Design and Analysis of Visual Insight of Child using IBM Multimedia Analysis and Retrieval System

Sangeeta Singh
Department of
computer engineering
MIET Institute of engineering
Meerut ,UP

Mukesh Rawat, PhD
Department of
computer engineering
MIET Institute of engineering
Meerut ,UP

ABSTRACT
Humans recognize objects with astonishing ease and speed; here we are using behavioral methods to investigate the sequence of processes involved in visual object recognition in natural scenes. The greatest challenge of our times is to understand how our brains function. Images play a big role in representing the inner feeling of the child or a person. Images are created in his/her memory of the daily things happening in his/her life. Then the images selected by the child reflect the visual insight of the child. The system developed is helpful for the same purpose. It shows the result in terms of accuracy compared with the other face recognition system.

Keywords
Visual insight, images, faces recognition.

1. INTRODUCTION
In our busy life, most of us pay less attention towards our children. Most of the parents do not want to know what the interest of their child is and what he or she want to do? They weighted the child with their own aim and continuously questioning him/her about their studies more than required. They want that their child spend 8-10 hours daily for his studies. When this happens regularly the child get depressed and in last try to end his/her life. So this is the essential duty of the parents and the school , where the child spends 8 hours daily to look insight the child and give him proper counseling if there is some problem with the child. [1] Here an idea is proposed in which we show child various images of various activities, things, scenery etc., such as quarrel between the parents, sea, rising sun, playing with parents and friends, abusing by parents and teachers etc. and told child to select images what he or she do and sees in his/her daily life .The images selected by the child reflect the visual insight of the child and by analyzing the images, give proper counseling to the child.

2. ISSUES AND CHALLENGES OF THE SYSTEM:
2.1 Recognition of Image
The systems developed using Matlab, ImageJ, NI Vision Builder, Halcon, OpenGL; Delphi etc are not too accurate.

SOLUTION: Where as systems developed using IMARS has visual feature-based machine-learning framework for large-scale semantic modeling and classification of image and video content. It can also work at large-scale for high-volume batch processing of millions and images and videos per day.

2.2 Time Required To Identify
Most of the systems developed using Matlab , ImageJ etc are more time consuming and takes large amount of time in recognizing the image .

SOLUTION: The system developed using IBM take less time to accurately identify the image.

2.3 Handling The Number Of Client Requests
Most of the systems developed using   MATLAB, ImageJ, OpenGL, Delphi etc work mainly in a uniprocessor system, handling one client at a time.

SOLUTIONS: The system developed using IBM Visual recognition API, can handle number of client requests. The reason for this is that, here the system is deployed in a cloud based environment which can be easily accessible to the clients and they can further process their requests easily using API.

3. IMAGE PROCESSING USING IBM BLUEMIX CLOUD
[2, 3]The IBM Watson Visual Recognition service uses deep learning algorithms to analyze images (.jpg, or .png) for scenes, objects, faces, and other content, and return keywords that provide information about that content. It can also create custom collections of your own images, and then upload an image to search the collection for similar images. A custom classifier is created which is a group of classes that are trained against each other. This allows you to create a multi-faceted classifier that can identify highly specialized subjects, while also providing a score for each individual class.

It allows users to understand the contents of an image or video frame, answering the question: "What is in this image?" Submit an image, and the service returns scores for relevant classifiers representing things such as objects, events and settings.

4. IMARS
[4]IBM Multimedia Analysis and Retrieval System (IMARS) is a powerful system that can be used to automatically index, classify, and search large collections of digital images and videos. IMARS works by applying computer-based algorithms that analyze visual features of the images and videos, and subsequently allows them to be automatically organized and searched based on their visual content. IMARS is comprised of the IMARS extraction tool and the IMARS search tool. The IMARS extraction tool takes a collection of images and videos from the user, and produces indexes based on mathematical analyses of each piece of content. These indexes organize the results of the analyses for the IMARS search tool.

IMARS is supported by a broad array of pre-trained semantic classifiers that automatically identify whether each new image and video belongs to one or more of the pre-defined semantic categories in the taxonomy based on its extracted visual descriptors. IMARS provides additional capabilities based on unsupervised classification that cluster the images and videos purely based on their extracted visual descriptors, without assigning them any label, and allow searching based on visual

similarity. The IMARS search tool provides a graphical interface which allows the user to search, browse and navigate the collection based on the values produced by the analyses performed by the IMARS extraction tool.

5. ARCHITECTURE OF VISUAL RECOGNITION SYSTEM

As shown below (see Figure 1) in the architecture the web based system is hosted in cloud through which various users can work simultaneously with the system .The user firstly browses to the website and upload the images. The images selected by the user are sent to IBM Visual recognition API for processing. And once the image is processed by the API it returns the score back in the json (java script object notation) form. The score is then sent back to the user. Now the user is well aware of the psychological behavior of the child. The JSON format is syntactically identical to the code for creating JavaScript objects. Because of this similarity we use JSON, instead of using a parser (like XML does).

