GUJARAT TECHNOLOGICAL UNIVERSITY

SARVAJANIK COLLEGE OF ENGINEERING AND TECHNOLOGY

Report on:

OPTICAL CHARACTER RECOGNITION BASED CHARACTER SCANNING

Under subject of

DESIGN ENGINEERING - 2A
B.E. III, Semester - V
Electronics and Communication Engg.

Team id: 22269

Submitted by:-

Sr.	Name	Enrolment Number
1	**Ahmed Godil**	**160420111003**
2	**Archana Bajari**	**160420111006**
3	**Parantap Vakharwala**	**160420111031**
4	**Lakshesh Vyas**	**160420111060**

PROF. Sarosh Dastoor
faculty guide

PROF. Dr. Maulin Joshi
Head of Department

Academic Year
(2018-19)

CONTENTS

TABLE OF FIGURES

ACKNOWLEDGEMENT

We would like to express our deepest appreciation to all those who provided us the support to complete this project. A special gratitude to our guide, Prof. Sarosh Datoor whose contribution to stimulating suggestions and encouragement, helped us to coordinate our project.

We respect and thank our project guide for giving us an opportunity to do the project work in optical character recognition based character scanning and providing us all support and guidance which made us to complete the work on time.

Last but not the least, many thanks go to the head of the project who have given the full effort in guiding the team in achieving the goal as well as an encouragement to maintain our progress in track. And we are thankful that he has improved our presentation skills by his comments and tips.

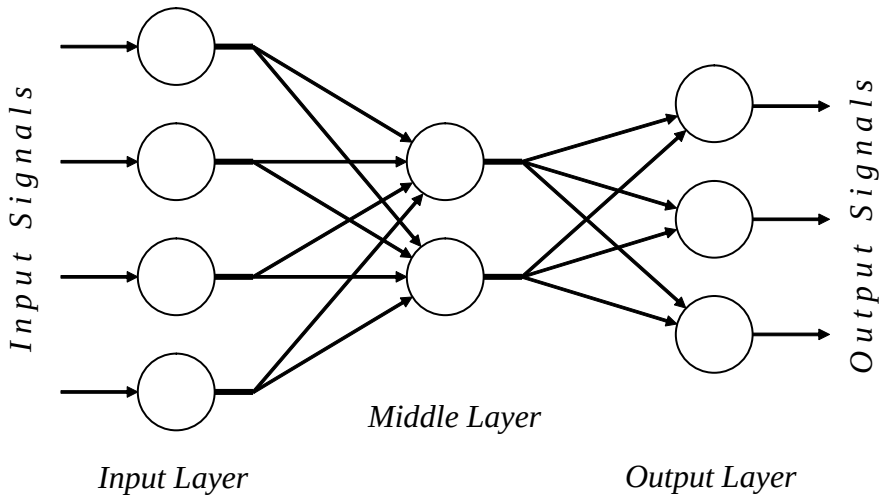
1. ABSTRACT

As with the advent of computer science we have increasing availability to more powerful computers. This got us to believe that by the help of neural network we could train a system to interpret alpha numeric characters.

Use of Artificial Neural Network will aid us in generating and developing a system for both training and test phases of the project. Matlab has been used for the application as it has all the libraries and functions required to carry out the tasks.

2. INTRODUCTION

1) Optical Character Recognition (OCR) is to classify optical patterns corresponding to alphanumeric or other characters.
2) It type of document image analysis where a scanned digital image that contains either machine printed or handwritten script is input into an OCR software engine.
3) OCR works by first pre-processing the digital page image into its smallest component parts with layout analysis to find text blocks, sentence/line blocks, word blocks and character blocks
4) A **neural network** can be defined as a model of reasoning based on the human brain. The brain consists of a densely interconnected set of nerve cells, or basic information-processing units, called **neurons**.
5) An artificial neural network consists of a number of very simple processors, also called **neurons**, which are analogous to the biological neurons in the brain.



6) The neurons are connected by weighted links passing signals from one neuron to another.
7) The neuron computes the weighted sum of the input signals and compares the result with a **threshold value**,

8) The neuron uses the following transfer or **activation function**:

$$X = \sum_{i=1}^{n} x_i w_i \quad Y = \begin{cases} +1, & \text{if } X \geq \theta \\ -1, & \text{if } X < \theta \end{cases}$$

,

Feedforward neural network:

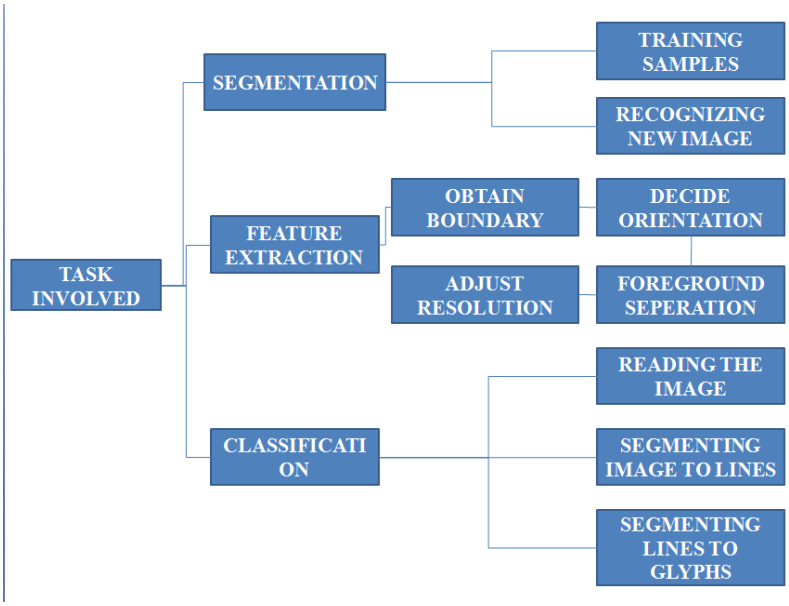
1) A **feedforward neural network** is an artificial neural network wherein connections between the nodes do *not* form a cycle.
2) In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes.

SUPERVISED LEARNING:

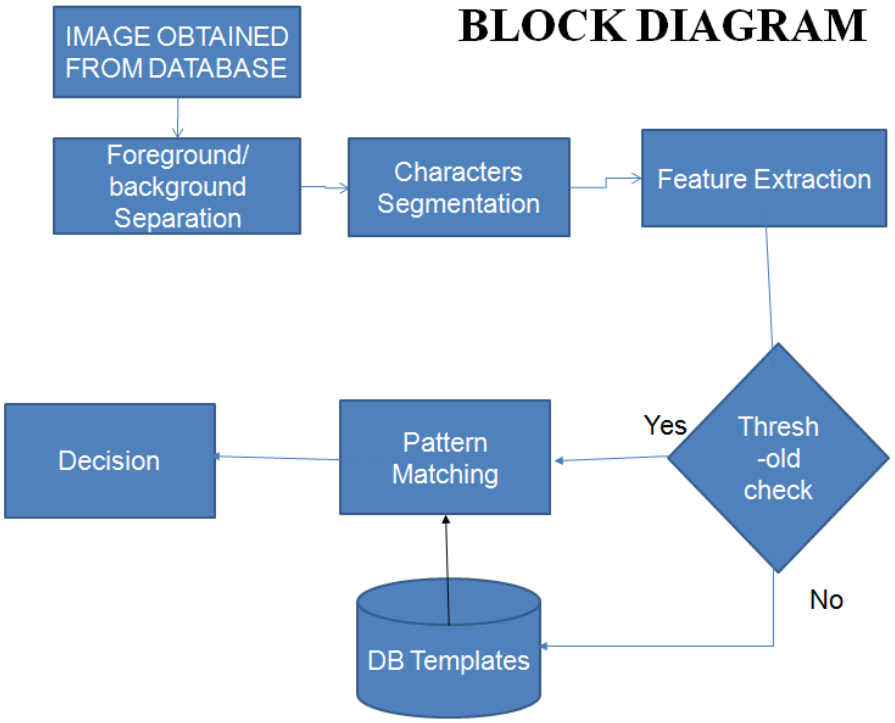
Supervised learning is where you have input variables (x) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output.

It is called supervised learning because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process.

THINGS TO DO:



BLOCK DIAGRAM



3. AEIOU FRAMEWORK

1. Activities
2. Environment
3. Interaction
4. Object
5. User

3.1 ACTIVITIES

These are **goal-directed sets of actions**—paths towards things people want to accomplish. What are the modes people work in, and the specific activities and processes they go through. Activities are goal directed sets of actions—paths towards things people want to accomplish. It shows the activities observed on the location. It also specifies the features, elements and special notes.

