EMD and Horizontal Visibility Graph Based Disease Tagging for Covid-Positive Chest Radiographs

Niranjan Chavan, Priya Ranjan, Kumar Dron Shrivastav, and Rajiv Janardhanan

Abstract This work describes preliminary steps in the ongoing implementation of horizontal visibility graphs (HVG) and related Hamming-Ipsen-Mikhailov (HIM) network similarity (distance) metric to provide automatic disease tag for normal and COVID-positive chest radiographs. A detailed exploration in transformation of a normal or COVID positive chest radiograph to a horizontal visibility graph and its network/graph-theoretic analysis and visualization in R computational environment is presented. Further, HIM network similarity metric is illustrated and its usage in generating automatic disease tag based on test radiograph's HIM-distance from healthy and COVID positive representative radiographs is presented. Finally, statistically success rate of 60% is observed despite of low quality and mismatched (Normal and COVID positive radiographs are not from same patients) using HVG–HIM and 30% using EMD, which augurs well for the development of this system as a quick disease tag device. Difference in drastic performance is owing to serious computational investment in HVG–HIM. A webservice based portal for automated diseases tagging of chest radiograph is proposed, built and illustrated to take basic clinical services to the poorest of the poor in the LMICs and in African countries. It can be used by primary health care centers (PHCs) for a first aid scan and then patient can be referred to specialists. On a macro scale where patients overwhelm medical

P. Ranjan
Bhubaneswar Institute of Technology, Info Valley, Harapur, Odisha, India

N. Chavan
Institute for Thermal Energy Technology and Safety (ITES), Karlsruher Institute for Technology, 76344 Karlsruhe, Germany
e-mail: niranjan.chavan@kit.edu

R. Janardhanan (✉)
SRM Institute of Science and Technology, Tamil Nadu, SRM Nagar, Kattankulathur-603203, Chengalpattu District, India
e-mail: rajiv@srmist.edu.in

K. D. Shrivastav
Health Data Analytics and Visualization Environment, Amity Institute of Public Health, Amity University Uttar Pradesh, Sector 125, Noida, India
e-mail: kdshrivastav@amity.edu

S. Adibi et al. (eds.), *The Science behind the COVID Pandemic and Healthcare Technology Solutions*, Springer Series on Bio- and Neurosystems 15,
https://doi.org/10.1007/978-3-031-10031-4_14

facilities due to astronomical numbers involved, this kind of system can relieve the suffering of humanity to some extent. We also reflect on our programmatic and static computational approach as compared to nonlinear dynamical and often unstable, energy-hogging deep learning.

Keywords COVID-29 · Earth Mover's Distance (EMD) · Horizontal visibility graphs (HVG) · Chest radio-graphs

1 Introduction

As the healthcare systems across the world grapple with the grim realities of changed landscape of disease burden in the post-covid era, it is becoming imperative to develop fruganomic healthcare technologies capable of gathering evidence from the community to augment the hospital based registries in providing reasonably accurate estimates of the disease burden which will definitely help the health policy administrators to get a unequivocal narrative in a nuanced manner to develop niche specific disease surveillance and forecasting systems. This will indeed help the Healthcare administrators to develop and deploy novel, affordable and more importantly accessible instruments such as policies and programmes aimed at evaluating the health status of the community at large particularly in resource limited healthcare systems prevalent in low-and lower-middle-income countries (LLMICs) such as the Indian sub-continent. India endowed with unique geological relief structures and divergent genetic base provides a unique landscape of disease burden necessitating the need to develop tools capable of being ported into mobile platforms (iOS/HTC/Android), since mobiles have a good penetration in the rural milieus of the Indian sub-continent. This is further compounded by the fact that Non-communicable Diseases (NCDs) disproportionately affect people living LLMICs [1–3], accounting for three quarters of the mortalities within LLMICs [4]. The relationship among NCDs, poverty, social and economic development [5], is likely to pose a major challenge to development as well as attainment of Sustainable Development Goals (SDGs) by 2030 [6, 7]. The marginalized sections of the society in LLMICs are vulnerable to NCDs for many reasons, including socio-economic constraints, psychosocial stress, higher levels of risk behavior, unhealthy living conditions, limited access to high-quality health care along with reduced opportunity to prevent complications. The prevalence of unhealthy risk behaviors such as consumption of tobacco and alcohol products along with sedentary lifestyle will make these population vulnerable to the ravages of "NCD Epidemic" which has been hitherto underacknowledged and unaddressed until the advent of COVID pandemic. The fact that India is projected to experience more deaths from NCDs than any other country over the next decade, primarily due to the size of the population and worsening risk factor profile will significantly impact economic growth. The deeply entrenched social and economic disparities, with lack of affordable and accessible healthcare presents a pre-emptive scenario for

the emergence of several epidemics/ pandemics such as COVID-19 with devastating consequences.

In the last two decades, the translation of the fundamental concepts of precision medicine at a community level to understand the patterns and processes associated with the landscape of disease burden in the LLMICs having fractious and fractionated health-care ecosystems such as the Indian sub-continent necessitates the need to develop novel, cutting edge, and disruptive fruganomic community empowering solutions aimed alleviating the healthcare disparities. In the last two decades India has witnessed an epidemiological transition from communicable diseases to NCDs in the last two decades with cardiovascular diseases and Cancer accounting for a significant proportion of morbidities and mortalities, the un-finished agenda of communicable diseases has led to emergence of recent pandemics such as COVID-19 [8–10]. Although the World Health Organization (WHO) has recognized the outbreak of COVID-19 in January 2020 and declared it as a pandemic in March 2020, the statutory impact on the economy of all the countries including India has clearly affected the health outcomes of the populace belonging to the weaker socio-economic strata.

COVID-19 associated co-morbidity was observed in patients who had underlying risk factors of hypertension, diabetes, and chronic respiratory problems. Chronic respiratory problems account for 8% of the mortalities and India has 18% of the global population with an ever-increasing burden of chronic respiratory diseases including chronic obstructive pulmonary disease (COPD), asthma, pneumoconiosis, interstitial lung diseases, and pulmonary sarcoidosis [11–13]. Environmental factors such as air pollution, water pollution, and soil pollution to name a few are known to significantly contribute to premature mortality and disease burden globally, with the highest impact in low-income and middle-income countries such as the Indian subcontinent endowed with resource limited healthcare systems [14, 15].

Recent evidences suggest that use of multi-modal multi sensor fusion technologies along with big data enabled platform, would significantly contribute towards the strengthening resource-deprived healthcare systems prevalent in the Indian subcontinent. The technology (artificial intelligence AI) enabled transition of precision medicine to precision public health must be integrated into the existing framework of healthcare systems with a view administer the provision of affordable and accessible healthcare solutions intrinsic to the niche specific needs of LLMICs. The integration of the disruptive and cutting-edge healthcare solutions within the framework of the existing healthcare systems will significantly improve the health outcomes of the community at large.

A major drawback of the existing interpretation algorithms based on Artificial Neural Networks (ANN) is its black-box nature, which coupled with increased computational complexity leads to increased carbon foot-printing and thereby global warming [16, 17] but solutions are also emerging [18]. Apart from this, the process of obtaining a result is also difficult to understand as to why and how it arrived at the answer [19–21]. Further, the nonlinear dynamical behavior of deep neural networks is prone to chaotic nature and fundamental underlying unpredictability [22–24]. On the other hand, static and predictable algorithms like, Earth Movers

Distance (EMDs) and Visibility Graph perform image match by computing perceptual similarity and provide more meaningful and interpretable solutions to matching problems. Taken together, ANNs have a long and very well researched history of inherent instability and its auto-mated decisions can't be entrusted to make decisions critical to the survival of a patient afflicted with a severe case of a COVID-19 lung ailment [25–28].

Chest radiographs are still most common modality for diagnosing lung disease conditions and the development of tools and applications that can seamlessly evaluate lung health of LMICs such as India will significantly augment healthcare outcomes. Additionally, the lack of well-structured databases for referencing and analysis hinders the progression of research from aiding and optimizing processes and clinical decision making with the help of Artificial Intelligence (AI) [29, 30].

India endowed with diverse genetic base and socio-cultural norms, presents a unique landscape of disease burden necessitating the need for niche specific databases for enhancing the accuracy of AI-enabled tools. The socioeconomic impact and benefits of AI based automation of Chest Radiograph analysis for LMICs like India will significantly improve the clinical outcomes of patients afflicted with lung diseases and outweigh the challenges leading to its integration to the existing framework of the healthcare system [31, 23]. The potential application of AI-enabled platforms would provide a valuable, precision public health tool for better management of lung disease epidemic by improving the clinical outcomes thereby alleviating a significant burden on the national health spend.

The fundamental impact of integrating smart clinical devices, IoT, and Industry 4.0 with clinical software and closed-loop resource allocation is the ability to rapidly deploy medical infrastructure in challenging places during natural calamities like flood, earthquake, drought, Tsunamis etc., while drastically reducing costs and reacting to demands in patients' preferences, pharma industry changes, the supply chain, and technology upgrades.

Broadly, this work makes following contributions in moving towards clinical dis-ease tagging as a webservice:

- This work describes a natural approach for automated Covid positive chest radiograph tagging using computational ideas like perceptual similarity, Earth Mover's Distance (EMD) and converting the chest radiographs into a network/graph using horizontal visibility graph procedure and then computing similarity scores by HIM network distance metric. From all perspectives, it is a first work of its kind whose time has arrived due to Covid clinical and hence socio-economic emergency.
- This disease tagging is being presented as an app on any mobile platform of choice where all user/patient/doctor/medical professional must do is to upload their Covid chest radiograph and system will generate disease tag. This is possible due to advent of affordable smartphone technology and its accessibility across socio-economic spectrums.
- Same algorithmic ideas with intense computational engine are interfaced as web service, where disease tagging can be done in large batches since one of the major issues in Covid waves, large numbers of people getting infected in a very short

time span. These is a clear need for such a system which can cope up with this kind of Covid-infection load in an agile fashion.

