An Efficient Visual based Question Answering System using Machine Learning

D.Saranya , Assistant Professor, Department of Computer Science and Engineering, Mailam
Engineering College, Mailam,dsaranya93@gmail.com
Dr. M. Ramalingam, Professor, Department of Information Technology, Mailam Engineering
College, Mailam, rammalingamm_2005@yahoo.co.in,
AG.Noorul Julaiha, Assistant professor, Department of computer science and engineering,
Rajalakhmi Insitute of technology Chennai.
Dr.G.Nalinipriya, Professor/IT ,Saveetha Engineering College,,Chennai.nalini.anbu@gmail.com
Dr.T. Priyaradhikadevi. Head of the Department, Department of Computer Science and
Engineering, Mailam Engineering College, Mailam, hodcse@mailamengg.com.

ABSTRACT

In recent decades, enormous development has also been made in the fields of computer vision, object recognition, and natural language processing (NLP). Artificial Intelligence (AI) applications use NLP to provide the computer with a "comprehensive" capability, such as question-answering models. Natural language inquiries can be addressed by such a computer on any part of an unstructured document. An extension of this method is to integrate NLP with computer vision to perform the Visual Question Answering (VQA) mission, which is to construct systems that can respond to image questions in natural language. A variety of programs that use deep-learning frameworks and machine learning were developed for VQA. The research study implemented a VQA system that, by using a deep convolution neural network (CNN) that collects image attributes, achieves proper knowledge from images. More precisely, with this intent, functionality embedding from the output layer of the VGG19 models is used. Complex reasoning skills and comprehension of natural language are accomplished by our method so that the query can be understood accurately and an acceptable response can be returned. To acquire latent semantic embedding to retrieve information from the query, the Infer Sent model is used. To combine the picture and language models, different architectures are suggested. Our method ensures performance similar to the VQA dataset baseline programs.

1. INTRODUCTION

Recent developments in machine vision and deep learning research have made tremendous strides in many computer vision tasks, such as the identification of image classification artifacts and recognition of events. Given enough information, deep convolutionary neural networks (CNNs) rival the ability of humans to classify images. Similar findings can be expected for other-oriented computer vision issues with annotated databases quickly growing in scale due to crowd-sourcing. Nevertheless, these topics are limited in nature and do not entail holistic awareness of images. As humans, the objects in an image can be defined, the spatial locations of these objects can be recognized, their characteristics and relationships can be implied to each other, and with the surrounding context, we can also reason about the intent of each object. We may ask random picture questions and share the information obtained from them as well.

The overarching goal of VQA is to extract question-relevant semantic information from the images, which ranges from the detection of minute details to the inference of abstract scene while many computer vision issues include the retrieval of information from the images, as opposed to VQA; they are constrained in reach and generality. Recognition of objects, recognition of events, and classification of scenes may all be viewed as tasks for image classification, with today's best methods doing this using CNNs trained to classify images into particular semantic categories. Object detection, where algorithms today surpass humans in precision, is the most effective of all. Object identification, however, only makes it possible to locate the predominant object in an image while understanding its spatial position or its role in the larger image. Object detection involves the localization of specific semantic concepts (e.g., cars or people) by placing a bounding box around each instance of the object in an image. The best object detection methods all use deep CNNs. Semantic segmentation takes the task of localization a step further by classifying each pixel as belonging to a particular semantic class. Instance segmentation further builds upon localization by differentiating between separate instances of the same semantic class.

Five essential VQA datasets have been publicly available. These datasets make it easy to train and test VQA programs. As of this article, the main datasets for VQA are DAQUAR, COCO-QA, The VQA Dataset, FM-IQA, Visual7W, and Visual Genome. With exception of DAQUAR, all of the data sets include images from the Microsoft Common Objects in Context (COCO) dataset, In addition to the COCO images, Digital Genome and Visual7W use images through Flickr100M, which comprises 328,000 images, 91 highly similar groups at over 2 million labeled data, and an aggregate of There are simulated cartoon images in a section of The VQA Dataset, which we can refer to as SYNTH-VQA. The remainder of the VQA dataset, compatible with other articles, will be referred to since COCO-VQA, as it includes images from the COCO image dataset. Table 1 contains statistics for each of these datasets. To capture the heterogeneity in queries, photos, and ideas that exist in real-life situations, an optimal VQA dataset needs to be sufficiently large. It should also have a fair evaluation scheme that is difficult to 'game' and doing well on it indicates that an algorithm can answer a large variety of question types about images that have definitive answers. If a dataset contains easily exploitable biases in the distribution of the questions or answers, it may be possible for an algorithm to perform well on the dataset without really solving the VQA problem.