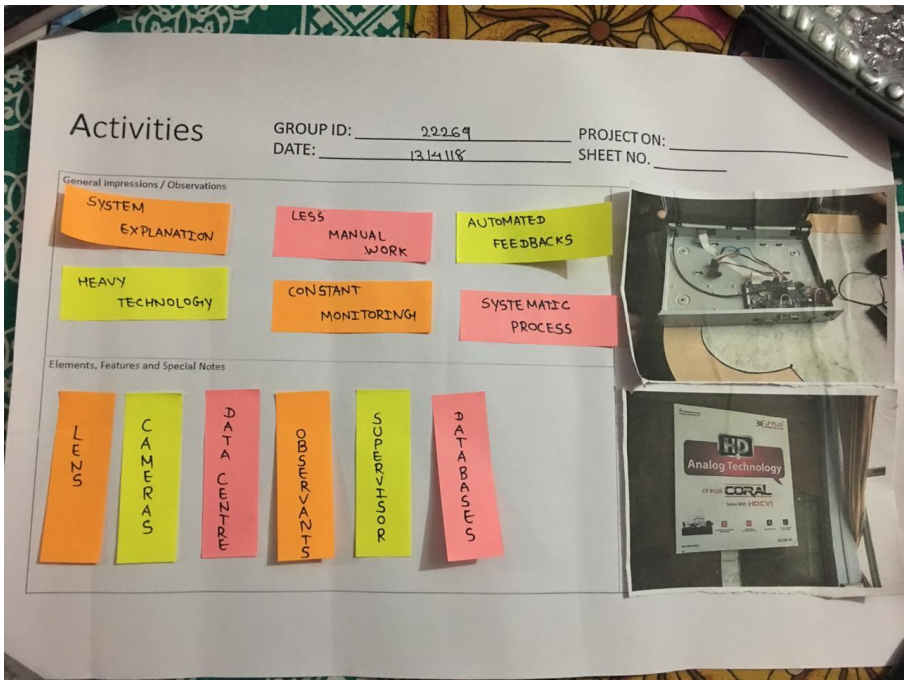


Figure 1 Activity

3.2 ENVIRONMENT

What is the character and function of the space overall, of each individual's spaces, and of shared spaces?

ENVIRONMENTS –THE SURROUNDINGS…

-Where (Activity) is happening?

-What are circumstances?

-Time effects, any?

-Seasonal effects, any?

-Person effects, any?

-Operational effects, any?

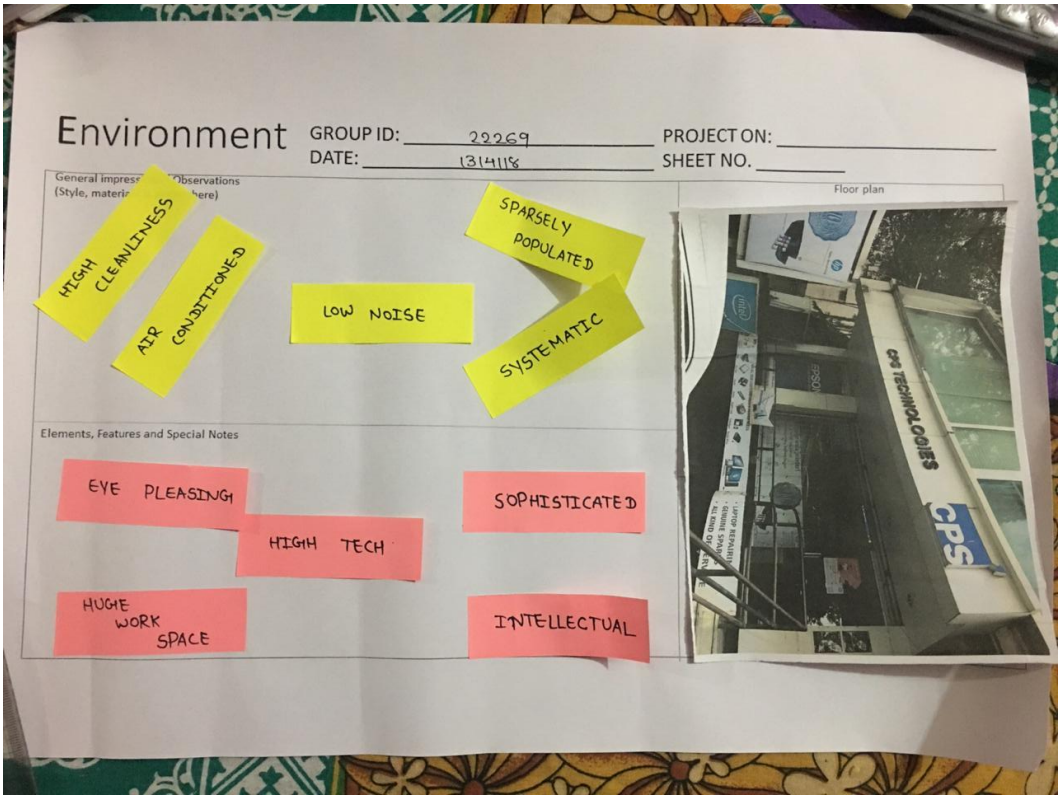


Figure 2 Environment

3.3INTERACTION

Interactions are between a person and someone or something else; they are the building blocks of activities. These are between a person and someone or something else; they are the building blocks of activities.

What is the nature of routine and special interactions between people; between people and objects in their environment, and across distances?

-Who are involved? Why? How?

-What (users) are holding/ operating/ using…

-How the contacts are in action?

-Action vs. Reaction…

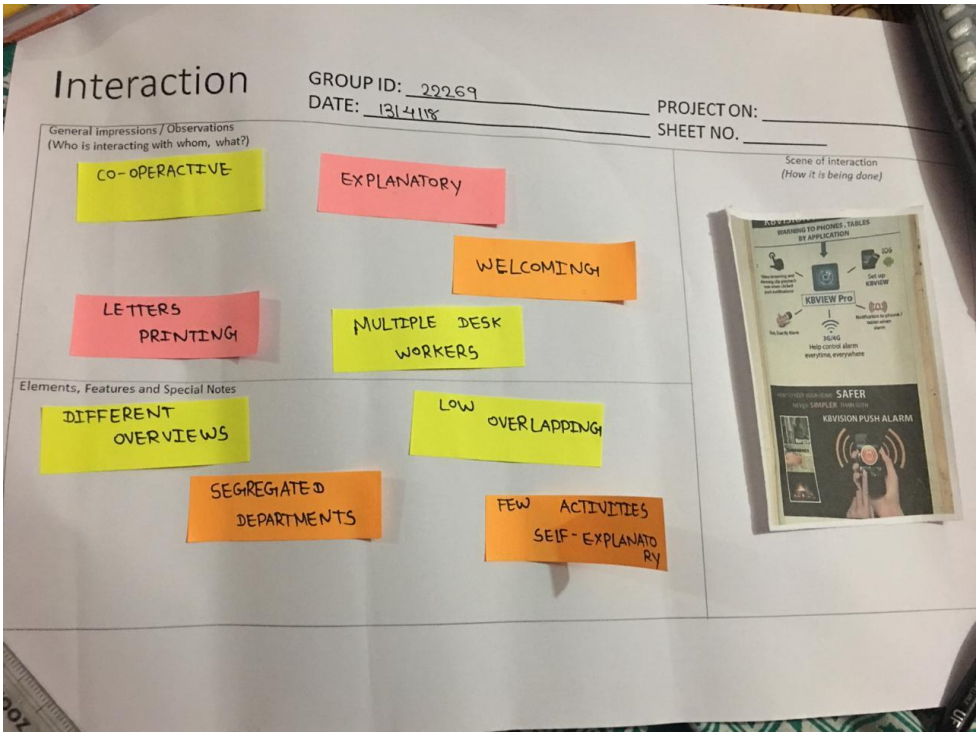


Figure 3 Interaction

3.4 OBJECT

These are building blocks of the environment, key elements sometimes put to complex or unintended uses (thus changing their function, meaning and context).

What are the objects and devices people have in their environments and how do they relate to their activities?

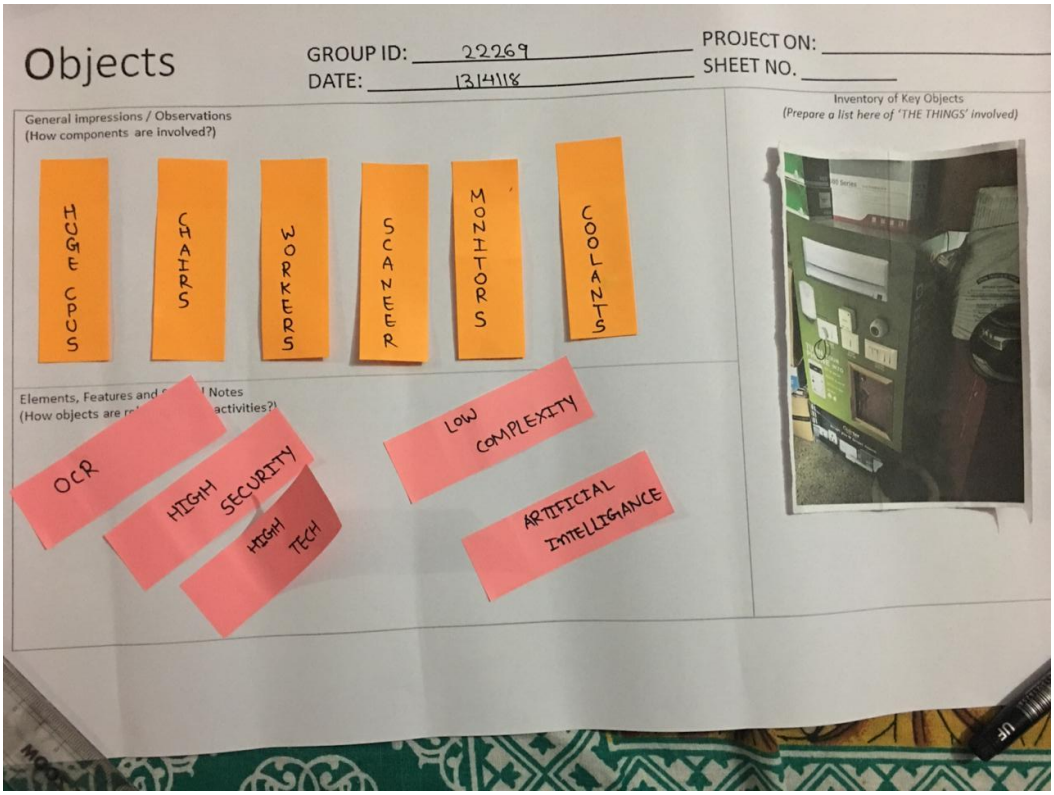


Figure 4 Object

3.5 USER

Users are the people whose behaviors, preferences, and needs are being observed. What are their roles and responsibilities, the persons co related to them, and features and special notes on the basis of them are taken into account.

