International Journal of Innovative Research in Engineering
Volume 3, Issue 2 (March-April 2022), PP: 235-240.
www.theijire.com
ISSN No: 2582-8746

COMPANION - An Application for Impaired Persons

Sayed Mohammad Umair Rashid[1], Khan Needa Anjum Ishar[2], Dr. Shabina Sayed[3]
[1,2,3]*Department Of Information Technology, M.H. Saboo Siddik College of Engineering, Mumbai, India*

How to cite this paper:
Sayed Mohammad Umair Rashid[1], Khan Needa Anjum Ishar[2], Dr. Shabina Sayed[3], "COMPANION - An Application for Impaired Persons", IJIRE-V3I02-235-240.

Abstract: *Communication barrier is one of the biggest challenges for hearing and speech impaired persons. One has to either convey their messages by sign language or by using familiar gestures. Our proposed system aims to tackle this issue to an extent. The main motive of this project is to design a computer interaction device that takes input in a gesture format and provides an output in the form of text, hence conveying the message in a readable format to normal persons. Thus, our project COMPANION converts sign language into text.*

Key Word*: Hand Gestures Recognition, Human-Computer Interaction, OpenCV*

I. INTRODUCTION

Communication is a means of sharing or exchanging information, thoughts, or ideas. To establish any kind of communication between two persons, they are required to know and understand a common language. It is easier for a normal person i.e., a person with no disability, but in the case of a hearing and speech impaired person, the means of exchanging communication are different. There are around 466 million people globally with hearing loss and around 34 million of these fall under the age of children. Their way of communication is done using Sign Language among themselves or with a normal person who has the knowledge of Sign Language.

Sign Language

Sign Languages are the native language and the prime means of communication of the deaf and dumb community and it is done by hand movements and gestures. There are different types of sign language such as British Sign Language (BSL), Indian Sign Language(ISL), Japanese Sign Language (JSL), American Sign Language (ASL), and German Sign Language (GSL). ISL uses both hands to form any gestures or signs and they are derived from British Sign Language(BSL) and French Sign Language(FSL). Most of the researchers in this field concentrate on ASL(American Sign Language) because, in this system, gestures and signs are formed using a single hand so the complexity is very less. Another important feature of ASL is that it already has a standard database, available for use.

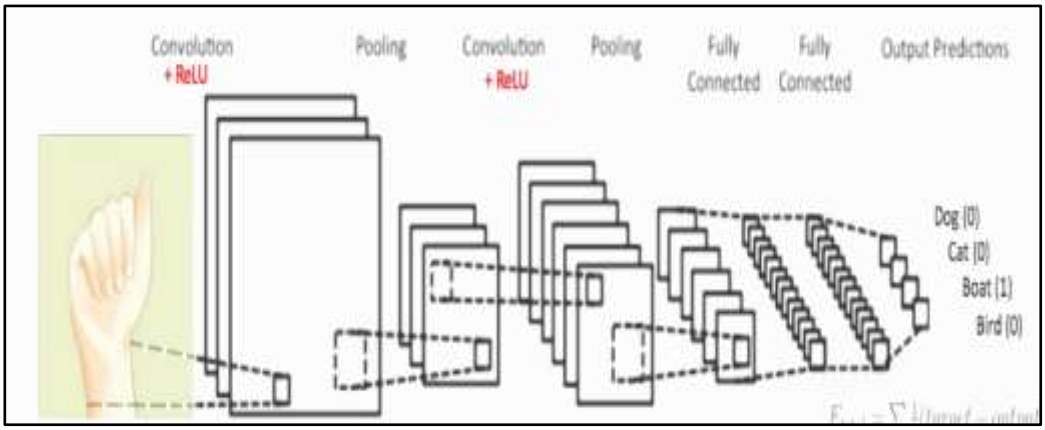

Fig. 1: Stages of CNN breakthrough steps

American Sign Language(ASL)

ASL is a natural language and it serves as the predominant sign language of the deaf and dumb community in the U.S. Besides North America, ASL is used in many countries around the world including Africa and Southeast Asia. There are very few people that can easily communicate using ASL or any other sign language. Therefore, it is necessary to find a way so that communication is possible between the majority of the hearing and mute community. The ASL recognition system is the new

way of understanding the thoughts and ideas of disabled people through sign gestures without having the knowledge of sign language.

