Review on Intelligent Spectacles to Prevent Road Accident

|  |  |  |
| --- | --- | --- |
| Madhuri Pal1  [madhuridec09@gmail.com](mailto:madhuridec09@gmail.com)  *Professor, Department of Computer Science and Engineering*  *J D College of Engineering and Management, Nagpur, India* | Sagar Kadwe2  [skadwe9@gmail.com](mailto:skadwe9@gmail.com)  *Student, Department of Computer Science and Engineering*  *J D College of Engineering and Management, Nagpur, India* | Nakul Gopal3  [nakulgopal13@gmail.com](mailto:nakulgopal13@gmail.com)  *Student, Department of Computer Science and Engineering*  *J D College of Engineering and Management, Nagpur, India* |
| Tejas Jaiswal4  [tejasjaiswal972@gmail.com](mailto:tejasjaiswal972@gmail.com)  *Student, Department of Computer Science and Engineering*  *J D College of Engineering and Management, Nagpur, India* | Mithun Chide5  mithunchide03[@gmail.com](mailto:harshalbhoyar830@gmail.com)  *Student, Department of Computer Science and Engineering*  *J D College of Engineering and Management, Nagpur, India* | Yash Paunikar6  [yashpaunikar474@gmail.com](mailto:yashpaunikar474@gmail.com)  *Student, Department of Computer Science and Engineering*  *J D College of Engineering and Management, Nagpur, India* |

***Abstract*— To overcome the problem of writing code efficiently the Artificial Read and Learn System is designed. The Artificial Read and Learn System uses deep learning approach known as Natural Language Processing. Which is used to learn and generate the code snippets from natural language description. The model will enable a developer to write efficiently just by providing the description in natural language (English). This paper describes the different techniques adopted in this era of AI. Additionally, it includes the objective, the scope of advancement and the approach for implementation for reducing   developer’s efforts. Artificial Read and Learn System is a machine learning model that provide source code for given problem statements by using self-attention based transformer architecture. Although we didn’t aim to achieve 100% accuracy because the natural language is very ambiguous and has very rich set of vocabulary. One such approach can give rise new era of development. We are proposing a potential research direction of creating a model by combining various inputs.**

Keywords— Augmented reality, Field of Vision, Intelligent Spectacles

# Introduction

Life, on earth we know that how life follows a defined cycle. Every organism evolve by learning from its experiences in real world. We humans learn different things from different sources, one of the source is books .As a human child learns new things from books. How interesting it is to think that, what if machines or just say computers can do the same thing? Sounds like fiction.

But I and my team have decided to bring this fiction into existence. Accordingly we have planned our AI program will learn new things from reading books.

This program will feature two main concept of AI. The major one is the Natural Language Processing and the other one is Machine learning. The program will read the text from the books provided books are to be in soft copy in any format. The text will be given as input to the NLP algorithm.The NLP algorithm consist of two sections NLU (Natural Language Understanding )and NLG(Natural Language Generation). The NLU algorithm will understand the meaning of the text .

# Awareness

In this Digital Era, many of the digital devices have been developed. At the time throughout the day, we were using electronic appliances as it became part of our daily life but still, we were unaware of many such devices. Some of the common devices we are familiar with are Digital Watches, Video Games, Smartphones, etc. that we use regularly in our daily life. As the world is moving in a digital era, all the technical appliances are being developed which saves time, reduces efforts and makes work easy. Using Video Games children can entertain themselves whereas smartphones which are the electronic handheld devices made life easier. People purchase many products but are unaware of features included in that because only some of the features are used by them. The technical world is progressing fast and normal people are unaware of all these technologies and rely on traditional methods. The majority of problems encountered are because of interacting usage. Many devices which are useful in certain cases are difficult to interact with. Any online portal forms a new and environment for common and new user. It is true in almost every case. The product needs not only to be efficient but also should provide an easy and simplistic interactive environment for the end-user. The development pace is increasing rapidly, and a lot of the products are being developed but because of unawareness many of them are failing to sustain in the market and being unsuccessful.