[5]The Visual Recognition services come with a set of built-in classes we can analyze images (.jpg, or .png) for scenes, objects, faces, and other content, and return keywords that provide information about that content. We can also train custom classifiers to create specialized classes, and create custom collections to search for similar images.

To perform better, sets of images are used to train new custom classifiers. (The service is based on the IBM Multimedia Analysis and Retrieval System, or IMARS).We can authenticate to the Visual Recognition API by providing the API key that is provided in the service credentials for the service instance that we want to use.

This visual feature extraction, works by using the computer to analyze the pixel-level contents of each image and video, and create a multidimensional vector description of its visual features. Since there are many important dimensions of visual contents, such as color, texture, shape and spatial layout, MARS utilizes a large set of visual feature extraction algorithms that extract descriptors across a wide array of visual dimensions. This result goes to visual semantic extraction, which works by applying machine learning techniques to the extracted visual descriptors.

The deep learning algorithms are used to analyze images that can give you insights into your visual content. It is possible to organize image libraries, understand an individual image, and create custom classifiers for specific results that are tailored to the needs. This system provides a graphical interface which allows the user to search, browse and navigate the collection based on the values produced for analysis. Hence this cloud based system provides an interface which can be used by multiple users at the simultaneously.

The result we get form Visual Recognition API is in JSON format. JSON (JavaScript Object Notation) is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. JSON is a text format that is completely language independent but uses conventions that are familiar to programmers of the C-family of languages, including C, C++, C#, Java, JavaScript, Perl, Python, and many others. These properties make JSON an ideal data-interchange language. It is built on two structures:

- A collection of name/value pairs. In various languages, this is realized as an *object*, record, struct, dictionary, hash table, keyed list, or associative array.

- An ordered list of values. In most languages, this is realized as an *array*, vector, list, or sequence.

These are universal data structures. Virtually all modern programming languages support them in one form or another. It makes sense that a data format that is interchangeable with programming languages also be based on these structures.
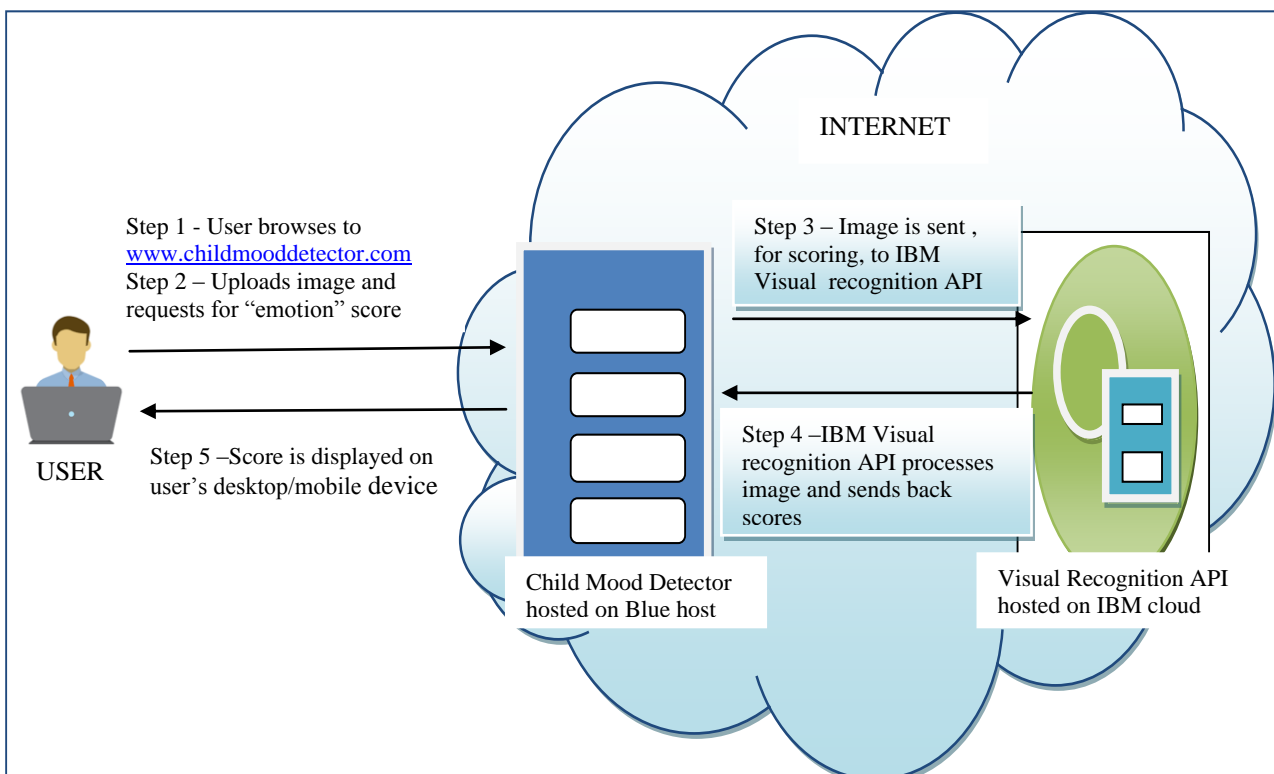

