

SYSTEM AND METHOD FOR PREDICTING CARDIOVASCULAR CONDITIONS IN SUBJECTS BY RETRIEVING AND ANALYZING HEART SOUNDS

DESCRIPTION

TECHNICAL FIELD

5 The present disclosure relates generally to the field of precision cardiology; more specifically, the present disclosure relates to a system and method for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds using statistical heuristics-based approaches and/or Machine Learning based approaches and/or deep learning-based modelling.

10 BACKGROUND

The detection of abnormalities in the frequencies associated with the heart sound is crucial for several reasons, as it provides valuable information about the cardiovascular system. Methods currently known for the diagnosis of cardiovascular diseases or pathologies are detected by recording the frequencies associated with the mechanical
15 vibrations due to the closing and opening of the heart valves. One of the first forms of diagnostic test based on the detection of the heartbeat and other acoustic signals is the auscultation of the chest, performed using the stethoscope. This technique is based on listening to sounds and noises coming from the chest, which can indicate the presence of any pathologies when they differ from the regular ones. More advanced techniques
20 for detecting this type of signal make use of more accurate instruments such as, for example, digital stethoscopes. Furthermore, the development of phonocardiography has made it possible to graphically represent the detected signals. These techniques require positioning the instrument close to the heart to more accurately detect sound signals.

Since the signals are also propagated in peripheral areas of the body, for example,
25 through the arteries, techniques have been developed that allow vibrations to be

detected in locations remote from the heart. This type of technique is based on the use of radar or visible light sources. The advantage lies in the detection that does not require direct contact with the patient. However, remote diagnostic techniques allow only a fraction of the vibration frequencies coming from the heart to be detected. Mostly low-
5 frequency signals are detected, i.e. those from which information can only be obtained about the rate of the heartbeat.

To overcome these limitations in the processing of the heart sound data obtained, many sound-based diagnostic approaches for health assessment have been disclosed in the prior art. However, the processing methods disclosed in the prior art are limited.
10 Therefore, there is a need to address the aforementioned technical drawbacks in existing technologies to efficiently quantify and analyse heart sounds using biophotonics and machine learning for precision cardiology.

SUMMARY

The present disclosure describes methods and systems for detection, screening,
15 diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds. Embodiments of the present invention overcome the problems encountered in the prior art by enabling a non-contact approach to capture digital biomarkers for the detection, screening, diagnosis, and prediction of cardiovascular conditions. By decoding heart sounds at high frequency and high signal-to-noise-ratio
20 and with automated calibration of laser device and camera unit to capture optimal signals, the system helps to provide tailored medical care, diagnostics, and treatments to the individual characteristics of each subject. The system and method of the disclosure allow for the detection, screening, diagnosis, and predictive/prognostic modelling of cardiovascular conditions which can be detected by assessing the sounds of the heart
25 created by the vibration caused by the mechanical contraction of the heart muscle, heart

valves closing and opening, and the cardiovascular system's laminar and turbulent blood flow.

According to a first aspect, the present disclosure provides a system for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject by
5 retrieving and analyzing heart sounds, wherein the system comprises:

a laser device configured to generate a first laser signal comprising a wavelength between 400 nm and 2500 nm with a power of at least 0.1 mW to at most 5 mW, wherein the first laser signal is directed to a region comprising neck area and/or chest area of the subject to obtain a second laser signal by the reflection of the first laser signal;

10 a camera unit configured to (i) acquire the second reflected laser signal periodically at a signal acquisition frequency of at least 600 Hz, and (ii) generate a plurality of image frame data comprising spots, wherein each of the acquisitions generates an image comprising spots;

a control unit configured to perform real-time calibration of the laser device and
15 the camera unit parameters;

a server communicatively connected to the laser device, the camera unit, and the control unit, wherein the server comprises a processor and a memory that stores a set of machine-readable instructions operable, when executed by the processor, to:

process the plurality of image frames to extract time-series heart sound
20 data by (i) detecting variation of the position of each of the spots between consecutive image frames, (ii) obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal comprises the overall variation of the image frames collected consecutively in a time interval, (iii) converting the raw signal into amplitude to obtain the time-series
25 heart sound data,

automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time series heart sound data using a statistical model and/or a machine learning model, and

predict the cardiovascular conditions of the subject using the machine
5 learning model on the heart sound data.