Rest of this chapter is organized as follows: Sect. 2 discusses related work in the context of HVG. Section 3 describes clinical and public health motivations behind deploying this technology. Section 4 presents basic mathematical definition of EMD. Section 5 describes HVG, HIM and related results. Section 6 illustrates different levels of processing in HVG for Covid positive chest radiographs. Section 7 reports on computational results from HVG–HIM implementation on chest radiographs for disease tagging. Section 8 demonstrates our basic implementations as a webservice for remote accessibilities in LMICs. Finally, Sect. 9 collects our insights and experiences in conclusion and suggests future directions for development.

2 Related Work

Motivation behind visibility graph generation was to develop simple and fast computational methods, which transform a time series into a network or a graph. This resulting visibility graph in turn inherits multiple features of the original timeseries in its spatial organization. For example, periodic timeseries transform into regular graphs, and random timeseries manifest themselves random graphs [32, 33]. Along these lines, horizontal visibility algorithm, a geometrically more intuitive and analytically tractable version of visibility graph algorithm, focusing on the transformation of timeseries into graphs [34], has been proposed. It turns out that, exact results on the topological properties of these horizontal visibility graphs, like, the degree distribution, the clustering coefficient, and the mean path length, can be obtained. The horizontal visibility algorithm can also be used as an intuitive method to discern between any two different timeseries. It is precisely this capability we leverage here to automatically categorize normal and Covid positive chest radiographs. HVG along with features like mean node degree and degree distribution has been used to categorize the sleep stages based on graph domain properties from a single-channel electroencephalogram (EEG) signal [17, 35].

Visibility graph methods have been found effective in describing the fractal properties of Geophysical time series [36]. The understanding of various graph-theoretical metrics pertaining to visibility graphs, their interdependent nature, and their sensitivity with respect to missing values and randomness are explored. Visibility graph algorithms have been applied to fMRI time series to simultaneously compute and process relevant dimensions of both local and global dynamics in a natural fashion, and to explore a transformation between time series and network theory in the context of network neuroscience [37]. It has been illustrated that the network architecture of the image visibility graphs represents important information on the organization of the image from which they are derived and potentially they can make good image filters [38]. Using HVG, a general class of predictors, which can be deployed to augment existing properties used in heart rate variability (HRV) analysis, and which

show high predictive power for multiple cardiovascular diseases, have been defined and validated [39]. Normalized weight vertical visibility algorithm (NWVVA) has been proposed to extract EMG-based features for myopathy and ALS detection [40]. In this algorithm, sampling points or nodes based on sampling theory are derived, and features are computed based on interrelations among the vertical visibility nodes with their amplitude differences as weights. The similarity graph algorithm are used to analyze the time series of motor activity, extracted from actigraph registrations over 12 days in depressed and schizophrenic patients. These were mapped into a graph and then techniques from graph theory were applied to describe these time series, searching for variations in complexity [41].

Visibility graph methods were deployed to analyze ECoG signals in rats [42]. Subsequently, typical metrics in network science (graph properties) were applied to compute network properties of topological structure of these graphs derived from ECoG signals. A family of Feigenbaum graphs, which are horizontal visibility graphs (HVGs) generated from the trajectories of one-parameter unimodal maps undergoing a period-doubling route to chaos (Feigenbaum scenario), have been analyzed [43]. It has been found that while the maximum eigenvalue of HVG can easily discern chaos from a white noise process, it is not a good metric to quantify the chaoticity of the process, and that the eigenvalue density is perhaps a better indicator for the same.

3 Motivation for Building This Tool and Methodology

This work is motivated by following two objectives.

1. Development and validation of an Intelligent Decision Support System for segregating Chest Radiographs to detect COVID-19 associated lung diseases in both tertiary care settings and extended community along with tracking of patients through low end mobile health applications [44–46].
2. Integration and validation of multi-modal tool in clinical practice involving automated processing of anonymized chest radiographs along with conventional molecular biomarkers [47, 48] of tissue hypoxia in both angiogenic and fibrotic phases of the lung disease progression forming the rationale of effective triage methods for prioritizing the most urgent conditions to wait listed ones.

The race and sex-specific variations in the levels of conventional biomarkers such as Angiogenesis/Fibrosis indeed necessitate the validation and confirmation by a non-invasive AI-enabled modality, which can seamlessly crunch a large amount of data in an affordable and accessible manner. Our fruganomic data intensive AI-enabled tool will not only facilitate the same by incorporating the clinical-epidemiological features of the subjects evaluated at tertiary care centers and the extended community but also upon integration with the digital signals from surrogate molecular markers will result in the creation of a multi-modal multi fusion sensor technology [27, 29, 30] which will aim at not only resolving the dogma of missed and misdiagnosis of Lung diseases such as Tuberculosis or Pneumonia at tertiary care centers and extended

community but also individualize the risk assessment of patients with suspected myocardial infarction or to categorize patients into low- or high-risk groups.

In Recent years, various computer-based tools have been developed which can be reliably used for computational disease tagging purposes. Healthcare Professionals with the help of such tools can accurately and computationally tag different disease conditions within a short time with a view to significantly improve the health outcomes of the community at large [49–57].

In the past people have prospected the use of deep learning models with limited efficiency to diagnose lung diseases which use X-ray images as a modality to evaluate lung health as well as predict the onset of diseases such as Covid-19 in the patients [31]. In this paper, we have explored the possibility to predict the lung ailment by applying Earth mover's Distance algorithm [58, 59] as our ongoing work along with Visibility Graph to the X-Ray images of the patients. EMD mimics the human perception of texture similarity whilst Horizontal visibility graph (HVG) and Hamming-Ipsen-Mikhailov (HIM) distance-based similarity approach forms a corner stone for automatically distinguishing clinical multimedia in an automated fashion. This stable and programmatic algorithmic capability can be leveraged to provide automated disease tagging where highly trained medical professional services are either too scarce or unaffordable. These observations when coupled together form the rationale for scalable automated clinical disease tagging for community-oriented health intervention.

3.1 Earthmover's Distance (EMD)

Earthmover's Distance (EMD) is a method to calculate the disparity between two multi-dimensional distribution in some space where a distance magnitude between single ones (ground distance) is given. Suppose the two distributions are there, one can be considered as the area with the mass of earth, and the other as a collection of holes in that same area. Then, the EMD is the measure of the least amount of work required to fill the holes with earth. Here the unit of work is the force needed in transporting unit earth by a unit of ground distance. So, it can also be defined as the minimum cost that must be provided to convert one histogram into other. Measuring of EMD is based on a solution of transportation problem [16]. For finding mathematical representation, firstly we formalized it as the following linear programming problem:

Let X be the first signature with n clusters, xi is the cluster representative, and wxi is the weight of cluster.

Let Y be the second signature with m clusters, yi is the cluster representative, and wyi is the weight of cluster.

Let D be the ground distance matrix, di j is the ground distance between clusters xi and y j.

Let F be the flow matrix and fi j is the between xi and y j.

Then,

$$X = \{(x_1, w_{x1}), (x_2, w_{x2}), (x_3, w_{x3}), ...(x_n, w_{xn})\}$$

$$Y = \{(y_1, w_{y1}), (y_2, w_{y2}), (y_3, w_{y3}), ...(y_m, w_{ym})\}$$

$$D = [d_{ij}]$$

$$F = [f_{ij}]$$

Now, the WORK (X, Y, F) $= \sum_{i=1}^{n}\sum_{i=1}^{m} f_{ij}d_{ij}$ Subject to constraints: (i) $f_{ij} \geq 0$, where $0 \leq i \leq n$, $0 \leq j \leq m$; (ii) $\sum_{j=1}^{m} f_{ij} \leq w_{xi}$, where $0 \leq i \leq n$; (iii) $\sum_{j=1}^{m} f_{ij} \leq w_{yj}$, where $0 \leq j \leq m$; (iv) $\sum_{i=1}^{n}\sum_{i=1}^{m} f_{ij} = \min \sum_{i=1}^{n} w_{xi}. \sum_{j=1}^{m} w_{yi}$ The constraint (i) enables mass moving from X to Y. (ii) and (iii) restricts the amount of mass that can be sent by the clusters in X to their weights and the clusters in Y to receive no more mass than their weights. (iv) One forces to move the maximum amount of mass possible. It is also known as the total flow. Once we solve the transportation problem, we will get the optimal flow F. Now the Earth Mover's Distance is defined as the work normalised by the total flow:

$$EMD(X, Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}d_{ij} \div \sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}$$

3.2 Horizontal Visibility Graph (HVG) and Its Application for X-ray Chest Radiograph Processing in R

The notion of visibility says that if two data points in a time series are in the line of sight without being obstructed by any other data points then they are visibible and hence they are connected in a visibility graph. This tranformation by visibility gives rise to the mapping of a timeseries into a network as per given specific geometric condition which is outlined below. Any two given data points (t1, i1) and (t2, i2) from timeseries obtained from covid or normal X-ray image matrix time series will be said to be visible and hence connected in the ensuing graph if for any other data point (t3, i3), for all t1 < t3 < t2 satisfies.

$$\text{i3} < \text{i1} + (\text{i2} - \text{i1})$$

$$\text{tc} - \text{t1}$$

$$t2 - t1$$

What it essentially means that all values yi for all t1 < ti < t2 should stay below the line drawn between i1 and i2. Limiting this notion of visibility to only horizontal direction, one can intuitively understand the notion of horizontal visibility where two data points are horizontally visible if one can draw a horizontal line between them or establish a line of sight while all other values between these two data points are staying below this line: ii, il > ik for all k such that j < k < l [33]. Clearly, as in the visibility case, horizontal visibility algorithm maps a sequence of data points/timeseries to a horizontal visibility graph (HVG). Once, HVG representation is obtained, massive analytic capabilities of network analysis and tools of network science and graph theory can be deployed to analyze the original sequence of datapoints combinatorically, resulting in hitherto unknown criteria for data sequence characterization. While there are large number of visibility graph applications in multiple multidisciplinary areas, this work leverages this method for classifying and distinguishing patients with a certain pathology from healthy controls, by using the network attributes of HVGs as feature-vectors for automatic disease-tagging. In particular, an analysis of automation classification of healthy and corona-positive patients is presented with digital lung-Xray modality [34].