2. RELATED WORKS

VQA has been posed as either an open-ended task, in which an algorithm generates a string to answer a question or as a multiple-choice question where it picks among choices. For numerous, basic precision is sometimes used to test, via an algorithm having a correct answer if it made the correct choice. Easy precision is also used for open-ended VQA. For this scenario, the expected response string of an algorithm must precisely fit the answer to the input image. Accuracy can be quite strict, though, and some faults are even worse than others. If the query was, for instance,' What animals are in the photo? Instead of the right mark' pets,' and a machine produces' puppy,' it is penalized as harshly as it will be if it outputs 'zebra.' Questions can even include many valid responses, e.g.,' What is in the tree? ' 'Bald eagle' could have been classified as the right answer to ground facts because a machine that outputs' eagle 'or' bird 'will be penalized as well as if it had' yes 'output as the answer. Several approaches to exact precision for testing open-ended VQA algorithms have been suggested because of these problems.

A large number of VQA algorithms have been proposed in the past three years. All existing methods consist of extracting image features (image featurization), extracting question features

(question featurization), and an algorithm that combines these features to produce an answer. For image features, most algorithms use CNNs that are pre-trained on ImageNet, with common examples being VGGNet, ResNet, and Google Net. A wider variety of question factorizations have been explored, including bag-of-words (BOW), long short-term memory (LSTM) encoders, gated recurrent units (GRU), and skip thought vectors. To generate an answer, the most common approach is to treat VQA as a classification problem. The picture and query characteristics are the input to the classification scheme in this context and each specific answer is regarded as a different category. As shown in the featurization scheme, greatly varying types may be taken by the classification system. These systems differ significantly in how they integrate the question and image features.

In common, we can describe a VQA method as a model that uses the image as an input and a query about the image as a natural language and produces a natural language response as the output. This is by definition an issue in multi-discipline science. Now let us consider, for instance, the queries concerning the prior picture. For at least two motives, we need NLP: knowing the query and producing the answer. In text-based Q&A, a very well topic in NLP, those were typical issues.

The VQA dataset is comparatively bigger when related to several other datasets. It contains 50,000 abstract cartoon images, in addition to 204,721 images from the COCO data set. There are three questions per picture and ten responses per query with about 10 million answers, which is an over 760 K question. To accomplish this, the queries were created by a group of Amazon Mechanical Turk staff and the responses were written by some other group. One interesting point is that for testing, they propose two kinds of answer modes: open-ended and multiple-choice. They suggest a different metric for perhaps the first method:

accu= min (# humans that provided that answer,1)/3

That ensures that if at least 3 employees gave that same answer, a response is assumed 100 percent correct. They generate 18 member responses (wrong or right) per query for the various modes:

- The Correct Answer**:** The most common answer given by the ten annotators.
- Plausible Answers**:** Three answers were collected from annotators without looking at the image.
- Popular Answers**:** The top ten most popular answers in the dataset (e.g. "yes", "no", "2", "1", "white", "3", "red", "blue", "4", "green" for real images)
- Random Answers**:** Randomly selected correct answers for other questions.

Shyam, who studied at Delhi University. He presented a paper about the topic of visual question answering. In this survey, we look at a list of papers that addressed some of these challenges. We go through the most well-known datasets for the VQA task. As the first dataset for the VQA, we describe DAQUAR, and as one of the most utilized and very well built datasets in this region, the VQA 1.0 dataset. We explain the shortcomings of the VQA dataset which motivated the design of VQA 2.0. Moreover, we also explain several other datasets for the VQA task, such as Visual Madlibs, Visual 7W, and CLEVR.

There are many potential applications for VQA. The most direct use, perhaps, is to support blind and visually disabled people. Details about a picture on the Web or other social network may be generated by a VQA device. Integrating VQA into image retrieval systems is another clear application. On social media or e-commerce, this will have a major effect. VQA can also be used for educational or recreational purposes.

