A Literature Review on Detection of Plant Diseases

***Prof. A. R. Bhagat Patil*[1], *Lokesh Sharma*[2], *Nishant Aochar*[3], *Rajat Gaidhane*[4], *Vikas Sawarkar*[5], *Dr Punit Fulzele*[6], *Dr. Gaurav Mishra*[7]**

[1]*Dept of Computer Technology, Yeshwantrao Chavan College of Engineering (YCCE), Nagpur*
[2]*Dept of Computer Technology, Yeshwantrao Chavan College of Engineering (YCCE), Nagpur*
[3]*Dept of Computer Technology, Yeshwantrao Chavan College of Engineering (YCCE), Nagpur*
[4]*Dept of Computer Technology, Yeshwantrao Chavan College of Engineering (YCCE), Nagpur*
[5]*Dept of Computer Technology, Yeshwantrao Chavan College of Engineering (YCCE), Nagpur*
[6]*Asso Prof, Dept of Pediatric Dentistry, Sharad Pawar Dental College, Datta Meghe Institute of Medical Sciences, Sawangi (M), Wardha*
[7]*Professor Department of Radiodiagnosis, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences, Sawangi (M), Wardha.*

Corresponding Author:
Dr Punit Fulzele, Asso. Prof, Dept of Pediatric Dentistry, Sharad Pawar Dental College, Datta Meghe Institute of medical Sciences, Sawangi (M), Wardha
Punitr007@gmail.com, 9890417646

Conflict of Interest: None
Funding: NewGen Innovation and Entrepreneurship Development Centre, National Science and Technology Entrepreneurship Development Board (NSTEDB), Department of Science & Technology (DST), Government of India

Abstract:
With increase in population the need for food is on rise, in such circumstances, plant diseases prove to be a major threat to agricultural produce and result in disastrous consequences for farmers. Early detection of plant disease can help in ensuring food security and controlling financial losses. The images of diseased plants can be used to identify the diseases. Classification abilities of Convolutional Neural Networks are used to obtain reliable output. Google's pretrained model 'Inception v3' is used. The Inception v3 model is trained over a dataset of diseased plants obtained from 'Plant Village Dataset'. The developed detection approach is evaluated on measures of F1 score, precision and recall.

Keywords: - Plant disease, image processing, image acquisition, segmentation, feature extraction, classification.

I. INTRODUCTION

Plant diseases are one of major reasons behind the production and economic losses in agriculture. Identifying disease correctly is a challenging task and requires expertise. Frequently the illnesses or its signs like colored spots or streaks can be observed on the plant leaves. The plants diseases is usually caused by microbes including fungi, bacteria, and viruses. There is a wide spectrum of signs and symptoms which differ because of the cause or etiology of the plant disease.

The neural networks have been an emerging application in numerous and diverse areas as examples of end to end learning. The nodes in a neural network are mathematical functions that take numerical inputs from the incoming edges and provide a numerical output as an outgoing edge.

The CNN may holds its applications in the agricultural field including identification of diseases and also to quantify the diseased area. Usually the diseases are identified by naked eye observation by an expert. This method involves huge time in vast farms or land. The use of convolutional neural network in recognition and detection of plant diseases early will be effective to increase the quality of products.

To develop such a precise image classifier aimed at diagnosis of diseases of plant, we need a large, processed and verified dataset containing various diseased and healthy plant images.
'Plant Village' project has collected thousands of plant images and made it open and free to use. The dataset is already processed and is available in three versions as Colored, gray scale and segmented.

II. LITERATURE SURVEY

In the paper "**Deep learning for Image-Based Plant detection" [1]** the authors Prasanna Mohanty et al**.,** has proposed an approach to detect disease in plants by training a convolutional neural network. The CNN model is trained to identify healthy and diseased plants of 14 species. The model achieved an accuracy of
99.35% on test set data. When using the model on images procured from trusted online sources, the model achieves an accuracy of 31.4%, while this is better than a simple model of random selection, a more diverse set of training data can aid to increase the accuracy. Also some other variations of model or neural network training may yield higher accuracy, thus paving path for making plant disease detection easily available to everyone.