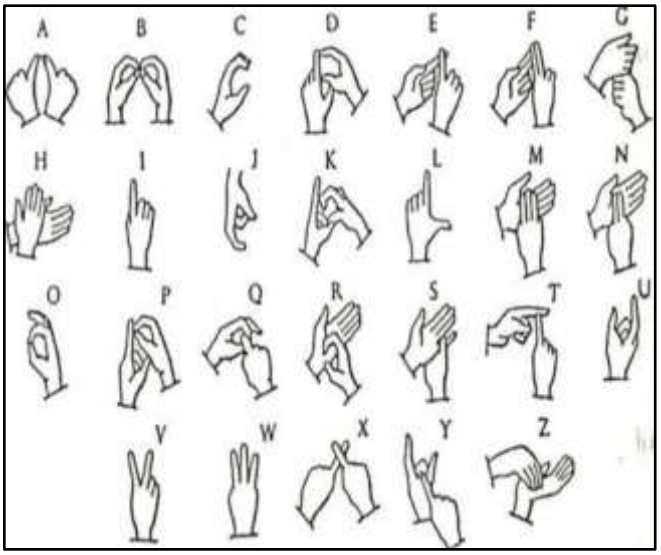

Fig. 2: American Sign Languages

II.RELATED WORKS

Real-Time Recognition of Indian Sign Language (978-1-5386-9471-8/19/2019 IEEE)

The real-time sign language recognition system is developed for recognizing the gestures of Indian Sign Language (ISL). Generally, sign languages consist of hand gestures and facial expressions. For recognizing the signs, the Regions of Interest (ROI) are identified and tracked using the skin segmentation feature of OpenCV. The training and prediction of hand gestures are performed by applying fuzzy c-means clustering machine learning algorithms. Gesture recognition has many applications such as gesture-controlled robots and automated homes, game control, Human-Computer Interaction (HCI), and sign language interpretation. The proposed system is used to recognize real-time signs. Hence it is very much useful for hearing and speech impaired people to communicate with normal people.

Methodology:

The proposed system has a camera unit for capturing the gestures of hearing and speech impaired people. The real-time sign language recognition system was designed as a portable unit for more convenience for the users. The raw videos taken in a dynamic background are given as input to the system. The image frames are resized to maintain equality among all the videos. OpenCV (Open-Source Library for Computer Vision) is used for feature extraction and video classification.

Sign Language Recognition Using Modified Convolutional Neural Network Model (978-1-5386-9422-0/18/2018 IEEE)

Sign Language is an interesting topic and similar to Action Recognition. Especially along with the great development of Deep Learning. Video-based Sign Language Recognition is our concern because we want to recognize a sign not only by the shape but also by the action the signer does. The problem is sign language is very complex and varies. The variation of sign language is making the system harder to recognize all the words accurately. Many researchers have been researching Sign Language Recognition for a long time. So many methods have been used to find out which one is the best method. Because of the similarity between Sign Language Recognition and Action Recognition, we are trying to implement one of the top-tier models in Action Recognition which is i3d inception. This model is also a new Action Recognition model with very high accuracy. So we can know if it is possible to adopt Action Recognition behavior into Sign Language Recognition. The goal of this paper is to implement the i3d inception model for Sign Language Recognition with the transfer learning method. From the test we've been doing, we got 100% accuracy on training with 10 words and 10 signers with 100 classes but the validation accuracy is pretty low. This model is too overfit.

Methodology:

i3d inception without modification is too overfit because of the results, we trained the model with 10 signers and 100 classes (500 videos) with 200 epochs, which have a good training accuracy, but very low validation accuracy. We can do a lot more things with this model, like freeze the layers, remove some inception modules, remove the transfer learning, and change the fully connected layer into another deep learning model. In our opinion, the fully connected layer is not much of a matter. We think we should be concerned about the convolutional neural network layer more since it's the detector.

The other reason for these results might be from the dataset, dataset. The differences in background lightning of LSA64 could be the cause of the overfit. Because the differences in background lightning could be a feature of machine learning. So, machine learning would learn from the wrong feature.