3.2.1 Hamming-Ipsen-Mikhailov (HIM) and Network Similarity Metric

Hamming distance is a simple metric which computes the number of slots where two strings of equal length differ [60]. Alternatively, it counts the number of edits or substitutions required to transform one representation into the other. Generally speaking, its edit distance between two strings and can be deployed as a local metric to compute two networks' similarity indices. Ipsen-Mikhailov distance was pioneered by Ipsen [61] for graph reconstruction problems. Jurman et.al. [62] expanded its usage to "graph-comparison" methods.

The Ipsen-Mikhailov (IM) distance is a spectral measure which models a topology of N molecules connected by flexible springs. These network topologies are organized by the underlying adjacency matrix. The global (spectral) metric IM is the Ipsen-Mikhailov distance pertaining to the square-root of the squared difference of the Laplacian spectrum for each graph. The Ipsen-Mikhailov distance outlines the difference between two graphs by comparing their respective spectral densities and not by the raw eigenvalues themselves.

To take the advantage of local nature of Hamming and global nature of IM, the Hamming-Ipsen-Mikhailov distance is proposed. It is is a weighted combination of the Ipsen-Mikhailov (IM) and the normalized Hamming (H). The Hamming-Ipsen Mikhailov (HIM) distance is an Euclidean metric on the space created by the Cartesian product of the metric space associated with H and IM. The contributions of global and local information is governed by a combination factor ξ used in the

formula. When ξ is one, local and global information are in balance; when ξ is tending to 0, it becomes (local) Hamming distance; and when it goes to ∞ it resembles the (global) Ipsen Mikhailov distance.

$$d_{HIM} = \frac{1}{\sqrt{1+\xi}}\sqrt{\xi IM^2 + H^2}$$

Like mentioned earlier, this distance benefits from the strengths of both the Hamming and the Ipsen-Mikhailov distances by leveraging local and global information. Further, since it combines two distances with a non-negative weight, it defines a proper network distance between graphs. The parameter ξ gives the control to the metric by letting the user favor one type of information over the other. However, empirically, it is well observed that this distance is computationally expensive, and thus costly to apply to the analysis of massive graphs and large datasets. For our purposes here, HIM distance is used to compare two horizontal visibility graphs (HVG) which are generated either from normal or covid positive x-ray radiographs. The ensuing network similarity helps us decide the appropriate disease tag as will be demonstrated in the computational results.

4 Dataset

Primary source of normal and Covid-positive chest radiographs have been sourced from Rajiv Gandhi Cancer Institute and Research Centre [63] where representative normal and unhealthy ECGs, were compared *Diseasetagging with Visibility Graph and EMD based analysis of training data.*

With the given 100 test ECGs. Similar process has been followed for Covid-positive disease tagging using EMD with 30% success rate. Following similar reasoning, if test chest radiograph is closer to normal radiograph i.e. its VG-HIM distance is smaller to normal one, then it is tagged as a normal chest radiograph and if it resembles Covid positive chest radiograph, i.e. its VG-HIM distance is small with respect to representative Covid positive chest radiograph, then it is tagged as a Covid positive chest radiograph as shown in Table 3 next. A success rate of 60% for HVG–HIM based disease tagging has been reported. The full process has been shown as a flowchart in Fig. 13. We compute that our success rate is 60 out of hundred or 60% which calls for multimodality and that is where biomarkers [47, 48] walk in as a natural basis of Covid positive classification to further enhance the automated tagging of Covid positive chest radiographs with enhanced confidence.

4.1 Transformation of a Chest Radiograph to a Horizontal Visibility Graph: Different Stage of Processing

To illustrate the complete process from starting with a chest radiograph to generating its horizontal visibility graph to get it ready for HIM distance computation is accomplished in multiple different computational processing stages. We are going to illustrate it using COVID positive training image, COVID Train7.jpeg as shown in Fig. 1. We can notice that compared with a normal chest radiograph it has more white cloud like structures which possibly might be due to Covid positive nature of the radiograph. To process it, chest radiograph is converted to a down sampled numerical matrix in R computational environment. In our case we have downsized it to 8×8 to present it in all clarity and show the relevance of different processing algorithms and visualization. This transformed Covid positive chest radiograph as a 8×8 matrix is displayed in Fig. 2 and different color intensities show different grey levels in original covid positive chest radiograph.

Fig. 1 Chest X-Ray

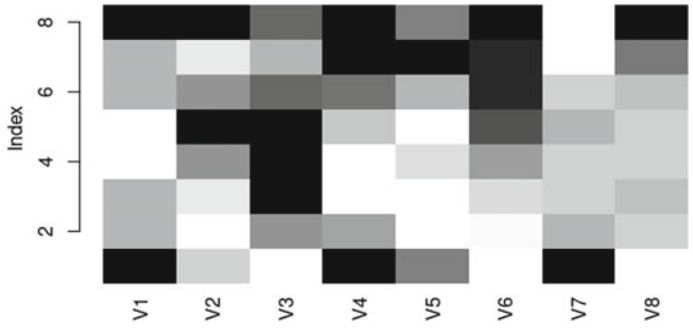

Fig. 2 Matrix Representation

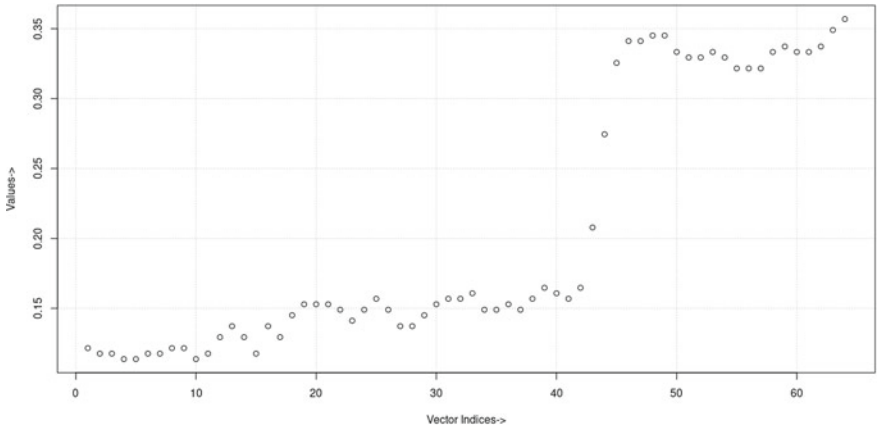

Fig. 3 Matrix to vector conversion of COVID positive training image, COVID_Train7 and corresponding plot of values

In next stage of transformation this 8 × 8 matrix is stacked as a vector of size 64 and is plotted as a time series in Fig. 3. Now stage is set for the transformation of this timeseries to a horizontal visibility graph. Once horizontal visibility graph algorithm processes this timeseries, a network is generated whose adjacency plot is shown in Fig. 4. Its largely sparse graph with few connectivity here and there as displayed by yellow-colored cells. Real network shape and connectivity patterns of horizontal visibility graph is demonstrated in Fig. 5. It can be visualized in multiple ways in R environment and exposes larger number of features and properties of horizontal visibility graphs resulting from chest radiographs.

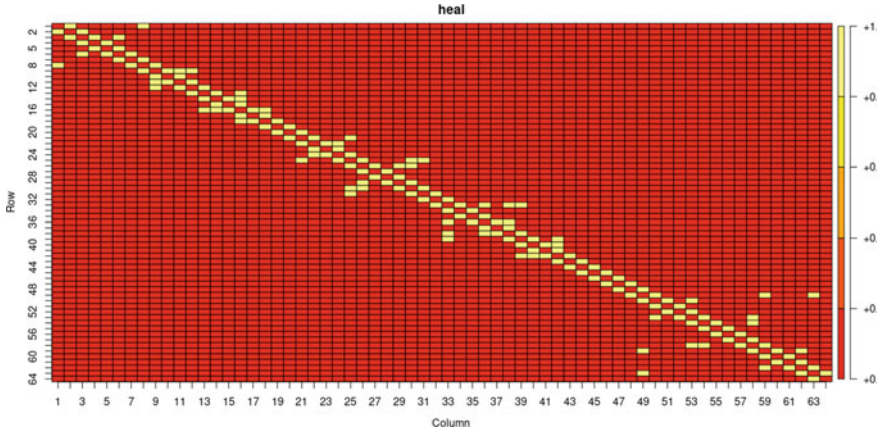

Fig. 4 Adjacency matrix plot of horizontal visibility graph generated from COVID positive training image, COVID_Train7.jpeg. Only yellow cells are one indicating connectivity and rest are disconnected

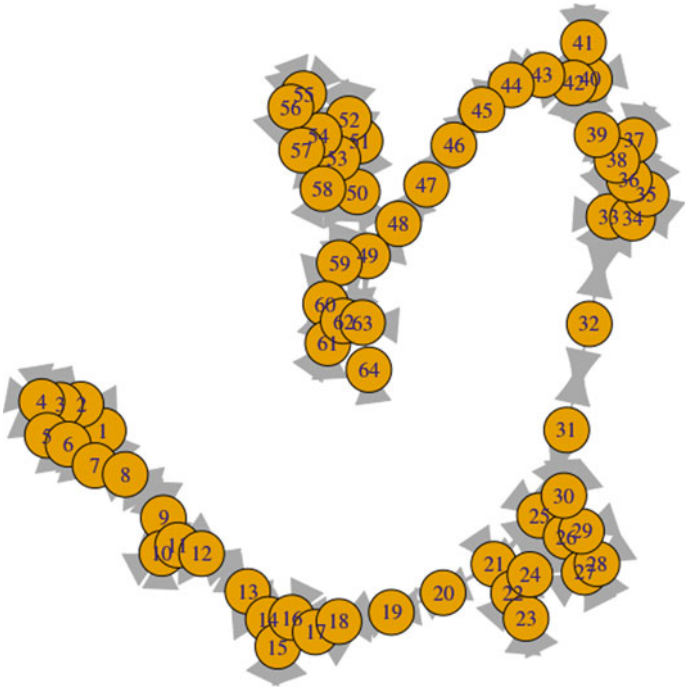

Fig. 5 Network plot of horizontal visibility graph generated from COVID positive training image, COVID_Train7

Before we move to next Fig. 6, we need to recollect the definition of heatmap. A heatmap is a two-dimensional grid kind of visual representation of data/information/signal in a colorful fashion. Heatmaps can aide the viewer in trying to make sense of a complex spatial distribution of information. What Fig. 6 shows is connectivity activity on a two-dimensional grid to communicate in a user-friendly fashion. Figure 7 provides another view of same horizontal visibility graph obtained from Covid positive image and Fig. 8 shows the same diagram with node size being proportional to 5th power of the degree of the node, i.e. highly connected nodes or hubs are depicted with larger circles as compared to sparsely connected nodes. Figures 9 and 10 show the degree distribution and cumulative degree.

