

Image Based Tool for Level 1 and Level 2 Autistic People

Anesh Kaul¹, Harsh Mittal², Krish Anawadia³, Mukesh Patel⁴,

Prof. Pankaj Sonawane⁵

^{1 2 3 4}U.G. Student, Department of Computer Engineering, D.J. Sanghvi College of Engineering, Maharashtra, India

⁵Professor, Department of Computer Engineering, D.J. Sanghvi College of Engineering, Maharashtra, India

Abstract - This paper proposes an assistive software tool especially for the special needs of people with autistic and mental disabilities. Using correct technology advances good prospects for especially abled individuals to partake in academic activities effectively. It is a rightful plus the meticulous duty of every citizen to empower especially abled ones with infirmities to gain access to and meet the educational and physical training as effortlessly and realistically as possible. The software tool proposed intends to give a channel to fill the gap between expressions and the real-world concepts via relevant images. The NLP unit serves as the mainstay of the system. It analyses natural language expressions into procedural semantics, networks through integration-based grammar to provide semantic flexibility for the people. The Linguistic input list of words and statements are diagrammed to significant pictures from the data via the arbitration of cosine similarity. The system uses verbal processing for these image embeddings to represent the gist of images and functions realistically well.

Key Words: Natural Language Processing (NLP), Linguistic Theory, Verbal Processing, Cosine Similarity, Image Processing

1. INTRODUCTION

Information retrieval is the space of study involved with checking the documents, for useful data among documents, and for data concerning documents and finding material of an unstructured nature that satisfies an information wanted from a big storage. Information retrieval is additionally used to assist semi structured search like searching a document wherever the title contains 'Java' and therefore the body contains threading and agglomeration that is the task of springing up with an honest clustering of the documents supporting their content. An IR method starts once a user inputs a question into the system. Queries are unit formal statements of knowledge requirements, for instance search strings in internet search engines. In IR, a question doesn't unambiguously determine an entity within the assortment. Instead, many objects might match the question, maybe with totally separate levels of relevance. Current technologies are facilitating the approach with

which we tend to understand and express. Through creative means, unfit individuals will browse articles, converse with machines.

New techniques are currently thought to grasp the precise requirements of these individuals. The motto here is to appeal for straightforward methods for such individuals to interpret useful data, a lot of significantly relevant data extraction from the net.

'Learning styles' may prove to be an idea that tries to explain the ways by which individuals gain details concerning their surroundings which facilitate somebody to understand and remember an ability or truth. Several people think that special ways are the sole to be thanked to escort these individuals however others think that the sole thanks to dispatch skill coaching is through using phrases.

2. REVIEW OF LIRTERATURE

[1] Kilicaslan, Y., Ucar, O., Guner, E.S., Bal, designed and developed an application interface for people with autistic and intellectual disabilities the use of NLP technology. The gadget enter is a Turkish text that is constantly processed and the end result is produced by way of a chain of text-related pics in order that the disabled character can without problems recognize. 4 distinct software program additives were used at the same time as encoding the components that make up the machine. The GUI and QG are fully encrypted the usage of JAVA. The IDB maintains a database of activities and JDBC is used to get entry to this database. the primary component is the SFG that's a logical machine and grammatical division. The SFG consists of most important parts of grammar and parser. Parser performs two key capabilities - ensuring that the words that make up the punctuation are linguistic and combines phrases and phrases into larger sentence devices that adhere strictly to grammatical rules. The output of the evaluation process is a language structure that includes each syntactic and semantic facts. In each instance there may be a complete digestive tract. The semantic framework generated by way of SFG is then given to QG, which interprets this framework into a site question and through the query,

reveals the proper photo on the web site this is without a doubt seen to people with autistic and highbrow disabilities relying on the enter

supplied, an appropriate set of images is obtained.

[2] In this paper, they suggested an ontology-based approach to entity-based IR. Through assigning a color to the picture segmentation algorithm, the process of putting a picture onto a website is performed and breaks down the number of regions. After this, a low-level setting set is calculated for each built-in area. These factors usually define the color, location, size and shape of a region. These features are translated into intermediate level adjectives. These make the ontology. Question output is a set of images that may be relevant, the relationship of which can be expressed in volume at this point. They are displayed to the user at random. The user then scans one or more pages of the picture, checking the appropriate "appropriate" boxes. After submitting this feedback, one or more vector machines are trained and later measured corresponding to the compatibility of all the regions returned to the original question. Photos are then sent to the user at a fixed rate. The proposed algorithms were tested on a collection of 5000 images from Corel's library. This creates a website that contains around 30000 regions.