The VQA Collaboration seems to have very detailed and helpful websites with VQA data, tools, and applications. They conduct both the VQA Contest and the VQA Solution Workshop, and it is

useful to look at the assignments, presentations, and articles although they offer a clear example of the field's potential path.

3. PROPOSED SYSTEM

We suggest the task of addressing available and open-ended visual queries (VQA). The job is to have an accurate natural language response in light of an image and a natural language query about the picture. Both the questions and responses are open-ended, mirroring real-world situations, such as assisting the visually challenged. Visual questions, including background information and underlying meaning, selectively target multiple areas of an image. As a consequence, a VQA efficient system usually requires a more comprehensive interpretation of the picture and complicated logic than a generic image generation system. Besides, since many open-ended responses include either a few words or a closed collection of responses that can be presented in a multiple-choice format, VQA is capable of automated evaluation. We have a dataset of ~0.25M images, ~0.76M requests, and ~10M replies, and address the information it contains. Multiple classifications and techniques are given and contrasted with individual output for VQA.

3.1. Algorithms for VQA

In the past three years, a significant number of VQA techniques have been suggested. All existing methods consist of

- Extracting image features
- Extracting question features (question featurization)
- An algorithm that combines these features to produce an answer.

For image features, most algorithms use CNNs that are pre-trained on ImageNet, with common examples being VGGNet, ResNet, and GoogLeNet. A wider variety of question featurization's have been explored, including bag-of-words (BOW), long short term memory (LSTM) encoders, gated recurrent units (GRU), and skip-thought vectors. The most popular approach is to determine VQA as a classification problem to produce a response. The picture and query characteristics are the input to the classification scheme in this context and each specific answer is regarded as a different category. As illustrated in the factorization scheme and the classification system can take widely varied forms. These structures vary greatly in how the query and picture characteristics are integrated.

To successfully address the task, there are four main modules: image feature extraction, question understanding, answer generation, and feature filters. For the project, the first 3 phases are important, which partly explain photos and query and justification potential responses to the correct outcome. Depend are various techniques, the fourth module is applied to improve the final accuracies. These possible techniques include normalization, BOW, episodic memory network, etc. Due to the time limit, only apply the episodic memory network in this project.

3.2. Visual Feature Extraction

In this project, a pre-trained convolutional neural network-based model VGGNet is applied to extract image features. Depend on the different question answering model's architecture; the feature layer is selected accordingly. For instance, the fully connected layer "fc3" is selected for CNN+LSTM. This layer has 4096 parameters and could be input to answer the generation module directly. For the DMN model, the "conv5 3" layer is selected. This layer has 14×14×512 parameters. During the "conv5 3" layer extraction execution, different difficulties are encountered: a very huge

vector could not be written out of storage, out of hard drive space. During runtime, We have focused on extracting functions as well. However, this results from unreasonable long training time.

3.3. Question Understanding

LSTM is used by the system to derive query characteristics q (the final hidden state of LSTM). To reflect each term, the pre-trained Glove is added.

3.4. Answer Generation

The answer generation module receives both question feature (or filter question feature) and image feature (or filtered image feature). These two characteristics are then normalized into a series of words and entered into an LSTM module. To build the model, cross-entropy loss on answers is added.

3.5. Episodic Memory Network

There are various filters for images and questions. In this project, the episodic memory networks applied that functions as the attention mechanism on image input.

4. MODELS FOR VISUAL QUESTION ANSWERING

4.1. Baseline Models

Baseline methods help determine the difficulty of a dataset and establish the minimal level of performance that more sophisticated algorithms should exceed. For VQA, the simplest baselines are random guessing and guessing the most repeated answers. A widely used baseline classification system is to apply a linear or non-linear, e.g., multi-layer perceptron (MLP), classifier to the image and question features after they have been combined into a They found that an MLP model with two hidden layers trained on these off-the-shelf features worked well for all datasets. However, in their work, a linear classifier outperformed the MLP model on smaller datasets, likely due to the MLP model overfitting.