Users are the people whose behaviours, preferences, and needs are being observed. Who is there? What are their roles and relationships? What are their values and prejudices?

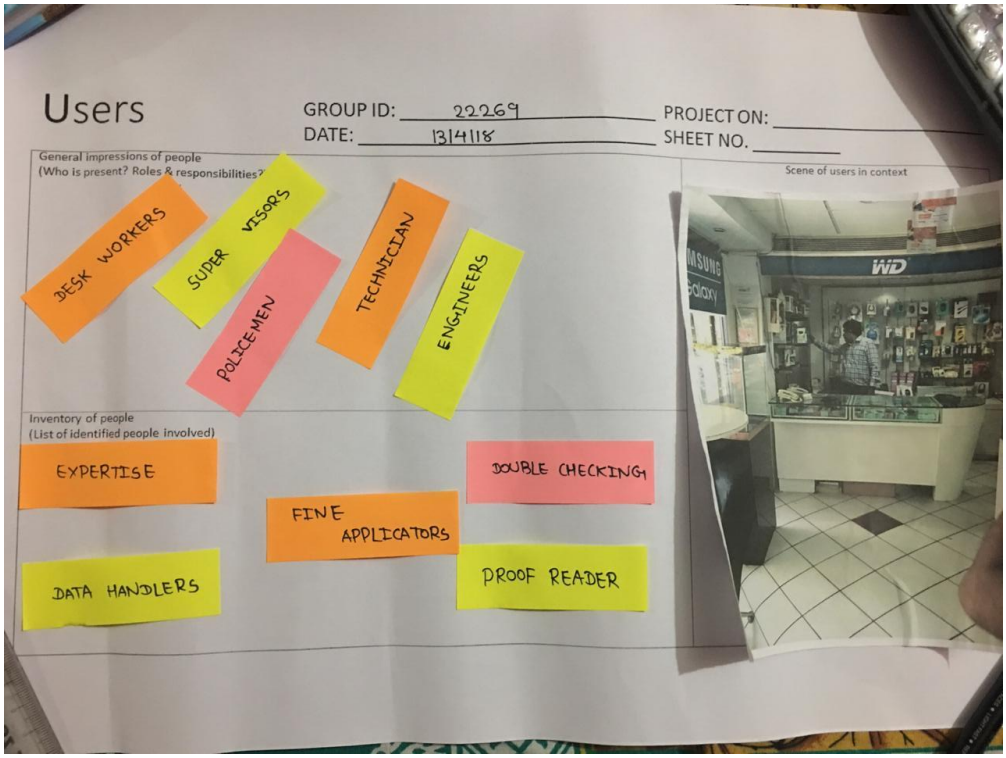


Figure 5 User

4. EMPATHY MAPPING

With the concept of design engineering we started to think of various problems occurring worldwide which we can resolve by our technical skills to contribute to our society as budding engineers.

4.1 USERS FOR OUR PROPOSED IDEA OF THE PROBLEM

Figure 6 Empathy-User

4.2 STAKEHOLDERS

Figure 7Emapthy-stakeholders

4.3 ACTIVITIES

Figure 8 Empathy-activities

4.4 STORY BOARDING

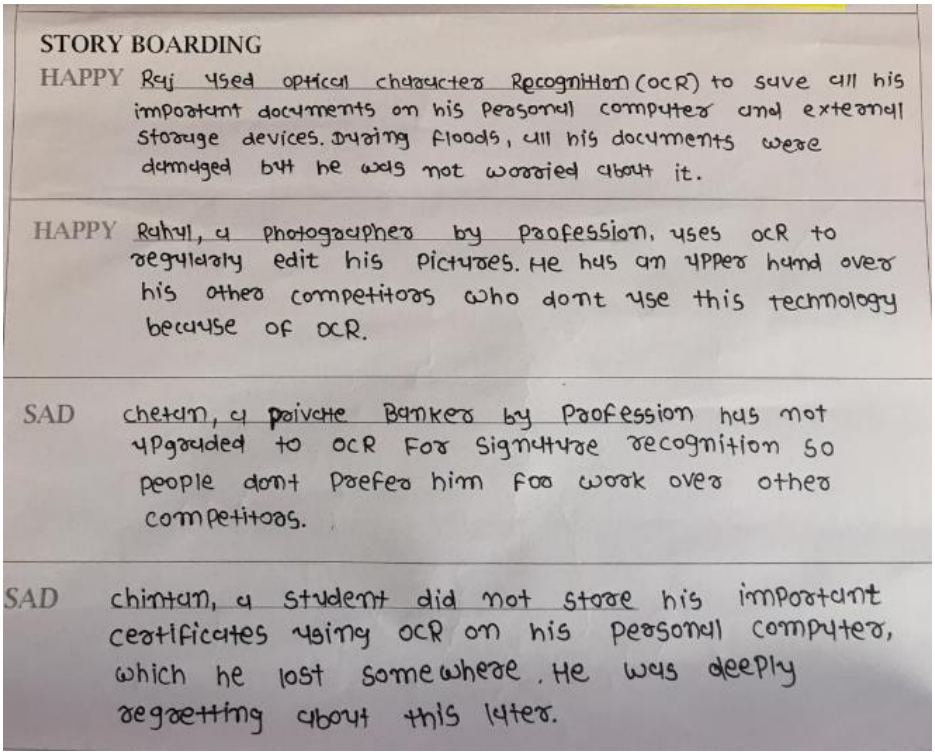


Figure 9 Empathy-storyboardin

4.5 CANVAS PICTURE

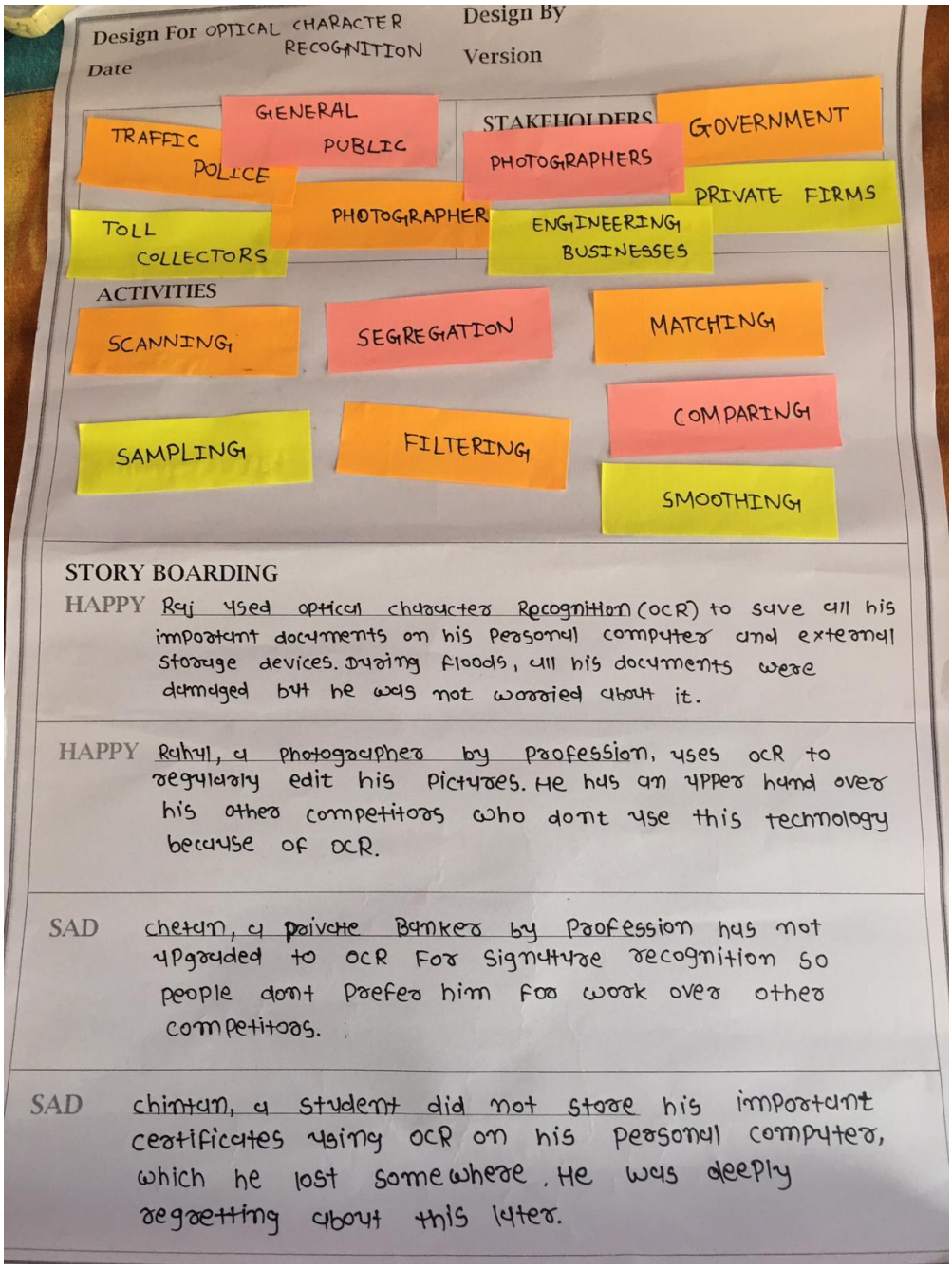


Figure 10 Empathy Canvas

5. IDEATION CANVAS

5.1 PEOPLE

This section can have all the people involved and related to the domain of 'OCR'. This may also have all the people related to the users as well as stakeholders of the domain 'OCR'.