[3] In this paper, they propose a model for processing mathematical language for picture acquisition, which combines semantic information provided by WordNet with an online dictionary reference system, and low-level pictorial attributes. In this case, the semantic sequence of phrase sensors from WordNet are utilized to enhance the relationship between image and descriptive meaning of the context. The mathematical algorithm for choosing a keyword is used to describe those conceptual images. The model was tested on a geographical picture website with 10 separate words. The conclusions prove that it will significantly improve recovery performance. This also demonstrates the high power of the method of constructing information ontologies. In the next section, reference is made to methods for learning the semantics of phrases and images as well as the natural mathematical modeling model presented.

3. PROPOSED ARCHITECTURE

The work on the model is that the work wiped out the sphere of the language communication process and image process that has long unreal humanity and has been gaining within the last decade for technological developments. The model for unfit folks is gaining additional and additional importance during this technology- driven world. Numerous user friendly and economical tools were developed for the help of especially abled folks like GUI, that was created by Stephen Arnold Douglas Engelbart, a talking serial printer for visually

impaired folks, a talking show terminal, a screen reader for the visually impaired, a speech recognizer (Via Voice), a talking application program (Home Page Reader) that were developed by IBM and plenty of helpful tools was created by Microsoft. User is curious about the centered results relating to a product with some special attributes.

The model also allows one to enter the statements vocally via voice input function which in turn uses Speech Recognition to identify the exact words spoken by the user.

3.1 Architectural Diagram

The model comprises of four important working modules: Graphical User Interface (GUI), NLP unit, Query Function, and Image dataset. All of these have been developed using different software environments which helped to get the desired system components.

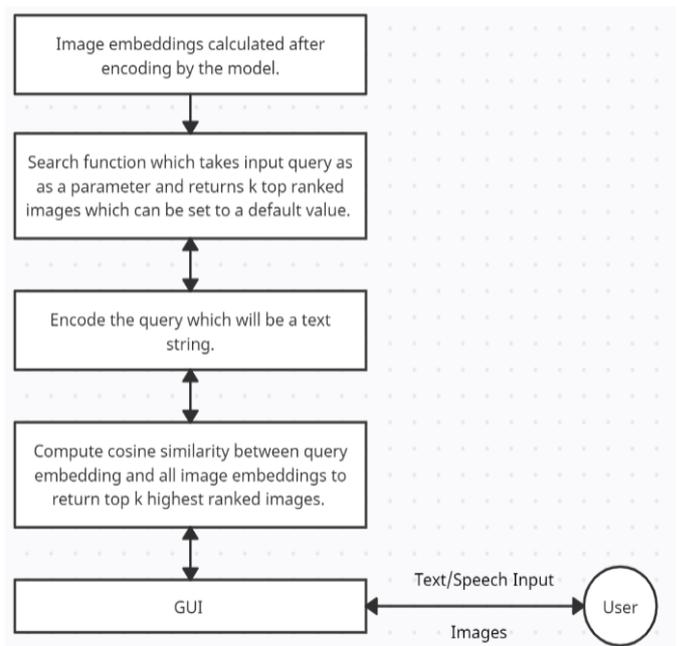


Fig -1: Architecture

3.2 Software Design

Almost every user is getting bored with the existing system, they want something new which can fulfill their needs and can also work as basic applications. It looks great if the assistive tool which we use can display more than just a black and white image i.e., learning tool. The main point for building the system is to make a tool that can be used in home conditions as well as in business space. The additional features we have added are English input, cosine similarity, query, colored images, three input matching images, and more for future reference. The

proposed system supports additional flexibility and comfort.

3.3 Detailed Design

In this proposed study, the main components which are used are the input description in English, NLP unit, Query function, GUI for future reference. The dataset stores color images of various queries according to user input.