One of the products is Google Glass. It is very popular because of its advanced features and adopted by most of the countries. Their features are so advanced & cost price is so high that the Glasses can only be used by economy class. Whereas in a country like India, such an economy class is hard to find and it is the main reason, such technology is not spreading far. Moreover, people can use Smartphones rather than using smart glasses which are more useful than glasses. Some of the major areas which use the technology of google glass are industry, medical field, the business sector where time is highly valued. Till 2018, 21 million units of google glasses were expected to sold as forecasted by Google. The statistics have not been disclosed yet by Google though. Looking at the current market scenario, it is hard to even say that Google may have achieved the forecasted goal.

Most of the people are unaware of:

* Many devices are in development but are not as handy as Google Glass
* Google glass supports the trend towards the passive consumption of content and technologies that fulfil the people's need
* It gives immediate notification to people regarding search operations and other Google services
* Almost all of the Google Glass features are available in smartphones, but it cannot be used every time and everywhere
* Google Glass has the features to get reservations, appointments, the reminders for important dates and events

Google is the first to move towards the digitalization of spectacles. Many products have been developed and some are still in development. Many have features which are difficult to handle in certain conditions like while driving, it can be dangerous to use smartphones. In digitalization, one of the approaches to solve similar problems is to use a head-mounted wearable device, so it can be easily accessible without having any problem. All that being said, after our research, we found that this technology can be really useful in avoiding road accidents if it is implemented the right way. To overcome this, we are preparing a module to overcome the problems that occurs in already developed modules.

# Objectives

Artificial read and learn system is developed many advanced technologies having an interest in automation we are planning to develop a system for smart coding with following objective in our mind:

* To reduce human efforts faced by developers while programming.
* To provide assistance to developers for code Generation.
* To keep the Software development Feasible.

# Tools and data collection

When the google glass was introduced, it had very few and specific features. But after further development, the smart glasses first appeared in the year 2002 by Google brought various possibilities in this area. With developing technology, there came many possibilities for advancements. The features like below can be improved with the aid of new technologies as well.

1. Image Recognition & Object Detection:- In the field of Image Recognition, development is still undergoing to improve efficiency. Many novel algorithms have been developed. There are many applications of object detection and image recognition like security or information gain.
2. Voice Assistance:- Google glass has the feature of a voice assistant. For the best accuracy and recognition, the voice commands are carefully turned and designed. This improves interaction with a product.
3. Navigation:- Google Glass provides the navigation feature also for travelers and presents directions on the display in front of the glass frame. It can be used at the time of traveling when directions for destination are unknown. Navigation can be provided by the reference of google map.
4. Google Search:- It also provides the search engine which operates with voice and works on the internet.

# Research methodolgy

The intelligent spectacle will be used to prevent road accidents. The time delay must be avoided. The major problems are the high processor usage for efficiency and keeping the expenditure nominal. Other than these two, heavy processor usage will cause major battery drain and the storage requirement will be also high. APIs are usable in this case to move data between cloud and spectacle. The cloud also brings the possibility of high GPUs. Many cloud services provide environments with high GPUs. APIs will speed up the data transfer process. Both these techniques are heavily used today in various applications. It will also resolve the problem of battery drainage.

# Process description

As the spectacle is going to be used while driving, the entire process must be efficient. The product will work on various models to facilitate the user with various helpful features. The voice command from the user will act as an input which in turn will trigger the specific model to commence working on that input. Then the output will be produced and will be dictated to the user. There are various algorithms developed for object detection, speech recognition. Our major objective is to train one of these algorithms as per our product’s requirement, using various data analysis techniques like feature extraction and data preprocessing. Most of the Machine Learning models need a high GPU requirement for smooth and fast processing. Various cloud services provide the environment with high GPUs. The Cloud Service Platforms like IBM Watson, Amazon Sagemaker, etc are mostly used for such tasks. Using this service in our spectacle will demand the prime need for High-Speed Internet Connectivity.