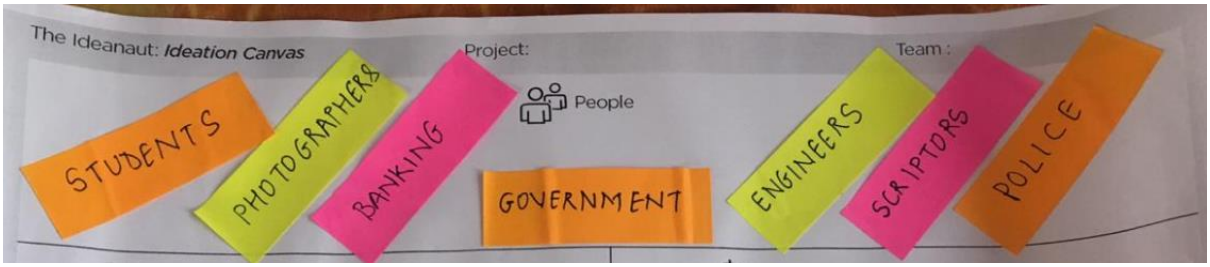


Figure 11 Ideation- people

5.2 ACTIVITIES

This section has activities of all the people mentioned above in the “People” section in the ideation Canvas. What they may do on the roads and on the traffic signal is noted in form of keyword.

Figure 12 Ideation- activities

5.3 SITUATION/CONTEXT/LOCATION

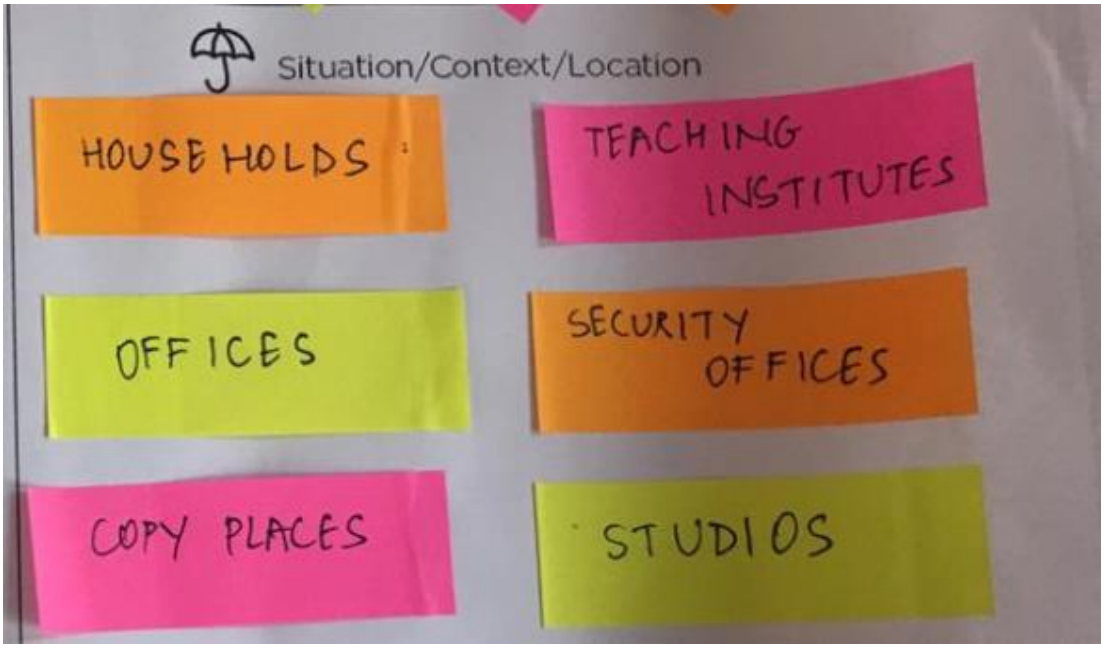


Figure 13 Ideation- situation

5.4 PROPS/POSSIBLE SOLUTIONS

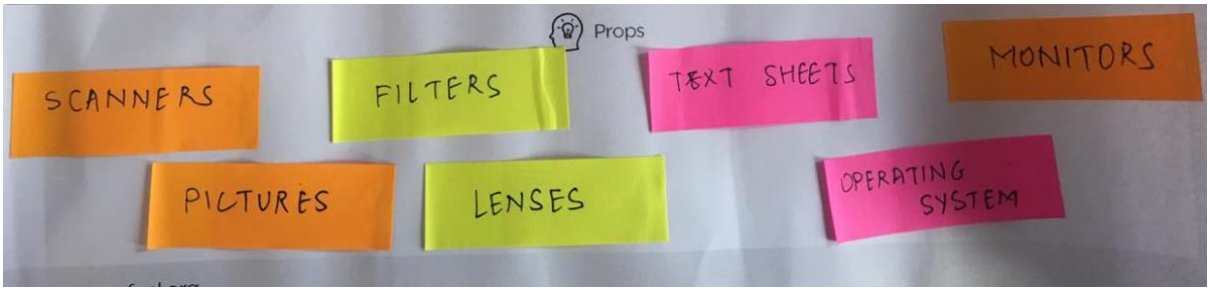


Figure 14 Ideation- props

5.5 CANVAS PICTURE

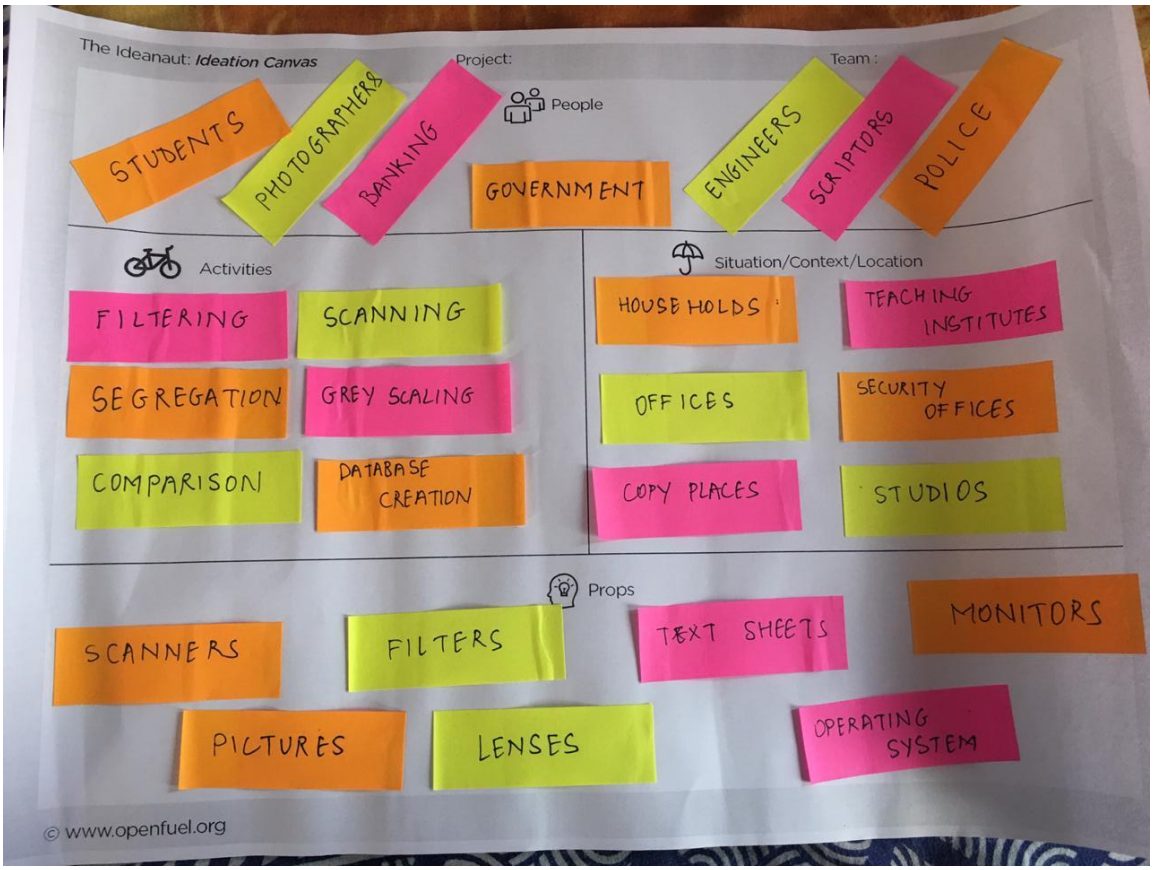


Figure 15 Ideation Canvas

6. PRODUCT DEVELOPMENT CANVAS

6.1 PURPOSE

Figure 16 Product- purpose

6.2 PEOPLE

Figure 17 Product- people

6.3 PRODUCT EXPERIENCE

- It ensures perfect character match
- Vast number of hidden layers ensures error less charcter match
- Traning set can be customised according to need.