Table 1: Comparison of the above Research Papers

Factor	Research Paper	Key points
Accuracy in the training dataset	Gnosis	69%
	Just Be My Eyes	78%
Speed in Detection	Gnosis	69%
	Just Be My Eyes	64%
Security	Gnosis	45%
	Just Be My Eyes	74%
Detection in Noise	Gnosis	88%
	Just Be My Eyes	90%

Procedure methodology

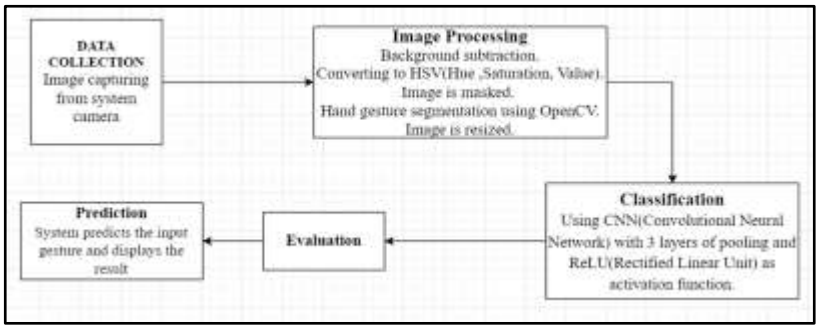

Fig 3. Methodology of System

1. DATA COLLECTION

Data collection is an essential part of our proposed system as our result highly depends on the dataset collected. Therefore, we have created our own dataset of ASL (American Sign Language) having 1750 images of each alphabet. This dataset folder of collected sign gestures is further split into training and testing folders. All the captured images are stored in the PNG format because there is no loss in quality whenever the image is opened or closed many times, also, PNG is good in handling high contrast images. The system camera will capture the images in the RGB color space.

2. IMAGE PROCESSING

I.Background Subtraction and HSV color space

The images that are captured are in RGB color spaces, it becomes troublesome to perform hand segmentation based on skin color only. Therefore, we tend to rework these images in HSV colorspace. HSV is a model that splits the color of an image into 3 components i.e.Hue, Saturation, and Value. It is a strong tool to improve the stability of the pictures by setting the brightness associated, it is quite the same as how humans understand color. In HSV, Hue represents color, it is an angle from zero degrees to 360 degrees and it is unaffected by any reasonable shadings and illumination, therefore it will be used as background elimination. Saturation indicates the range of grey within color space. Value is that the brightness of the color varies with color saturation. A trackbar has H starting from 0 to 179, S from 0 to 255, and V ranging from 0 to 255. This can facilitate setting the background to black and detecting the hand gesture in real-time. The below image is obtained when applying a mask.

II.Hand Gesture Segmentation

Gesture segmentation is an important part of sign detection. The captured image is then transformed into a grayscale image. In this process, there will be a loss of color of the skin gesture, but it will also improve the robustness of the system to the changes in lightning or illumination. The image frame is resized to 64 x 64 pixels. At the end of this process, binary images of 64 x 64 are obtained where the white area represents the hand sign and the black area represents the background.

III.Feature Extraction

Extracting the features from the image is very important in image processing. It refers to transforming raw data into numerical features which can be processed while keeping the information in the original data set. It provides better results than applying machine learning algorithms directly to the raw data. Raw images when captured and stored as a dataset take up a lot of system space because they are composed of lots of data. With the help of this process, we can reduce this problem after extracting the most important features.

- Input layer: - In this layer, input is given to the CNN model. The number of neurons present in the input layer equals the total number of features the given data has.
- Hidden layer: - Output of the above layer is passed to the hidden layers. Depending on the dataset and model, there can be many hidden layers and each layer will have a different number of artificial neurons which are in higher numbers than the features present in the data. The output from every layer is computed by matrix multiplication of output of the previous layer with analysis in a position weights of that layer and then by way of the addition of study capable biases accompanied via activation characteristic which makes the network nonlinear.
- Output layer: - In this layer, the output of the hidden layer is fed into a logistic function which converts the output of each class into the probability score of the respective class.

III.RESULT

The user needs to start the Companion application and is prompted to a video showing different sign languages. Either he/she needs to finish the video or needs to press the letter 'q' to skip the video. Once the application interface is popped up, the user needs to select any operation such as scanning any gesture and identify the gesture, form a sentence, create a new local gesture or export the gesture file in a pdf file format.

