IV. Data considerations and compliances

For the development of this algorithm, we need ARMD affected patient's data. Datasets consist of various personal information such as age, gender, race etc., biological information such as height, weight etc., habitual information such as diet, smoker, drinker etc. and eye images. We add similar information of normal people or people that may be tested negative for ARMD to the datasets to increase the reliability of Bayesian Network. While predicting the probability the sample space needs to be comprehensive and random. Fortunately, over the time hospitals have generated enough data and in clinical practices, it is common to collect the information mentioned above. As an alternate, the organizations that conduct clinical studies or support clinical studies by providing Statistical Analyzing Services (SAS) possess large and comprehensive datasets. Since our work involves an automated learning, it can utilize multiple formats of data (like csv, SAS etc.) [11].

Datasets are large, rich and varied. We emphasize large datasets because machine learning algorithms make use of the rich, varied datasets and relate them by finding high dimensional interactions among datasets [10]. Sample sizes are large enough to include variety of data consisting various information about the AMD affected and the unaffected. Automated learning in the algorithm requires a bit more data to overcome missing data problem in datasets while constructing Bayesian Network [11]. When it comes to applications that require more data, security considerations are high. As a measure the following steps are taken. 1) Data are anonymous 2) No prospect data collection from patients 3) Only necessary data are collected 4) Encrypted and secure Storage 5) Destroy data when no longer needed 6) Compliance with clinical terminologies 7) Compliance with security, data transfer standards and laws that vary country to country.

V. Conclusion

Through this work we expressed an algorithm for early detection of ARMD that makes the best use of Bayesian Networks with an accuracy of over 90% through our initial training sets. There are very less literature available in the area of machine learning when it comes to clinical and biological solutions. On that interest, we still have a very long way to go. It is our aim to study, analyze and provide more exploratory solutions especially for clinical problems in our future works. We continue to refine the proposed algorithm and the system in our future works. The outcomes of the system are to be carefully studied. Usually machine learning systems get better as time progresses and so the outcomes are more accurate with time. Also, we continue to study ARMD and how machine learning could be equipped to provide better solutions for biological and clinical problems.

Acknowledgement

The Authors thank the Management and the Head of the Department of CHRIST (Deemed to be University), Bangalore for permitting and supporting us to carry out the research. We would also like to thank Mr. Suriya Prakash B, Software Engineer and Mr. Ullas, Little Flower Hospital, for helping us in our research.

References

[1] Bressler NM, "Age-related macular degeneration is the leading cause of blindness" JAMA 2004; 291(15):1900-1901 DOI:10.1001/jama.291.15.1900.

[2] Global Burden of disease Study 2013 Collaborators "Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013", The Lancet. 386 (9995): 743–800. August 2015, DOI: 10.1016/S0140-6736(15)60692-4.

[3] GBD 2015 Disease and Injury Incidence and Prevalence Collaborators (October 2016), "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015", The Lancet. 388 (10053):1545–1602, DOI:https://doi.org/10.1016/S0140-6736(16)31678-6, PMID 27733282.

[4] Mehta S, "Age-Related Macular Degeneration", Prim Care, 2015 September;42(3):377-91, DOI: 10.1016/j.pop.2015.05.009, PMID 26319344.

[5] Ratnapriya, R, and E Y Chew, "Age-Related Macular Degeneration– Clinical Review and Genetics Update", Clinical Genetics 2013 August;84(2):160-6, DOI: 10.1111/cge.12206 PMC. Web. 22 Apr. 2016.

[6] P. Burlina, D.E.Freund, N.Joshi, Y.Wolfson, N.M.Bressler, "Detection of Age-related Macular Degeneration via Deep Learning", IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2016 - Proceedings (Vol. 2016-June, pp. 184-188). [7493240] IEEE Computer Society, DOI: https://doi.org/10.1109/ISBI.2016.7493240.

[7] Liang Z, Wong DW, Liu J, Chan KL, Wong TY, "Towards automatic detection of age-related macular degeneration in retinal fundus image", 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, August 31 - September4,2010,**DOI:** 10.1109/IEMBS.2010.5627289.

[8] S Albert Jerome, Vyshakh Asokan, "Computer Aided Approach for Detection of Age Related Macular Degeneration from Retinal Fundus Images", 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT] March 2016, DOI:10.1109/ICCPCT.2016.7530348.

[9] Murphy K. (1998), "A Brief Introduction to Graphical Models and Bayesian Networks".

[10] Kevin P Murthy, "Machine Learning: A Probabilistic Perspective", Adaptive Computation and Machine Learning Series. Cambridge, Massachusetts, MIT Press; 2012.

[11] Michal Horný, "Bayesian Networks", Boston University, April 18, 2014.

[12] Lucas P, "Bayesian Networks in Medicine: a Model-based Approach to Medical Decision Making", University of Aberdeen, December 2001.

[13] Jiebo Luoa, Andreas E. Savakisb, Amit Singhal, "A Bayesian network based framework for semantic image understanding", Pattern Recognition 38 (2005), pp:919 – 934.