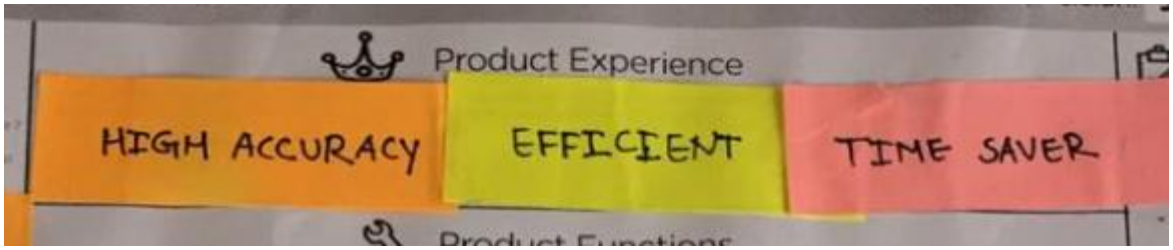


Figure 18 Product Experience

6.4 PRODUCT FUNCTIONS

- Neural Network based charcter recognition.
- Can be further trained to identify various fonts aswell.

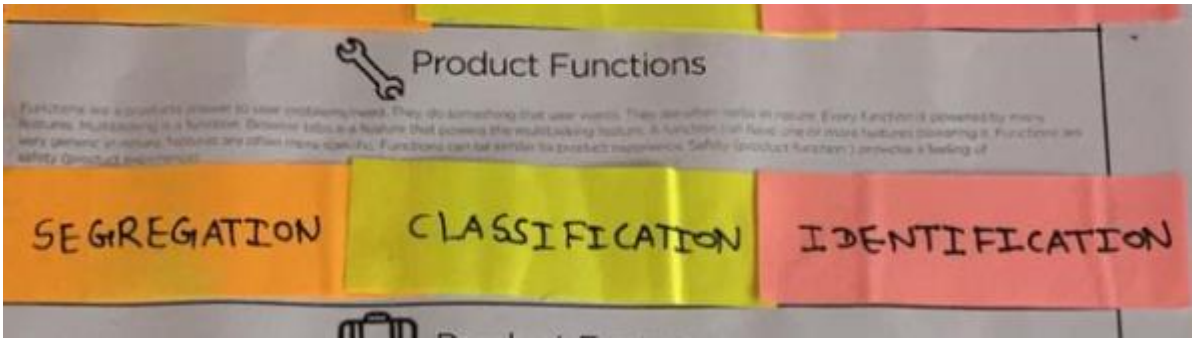


Figure 19 Product Function

6.5 PRODUCT FEATURES

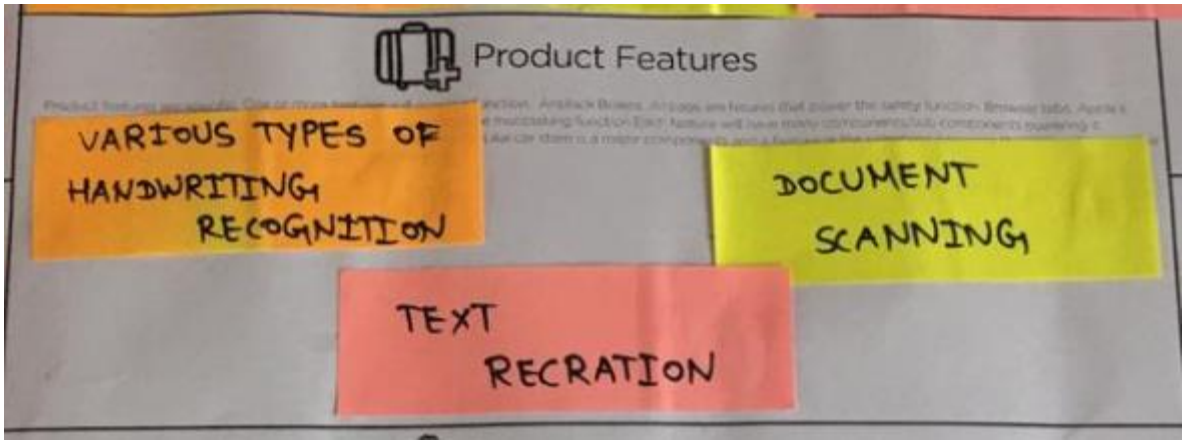


Figure 20 Product Features

6.6 COMPONENTS

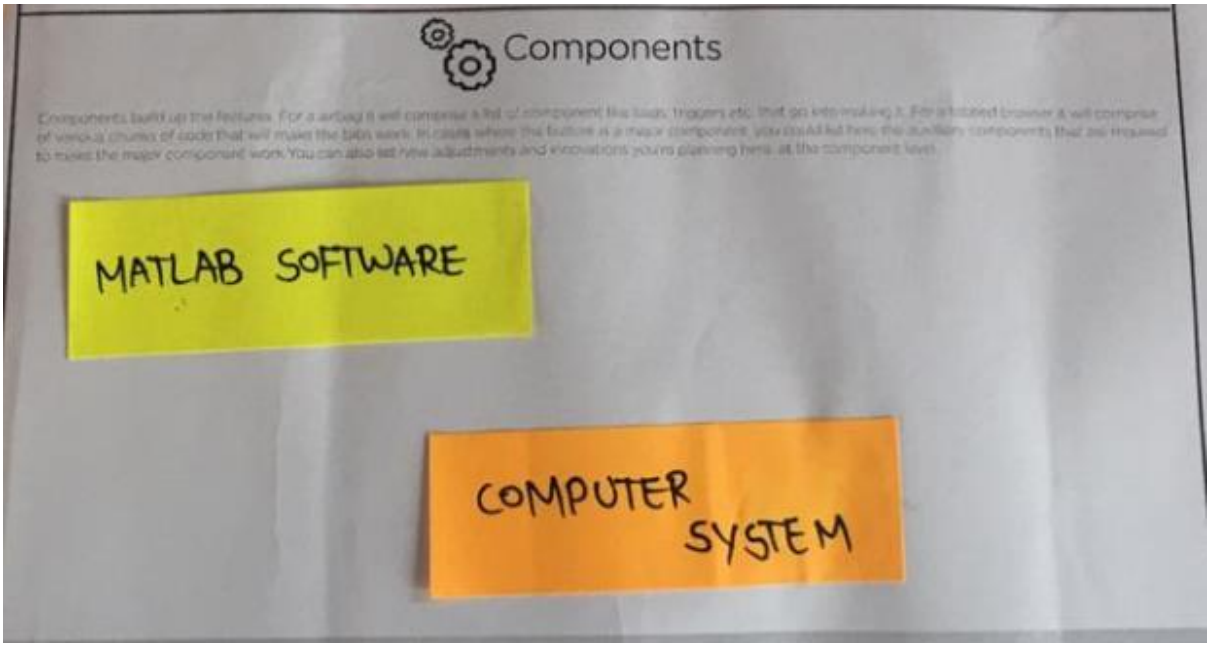


Figure 21 Product Components

6.7 CUSTOMER REVALIDATION

Figure 22 Customer Revalidation

6.8 REJECTIONS OR REDESIGN OR RETAIN

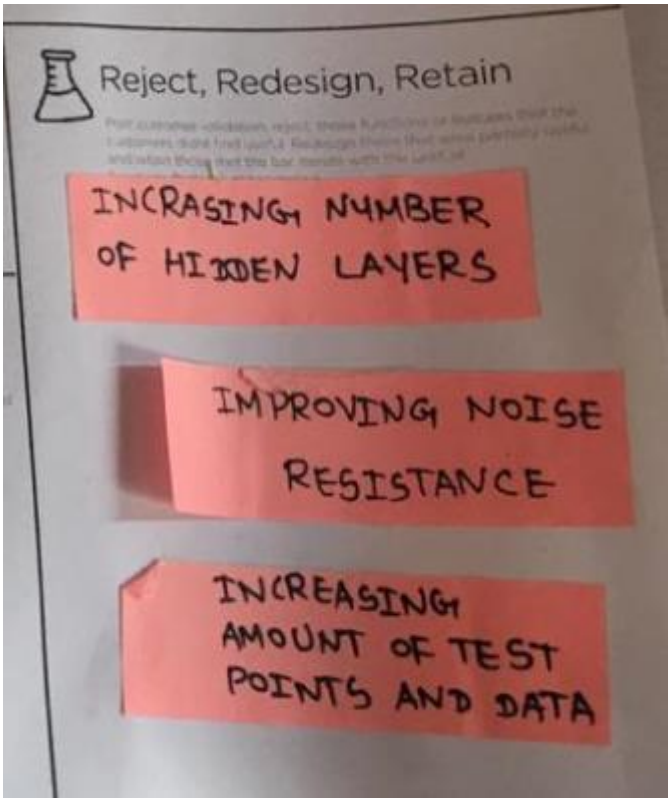


Figure 23 Rejection or Retain

6.10 CANVAS PICTURE

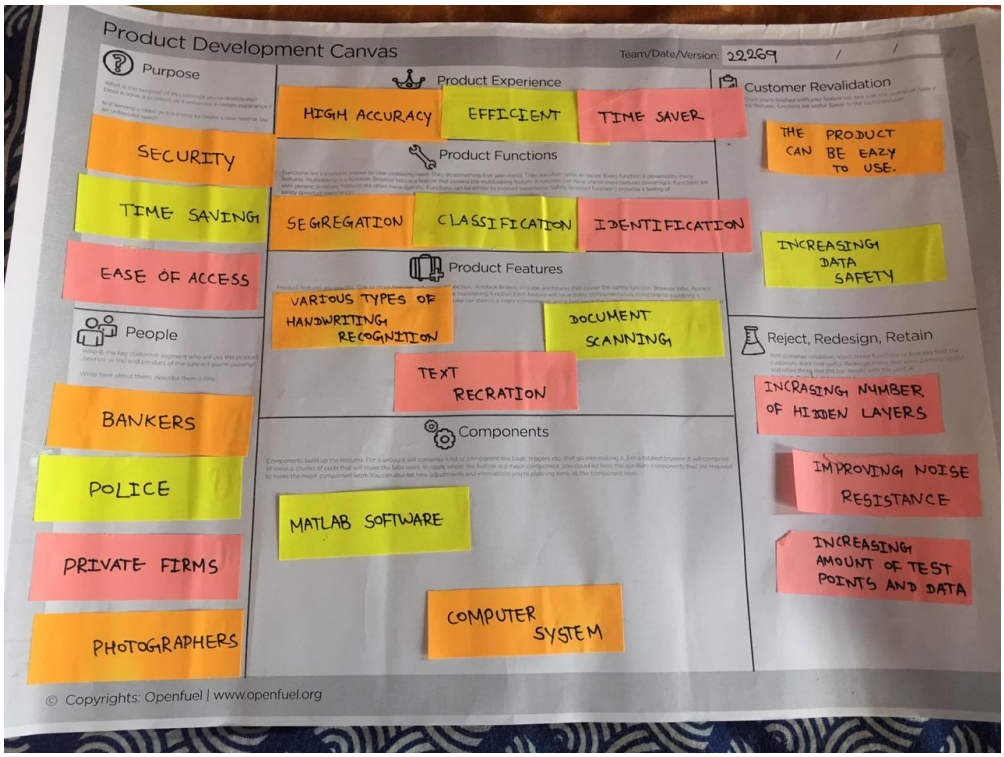


Figure 24 Product Canvas

7. MIND MAPPING

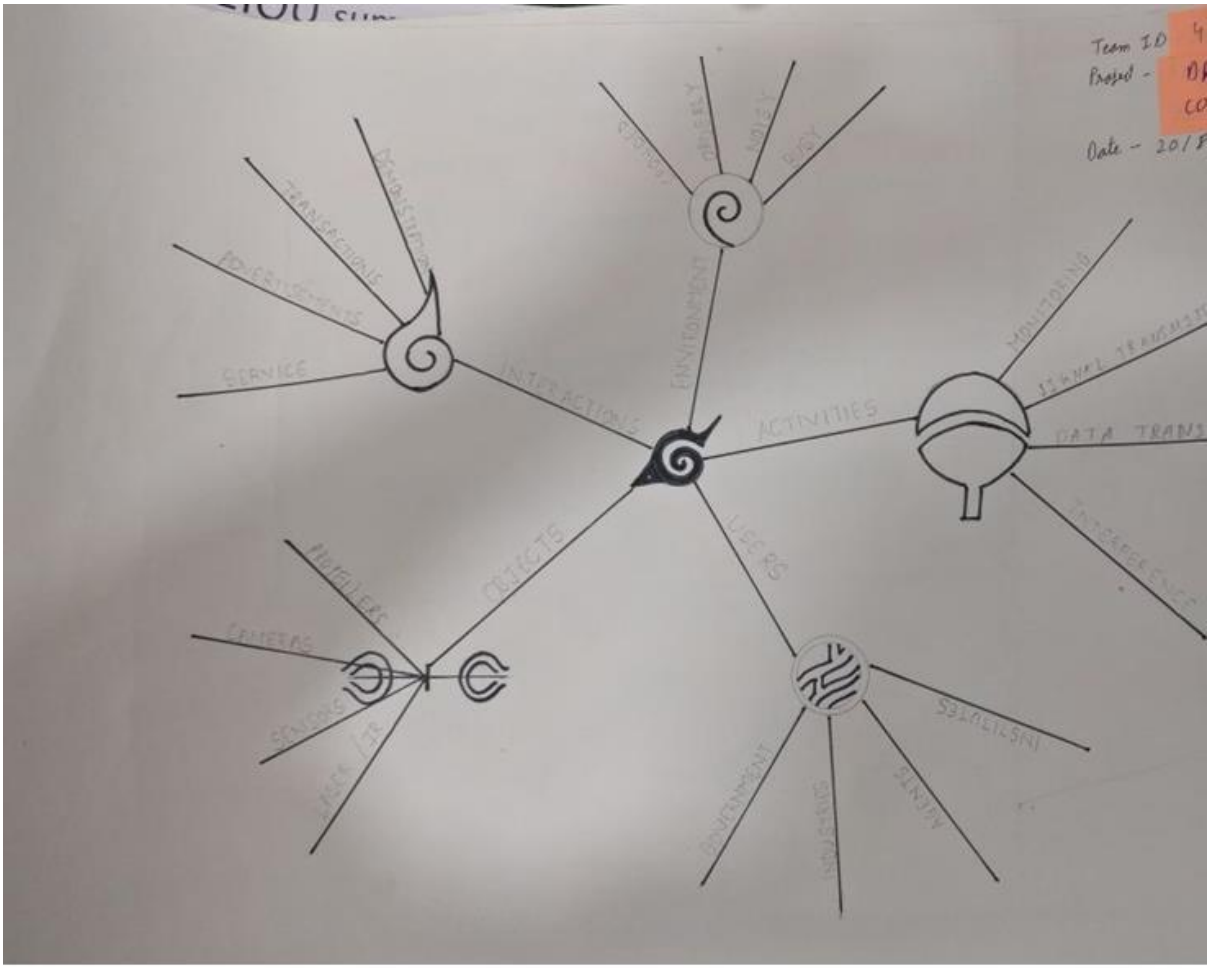


Figure 25 Mind Mapping

8. Summary of Research Paper

1) There are several OCR mobile applications in the market running on both android and iOS platforms. This is possible due to great cameras and processors compatible with usage of ocrs. This paper proposes a framework of Optical Character Recognition (OCR) on mobile device using server-based processing. Comparison methods proposed by this paper by conducting a series of tests using standalone and server-based OCR on mobile devices, and compare the results of the accuracy and time required for the entire OCR processing. Server-based mobile OCR obtains 5% higher character recognition accuracy than the standalone OCR and its format recognition accuracy is 99.8%. The framework tries to overcome the limitation of mobile device capability process, so the devices can do the computationally intensive application more quickly.

2) The inclination of optical technologies like OCR lies in achieving higher recognition rates with optimal or reduced computational complexities. At present there exist optical technologies for automation of reading the text from document images with almost nearing to 100% accuracy. Especially, the Roman language OCR's are reliable and robust enough in producing higher accuracies by being able to recognize varying font styles of varying sizes. However for the font style/ size independent OCR's one of the main aspect is its computational complexity. It is significant concern to reduce the computational complexities involved in the process of character recognition through a font style / size independent OCR. In this paper, a technique for classification of the font style based on character image is proposed by