4.2. Bayesian and Question-Aware Models

VQA requires drawing inferences and modeling relationships between the question and the image. Once the questions and images are futurized, modeling co-occurrence statistics of the question and image features can help draw inferences about the correct answers. Two major Bayesian VQA frameworks have explored modeling these relationships. In, the first Bayesian framework for VQA was proposed. The authors used semantic segmentation to identify the objects in an image and their positions. Then, a Bayesian algorithm was trained to model the spatial relationships of the objects, which was used to compute each answer's probability. This was the earliest known algorithm for VQA, but its efficacy is surpassed by simple baseline models. This is partially due to it being dependent on the results of the semantic segmentation, which was imperfect. A very different Bayesian model was proposed. The model exploited the fact that the type of answer can be predicted using solely the question. For example, 'What color is the flower?' would be assigned as a color question by the model, essentially turning the open-ended problem into a multiple-choice one. To do this, the model used a variant of quadratic discriminant analysis, which modeled the probability of image features given the question features and the answer type. ResNet-152 was used for the image features, and skip-thought vectors were used to represent the question.

4.3. Attention Based Models

Using global features alone may obscure task-relevant regions of the input space. Attentive models attempt to overcome this limitation. These models learn to 'attend' to the most relevant regions of the input space. Attention models have shown great successes in other vision and NLP tasks, such as object recognition, captioning, and machine translation. In VQA, numerous models have used spatial attention to create region-specific CNN features, rather than using global features from the entire image. Fewer models have also explored incorporating attention into the text representation. The basic idea behind all these models is that certain visual regions in an image and certain words in a question are more informative than others for answering a given question. For example, for a system answering 'What color is the umbrella?' the image region containing the umbrella is more informative than other image regions. Similarly, 'color' and 'umbrella' are the textual inputs that need to be addressed more directly than the others. Global image features, e.g., the last hidden layer of a CNN, and global text features, e.g., bag-of-words, skip-thoughts, etc. may not be granular enough to address region specific questions.

4.4. Bilinear Pooling Methods

VQA relies on jointly analyzing the image and the question. Early models did this by combining their respective features using simple methods, e.g., concatenation or using an element-wise product between the question and image features, but more complex interactions would be possible with an outer-product between these two streams of information. Similar ideas were shown to work well for improving fine-grained image recognition. Below, we describe the two most prominent VQA methods that have used bilinear pooling.

5. ALGORITHM USED IN VISUAL QUESTION ANSWERING

In Visual question answering which contains two Algorithms, the algorithm is listed below.

- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)

5.1. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a specific class of neural networks that perform exceptionally well in image classification and recognition and other computer vision tasks. Convolutional Networks have been very effective in identifying objects, human faces, and traffic signals apart from rendering vision in self-driving cars and robots. All of the convolutional models will consist of the following four operations, which can be considered as the building blocks of CNNs:

1. Convolution
2. Non-Linearity (ReLU)
3. Pooling or Sub Sampling
4. Fully-connected Layer

5.2. Recurrent Neural Networks

Recurrent neural networks (RNNs) are a special class of neural networks that were modeled to recognize patterns in sequences of data, such as handwriting, text, genomes, numerical data, or time-series data generated from stock markets, sensors, etc. Recurrent nets are arguably the most powerful class of networks modeled to identify structures in data sequences, such as handwriting, time-series

data, text, speech, genomes, or information generated from sensors, stock markets, IoT devices, etc. Recurrent neural networks not only take the input at the current step as input but also what they have perceived in the previous step.

The decision a recurrent neural network will reach at time t is affected by the decision it has reached at time t-1. We can say that there are two sources of input to a recurrent neural network, the present and the most recent past, based on which they decide on the new data. These networks differ from feed-forward neural networks by the feedback loop connected to their previous decisions; feeding the outputs back as input continuously. Recurrent networks make use of the information or knowledge present in the sequence itself to achieve tasks that feedforward networks cannot. Adding memory to the network makes all the difference.
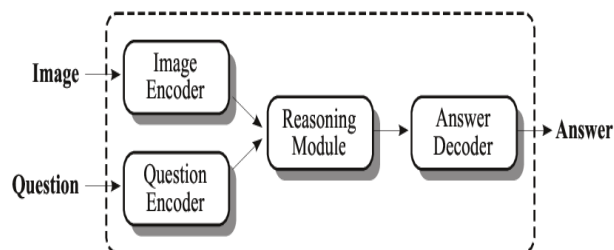