Malvika Ranjan et al. in the paper "**Detection and Classification of leaf disease using Artificial Neural Network"** proposed an approach to detect diseases in plant utilizing the captured image of the diseased leaf. Artificial Neural Network (ANN) is trained by properly choosing feature values to distinguish diseased plants and healthy samples. The ANN model achieves an accuracy of 80%.

According to paper "**Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features" [3]** by S. Arivazhagan, disease identification process includes four main steps as follows: first, a color transformation structure is taken for the input RGB image, and then by means of a specific threshold value, the green pixels are detected and uninvolved, which is followed by segmentation process, and for obtaining beneficial segments the texture statistics are computed. At last, classifier is used for the features that are extracted to classify the disease..

Kulkarni et al. in the paper "**Applying image processing technique to detect plant diseases" [4],** a methodology for early and accurately plant diseases detection, using artificial neural network (ANN) and diverse image processing techniques. As the proposed approach is based on ANN classifier for classification and Gabor filter for feature extraction, it gives better results with a recognition rate of up to 91%.

In paper "**Plant disease detection using CNN and GAN" [5],** by Emaneul Cortes, an approach to detect plant disease using Generative Adversarial networks has been proposed. Background segmentation is used for ensuring proper feature extraction and output mapping. It is seen that using Gans may hold promise to classify diseases in plants, however segmenting based on background did not improve accuracy.

In the paper "**Convolutional Neural Network based Inception v3 Model for Animal Classification**" **[6]**, Jyotsna Bankar et al. have proposed use of inception v3 model in classifying animals in different species. Inception v3 can be used to classify objects as well as to categorize them, this capability of inception v3 makes it instrumental in various image classifiers.

III. PROPOSED METHOD

- Dataset Classification
- Building the CNN using transfer learning
- Training our Network
- Testing

1) Dataset Classification

Selection of proper set of images for training of model is a significant task.

Centroid of each image is calculated to retrieve select images. Centroid can be calculated by use of contours.

Contour is a curve that joins all the points along the periphery of a shape. Contours can much be detected much precisely on binary images. Hence, every image has to be converted to grayscale with a threshold applied on it. 'find Contours' function can be use for this purpose. Three arguments provided to this function are Source image, contour retrieval mode and contour approximation method. Output of the function contains the images, contours and hierarchy. Output contains all contours in the image. Every contour is array of (x,y) coordinates of boundary points. Contour Approximation Method is used to specify the coordinates to be stored. CHAIN_APPROX_NONE stores all boundary points. But all boundary points are not required. Because for finding contour of a straight line all points are not needed only two points are sufficient. CHAIN_APPROX_SIMPLE provides this kind of output by eliminating all redundant points and compressing the contour[7].

Having found the contours, the image moments are calculated. Image moments are used to calculate the centre of mass or the centroid of an object. The function cv2. Moments return a dictionary of all moment values. From this moments one can extract data such as centroid, area, etc. As we only need centroid of the image, it is given by the relations,

Cx = (M["m10"]/M["m00"]) and
Cy = (M["m01"] / M["m00"])

here, M is the dictionary of moments.

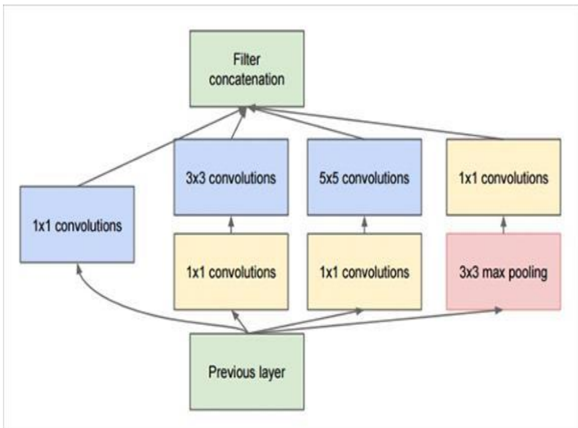
After calculating the centroid of each image using the above method, we set a definite range for (x, y) coordinates after overlooking all the centroids. The images falling between the range are selected for further processing.

2) Building the CNN using transfer learning

Image identification has become feasible with the advent of Convolutional Neural Networks. But designing a CNN that identifies objects and classifies them into distinct classes is a complex task. By making use of transfer learning it can be simplified. In transfer learning we have trained our model that has been trained on Plant Village dataset using 12GB TESLA k80 GPU. Also Transfer learning significantly reduces training time and gives much better performance for relatively small dataset [8].