employing the distance profile features with respect to left, right and diagonal directions of a character image. The major objective of this work is to reduce the complexity of the ge ne ri c O CR syste ms by font styl e re cogni ti on. The di s t a n c e profile features of character images are fed to a support vector machine classifier. For experimentation, the training data sets are comprised of around 10 widely used font styles of upper case letters as well as lower case letters. The testing is conducted wi th the character images that are extracted from various non editable document sources comprising of 5 different font styles. The performance of algorithm is found to be satisfactory with an accuracy of 80%.

3) It is very difficult to retrieve image from large no of database which contain some message on it. The OCR techniques are becoming very efficient techniques for external and fast retrievals. The OCR technique search image based on data or text written on image. It search image or message contain in the image OCR the mechanical or electronic conversion of images of typed, handwritten or printed text, whether from scanned document, a scene-photo. For that Tesseract will be used. After recognizing text from image by OCR, it will store that message in any file. After that this has used Boyer–Moore string search algorithm, for searching string stored in file.

4) one critical procedure in OCR is to detect text characters from a document image. However, some documents might come with embedded background images which often mislead the algorithms of character detection. For example, small dots or sharp edges from the background image are often bound- boxed as characters and passed to the next stage of the OCR pipeline, which causes an error chain. Motivated by this

observation, we present a novel and cost-effective image preprocessing method to accomplish the task. We first enhance the document images before OCR by utilizing the brightness and chromaticity as contrast parameters. Then we convert color images to gray and threshold it. This way, background images can be removed effectively without losing the quality of text characters. The method was tested using Tesseract (an open source OCR engine) and compared with two commercial OCR software ABBYY Finereader and HANWANG (OCR software for Chinese characters). The experimental results show that the recognition accuracies are improved significantly after removing background images.