A histogram type horizontal visibility graph is demonstrated in Fig. 11 obtained from Covid positive chest radiograph data. Its curvy form is demonstrated in Fig. 12. Both demonstrate interesting connectivity patterns. At this stage Covid positive chest radiograph's HVG is ready to be used by HIM-distance metric to compute similairty among different HVGs generated from normal and Covid positive chest radiographs.

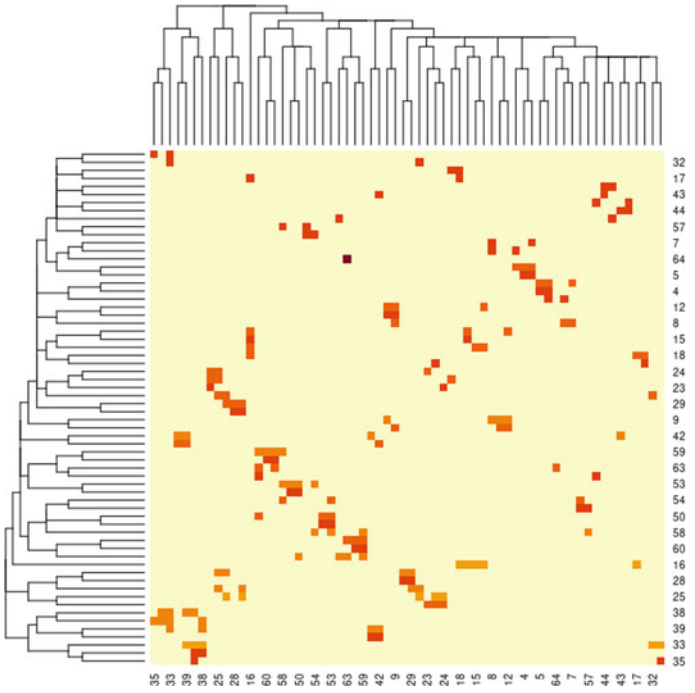

Fig. 6 Heatmap of horizontal visibility graph generated from COVID positive training image, COVID_Train7

4.2 Computational Infrastructure Deployed

Matlab has been used for performing geometrical part of the work. EMD aspect of this work has been performed in R software (Rstudio Version 1.3.1093 ©2009–2020 RStudio, PBC"Apricot Nasturtium" (aee44535, 2020-09-17) for Ubuntu Bionic Mozilla/5.0 (X11; Linux x86i 64) AppleWebKit/537.36 (KHTML, like Gecko) QtWebEngine/5.12.8 Chrome/69.0.3497.128 Safari/537.36) on a HP Probook laptop.

Laptop's operating system and other basic information from comand uname -a is given below:

Linux Krishna 5.4.0–48-generic #52-Ubuntu SMP Thu Sep 10 10:58:49 UTC 2020

× 86 64 × 86i 64 × 86 64 GNU/Linux

Output of hardware atributes of the laptop is as follows:

-memory

description: System memory

physical id: 0

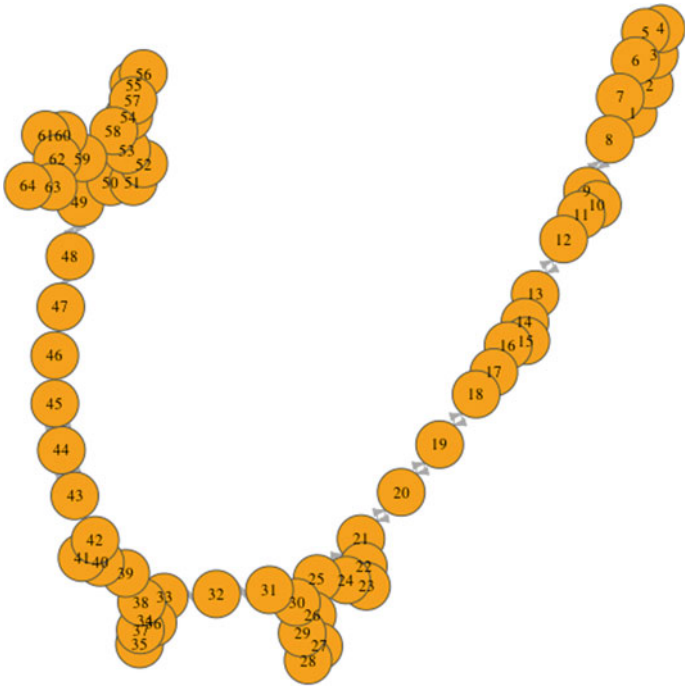

Fig. 7 Another network view of horizontal visibility graph generated from COVID positive training image, COVID_Train7

size: 8320MiB

-cpu

product: Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz

vendor: Intel Corp.

physical id: 1

bus info: cpu@0

size: 3304 MHz

capacity: 3400 MHz

Finally, Fig. 13 depicts the flow chart for Covid computational disease tagging algorithm using Visibility graph and network distance HIM in a sequential fashion. It summarizes all the computational steps used in different stages of processing at high level.

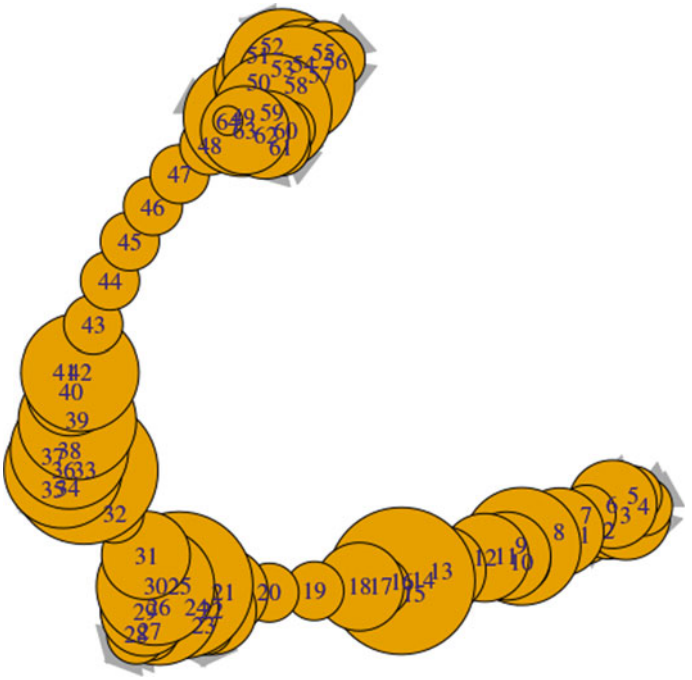

Fig. 8 Network view with node size 5th power of degree for horizontal visibility graph generated from COVID positive training image, COVID_Train7.jpeg

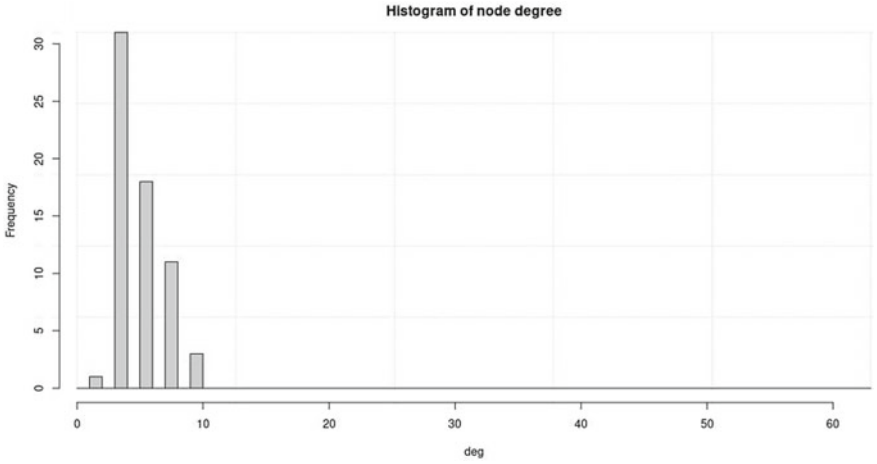

Fig. 9 Histogram of degree for horizontal visibility graph generated from COVID positive training chest radiograph, COVID Train7.jpeg. Moderately connected

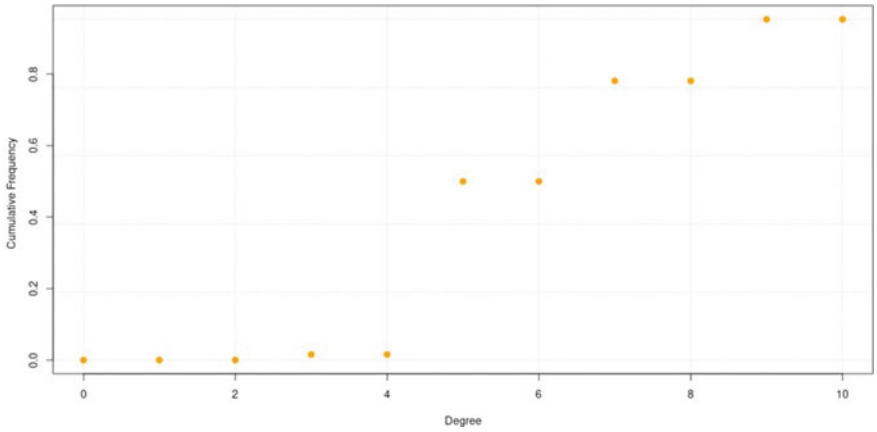

Fig. 10 Cumulative degree distribution for horizontal visibility graph generated from COVID positive training chest radiograph, COVID Train7.jpeg. Moderately connected