# Algorithm

## Training/Development Phase

Select Necessary and Efficient Models

Train the models on High GPU Environment

Deploy the trained models on embedded system directly OR Deploy the trained models on server & access via REST API on embedded system

## After Training/Deployment of Models Phase

READ(voice\_input)

Fuction1(voice\_input){

Process the voice input on GMM/HMM and fetch the text data

return voice\_text

end

}

Function2(voice\_text){

Switch(voice\_text):

Case objectDetection:

Run model for Object Detection

Case nightVision:

Run model for Night Vision

Case navigation:

Run model for Navigation

Case notificationsAndUpdates:

Run model for Weather Updates and Mobile Notifications

}

Function3(output){

Process the output and present it in form of voice

return voice\_output

}

The entire process is divided into 2 phases. The training phase will include the data collection, its analysis, models implementation, their training and deployment. The later phase is the part of deployment process which presents how the user will access the trained models based on various inputs and commands.

If ‘m’ is average time assumed for generating output by each model

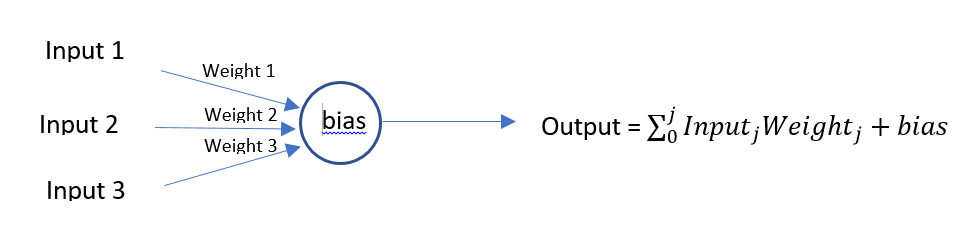
In worst case the time complexity would be O(3m + 4).

In best case it would be Ω(3m + 1).

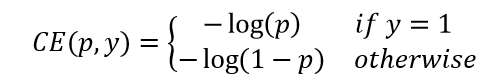
# Mathematical model

The model is going to use RetinaNet framework whose mathematical working is explained. The inputs, weights, nodes and outputs form a neural network.

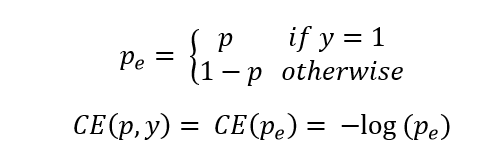
Most basic neural model would look like:



RetinaNet introduced the Focal Loss Function for improvement in Object Detection Algorithm. This Focal Loss Function is used over Cross Entropy Function. It is given by:

[2]

Simplified as:

 [2]

Where (p,y) is the set of probability ‘p’ of predicted class ‘y’

Most of the time in Machine Learning, the datasets we have are imbalanced. To generate an unbiased result, it is important for the datasets to be unbiased.

For example:

We have a dataset including 10000 images of cars, 10 images of bikes

Such dataset is biased towards cars and may not give accurate predictions.

To balance the datasets, we can either increase the weights of classes present in minority or lower the weights of classes present in majority. The Focal Loss Function uses the later one.

C:\Users\Avinash Barve\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\34D7BD64.tmp

Where γ is the weighing parameter used for balancing.