Google has released pertained models on tensor flow's official website. **"Inception v3"** is one of such models that is trained on ImageNet dataset and can identify 1000 classes such as television, keyboard, car and some animals. It is one of the most widely used model for image classification[9].

The Inception v3 network is 48 layers deep and has an input image size of 299 by 299. The network takes image as input and gives label as output. The characteristic of Inception v3 is factorization. The purpose of factorizing convolutions is to decrease the parameters and connections while retaining the efficiency of the network.



a. Factorizing Convolutions

By means of factorizing convolutions the no. of connections and parameters are reduced to a considerable degree without adversly affecting the efficiency of the system.

Factorization can be into smaller convolutions such as, two 3 by 3 convolutions replace one 5 by 5 convolution; or assymetric convolutions such as 3 by
1 convolution followed by 1 by 3 replaces 3 by 3 convolution.

b. Auxilary classifier

In Inception-v3, auxiliary classifier is used as regularizer. Batch normalization, introduced in Inception v2, is also used in the auxiliary classifier.

c. Efficient Grid Size Reduction

Usually feature map downsizing is done by max pooling. But the approach either tends to be too greedy or too expensive. In inception v3 320 feature maps are obatined by max pooling and these are concatenated to obatin 640 feature maps. Efficient grid size reduction in Inception v3 produces inexpensive yet effective network. *3) Training the network*

The deep convolutional model can be used to classify labels specific to the task at hand. For this the Inception v3 model is loaded. New classes to be recognised are specified and Inception v3 model is trained over different batches for certain number of epoches, thus harnessing the image classifying abilities of Inception v3 to classify diseased plants.

Retrain script is an important medium to generate custom image classifier using inception v3. It traines a new layer which performs task of classifying the custom specified classes. The script can be can be altered for parameters such as:- image_dir, intermediate_output_graphs_dir, output_graph, output_labels, ditortion feature, number of training steps(epoches), learning rate, etc.

The dataset contains labelled folders with images, the path to these folders is provided to the retrain script. A portion of these images is retained for testing.

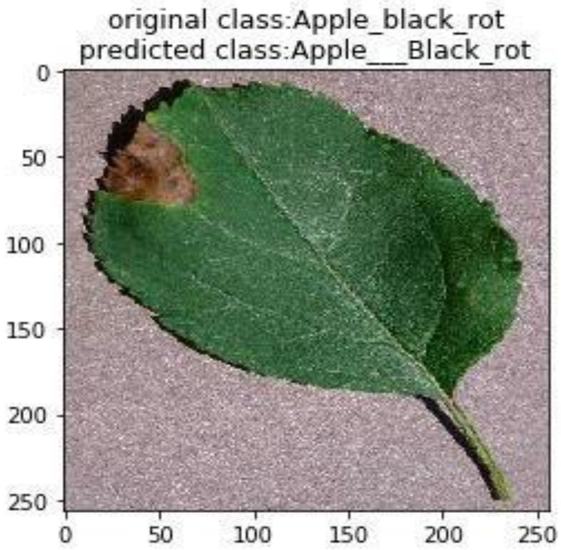
The Inception v3 model is iteratively trained over differrent batches for certain number of epochs. Labels for each of the disease is provided to the network along with the images