5) The onset of wireless communication technology along with other hi-fi communication techniques has facilitated e- commerce with many aspects, such as time complexity and its easy reach to mankind. The modern shopping style has helped e- commerce in gaining its popularity with its comfort and convenience. The key idea here is to assist a person while shopping, by paying the exact bill for the very product he wants. Barcode reader and editor is such an idea with which the barcodes of the product can be read and can be corrected as well in necessary condition. This work deals with three major steps in materializing barcode reading and editing process. First of all the image of the barcode is enhanced by image enhancement techniques. Secondly, the barcode is read and finally the barcode is saved in a database for editing purpose. Optical Character Recognition (OCR) technique is used in the decoding process. At the output the barcode image is displayed.

9. LEARNING NEED MATRIX

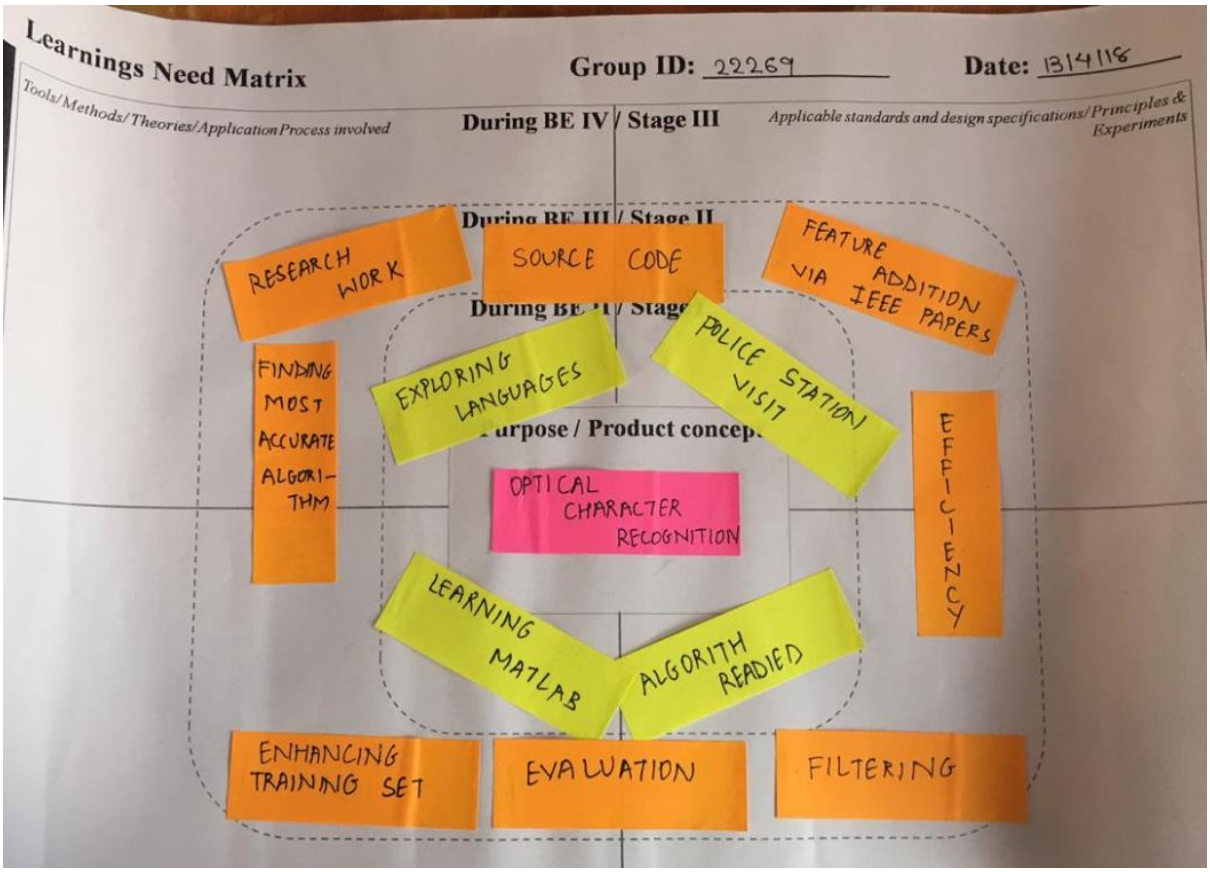


Fig27. LEARNING NEED MATRIX

10. SCRAMPER TOOLS

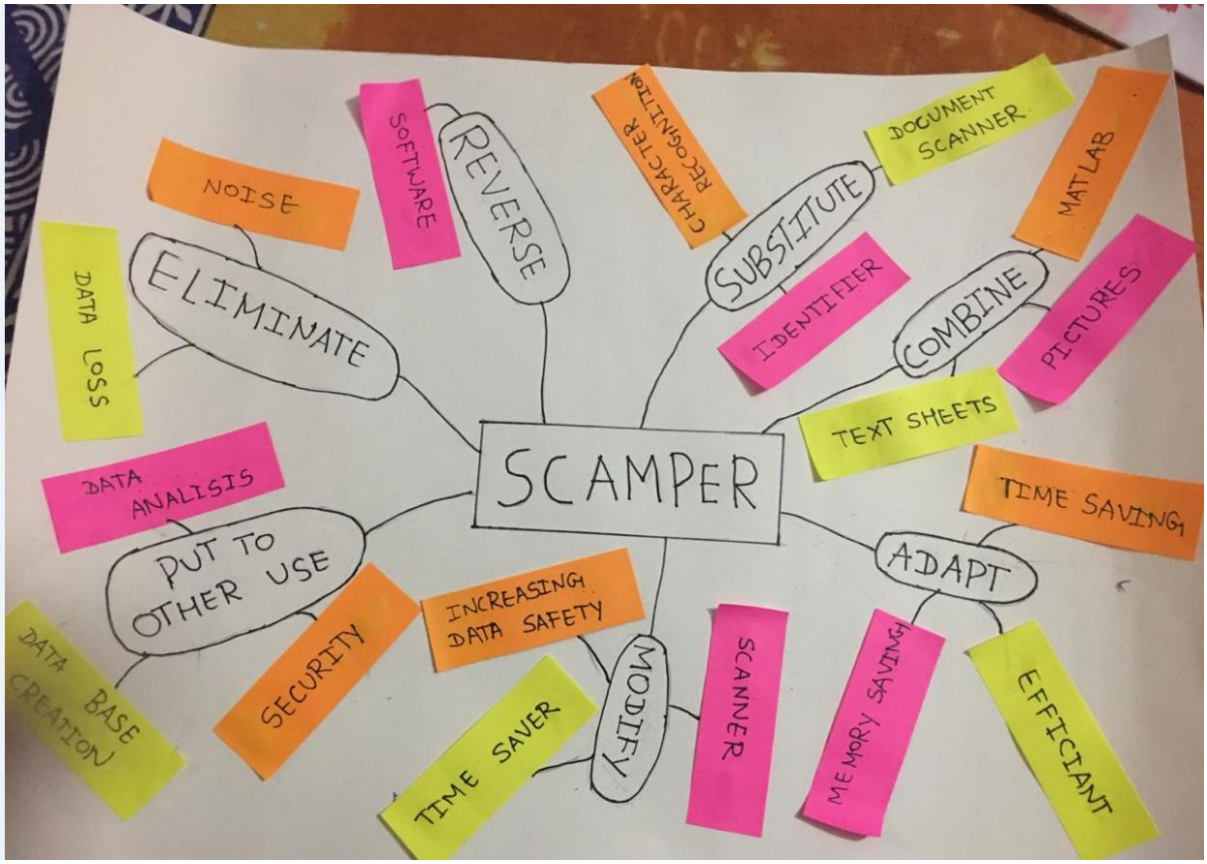


Fig. Scramper Tool

11. Prototype Code:

```
%The script prprob defines a matrix X with 26 columns, one for each letter of the alphabet.
%Each column has 35 values which can either be 1 or 0.
%Each column of 35 values defines a 5x7 bitmap of a letter.
%The matrix T is a 26x26 identity matrix which maps the 26 input vectors to the 26 classes.

[X,T] = prprob;
plotchar(X(:,1)) %Here A, the first letter, is plotted as a bit map.

%To solve this problem we will use a feedforward neural network set up for
%pattern recognition with 25 hidden neurons.
%Since the neural network is initialized with random initial weights,
%the results after training vary slightly every time the example is run.
%To avoid this randomness, the random seed is set to reproduce the same results every time.
%This is not necessary for your own applications.

setdemorandstream(pi);
net1 = feedforwardnet(25);
view(net1)

net1.divideFcn = '';
net1 = train(net1,X,T,nnMATLAB); %The function train divides up the data into training,
validation and test sets.
%Training stops when the network is no longer likely to improve on the training or validation
sets.

%Training the Second Neural Network
%network to also reconise noisy versions of the letters.
numNoise = 30; %30 noisy copies of each letter Xn are created. Values are limited between 0
and 1. The corresponding targets Tn are also defined.
Xn = min(max(repmat(X,1,numNoise)+randn(35,26*numNoise)*0.2,0),1);
Tn = repmat(T,1,numNoise);

figure
plotchar(Xn(:,1))

%Create second network and train
net2 = feedforwardnet(25);
net2 = train(net2,Xn,Tn,nnMATLAB);

%Testing Both Neural Networks
noiseLevels = 0:.05:1;
numLevels = length(noiseLevels);
percError1 = zeros(1,numLevels);
percError2 = zeros(1,numLevels);
for i = 1:numLevels
  Xtest = min(max(repmat(X,1,numNoise)+randn(35,26*numNoise)*noiseLevels(i),0),1);
  Y1 = net1(Xtest);
```

```
  percError1(i) = sum(sum(abs(Tn-compet(Y1))))/(26*numNoise*2);
  Y2 = net2(Xtest);
  percError2(i) = sum(sum(abs(Tn-compet(Y2))))/(26*numNoise*2);
end

figure
plot(noiseLevels,percError1*100,'--',noiseLevels,percError2*100);
```

