, "Artificial Neural Network(The Multilayer perception)- A review of application in the atmospheric sciences", Pergamon, pp. 2627-2636 , June 1998.

[9] Gaurang Panchal, Amit Ganatra, Y P Kosta and Devyani Panchal, "Behaviour Analysis of Multilayer Perceptron with Multiple Hidden Neurons and Hidden Layers", IJCTE, Vol. 3 Issue 2, April 2011.

[10] Jao A. Mattar, "Constructivism and Connectivism in Education Technology: Active, Situated, Authentic, Experiential, and Anchored Learning", Boise State University, 08/2010.