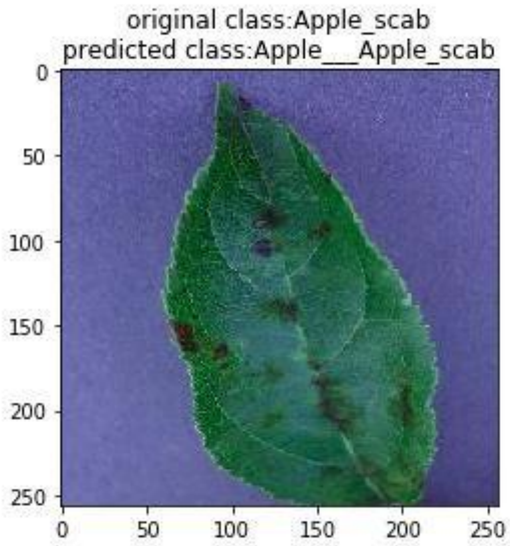
ISSN 2515-8260 Volume 7, Issue 07, 2020

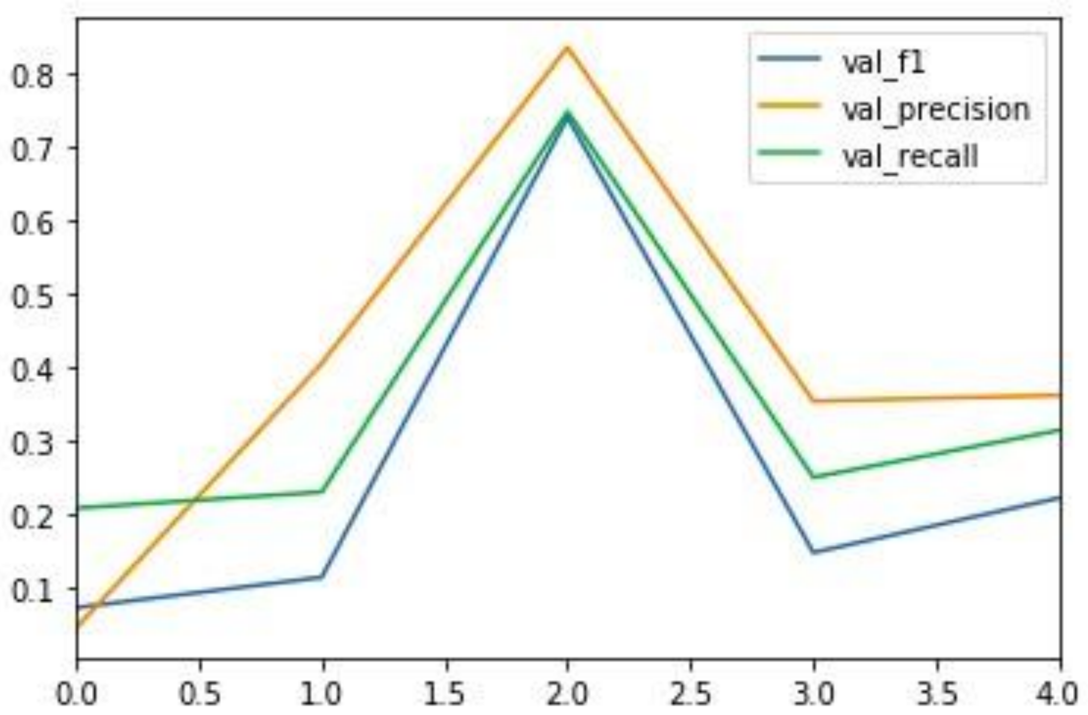
belonging to the label. A separate set of images is kept aside for testing. Callback functions are used to obtain statistics during model traning. Parameters such as loss, validation loss and accuracy are obtained with callbacks. We use these values to monitor the performance of the model through different epoches. Callbacks enable us to interact with the model during the training of model. Callbacks can also be used to inhibit training after a certain level of required effeciency is achieved to stop model from overfitting.

4) Testing

The trained model is tested on a set of images. Random images are introduced to the network and output label is compared to the original known label of the image. Parameters used for evaluation are F1 score, precision and recall. Precision is the proportion of predicted positives that are truly positives. Recall gives the proportion of actual positives correctively classified. F1 score helps in maintaining a balance between precision and recall. Evidences for needs of such innovations are event from GBD surveys[10-17]. A number of studies on related aspects have been reported by different authors[18-27].






val_f1
val_precision
val_recall
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0

IV. CONCLUSION

This paper proposes a CNN based method for plant disease classification using the leaves of diseased plants. Building such a neural network with high efficiency is a complex task. Transfer learning can be employed to achieve greater efficiency . Inception v3 is one of the models available that inherently have the capability to classify images and further can be trained to identify different classes. Thus, use of Inception v3 can play key role in obtaining fast and effective plant disease identifiers. Also by dataset classification using contour method, the training set can be chosen to ensure proper training of model for all features. This provides better feature extraction than randomly classifying the dataset. Optimal results were obtained by employing the methods specified in the paper. Thus, with implementation and use of these methods for plant disease classification losses in agriculture can be reduced.