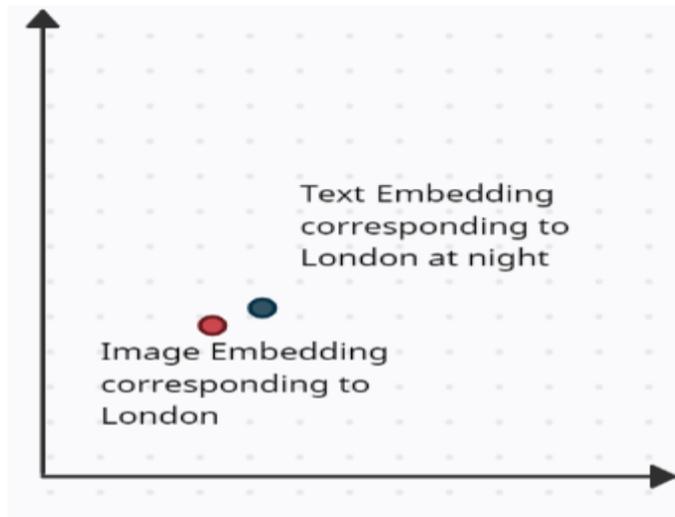


Chart -1: An example of a cosine similarity

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Fig -2: Formula to calculate the cosine similarity

The NLP Unit is the main portion of the system. It is the essential base for the project. NLP unit consists of two main parts which is Encoder and Searcher. The Encoder calculates embeddings for all images in the dataset. It also computes embeddings for input queries. Searcher takes the input query as a parameter. First it computes cosine-similarity among the query embedding and all picture embeddings. It then returns the top k highest ranked pictures, that are displayed.

Relevant pictures are been produced to users, so that given user can easily process it through his/her eyes. Users acts as an external agent which provides input in the system through text or with his/her voice. The input provided to the system by the user must be in English language only. There is also a GUI which acts as an interface among the given user and the application which can be used for future reference. GUI will be kept in front

of person's screen for the dual purpose. User can see the data clearly and accurately and can input his/her query.

4. IMPLEMENTATION

In the system implementation, we display relevant images on the monitor which gets displayed through the NLP unit.

Both text input and speech input are being supported. After the user gives a text or speech input, the model computes embeddings, encodes the queries, computes cosine-similarities, etc. and finally returns top k highest ranked images as output. Working with the current dataset yields results with 95.13% accuracy.

5. CONCLUSIONS

Many youngsters and men with syndrome suffer from speech and language issues. It's widely recognized that through the help of ways of pictorial learning, language obstacles are going to be down, and learning and understanding with all senses are going to be encouraged. Unfit youngsters will speak while not catching the meanings of the phrases they use. They'll gain expertise in establishing the relation between a phrase and its meaning. A kid hearing the phrase 'Red' might not be able to use that phrase to associate to the real red color which him/her views. In only these cases, machine package linking expressions or footage or icons depicting the meanings of those expressions will return to scene to assist, fasten and assist the education of those youngsters. Of course, all these facts conjointly apply to people. World Health Organization has similar issues. Hopefully, our present work and attainable enhancements side to that within the future can contribute to the education of these youngsters.

Till now, we've found an honest variety of datasets for our purpose. however, within the coming near future, we are going to add different language like Italian with larger dataset, conjointly we are going to create back-end logic advanced and take a look at how to integrate a user interface.

6. FUTURE SCOPE

A proper GUI to facilitate the input process of entering sentences/queries using voice or text is one of the major aspects in the future scope of this project. The GUI will also make the output screens look better. Apart from GUI, dataset(s) containing more animated/cartoon images might prove to be better than the existing dataset being used. One of the main reasons for this is that the younger percentage of users can benefit more by means of these kinds of images rather than real life/real world images respectively.

REFERENCES

- [1] Kilicaslan, Y., Ucar, O., Guner, E.S., Bal, K.: An NLP-Based Assistive Tool For Autistic And Mentally Retarded Children: An Initial Attempt, Trakya University, Faculty of Engineering and Architecture, 1-8 (2006)

<https://www.researchgate.net/publication/242250903>
- [2] Vasileios Mezaris, Ioannis Kompatsiaris, Michael Strintzis : An Ontology Approach To Object-Based Image Retrieval (IEEE 2003)

<https://ieeexplore.ieee.org/document/1246729>
- [3] Shivakumar Shashank, Nikhil Goel, P Rajaram, Dr. Ananthanarayana V.S., Cilson Cyriac: Semantic Image Retrieval System Based on Object Relationships (IEEE 2013)

<https://ieeexplore.ieee.org/abstract/document/6707598>
- [4] CRICK software package LTD., Clicker5,
<http://www.cricksoft.com/uk/products/clicker/>
- [5] FILLMORE C. J. (1968) The Case for Case. In E. Bach and R. Harms (eds.), *Universals in Linguistic Theory*, 1-88, New York: Holt Rinehart and Winston.
- [6] FILLMORE C. J. (1982) Frame linguistics. In *Linguistics within the Morning Calm*, 111-137, Seoul: Hanshin.