Fig 1: Architecture for Visual question Answering

6. IMPLEMENTATION

1. Get the input image
2. Generate the caption from the input image
3. Get the question from the user
4. Answering to the user from generated caption

6.1. Get the Input Image

This step is considered to be the first input, where the image that contains a certain object considered is inserted. The image can be either colored or gray image file of the human face with the most-used image extension (jpg, BMP, tiff ...etc.).We don't directly use the image as input into the model. The picture is scaled to 224 by 224, and then the activations are removed from the last VGGNet19 layer of CONV. These [512 x 7 x 7] shape activations are used as input features for the images.

6.2. Generate Caption from Input Image

A common body of basic understanding linking communication and perception is needed for visual question answering (VQA) and image captioning. We propose a new approach to improving the efficiency of VQA that takes advantage of this relation by jointly producing captions aimed at helping to address a particular visual query. The model is trained using an existing caption dataset by automatically determining question-relevant captions using an online gradient-based method.

6.3. Get the Question from the User

Get the input image-related question from the user. And we send the input question to the trained dataset. A single image can contain various answers to a related user question. This model has been very faster because of using the neural network concept.

6.4. Answering to User from Generated Caption

Next, we discuss the distinction between using captions models that can be created as well as using given by human annotators. In particular, we train our model with generated question-agnostic captions using the Up-Down (Anderson et al., 2018) caption, question-relevant captions from our caption generation module, and human-annotated captions from the COCO dataset.
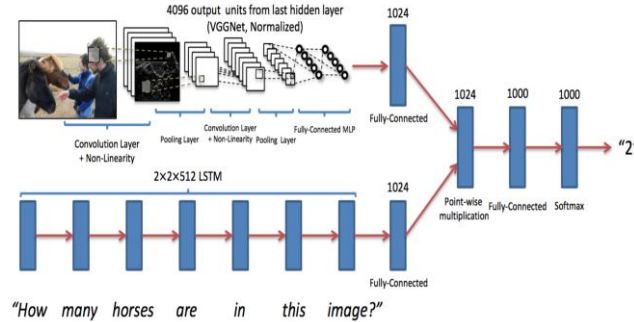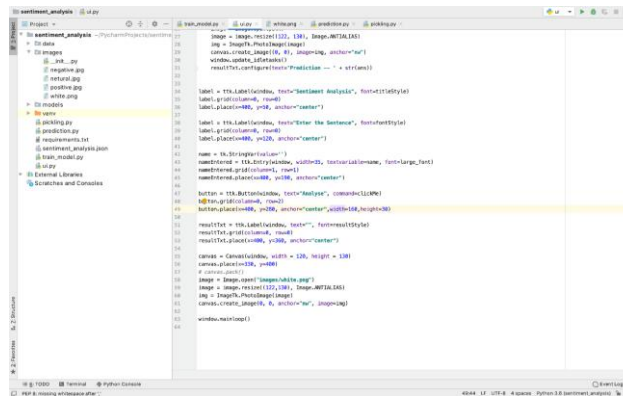