# Working Model

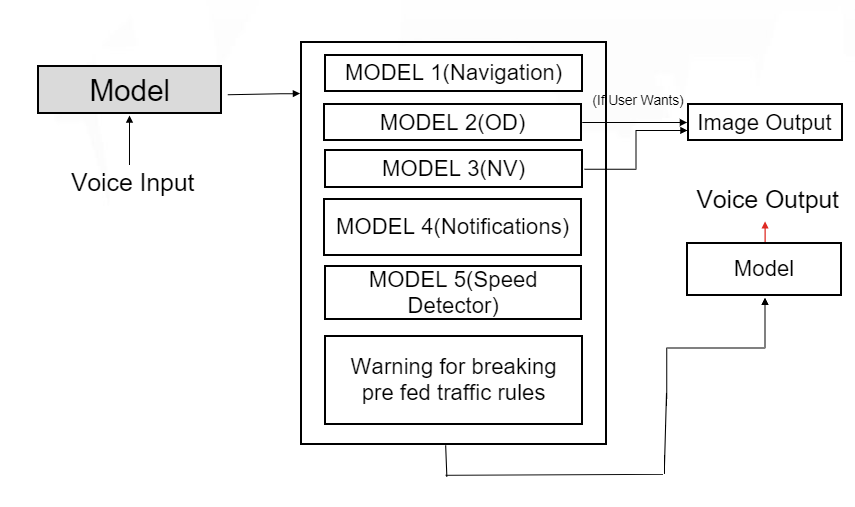


Figure 1

# Working Model Explaination

The voice command will be fetched from user using speech recognition which will trigger one of the listed models. The first model is for providing navigation instructions to user whereas second model is for object detection model which will help in detection of potholes, pedestrians, breakers, and vehicles. Model 3 will be trained for night vision and Model 4 for fetching notifications from user mobile. Model 5 will get triggered if the user exceeds the speed limit and the last one will warn before breaking any traffic rules.

# Future scope & Conclusion

Our model will return exact solution to your problem statement. It will reduce your efforts while writing a code. When we need to implement any new program or task for machines we have to program it as always. But, by using Natural Language Processing and Deep Learning we can let machines to learn from the given text data.To overcome the problem of always coding for every problem we are building a model which will read the text from the given Problem statement by using NLP and Transformer Architecture.Also the model will be able to code for a problem if we give it a problem statement.The problem of developers to generate huge codes can be resolved with the development of upcoming technologies. Many algorithms are still being improved to increase efficiency, which can greatly help in forming an ideal model..

# References

1. Jacob Devlin, Ming- Wei Chang, Kenton Lee, Kristina Toutanova,2019, IEEE, BERT: Pre- Training of Deep Bidirectional Transformers for Language Understanding.
2. Adams Wei Yu, David Dohan, Minh-Thang Luong, Rui Zhao, Kai Chen, Mohammad Norouzi, and Quoc V Le. 2018. QANet: Combining local convolution with global self-attention for reading comprehension. In ICLR.
3. Rowan Zellers, Yonatan Bisk, Roy Schwartz, and Yejin Choi. 2018. Swag: A large-scale adversarial dataset for grounded commonsense inference. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)
4. Zhang, V. Zhong, D. Chen, G. Angeli, and C. D. Manning. 2017. Position-aware attention and supervised data improve slot filling. In Empirical Methods in Natural Language Processing (EMNLP).
5. Huang, C. Zhu, Y. Shen, and W. Chen. 2018. Fusionnet: Fusing via fully-aware attention with application to machine comprehension. In International Conference on Learning Representations (ICLR)
6. <https://blogs.microsoft.com/ai/microsoft-creates-ai-can-read-document-answer-questions-well-person/>
7. Attention Is All You Need , [Ashish Vaswani](https://arxiv.org/search/cs?searchtype=author&query=Vaswani,+A), [Noam Shazeer](https://arxiv.org/search/cs?searchtype=author&query=Shazeer,+N), [Niki Parmar](https://arxiv.org/search/cs?searchtype=author&query=Parmar,+N), [Jakob Uszkoreit](https://arxiv.org/search/cs?searchtype=author&query=Uszkoreit,+J), [Llion Jones](https://arxiv.org/search/cs?searchtype=author&query=Jones,+L), [Aidan N. Gomez](https://arxiv.org/search/cs?searchtype=author&query=Gomez,+A+N), [Lukasz Kaiser](https://arxiv.org/search/cs?searchtype=author&query=Kaiser,+L), [Illia Polosukhin](https://arxiv.org/search/cs?searchtype=author&query=Polosukhin,+I)
8. Text generation with an RNN , https://www.tensorflow.org/tutorials/text/text\_generation