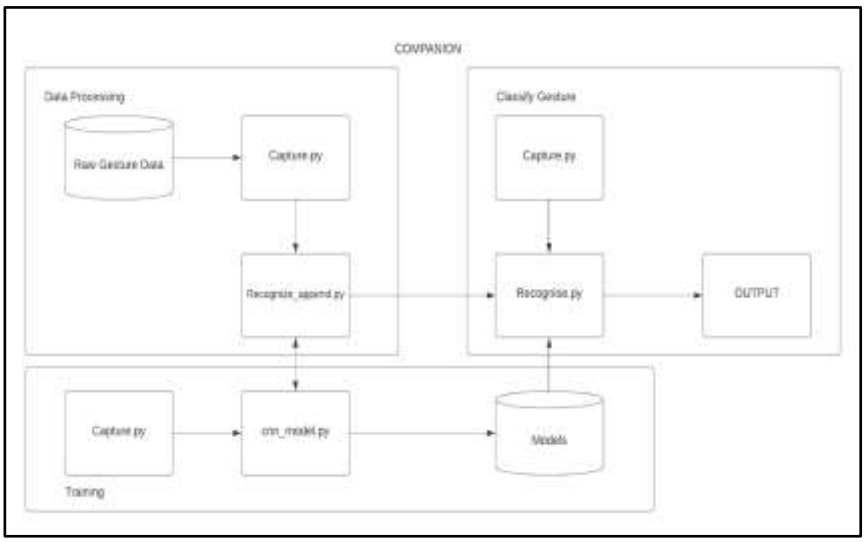

Fig. 4: System Architecture

When the user selects the operation for scanning a gesture, a visual mask for the application appears. By adjusting the local environment by the defined scroller, and being in a plain background, the user then shows any sign from the predefined ASL and the application predicts the sign.
If the user wants to add or customize any local gesture, he needs to switch the operation to create a gesture and add any gesture according to the user's comfort.
Now to communicate with a normal person, the user needs to select the operation for forming sentences. The user again adjusts the display settings similarly and then forms a combination of letters and then by using a space gesture, discontinuation of words is formed and thus making a complete sentence.

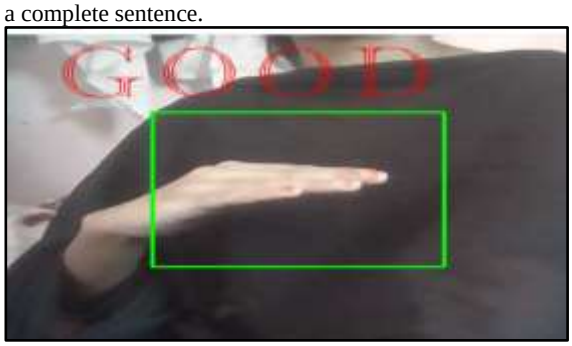

Fig. 5: Application Workflow

Fig. 6: Conversion of the loaded image into grayscale

The user can also export its local database by selecting the operation of converting the file into a pdf format, which makes the application reliable in case the application is formatted or used in a different device.