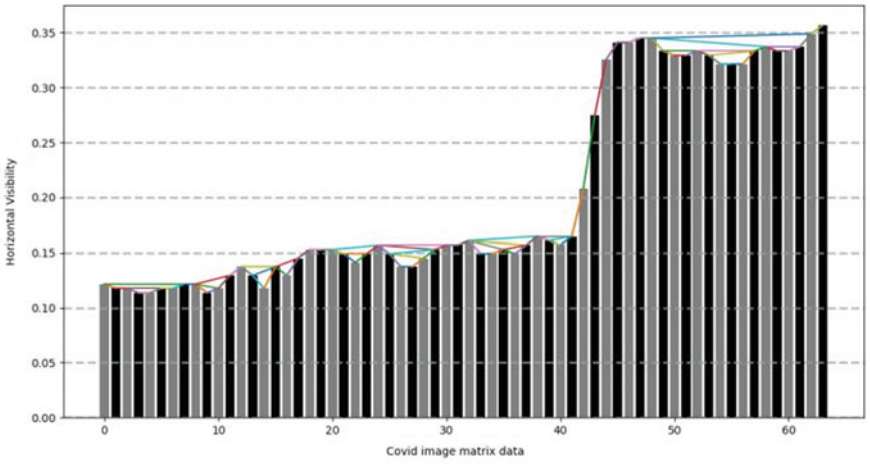

Fig. 11 Covid image matrix data

5 Experimental Results

5.1 Computational Experiment

This part describes the result of automated disease tagging using horizontal visibility graph and HIM based network similarity (distance) computation. Chest radiographs used here are sourced from Rajiv Gandhi Cancer Institute and Research Centre [64, 63].

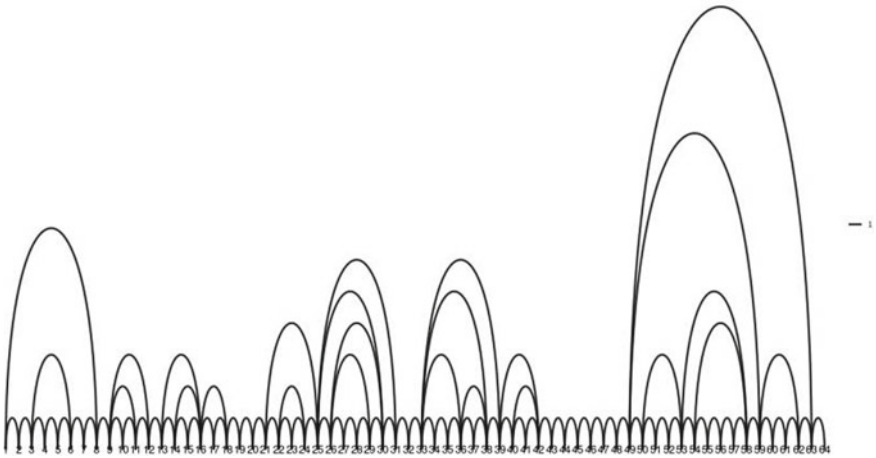

Fig. 12 Visibility graph generated from COVID positive training chest radiograph, COVID Train7.jpeg

Data Preprocessing To keep the computation of images and their processing commensurate to hardware platform capabilities, all the radiographs acquired are converted into the JPG format. For fast processing and declaration of results in almost-real-time, radiographs are down sampled to the 32×32 pixel size irrespective of their original size.

Radiograph data is grouped into two main groups, training and testing. Training group has 10 normal and 10 covid positive radiographs. Normal radiographs are compared amongst each other using HVG–HIM algorithm and representative normal radiograph is computed like our previous work using EMD as shown in Table 1. Same process is followed for the covid positive radiograph and a covid-positive representative radiograph is obtained. Out of multiple network distance available, Hamming-Ipsen Mikhailov (HIM) network distance is used for comparing the visibility graphs because of its balanced nature as a both global and local network distance or similarity metric.

Training Using Normal Chest Radiographs As shown in the flowchart in Fig. 13, we begin with evaluating the normal representative chest radiograph (Normal-Rep) among normal ensemble of training chest radiographs. This Normal-Rep will be used to compare the test chest radiograph with, to decide if test chest radiograph can be tagged normal or Covid-positive. This process of computing Normal-Rep is by converting all the normal training chest radiographs into visibility graphs and measuring their computational similarity with HIM metric. The Table 1 for deciding Normal-Rep is given below where third normal training chest radiograph has been designated Normal-Rep for this ensemble of ten normal training chest radiographs due to its highest similarity (Hence lowest column sum score) with all other normal training chest radiographs. Its score is indicated in third column and sum row in bold.

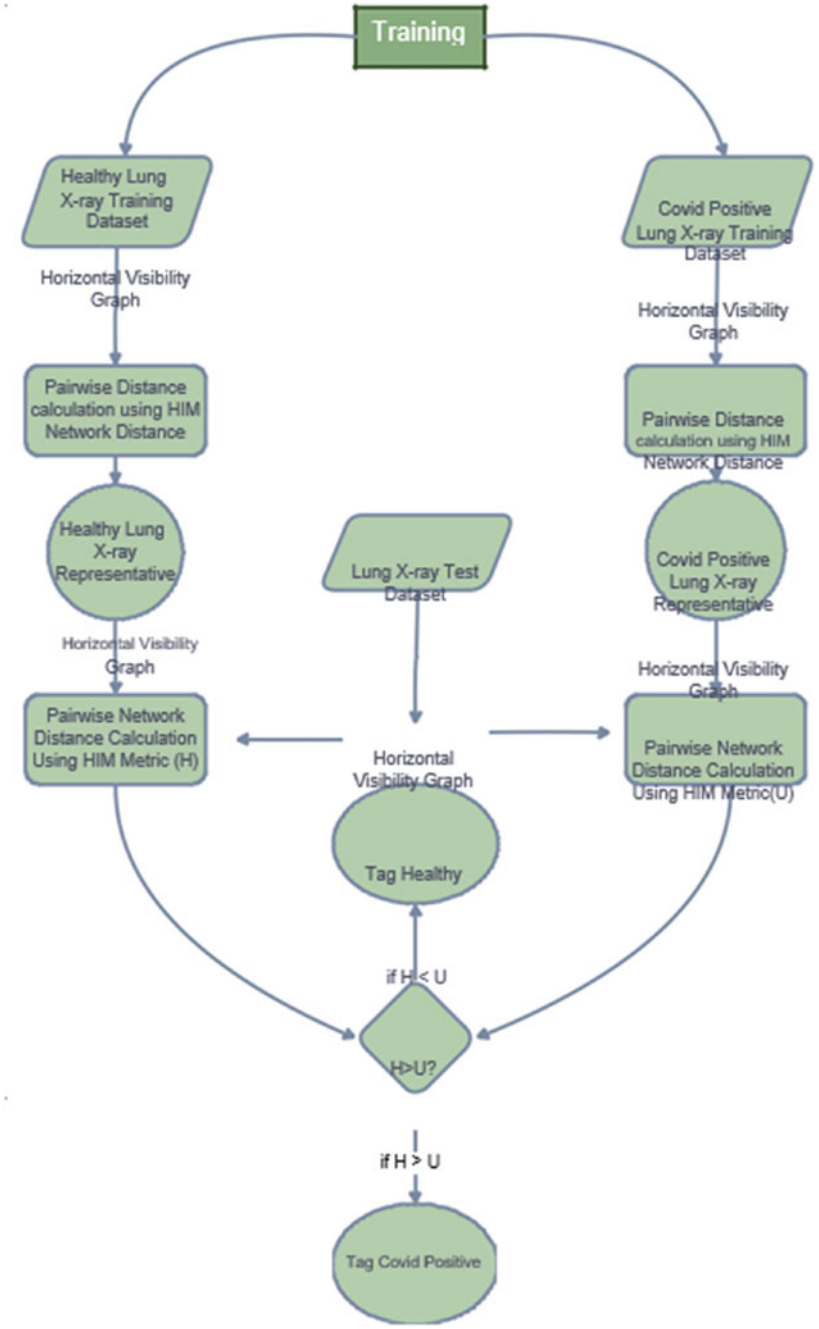

Fig. 13 Flow chart for Covid computational disease tagging algorithm using Visibility graph and Network Distance HIM

Table 1 HIM score table for computing HIM distance among normal chest radio-graphs and locating normal-Rep

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
1	0	0.002114	0.002061	0.002169	0.001873	0.002119	0.001924	0.001924	0.002170	0.001863
2	0.002114	0	0.001683	0.002144	0.002102	0.002152	0.002041	0.002028	0.002145	0.002044
3	0.002061	0.001683	0	0.002087	0.002007	0.001991	0.001995	0.001921	0.001966	0.002038
4	0.002169	0.002144	0.002087	0	0.002027	0.001768	0.002079	0.002110	0.002246	0.002099
5	0.001873	0.002102	0.002007	0.002027	0	0.002071	0.001885	0.001963	0.002217	0.001893
6	0.002119	0.002152	0.001991	0.001768	0.002071	0	0.002091	0.002117	0.002258	0.002079
7	0.001924	0.002041	0.001995	0.002079	0.001885	0.002091	0	0.001881	0.002090	0.001935
8	0.001924	0.002028	0.001921	0.002110	0.001963	0.002117	0.001881	0	0.002115	0.001783
9	0.002170	0.002145	0.001966	0.002246	0.002217	0.002258	0.002090	0.002115	0	0.002147
10	0.001863	0.002044	0.002038	0.002099	0.001893	0.002079	0.001935	0.001783	0.002147	0
Sum	0.018220	0.018457	0.017754	0.018733	0.018043	0.018649	0.017924	0.017846	0.019357	0.017886

Training Using Covid Positive Chest Radiographs Following the flowchart in Fig. 1. We begin with evaluating the Covid Positive representative chest radiograph (Covid-PositveRep) among Covid positive ensemble of training chest radiographs. This Covid Positive-Rep will be used to compare the test chest radiograph with, to decide if test chest radiograph can be tagged normal or Covid-positive. This process of computing CovidPositiveRep is realized by converting all the Covid Positive training chest radiographs into visibility graphs and measuring their computational similarity with HIM metric. The Table 2 for deciding Covid Positive-Rep is given below where fourth Covid positive training chest radiograph has been designated Covid Positive-Rep for this ensemble of ten Covid positive training chest radiographs due to its highest similarity (Hence lowest column sum score) with all other Covid positive training chest radiographs. Its score is indicated in fourth column and sum row in bold.