References:

[1] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7(September), [1419].https://doi.org/10.3389/fpls.2016.01419

[2] Ser Serawork Wallelign, Mihai Polceanu, Cédric Buche. Soybean Plant Disease Identification Using Convolutional Neural Network. FLAIRS-31, May 2018, Melbourne, United States. https://hal.archives-ouvertes.fr/hal-01807760

[3] S.Arivazhagan, R. Newlin Shebiah, S.Ananthi, S.Vishnu Varthini. 2013. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agric Eng Int: CIGR Journal.

[4] Huu Quan Cap, Katsumasa Suwa, Erika Fujita, Satoshi Kagiwada, Hiroyuki Uga, and Hitoshi Iyatomi. A deep learning approach for on-site plant leaf detection. 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA).

[5] Santhosh Kumar S, and B. K. Raghavendra. Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review. 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)

[6] Jyotsna Bankar, and Nitin R Gavai. Convolutional Neural Network based Inception v3 Model for Animal Classification. International Journal of Advanced Research in Computer and communication Engineering, May 2018

[7] Abirami Devaraj, Karunya Rathan, Sarvepalli Jaahnavi, and K Indira. Identification of Plant Disease using Image Processing Technique. 2019 International Conference on Communication and Signal Processing (ICCSP).

[8] Mercelin Francis, and C. Deisy. Disease Detection and Classification in Agricultural Plants Using Convolutional Neural Networks — A Visual Understanding. 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN).

[9] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, "A review of advanced techniques for detecting plant diseases," Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1–13, 2010.

[10] James, Spencer L, Chris D Castle, Zachary V Dingels, Jack T Fox, Erin B Hamilton, Zichen Liu, Nicholas L S Roberts, et al. "Estimating Global Injuries Morbidity and Mortality: Methods and Data Used in the Global Burden of Disease 2017 Study." *Injury Prevention* 26, no. Supp 1 (October 2020): i125–53. https://doi.org/10.1136/injuryprev-2019-043531.

[11] James, Spencer L, Chris D Castle, Zachary V Dingels, Jack T Fox, Erin B Hamilton, Zichen Liu, Nicholas L S Roberts, et al. "Global Injury Morbidity and Mortality from 1990 to 2017: Results from the Global Burden of Disease Study 2017." *Injury Prevention* 26, no. Supp 1 (October 2020): i96–114. https://doi.org/10.1136/injuryprev-2019-043494.

[12] Murray, Christopher J L, Cristiana Abbafati, Kaja M Abbas, Mohammad Abbasi, Mohsen Abbasi-Kangevari, Foad Abd-Allah, Mohammad Abdollahi, et al. "Five Insights from the Global Burden of Disease Study 2019." *The Lancet* 396, no. 10258 (October 2020): 1135–59. https://doi.org/10.1016/S0140-6736(20)31404-5.

[13] Murray, Christopher J L, Aleksandr Y Aravkin, Peng Zheng, Cristiana Abbafati, Kaja M Abbas, Mohsen Abbasi-Kangevari, Foad Abd-Allah, et al. "Global Burden of 87 Risk Factors in 204 Countries and Territories, 1990–2019: A Systematic Analysis for the Global Burden of Disease Study 2019." *The Lancet* 396, no. 10258 (October 2020): 1223–49. https://doi.org/10.1016/S0140-6736(20)30752-2.

[14] Vos, Theo, Stephen S Lim, Cristiana Abbafati, Kaja M Abbas, Mohammad Abbasi, Mitra Abbasifard, Mohsen Abbasi-Kangevari, et al. "Global Burden of 369 Diseases and Injuries in 204 Countries and Territories, 1990–2019: A Systematic Analysis for the Global Burden of Disease Study 2019." *The Lancet* 396, no. 10258 (October 2020): 1204–22. https://doi.org/10.1016/S0140-6736(20)30925-9.

[15] Wang, Haidong, Kaja M Abbas, Mitra Abbasifard, Mohsen Abbasi-Kangevari, Hedayat Abbastabar, Foad Abd-Allah, Ahmed Abdelalim, et al. "Global Age-Sex-Specific Fertility, Mortality, Healthy Life Expectancy (HALE), and Population Estimates in 204 Countries and Territories, 1950–2019: A Comprehensive Demographic Analysis for the Global Burden of Disease Study 2019." *The Lancet* 396, no. 10258 (October 2020): 1160–1203. https://doi.org/10.1016/S0140-6736(20)30977-6.