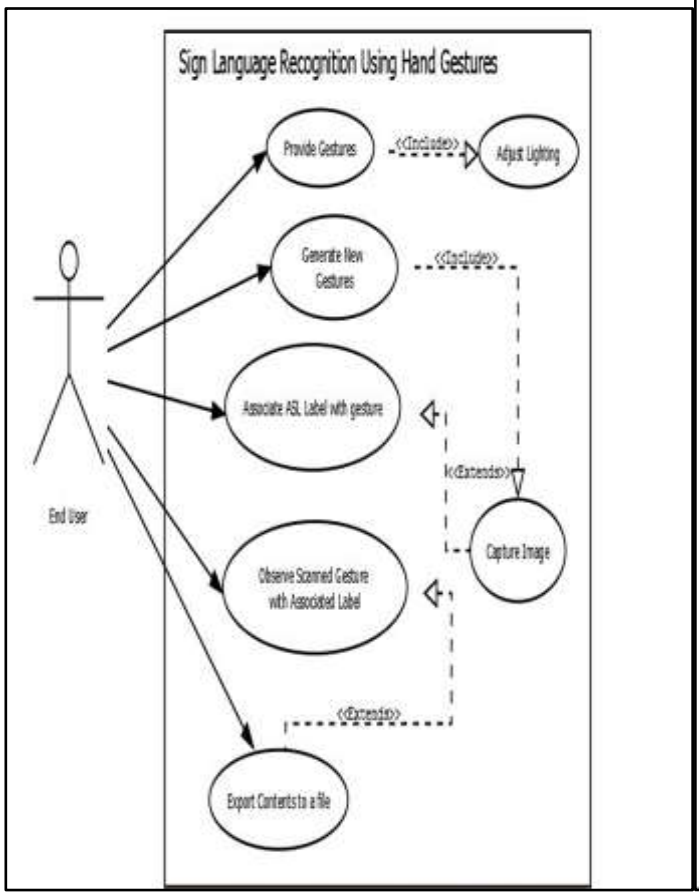

Fig. 7: Use Case Diagram

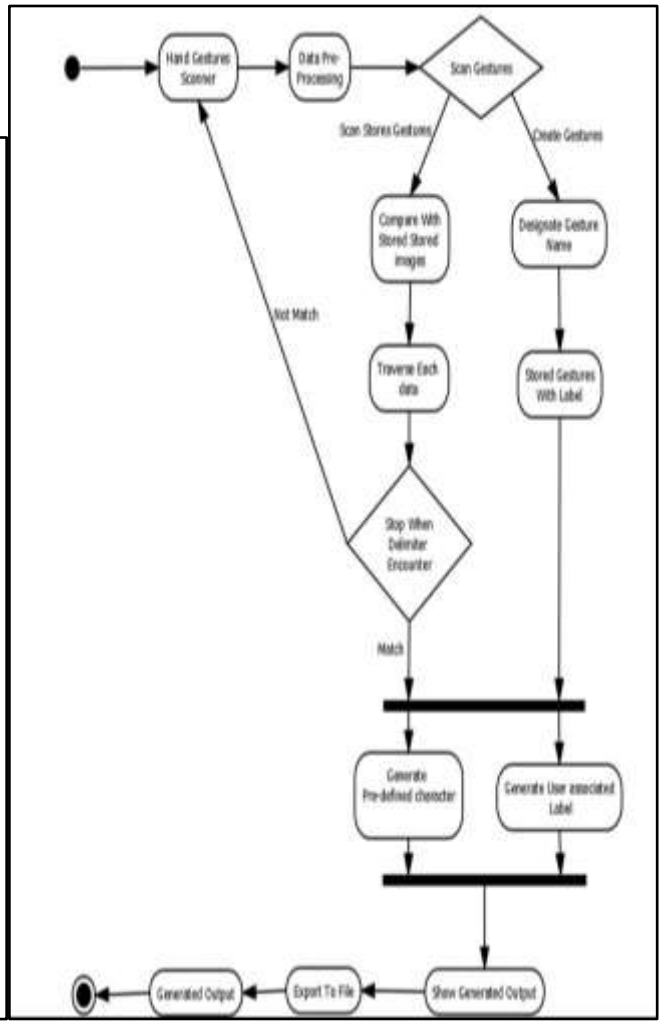

Fig. 8: Activity Diagram

IV.FUTURE SCOPE

1. Enabling TTS assistance to make the app more convenient.
2. Convert normal text to sign language
3. Introduce a two-way interface for visually impaired persons.
4. Project a mobile application to make it more reliable for users.

V.CONCLUSION

The application thus is made by removing the barrier of miscommunication between a speech-impaired person and a normal person. The user thus can easily interpret his language and form new gestures to locally expand the reach of communication to the audience.

References

1. *Buzzi, M. C., Buzzi, M., Leporini, B., & Trujillo, A. (2015). Design of Web-based tools to study blind people's touch-based interaction with smartphones. In Proceedings of the international peer-reviewed conference, Los Angeles, CA-USA (pp. 7–12). https://doi.org/10.1007/978-3-319-21380-4_2.*
2. *Kamble, K., & Kagalkar, R. (2014). A review: Translation of Text to Speech conversion for the Hindi Language. International Journal of Science and Research,3(11), 1027–1031.*
3. *Kumar, A. S., & Begum, S. (2014). Sign mobiles: An android app for specially-abled people. International Journal of Advanced Research in Computer Science and Software Engineering,4(9), 208–214.*
4. *Frey, B., Southern, C., & Romero, M. (2011). BRAILLETOUCH: Mobile texting for the visually impaired. Lecture Notes in Computer Science,6767, 19–25. https://doi.org/10.1007/978-3-642-21666-4_3.*
5. *Reddy, B. R., & Mahender, E. (2013). Speech Text conversion using the android platform. International Journal of Engineering Research and Applications,3(1), 253–258.*
6. *Kane, S. K., Jayant, C., Wobbroc, J. O., & Ladner, R. E. (2009) Freedom to roam: A study of mobile device adoption and accessibility for people with visual and motor disabilities. In Proceedings of the 11th international ACM SIGACCESS conference on computers and accessibility, Pittsburgh, Pennsylvania, USA (pp. 115-122). https://doi.org/10.1145/1639642.1639663.*
7. *Azenkot, S., Prasain, S., Borning, A., Fortuna, E., Ladner, R. E., & Wobbrock, J. O. (2011) Enhancing independence and safety for blind and deaf-blind public transit riders. Proceedings of the SIGCHI conference on human factors in computing systems, Vancouver, BC, Canada (pp. 3247–3256). https://doi.org/10.1145/1978942.1979424.*
8. *Kaur, S., & Dhindsa, K. S. (2015). Green computing—Saving the environment with intelligent use of computing. CSI Communication,39(5), 14–16.*
9. *Kane, S. K., Bigham, J. P., & Wobbrock, J. O. (2008). Slide rule: Making mobile touch screens accessible to blind people using multi-touch interaction techniques. In Proceedings of the 10th international ACM SIGACCESS conference on computers and accessibility, New York, NY, USA (pp. 73–80). https://doi.org/10.1145/1414471.1414487.*