Final Testing and Automated Disease Tagging for Test Chest Radiographs Using HVG–HIM In, the follow-up testing phase, both healthy and covid positive representative radiographs are compared using HIM distance with pretagged test dataset of 20 radiographs. This test dataset has both healthy and covid positive radiographs. The result of automated disease tagging is presented below. Let's define U as HIM Distance from Covid Positive representative and H as HIM Distance from Healthy representative. A simple observation tells us that this algorithm is able to tag the chest radiographs with overall accuracy of 60%. Healthy radiographs have been tagged with 60% accuracy and also covid positive radiographs are tagged with 60% accuracy. A natural future direction arises where other network distance metrics can be leveraged over larger datasets (Fig. 12, Table 3).

6 Final Testing and Automated Disease Tagging for Test Chest Radiographs with EMD

To draw a fair comparison between EMD and HVZ-VG, we run the same computational with direct perceptual similarity between chest radiographs-based evaluation and diseases tagging with EMD. To keep the computation of chest radiographs and their processing commensurate to hardware platform capabilities, all the radiographs acquired are converted into the JPG format. For fast processing and declaration of results in almost-real-time, radiographs are down sampled to the 32×32-pixel size irrespective of their original size. This is in consonance with the same computational experiment carried out HVG–HIM. Results of EMD-based disease tagging is shown in Table 4 where a meagre 30% accuracy is reported, and correct disease tag rows are highlighted in bold. This is in sharp contrast with accuracy of 60% achieved with HVZ–HIM, albeit at a higher computational investment.

Table 2 HIM score table for computing HIM distance among covid positive chest radiographs and locating covid positive-Rep

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
1	0	0.001879	0.001896	0.001809	0.001768	0.001861	0.001897	0.001996	0.001989	0.002110
2	0.001879	0	0.001882	0.001690	0.001873	0.001820	0.001738	0.001784	0.001838	0.001814
3	0.001896	0.001882	0	0.001825	0.001755	0.001837	0.001948	0.001865	0.001738	0.001981
4	0.001809	0.001690	0.001825	0	0.001880	0.001991	0.001619	0.001845	0.001640	0.001760
5	0.001768	0.001873	0.001755	0.001880	0	0.001728	0.002055	0.002103	0.001936	0.002020
6	0.001861	0.001820	0.001837	0.001991	0.001728	0	0.001773	0.001903	0.002042	0.001873
7	0.001897	0.001738	0.001948	0.001619	0.002055	0.001773	0	0.001780	0.001783	0.001598
8	0.001996	0.001784	0.001865	0.001845	0.002103	0.001903	0.001780	0	0.001767	0.001875
9	0.001989	0.001838	0.001738	0.001640	0.001936	0.002042	0.001783	0.001767	0	0.001881
10	0.002110	0.001814	0.001981	0.001760	0.002020	0.001873	0.001598	0.0018758	0.001881	0
Sum	0.017209	0.016324	0.016731	0.016063	0.017122	0.016833	0.016195	0.016921	0.016618	0.016916

Table 3 HIM score based disease tagging table for test covid positive and normal chest radiographs using HVG–HIM

No	U	H	Real Tag	VG-HIM Tag
1	0.001738350	0.002092146092198	Covid positive	Covid positive
2	0.002082185	0.001840777512553	Covid positive	Normal
3	0.001950630	0.00211087540523	Covid positive	Covid positive
4	0.001840535	0.001930288747592	Covid positive	Covid positive
5	0.002241159	0.001995573933206	Covid positive	Normal
6	0.001916763	0.001973678120048	Covid positive	Covid positive
7	0.001955399	0.002161508372919	Covid positive	Covid positive
8	0.001919654	0.001948495673396	Covid positive	Covid positive
9	0.002331740	0.002206820660675	Covid positive	Normal
10	0.002041363	0.00186923269593	Covid positive	Normal
11	0.001997673	0.001997358118727	Normal	Normal
12	0.002224062	0.002132561182092	Normal	Normal
13	0.001828268	0.001921438899592	Normal	Covid positive
14	0.002111531	0	Normal	Normal
15	0.002106018	0.002075943450341	Normal	Normal
16	0.001789069	0.002062848422154	Normal	Covid positive
17	0.002076307	0.002162210255525	Normal	Covid positive
18	0.002154178	0.001964655126736	Normal	Normal
19	0.001989834	0.002014787858529	Normal	Covid positive
20	0.002301401	0.002209421305833	Normal	Normal

7 Reflections on HVG–HIM and EMD Similarity Metrics

This experiment on working with HVG–HIM and EMD has given us certain insights into the implementation of these algorithms. Earth Movers Distance (EMD) algorithm computes the discrepancy pixel by pixel in the chest radiographs and gives us the overall average difference between the chest radiographs as a similarity metric. In the case of horizontal visibility graph (HVZ), each pixel compares itself with all other pixels of the same chest radiograph and gives a graphical representation. This graphical representation of one chest radiograph is compared with other chest radiographs's graphical representation using the network distance metric. For calculating the difference in these graphs various network distance metrics can be used. Here, we have used HammingIpsen-Mikhailov (HIM) distance.

From a computational aspect, EMD does far fewer calculations than HVG–HIM metric does. EMD computes the results within the few seconds for given set of ten chest radiographs with similar size whereas for the same task HVG–HIM network

Table 4 EMD score based disease tagging table for test covid positive and normal chest radiographs

No	EMD from NormalRep	EMD from CovidPositive-Rep	Real-Tag	EMD-Tag
1	0.50817084312439	0.102157711982727	Normal	CovidPositive
2	3.02615809440613	0.244212314486504	Normal	CovidPositive
3	0.554616451263428	0.10110604763031	Normal	CovidPositive
4	2.75559902191162	1.02294194698334	Normal	CovidPositive
5	1.33040499687195	0.4100721180439	Normal	CovidPositive
6	0.52178567647934	0.358028054237366	Normal	CovidPositive
7	1.14510023593903	4.19615602493286	Normal	Normal
8	0.391091376543045	0.514602303504944	Normal	Normal
9	1.55246329307556	0.12182080745697	Normal	CovidPositive
10	0.939411997795105	0.931683301925659	Normal	CovidPositive
11	1.28701484203339	0.468866437673569	CovidPositive	CovidPositive
12	1.32051146030426	1.4063401222229	CovidPositive	Normal
13	1.10182595252991	3.01665210723877	CovidPositive	Normal
14	0.449742645025253	1.7027291059494	‘ CovidPositive	Normal
15	0.765507996082306	1.09815609455109	CovidPositive	Normal
16	0.297308087348938	0.422655075788498	CovidPositive	Normal
17	‘0.361823529005051	0.790895044803619	CovidPositive	Normal
18	0.480859369039536	0.396273583173752	CovidPositive	CovidPositive
19	0.886679291725159	0.466326594352722	CovidPositive	CovidPositive
20	1.08181118965149	0.41103208065033	CovidPositive	CovidPositive

distance takes several minutes. Clearly, there is a learning that details matter. HVG–HIM is giving twice the accuracy of 60% compared to EMD which gives the accuracy of 30% for the same task. Naturally, HVG–HIM is achieving this performance because of large computational investment. This leads to an interesting deployment choice as in, where chances of Covid-positive prevalence is extremely low and high accuracy is not needed, one can deploy EMD based procedures but for regions where prevalence is higher and accuracy is of paramount importance, HVZ-HIM with serious computational infrastructure will be needed.

8 Towards a Web-Service Based Implementation

Covid has emerged as an unprecedented global pandemic with serious impact on every individual. Provision of immediate and adequate health infrastructure for covid patients visiting a health service facility or practicing tele consultancy based on pathological examination of chest radiograph is the need of the hour. After centuries

of advancements and developments in different health practices like Allopathy, Ayurved, Homeopathy and other forms of treatment strategies, human efforts against Covid has been dwarfed. The whole medical community is fighting with all available resources to tackle the situation and treat the patients. Having said that it is hard to deny in comparison to covid patients, the number of skilled and trained health workers like doctors, nurses and other health service providers is not sufficient, owing to this provisioning gap, high mortality and prolonged morbidity is recorded, especially in LMICs (Low- and Middle-income Countries.)

Today, we are living in a digital world where enormous amount of technology enabled health services are being practiced all across the world specially in the form of digital health encompassing—tele consultancy, telemedicine, telehealth, big data etc. They help in gathering meaningful information, processing it and producing the report in almost real-time so that the policy makers can formulate evidence-based health strategies leading to follow-up of patients more effectively.

Keeping this urgent need in mind, we have developed our own web-portal which is capable of collaborating with all the hospitals and individual medical practioners/patients through a centralized server. This server is designed in such a manner that any individual or hospital can access the server after proper validation and store their relevant information related to patients. Data security and confidentiality has been maintained by the server strictly. Edit access has been limited to information owners only on the portal. The purpose of this webservice is to store the information, process it and produce the result in form of scientific evidence which can be significantly utilized by policy makers for better decision making. At the same time, by accessing the portal, one can get all the relevant and accurate information related to Covid in form of text, presentation and multimedia (audio/video/images). This portal will also provide an individual specific service like tele consultancy to register and forward the unanswered queries directly to the specialist doctors and back to the query generator (possibly patient or someone curious about a medical condition) (Figs. 14 and 15).