Figure 1: Architecture of Visual Recognition System

6. ALGORITHM FOR SYSTEM LEARNING

The code shown below shows how the learning for images are done using IBM visual recognition API:

```
var    url   ←'https  ://gateway-a.watsonplatform.net/visual-recognition/api/v3/classifiers?api_key

var  curl ← curl_init();

var data←array("happymood_positive_examples","sadmood_positive_examples", "negative_examples" )

 curl_setopt (var curl, CURLOPT_URL, var url)

 curl_setopt (var curl, CURLOPT_POSTFIELDS, var data)

 curl_setopt ( var curl, CURLOPT_FOLLOWLOCATION, true)

 curl_setopt ( var curl, CURLOPT_VERBOSE, TRUE)

 curl_setopt ( var curl, CURLOPT_RETURNTRANSFER, TRUE)

var vr_exec ← curl_exec( var curl)

var er ←curl_error( var curl)

print_r( var er)

 curl_close(var curl)

var re ← json_decode(var vr_exec,true)

 print_r(var re)

 unlink()
```

7. RESULT ANALYSIS

The three inputs used by the systems learning process consists of two positive repository and one negative repository. When the system completes its learning process it returns the classifier ID.

Positive training set = {014.jpg, 321.jpg, a1.jpg, p11.jpg…}

Positive training set= {s23.jpg, 1144.jpeg, 54.jpeg……..}

Negative training set= {n11.jpeg, 43.jpeg, 116.jpeg……}

Classifier = 1

Statistical representation of the system compared to other image processing systems is shown below (see Table 1):

Table 1: Comparison with other systems

Number of images	Scores generated by IMARS system	Correctly identified by IMARS	Correctly identified by other image processing system using MATLAB
4	88-92	3	2
8	87-94	7	5
12	90-95	10	9
16	88-95	14	12
20	87-94	18	16
24	91-98	23	21

As shown in Table1, suppose there are 4 images then using IMARS system ,three images are accurately recognized with

88-92% score .It is more than enough for accurately recognizing an image.

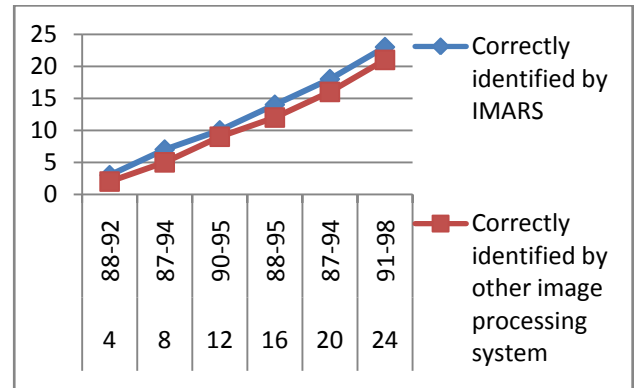Graphical representation of the above comparative data is shown (see Figure 2).

Figure 2: Comparison with other systems

The tabular as well as graphical representation of the comparisons shown above gives the idea of the accuracy of the IMARS system as compared to other systems. It shows the comparison chart on the basis of score generated when number of images request uploaded by the clients by the IMARS system as well as other systems.

8. CONCLUSION

This proposed system with Visual Recognition service using IMARS provides a unique feature to automatically annotate images based on visual content alone. The processing of pixels of an image is done using visual recognition API. It extracts visual insights related to activities, places, interests and people to deliver a comprehensive view of what users are communicating through visual media. It will be useful to our future generation to be perfectly counselled when needed. This paper also shows the detailed comparison of the system with other system of visual recognition .The comparisons are on the basis of time, accuracy; number of clients handled by the system is shown. Its future scope is that, the way we are recognizing static images, videos can also be analyzed using similar concepts. Hence the concept may also be useful for video recognition.

Thus the proposed work increases the efficiency to automate the process for discovering the behaviour change among pupils.

9. REFERENCES

[1] Bundesen and Habekost, 2008, Principles of Visual Attention: Linking Mind and Brain,Oxford University Press, UK, 378 p.

[2] Rafael C Gonzalez and Richard E Woods , 2008 , Digital Image Processing , Pearson , UK ,976p.

[3] Jain A.K, Fundamentals of Digital Image Processing, 1994, PHI, Delhi, 304p.

[4] "https://www.ibm.com/watson/developercloud/doc/visual-insights/ ", OCT 17, 2016

[5] "http:/ /www.programmableweb.com/api/ibm-watson-visual-insights ", OCT 17, 2016

[6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2885157/ ", OCT 17, 2016

[7] "http://www.globaleducationmagazine.com/education-suicide/ ", OCT 17, 2016

[8] "https://www.cdc.gov/violenceprevention/pdf/suicide-datasheet-a.pdf ", OCT 17, 2016