[16] Kinyoki DK, Ross JM, Lazzar-Atwood A, Munro SB, Schaeffer LE, Abbasalizad-Farhangi M, et al. Mapping local patterns of childhood overweight and wasting in low- and middle-income countries between 2000 and 2017. Nat Med 2020;26(5):750-759.

[17] Lozano R, Fullman N, Mumford JE, Knight M, Barthelemy CM, Abbafati C, et al. Measuring universal health coverage based on an index of effective coverage of health services in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020.

[18] Gaidhane, A., A. Sinha, M. Khatib, P. Simkhada, P. Behere, D. Saxena, B. Unnikrishnan, M. Khatib, M. Ahmed, and Q.S. Zahiruddin. "A Systematic Review on Effect of Electronic Media on Diet, Exercise, and Sexual Activity among Adolescents." *Indian Journal of Community Medicine* 43, no. 5 (2018): S56–65. https://doi.org/10.4103/ijcm.IJCM_143_18.

[19] Gaidhane, S., N. Khatib, Q.S. Zahiruddin, A. Gaidhane, S. Telrandhe, and P. Godhiwal. "Depression, Anxiety and Stress among the General Population in the Time of COVID-19 Lockdown: A Cross-Sectional Study Protocol." *International Journal of Research in Pharmaceutical Sciences* 11, no. Special Issue 1 (2020): 360–64. https://doi.org/10.26452/ijrps.v11iSPL1.2726.

[20] Khatib, M.N., A. Gaidhane, S. Gaidhane, and Z.S. Quazi. "Ghrelin as a Promising Therapeutic Option for Cancer Cachexia." *Cellular Physiology and Biochemistry* 48, no. 5 (2018): 2172–88. https://doi.org/10.1159/000492559.

[21] Khatib, M.N., S. Gaidhane, M. Khatib, M. Ahmed, A. Gaidhane, and Z.Q. Syed. "SARS-CoV and SARS-CoV-2: Similar Viruses with Different Trajectories." *Wutan Huatan Jisuan Jishu* 16, no. 5 (2020): 544–48.

[22] Khatib, M.N., A.H. Shankar, R. Kirubakaran, A. Gaidhane, S. Gaidhane, P. Simkhada, and S.Z. Quazi. "Ghrelin for the Management of Cachexia Associated with Cancer." *Cochrane Database of Systematic Reviews* 2018, no. 2 (2018). https://doi.org/10.1002/14651858.CD012229.pub2.

[23] Zodpey, S., A. Sharma, Q.S. Zahiruddin, A. Gaidhane, and S. Shrikhande. "Allopathic Doctors in India: Estimates, Norms and Projections." *Journal of Health Management* 20, no. 2 (2018): 151–63. https://doi.org/10.1177/0972063418763651.

[24] Khatib MN, Kirubakaran R, Gaidhane S, Shankar AH, Quazi Syed Z. Yoga for improving functional capacity, quality of life and cardiovascular outcomes in people with heart failure. Cochrane Database of Systematic Reviews [Internet]. 2017;2017(7). https://doi.org/10.1002/14651858.CD012015.pub2.

[25] Puri S, Fernandez S, Puranik A, Anand D, Gaidhane A, Quazi Syed Z, et al. Policy content and stakeholder network analysis for infant and young child feeding in India. BMC Public Health [Internet]. 2017;17. https://doi.org/10.1186/s12889-017-4339-z.

[26] Uddin S, Mahmood H, Senarath U, Zahiruddin Q, Karn S, Rasheed S, et al. Analysis of stakeholders networks of infant and young child nutrition programmes in Sri Lanka, India, Nepal, Bangladesh and Pakistan. BMC Public Health [Internet]. 2017;17. https://doi.org/10.1186/s12889-017-4337-1.

[27] Regmi PR, van Teijlingen E, Mahato P, Aryal N, Jadhav N, Simkhada P, et al. The health of nepali migrants in India: A qualitative study of lifestyles and risks. International Journal of Environmental Research and Public Health [Internet]. 2019;16(19). https://doi.org/10.3390/ijerph16193655.