8.1 Webservice Methodology

The web-based healthcare management system for Covid patients is poised with the latest front and web-page development language—PhP 7.3.28 and the core technology used is MVC (Model View Controller). There are different modules in the website which incorporate API (Application Programming Interface) to interact with dedicated servers for dedicated processing of HVG and HIM-distance in terms of image analysis using R (RStudio 3.6.3) and other statistical packages and report generation with an attractive and effective graphical representation for the available dataset. MySQL server is used as the backend RDBMS (Relational database Management System) for data input, process, and output to APIs and individuals for usage downstream. We have integrated these technologies because they are open-source and compatible for design, development, and deployment. Further, they are

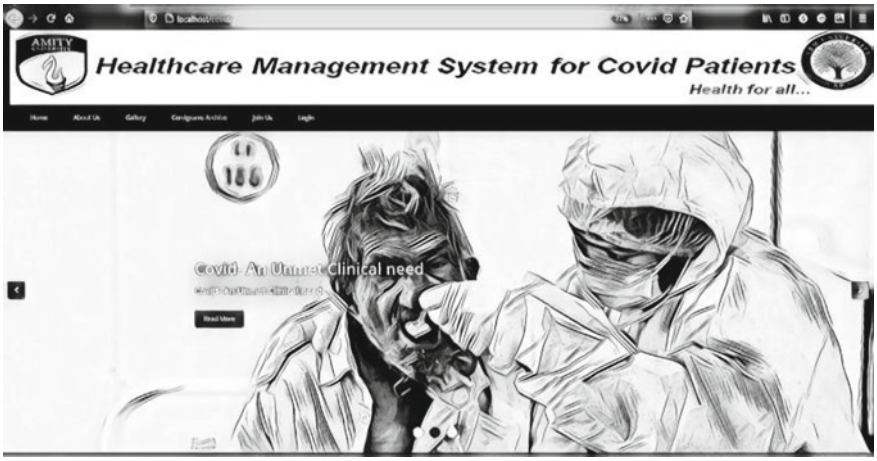

Fig. 14 Transiton from research to service provision: A pilot webservice for automated COVID disease tagging

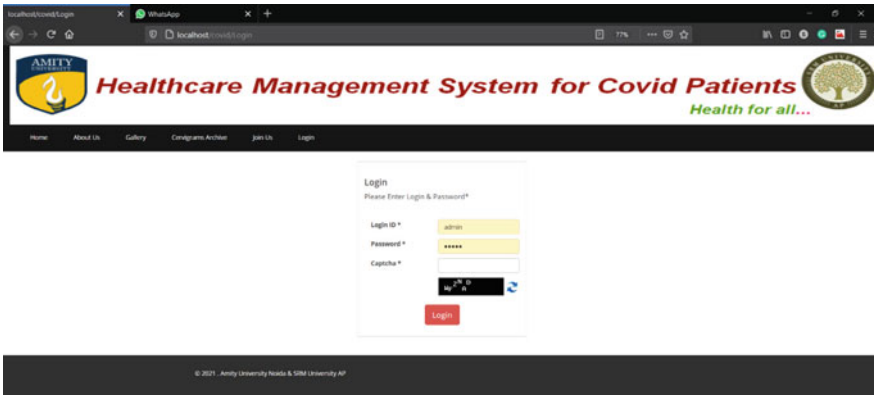

Fig. 15 Resource for Automated disease tagging with chest radiographs in LMICs and resource challenged African Countries

customizable as per the medical data-keeping requirements of this project. Security and confidentiality are maintained at all levels of data flow starting from information gathering to report generation. A brief architecture of webservice technology has been shown in Fig. 16.

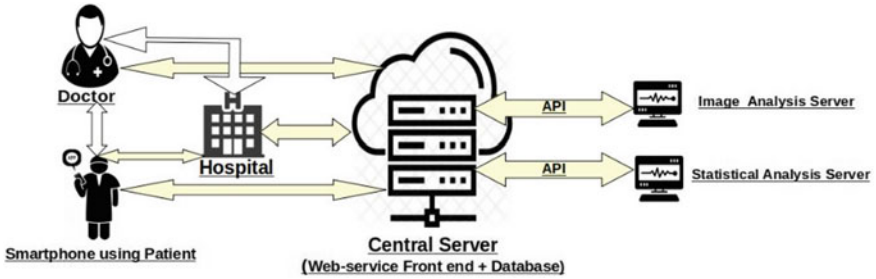

Fig. 16 Web Service Architecture for Covid Positive Chest Radiograph based Disease Tagging

9 Conclusions and Future Directions

Poor lung health is known to play a statutory role causing increased susceptibility related to COVID-19. Lifestyle choices including smoking and sedentary lifestyle leading to obesity aren't the only factors that influences lung health, environmental factors such as air pollution also exert a considerable effect. Researchers believe that while consumption of tobacco products (both smoking and smokeless), along with occupational hazards such as exposure to indoor and outdoor pollution makes people more susceptible to the infection that causes COVID-19 and its complications because these environmental factors also significantly damage the body's natural defenses against some bacteria and viruses. A large number of countries coming under the umbrella of LLMICs having populations endowed with poor lung function and consequently poor lung health reflect their health outcomes as poor.

The use of extant deep learning technologies is not necessarily solving the problem of integrating the evidences from the community level in resource limited healthcare systems as they are intensive and energy-hogging with respect to computational resources leading to increased carbon foot-printing and hence global warming.

To this end use of static algorithms such as Earth Movers Distance (EMD) and Horizontal Visibility Graph (HVG) add value as they require significantly lesser investment of computational resources and dispel the black box nature of the deep learning algorithms with the glass box nature with more transparency with respect to big data computation and analytics [65]. We propose to extend our studies on the use of EMD and HVG based time series analysis, in which dynamic timeseries and clinical multimedia segments are mapped to visibility graphs as being descriptions of the corresponding states and the successively occurring states are linked. This procedure capable of converting a dynamic time series to a temporal network and at the same time a network of networks could be provide us rich information benefiting short-term and long-term predictions about lung of an individual or community at large, thereby providing the policy administrators at local, regional and global level nuanced data for developing comprehensive niche specific solutions aimed at alleviating the lung-health disparities.

Use of multi-modal multi sensor fusion technologies combined with big data enabled platforms will go a long way in strengthening resource-deprived healthcare systems. Our proposed disruptive AI-enabled point of care solution aims to gather evidences from the community level so as to augment catering to the creation of affordable and accessible healthcare technologies will focus on the application of innovative concepts to improve health outcomes in an affordable and equitable manner to overcome healthcare disparities but also inculcate capacity building through the provision of unique platform to individuals /organizations to validate their proof of concepts to scale-ups and ultimately commercially viable sustainable solutions.

10 Device Utility

1. Has potential application as an Adjunct Clinical Aid for the Pulmonologist/Medical Professionals.
2. Automatic Classification of X-ray chest radiographs facilitating large scale screening of subjects in remote health camps.
3. Easy, fast and robust technology with capabilities to be implemented in web-based, desktop-based and smartphone-based applications when coupled with X-ray device on the internet.
4. It has potential of turning Covid disease management as a self-care exercise. Control moves from the hands of expensive hospital to cheap and affordable selfcare devices.

References

1. United Nations: Report of the director-general of the World Health Organization on the prevention and control of non-communicable diseases. Document A/R68/650 (2013). Accessed 15 Nov 2020
2. Dutta, U., Nagi, B., Garg, P.K., Sinha, S.K., Singh, K., Tandon, R.K.: Patients with gallstones develop gallbladder cancer at an earlier age. Eur. J. Cancer Prev. **14**, 381–385 (2005). https://doi.org/10.1097/00008469-200508000-00011
3. Gajalakshmi, C.K., Shanta, V., Swaminathan, R., Sankaranarayanan, R., Black, R.J.: Apopulation-based survival study on female breast cancer in Madras. Br. J. Cancer Res. **75**, 771–5 (1997). https://doi.org/10.1038/bjc.1997.137
4. Mendis, S.: Global status report on noncommunicable diseases 2014 (2014). http://www.who.int/nmh/publications/ncd-status-report-2014/en/. Accessed 15 Nov 2020
5. United Nations: Outcome document of the high-level meeting of the General Assembly on the comprehensive review and assessment of the progress achieved in the prevention and control of non-communicable diseases. A/RES/68/300 (2014). Accessed 15 Nov 2020
6. World Health Organization. Novel coronavirus—China. http://www.who.int/csr/zxcvXDdon/12-january-2020-novel-coronavirus-china/en/. Accessed: 15 Nov 2020

7. World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it. https://www.who.int/emergencies/diseases/novelcorona-virus2019/technical-guidance/naming-the-coronavirus-disease-(COVID-2019)-andthe- virus-that-causes-it. Accessed 15 Nov 2020
8. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al.: Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet **395**, 497–506 (2020)
9. World health organization (WHO): COVID 19 and NCDs. https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed 15 Nov 2020
10. Leung, C.: Clinical features of deaths in the novel coronavirus epidemic in China. Rev. Med. Virol. **30**(3), e2103 (2020)
11. Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., et al.: Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet **395**, 507–513 (2020)
12. GBD: Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir. Med. **5**, 691–706 (2017)
13. India State-Level Disease Burden Initiative CRD Collaborators: The burden of chronic respiratory diseases and their heterogeneity across the states of India: The Global Burden of Disease Study 1990–2016. Lancet Global Health **6**(12), e1363–e1374 (2018). https://doi.org/10.1016/S2214-109X(18)30409-1.28N.Chavanetal
14. Cohen, A.J., Brauer, M., Burnett, R., et al.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet **389**, 1907–1918 (2017)
15. Landrigan, P.J., Fuller, R., Acosta, N.J.R., et al.: The lancet commission on pollution and health. Lancet **391**, 462–512 (2018)
16. Rubner, Y., Tomasi, C.: The earth movers distance perceptual metrics for image database navigation. (2001)
17. Ji, H., Xu, T., Wu, W., Wang, J.: Visibility graph analysis on EEG signal. In: 2016 9th International Congress on Image and Signal Processing, pp. 1557–1561. BioMedical Engineering and Informatics (CISP-BMEI), Datong (2016). https://doi.org/10.1109/CISP-BMEI.2016.7852963
18. Thompson, N.C., Greenewald, K., Lee, K., Manso, G.F.: The Computational Limits of Deep Learning (2020). arxiv:2007.05558
19. Yampolskiy, R.V.: Unpredictability of AI (2019). Arxiv:1905.13053
20. Yampolskiy, R.V.: Unexplainability and incomprehensibility of artificial intelligence (2019). Arxiv:1907.03869
21. Sohl-Dickstein, J., Weiss, E.A., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics (2018). Arxiv:1503.03585
22. Li, H.: Analysis on the nonlinear dynamics of deep neural networks: topological entropy and chaos (2018). Arxiv:1804.03987
23. Shrivastav, K.D., Taneja, N., Das, A.M., Rana, S., Ranjan, P., Singh, H., Jaggi, V.K., Janardhanan, R.: Socio-demographic and clinico-pathological profile of cervical cancer patients at a tertiary care centre in New Delhi: a five-year retrospective analysis. Indian J. Community Health **33**(4) (2021)
24. Shrivastav, K.D., Arambam, P., Batra, S., Bhatia, V., Singh, H., Jaggi, V.K., Ranjan, P., Abed, E.H., Janardhanan, R.: Earth mover's distance-based tool for rapid screening of cervical cancer using cervigrams. Appl Sci **12**(9), 4661 (2022)
25. https://venturebeat.com/2020/07/15/mit-researchers-warn-that-deep-learning-isapproaching computational-limits. Published 15 July 2020
26. Scassellati, C., Bonvicini, C., Benussi, L., Ghidoni, R., Squitti, R.: Neurodevelopmental disorders: Metallomics studies for the identification of potential biomarkers associated to diagnosis and treatment. J. Trace Elem.Ts Med. Biol. **60**, 126499 (2020). https://doi.org/10.1016/j.jtemb.2020.126499. ISSN 0946-672X