10. *Baravkar, S. R., Borde, M. R., & Nivangune, M. K. (2013). Android text messaging application for visually impaired people. IRACST-Engineering Science and Technology: An International Journal,3(1), 58–61.*
11. *Swamy, S. P., & Gatwadi, M. (2014). Indian sign language interpreter with android implementation. International Journal of Computer Applications,97(13), 36–41.*
12. *Jadhav, A., & Patil, A. (2012). Android Speech to Text converter for SMS application. IOSR Journal of Engineering,2(3), 420–423. https://doi.org/10.5120/17067-7484.*
13. *Mache, S. R., Baheti, M. R., & Mahender, C. M. (2015). Review on text-to-speech. International Journal of Advanced Research in Computer and Communication,4(8), 54–59. https://doi.org/10.17148/IJARCCE.2015.4812.*
14. *Malgaonkar, S., Shah, P., Panchal, D., & Pradhan, S. (2016). AWAZ: A bridge between android phones and the visually impaired. International Journal of Computer Applications,134(1), 42–47. https://doi.org/10.5120/ijca2016907793.*
15. *Pampattiwar, S. R., & Chhangani, A. Z. (2014). Smartphone accessibility application for visually impaired. International Journal of Research in Advent Technology,2(4), 377–380.*
16. *Aher, A., Musale, K., Pagar, S., & Morwal, S. (2014). Implementation of a smart mobile app for the blind and deaf person using morse code. International Journal of Research in Advent Technology,2(2), 151–154.*
17. *El Gayyar, M., Yamany, H. F., Gaber, T., & Hassanien, A. E. (2013) Social network framework for deaf and blind people based on cloud computing. Proceedings of the 2013 Federated conference on computer science and information systems, Krakow, Poland, IEEE Xplore, Electronic ISBN: 978-83-60810-52-1 (pp. 1301–1307).*
18. *Yue, W. S., Zin, N. A. M. (2013). Voice recognition and visualization mobile apps game for training and teaching hearing handicapped children. In The 4th international conference on electrical engineering and informatics, Selangor, Malaysia (Vol. 11, pp. 479–486). https://doi.org/10.1016/j.protcy.2013.12.218.*
19. *Mukhopadhyay, D., & Kulkarni, S. (2015). MATHWAY: Offering a mathematical app for visually challenged users. In 4th international conference on advances in computing, communication, and control, Mumbai (pp. 1–6).*
20. *Hussain, A., Jomhari, N., Kamal, F. M., & Mohamad, N. (2014). making: Modeling mobile learning game to recite Quran for deaf children. International Journal on Islamic Application in Computer Science and Technology,2(2), 8–15.*
21. *Fakrudeen, M., Yousef, S., & Miraz, M. H. (2014). The finger-based technique(FBT): An innovative system for improved usability for blind users' dynamic interaction with mobile touch screen devices. In Proceedings of the World Congress on Engineering, London, UK (Vol. 1, pp. 128–133). https://doi.org/10.13140/2.1.4190.7529.*
22. *Patel, P. (2014). Why is android not accessible and what is accessibility anyway? https://www.linkedin.com. Accessed on November 09, 2016.*
23. *Hofstader, C. (2014). Testing android accessibility: I give up. http://chrishofstader.com. Accessed on November 09, 2016.*
24. *Sharma, D., & Kanwar, R. (2015). Text To Speech conversion with language translator under android environment. International Journal of Emerging Research in Management and Technology,4(6), 96–100.*
25. *Vithani, T., & Kumar, A. (2014). Modeling the mobile application development lifecycle. In Proceedings of the international multiconference of engineers and computer scientists, Hong Kong (pp. 596–600). ISBN: 978-988-19252-5-1, ISSN: 2078-0966 (Online).*