Fig 2: the process of visual question answering

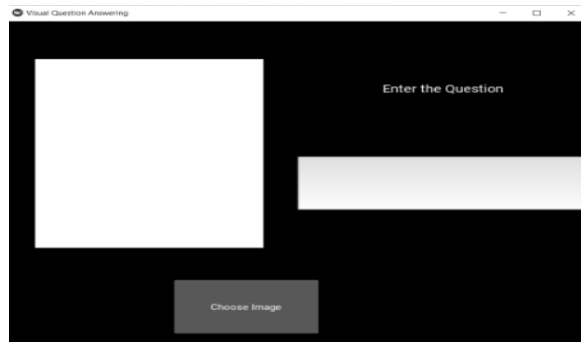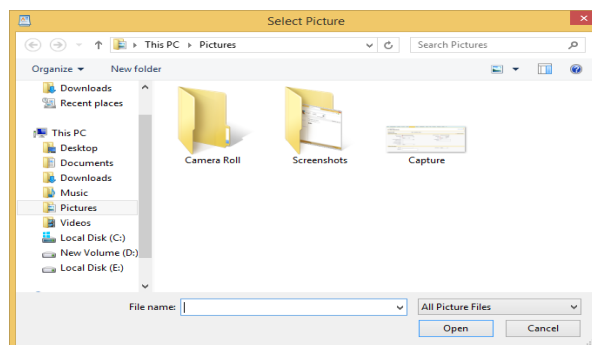

Fig 3: Prediction of data code

Fig 4: Main page

Fig 5: Image choosing

Fig 6: Answer for the Question

7. CONCLUSION AND FUTURE WORK

VQA is an important basic research problem in computer vision and natural language processing that requires a system to do much more than task specific algorithms, such as object recognition and object detection. A breakthrough in artificial intelligence will be an algorithm that can address random questions about pictures. We assume that every visual Turing test should be a mandatory part of VQA. We critically analyzed current databases and algorithms for VQA in this article. We discussed the challenges of evaluating answers generated by algorithms, especially multiword answers. We explained how current databases are afflicted by bias and other concerns. This is a major problem, and the field needs a dataset that evaluates the important characteristics of a VQA algorithm, so that if an algorithm performs well on that dataset then it means it is doing well on VQA in general.

Future work on VQA includes the creation of wider and far more varied datasets. Bias in these datasets will be difficult to overcome, but evaluating different kinds of questions individually in a nuanced manner, rather than using naive accuracy alone, will help significantly. Further work will be needed to develop VQA algorithms that can reason about image content, but these algorithms may lead to significant new areas of research.

REFERENCES

[1] B. M. Marlin, R. S. Zemel, S. Rowesis, and M. Slaney, Collaborative filtering and the missing at random assumption in proc.23rd conf.uncertainly Arif.Intell. , 2007, pp.267-275.

[2] R.Salakhutdinov and A.Mnih, Probabilistic matrix factorization, in Proc.Int.conf. Neural Inf. process.Syst., vol. Volume 20, pp.7-8, 2011.

[3] D. lee and H. S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature, vol.Volume 401, no.issue 6755, pp.788-791, 1999.

[4] K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp.770–778, 2016.

[5] Hasan, S.A., Ling, Y., Farri, O., Liu, J., Lungren, M., Muller, H.: Overview of the ImageCLEF 2018 medical domain visual question answering task pp.10-14 2018.

[6] D. Geman, S. Geman, N. Hallonquist, and L. Younes, "Visual Turing test for computer vision systems," Proceedings of the National Academy of Sciences, vol. 112, no. 12, 2015.

[7] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in Advances in Neural Information Processing Systems (NIPS), 2013.

[8] R. Bernardi, R. Cakici, D. Elliott, A. Erdem, E. Erdem, N. Ikizler-Cinbis, F. Keller, A. Muscat, and B. Plank, "Automatic description generation from images: A survey of models, datasets, and evaluation measures," Journal of Artificial Intelligence Research, vol. 55, pp. 409–442, 2016.

[9] Lu, Pan, and Ji, Lei and Zhang, Wei and Duan, Nan and Zhou, Ming and Wang, Jianyong. "R-VQA: Learning Visual Relation Facts with Semantic Attention for Visual Question Answering". arXiv preprint arXiv:1805.09701, 2018.

[10] I. Ilievski, S. Yan, and J. Feng. "A focused dynamic attention model for visual question answering." arXiv preprint arXiv:1604.01485, 2016.

[11] Wu, Qi and Shen, Chunhua and Wang, Peng and Dick, Anthony and van den Hengel, Anton. "Image captioning and visual question answering based on attributes and external knowledge". IEEE transactions on pattern analysis and machine intelligence, vol. 40,pages 1367- 1381, 2018.

[12] Z. Yang, X. He, J. Gao, L. Deng, and A. J. Smola." Stacked attention networks for image question answering." in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. B. Thomee, D. A. Shamma, G. Friedland, B. Elizalde, K. Ni, D. Poland, D. Borth, and L.-J. Li. "Yfcc100m: The new data in multimedia research". Communications of the ACM, vol. 59, no. 2, pp. 2016.

[13] Simonyan, Karen and Zisserman, Andrew. "Very deep convolutional networks for large-scale image recognition". arXiv preprint arXiv:1409.1556. 2014.