27. Chen, R.J., Lu, M.Y., Wang, J., Williamson, D.F.K., Rodig, S.J., Lindeman, N.I., Mahmood, F.: Pathomic Fusion: An Integrated Framework for Fusing Histopathology and Genomic Features for Cancer Diagnosis and Prognosis (2019). https://arxiv.org/abs/1912.08937
28. Carmichael, I., Calhoun, B.C., Hoadley, K.A., Troester, M.A., Geradts, J., Couture, H.D., Olsson, L., Perou, C.M., Niethammer, M., Hannig, J., Marron, J.S.: Joint and individual analysis of breast cancer histologic images and genomic covariates (2019). https://arxiv.org/abs/1912.00434
29. Carmichael, I., Calhoun, B.C., Hoadley, K.A., Troester, M.A., Geradts, J., Heather, D.C., Olsson, L., Perou, C.M., Niethammer, M., Hannig, J., Marron, J.S.: Joint and individual analysis of breast cancer histologic images and genomic covariates (2019). https://arxiv.org/abs/1912.00434
30. Guo, A., Chen, Z., Li, F., Li, W., Luo, Q.: Towards more reliable unsupervised tissue segmentation via integrating mass spectrometry imaging and hematoxylin-erosin stained histopathological image (2020). https://www.biorxiv.org/content/https://doi.org/10.1101/2020.07.17.208025v1
31. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.: Automated detection of COVID-19 cases using deep neural networks with. Comput. Biol. Med. **1**, 11 (2020)
32. Guo, A., Chen, Z., Li, F., Li, W., Luo, Q.: Towards more reliable unsupervised tissue segmentation via integrating mass spectrometry imaging and hematoxylin-erosin stained histopathological image (2020). https://www.biorxiv.org/content/https://doi.org/10.1101/2020.07.17.208025v1
33. Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C.: From time series to complex networks: the visibility graph. In: Proceedings of the National Academy of Sciences, vol. 105, 13th edn, pp. 4972–4975 (2008)
34. Luque, B., Lacasa, L., Ballesteros, F., Luque, J.: Horizontal visibility graphs: exact results for random time series. Phys. Rev. E **80**(4), 046103 (2009)
35. Zhu, G., Li, Y., Wen, P.: Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal. IEEE J. Biomed. Health Inform. **18**(6), 1813–1821 (2014)
36. Donner, R.V., Donges, J.F.: Visibility graph analysis of geophysical time series: potentials and possible pitfalls. Acta Geophys. **60**(3), 589–623 (2012)
37. Sannino, S., Stramaglia, S., Lacasa, L., Marinazzo, D.: Visibility graphs for fMRI data: Multiplex temporal graphs and their modulations across resting-state networks. Netw. Neurosci. **1**(3), 208–221 (2017)
38. Iacovacci, J., Lacasa, L.: Visibility graphs for image processing. IEEE Trans. Pattern Anal. Mach. Intell. **42**(4), 974–987 (2019)
39. Madl, T.: Network analysis of heart beat intervals using horizontal visibility graphs. In: 2016 Computing in Cardiology Conference (CinC), pp. 733–736. IEEE (2016)
40. Artameeyanant, P., Sultornsanee, S., Chamnongthai, K.: An EMG-based feature extraction method using a normalized weight vertical visibility algorithm for myopathy and neuropathy detection. Springerplus **5**(1), 1–26 (2016)
41. Fasmer, E.E., Fasmer, O.B., Berle, J.Ø., Oedegaard, K.J., Hauge, E.R.: Graph theory applied to the analysis of motor activity in patients with schizophrenia and depression. PloS one **13**(4), e0194791 (2018)
42. Mohammadpoory, Z., Nasrolahzadeh, M., Mahmoodian, N., Sayyah, M., Haddadnia, J.: Complex network based models of ECoG signals for detection of induced epileptic seizures in rats. Cogn. Neurodyn. **13**(4), 325–339 (2019)
43. Flanagan, R., Lacasa, L., Nicosia, V.: On the spectral properties of Feigenbaum graphs. J. Phys. A: Math. Theor. **53**(2), 025702 (2019)
44. https://onezero.medium.com/training-a-neural-network-can-emit-more-than-600-000-pounds-of-co2-but-not-for-long-d3074c1383fd
45. https://www.nature.com/articles/d41586-019-03013-5

46. Ranjan, P., Shrivastav, K.D., Vadlamani, S., Janardhanan, R.: HRIDAI:A tale of two categories of ECGs. In: International Symposium on Signal Processing and Intelligent Recognition Systems, pp. 243–263. Springer, Singapore (2020)
47. Kermali, M., Khalsa, R.K., Pillai, K., Ismail, Z., Harky, A.: The role of biomarkers in diagnosis of COVID-19—a systematic review. Life Sci. **254**: 117788 (2020). https://doi.org/10.1016/j.lfs.2020.117788. Epub 2020 May 13. PMID: 32475810; PMCID: PMC7219356
48. Leulseged, T.W., Hassen, I.S., Ayele, B.T., Tsegay, Y.G., Abebe, D.S., Edo, M.G., Maru, E.H. et al.: Laboratory biomarkers of COVID-19 disease severity and outcome: Findings from a developing country. Plos one **16**(3), e0246087 (2021)
49. Shrivastav, K.D., Arambam, P., Das, A.M., Saeed, S., Kaul, U., Ranjan, P., Janardhanan, R.: Earth mover's distance-based automated geometric visualization/classification of electrocardiogram signals. Trends Commun. Cloud Big Data 75–85 (2020)
50. Taneja, A., Ujlayan, A., Janardhanan, R., Ranjan, P.: Pancreatic cancer detection by an integrated level set-based deep learning model. Big Data Artif. Intell. Healthc. Appl. (2021). https://doi.org/10.1201/9781003093770-8
51. Khatri, A., Jain, R., Vashista, H., Mittal, N., Ranjan, P., Janardhanan, R.: Pneumonia identification in chest X-ray images using EMD. Trends Commun. Cloud Big Data 87–98 (2020). 9780367698393 C008.indd 180 14/06/21 11:24 PM
52. Goyal, A., Tirumalasetty, S., Bathla, D., Arya, M.K., Agrawal, R., Ranjan, P., Hossain, G., Challoo, R.: A computational segmentation tool for processing patient brain MRI image data to automatically extract gray and white matter regions. In: Emerging Research in Computing, Information, Communication and Applications, pp. 1–16. Springer, Singapore (2019)
53. Shrivastav, K.D., Das, A.K., Singh, H., Ranjan, P., Janardhanan, R.: Classification of colposcopic cervigrams using EMD in R. In: International Symposium on Signal Processing and Intelligent Recognition Systems, pp. 298–308. Springer, Singapore, 2018
54. Taneja, A., Ranjan, P., Ujlayan, A.: Multi-cell nuclei segmentation in cervical cancer images by integrated feature vectors. Multimed. Tools Appl. **77**(8), 9271–9290 (2018)
55. Taneja, A., Ranjan, P., Ujlayan, A.: An efficient SOM and EM-based intravascular ultrasound blood vessel image segmentation approach. Int. J. Syst. Assur. Eng. Manag. **7**(4), 442–449 (2016)
56. Taneja, A., Ranjan, P., Ujlayan, A.: Novel texture pattern based multi-level set segmentation in cervical cancer image analysis. In: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), p. 76. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp) (2016)
57. Agarwal, A., Ali, F., Kopparthi, A., Shrivastava, K.D., Ranjan, P., Jananardhanan, R.: Specular reflection removal in cervigrams. In: Proc. of CCB. Springer-Nature (2020)
58. Chavan, N., Ranjan, P., Kumar, U., Shrivastav, K.D., Janardhanan, R.: Kloman Meter: an EMD based tool for triaging lung diseases. In: Proc. of CCB. Springer-Nature (2020)
59. Chavan, N., Ranjan., P., Kumar, U., Shrivastav, K.D., Janardhanan, R.: EMD-based triaging of pulmonary diseases using chest radiographs (X-Rays). In: Metaheuristic Algorithms in Industry 4.0. Routledge (2021)
60. Hamming, R.W.: Error detecting and error correcting codes. The bell system technical journal **29**(2), 147–160 (1950). https://doi.org/10.1002/j.1538-7305.1950.tb00463.x
61. Ipsen, M., Mikhailov, A.S.: Evolutionary reconstruction of networks. Phys. Rev. E-Stat., Nonlinear, Soft Matter Phys. **66**(4), 6–9 (2002)
62. Jurman, G., Visintainer, R., Filosi, M., Riccadonna, S., Furlanello, C.: The HIM glocal metric and kernel for network comparison and classification. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 1–10. IEEE (2015)
63. https://github.com/rgbnihal2/COVID-19-X-ray-Dataset
64. https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
65. Bhatia, V., Ranjan, P., Taneja, N., Singh, H., Janardhanan, R.: Early and precision-oriented detection of cervical cancer. Big Data Artif. Intell. Healthc. Appl. (2021). https://doi.org/10.1201/9781003093770-9