Test Output:

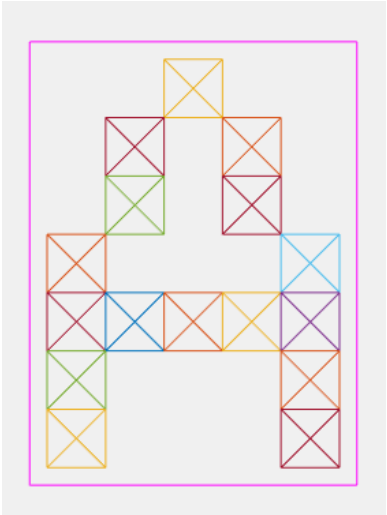


Figure 26 Input without noise

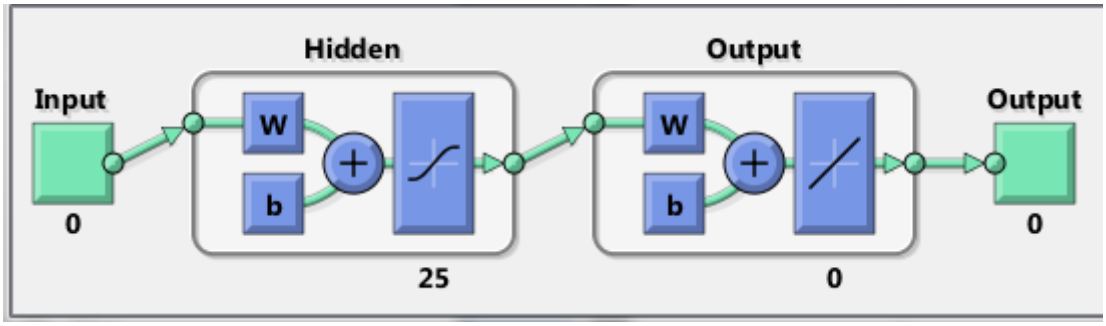


Figure 27 Neural Chart for 0 noise input

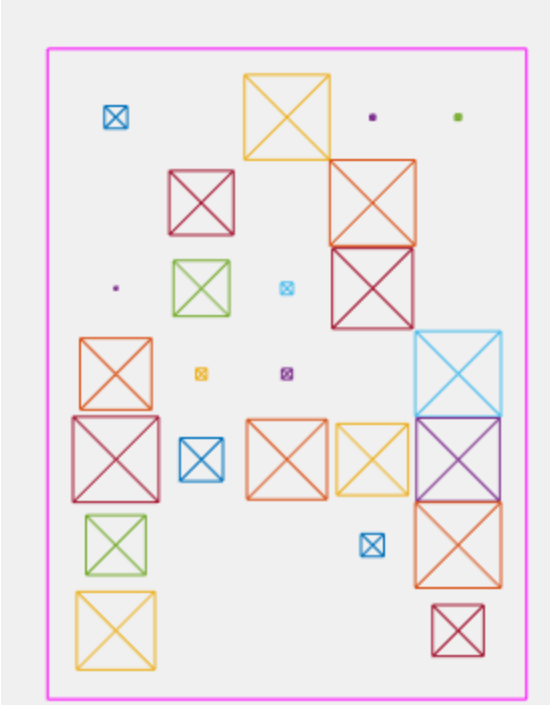


Figure 28 Input with noise

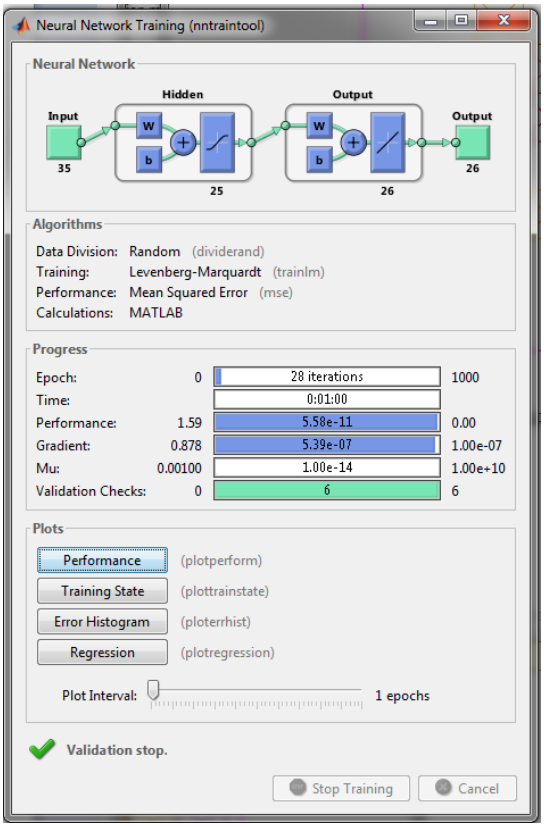


Figure 29 Neural chart for noise input

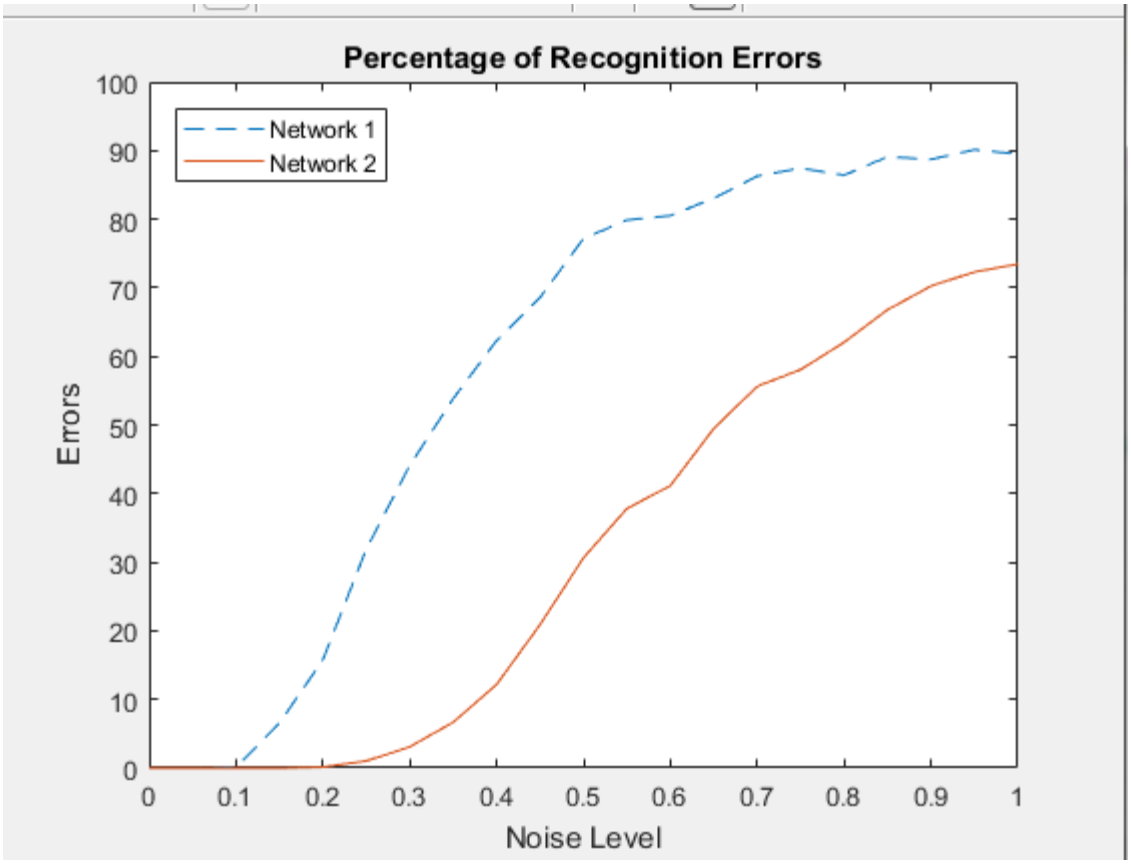


Figure 30 Gradient of recognition errors on application of noise

12. CONCLUSION

Through constructing the AEIOU frame work we have learned to construct design thinking and enhance our ideas through repeated and appropriate observations. These help us in selecting our topic from a vast space that we have selected.

It provides us capability to think on our future improvements on the project, and it also helps in putting academic and other regular college work into the practical application in context to our area of interest

Through constructing the three canvases we have constructed our view and idea of project in an efficient way.

By making the three canvases our communication, thinking, and technical skills have been developed, and we have represented our idea efficiently.

We also created a learning need matrix for getting a clear objective set for each period and a fixed goal to be achieved.

These activities have made us more promising thus we have enhanced our skills and shorted fees of the mistakes. All together it has been great experience.

Further we implemented various different algorithms for our Optical Character Recognition and then we fixed upon an efficient and reliable which is similar in huge proportions to artificial neural network and various other algorithms we implemented and made our algorithm efficient and accurate.

We have kept our program as of now limited to a limited number of fonts, that is limited to our comparison database and the number of fonts on which our program can work can be increased in future if needed.

13. REFERENCES

https://www.sciencedirect.com
https://ieeexplore.ieee.org
https://www.ericsson.com
https://bettstetter.com