[14] M. Malinowski and M. Fritz." A multi-world approach to question answering about real-world scenes based on uncertain input," in Advances in Neural Information Processing Systems (NIPS), 2014.

[15] Xu, Kelvin, and Ba, Jimmy and Kiros, Ryan and Cho, Kyunghyun and Courville, Aaron and Salakhudinov, Ruslan and Zemel, Rich and Bengio, Yoshua. "Show, attend and tell: Neural image caption generation with visual attention". Book International conference on machine learning, pages 2048–2057, 2015.

[16] M. M. Rahman, Y. Tan, J. Xue, L. Shao, and K. Lu, "3D object detection: learning 3D bounding boxes from scaled Down 2D bounding boxes in RGB-D images," *Information Sciences*, vol. 476, pp. 147–158, 2019.

[17] D. Tenney, L. Liu, and A. van den Hengel, "Graph-structured representations for visual question answering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Honolulu, HI, USA, July 2017.

[18] N. Xu, A.-A. Liu, J. Liu, W. Nie, and Y. Su, "Scene graph captioner: image captioning based on structural visual representation," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 477–485, 2019.

[19] D.Saranya, Dr.G.Nalinipriya, N.Kanagavalli, S.Arunkumar, G.Kavitha, "Deep CNN models for Driver Activity Recognition for Intelligent Vehicles", International Journal of Emerging Trends in Engineering Research, Vol.10, No.10 October 2020, pp. 7146- 7150.

[20] D.Saranya, S.Thulasidass, D.Gomathi, "Automatic Service Discovery using Ontology Learning Semantic Focused Crawler for Mining" *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* Volume 4 Issue 10, October(2015).

[21] Mrs. D. Gomathi, Dr. M. Ramalingam, Mr. D. Jeyakumar, Ms. D. Saranya, "The Proficient Context-Aware QoS framework for Reliable Web of Things*", International Journal of Advanced Research in Engineering and Technology (IJARET)* Volume 11, Issue 10, October 2020, pp. 1760-1768.

[22] T. Priyaradhikadevi, R.M.S Parvathi, and V. Chithra. Analyzing Mapping Technique in Ontology by Implementation of RUP, International Journal of Emerging Trends and Technology in Computer Science, vol. 1, no.4, pp.81-86, 2012. 14.

[23] T.Priyaradhikadevi, R.M.S Parvathi, G. Ezhilarasi, Implementation of Rank Based Semantics Association for Service Discovery and Composition, International Journal of Science and Research (IJSR) (Vol.2, No. 1), pp.188-191, 2013.

[24] Ramalingam, M., and Parvathi, R.M.S., 2013. A Novel Context for Fine-Grained Access in Semantic Web Services By Using Policy-Based Semantic Access Control. Asian Journal of Computer Science and Technology (AJCST), Vol1.No.1, pp.27-34.

[25] Ramalingam, M., and Parvathi, R.M.S., 2013. Secure Semantic Aware Middleware: a Security-Based Semantic Access Control for Web Services. International Review on Computers and Software (I.RE.CO.S.), Vol. 8, N. 9 ISSN 1828-6003.

[26] Ramalingam, M., and Parvathi, R.M.S., 2012. Policy-Based Semantic Access Control Framework for Fine-Grained Access in Semantic Web Services. European Journal of Scientific Research, ISSN 1450-216X Vol.74 No.1, pp.154-163.

[27] M. Ramalingam, D. Gomathi, R. Shankar Ram, A Framework for Road Transportation Modern Pedestrian Using Solar Energy, International Journal of Advanced Research in Engineering and Technology, October 2020.

[28] D. Gomathi, Dr. M. Ramalingam, Dr.G.Nalinipriya, N.Kanagavalli, D. Saranya, "Face Emotion Identification System for Visually Challenged Persons using Machine Learning", International Journal of Advanced Trends in Computer Science and Engineering, Volume 9, No.5, September - October 2020, pp- 8309-8314.

[29] M. Ramalingam, D. Gomathi, D. Saranya, Proactive Assistance System for Visually Challenged Persons Using IOT Based On Machine Learning, Journal of Computational and Theoretical Nanoscience, November 2020.