Embodiments of the present disclosure eliminate the aforementioned drawbacks in existing known approaches for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds. The advantage of the embodiments according to the present disclosure is that the
10 embodiments enable healthcare providers to optimize the prevention, diagnosis, and treatment of cardiovascular diseases by considering the unique characteristics of each subject. The system identifies and extracts heart sounds from regions with optimal signal quality by segmenting the second laser signal and compares amongst each other to assess the quality of the decoded data and if there is variation in Signal-to-Noise Ratio.
15 The system helps in the classification between subjects with cardiovascular disease and healthy subjects. The system further assesses the severity of the pathology by providing a probability score, where the higher probabilities indicate higher severity of the illness which can be further stratified into mild, moderate, and severe cardiovascular conditions. The system provides detection and subtyping labels for cardiology conditions such as
20 valvular heart diseases, arrhythmia, coronary artery disease, etc. It further quantifies the corresponding cardiology comorbidities present in the subject into mild, moderate, and severe, enabling better triaging and care.

One advantage of the system is that of carrying out a check on the subject without the need for direct contact with the subject and facilitate analysis to be carried out remotely.
25 The system also has the advantage of exploiting a less complex arrangement with

respect to known solutions, which can be exploited in different environments, not necessarily in an outpatient setting, with the advantage of being able to carry out an analysis at home, in the case of subjects who are unable to travel.

According to a second aspect, the present disclosure provides a method for detection,
5 screening, diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds, wherein the method comprises the steps of:

generating a first laser signal comprising a wavelength between 400 nm and 2500 nm with a power of at least 0.1 mW to at most 5 mW by a laser device, wherein the first laser signal is directed to a region comprising neck area and/or chest area of the subject
10 to obtain a second laser signal by the reflection of the first laser signal;

acquiring the second reflected laser signal periodically at a signal acquisition frequency of at least 600 Hz by a camera unit, wherein parameters of the laser device and the camera unit are calibrated in real-time through a control unit;

generate a plurality of image frame data comprising spots, wherein each of the
15 acquisitions generates an image comprising spots;

receiving the plurality of image frame data comprising the spots from the camera unit by a server, wherein the server is communicatively connected to the laser device, the camera unit and the control unit, wherein the server comprises a processor and a memory that stores a set of machine-readable instructions operable, when executed by
20 the processor, to:

process the plurality of image frames to extract time-series heart sound data by (i) detecting variation of the position of each of the spots between consecutive image frames, (ii) obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal
25 comprises the overall variation of the image frames collected consecutively in a

time interval, and (iii) converting the raw signal into amplitude to obtain the time-series heart sound data,

automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time series heart sound data using a statistical model and/or a machine learning model, and

predict the cardiovascular conditions of the subject using the machine learning model on the heart sound data.

The method identifies and extracts heart sounds from regions with optimal signal quality by segmenting the second laser signal and comparing amongst each other to assess the quality of the decoded data and if there is variation in Signal-to-Noise Ratio. The method helps in the classification between subjects with cardiovascular disease and healthy subjects. The method further assesses the severity of the pathology by providing a probability score, where the higher probabilities indicate higher severity of the illness - which can be further stratified into mild, moderate, and severe cardiovascular conditions.

The method provides detection and subtyping labels for cardiology conditions such as aortic stenosis, arrhythmia, coronary artery disease, etc. The predicted cardiovascular conditions by the method of the present disclosure leading to a cardiovascular pathology, using pharmacological and/or lifestyle change and/or other therapies can change the illness's (sound) signatures and nudge it towards the healthy baseline and/or make sure the condition does not deteriorate. This feedback loop is integral to enabling preventative medicine.

Additional aspects, advantages, features, and objects of the present disclosure are made apparent from the drawings and the detailed description of the illustrative embodiments construed in conjunction with the appended claims that follow. It will be appreciated that features of the present disclosure are susceptible to being combined in various

combinations without departing from the scope of the present disclosure as defined by the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The summary above, as well as the following detailed description of illustrative
5 embodiments, is better understood when read in conjunction with the appended drawings. To illustrate the present disclosure, exemplary constructions of the disclosure are shown in the drawings. However, the present disclosure is not limited to specific methods and instrumentalities disclosed herein. Moreover, those in the art will understand that the drawings are not to scale. Wherever possible, the same elements
10 have been indicated by identical numbers. Embodiments of the present disclosure will now be described, by way of example only, with reference to the following diagrams wherein:

FIG. 1 is a schematic illustration of system for predicting cardiovascular conditions in subjects by retrieving and analyzing heart sounds according to an embodiment of the
15 present disclosure;

FIG. 2 is a schematic illustration of the server of **FIG.1** including various modules according to an embodiment of the present disclosure;

FIGS. 3A-C is a flowchart illustrating a method for predicting cardiovascular conditions in subjects by retrieving and analyzing heart sounds according to an embodiment of the
20 present disclosure; and

FIG. 4 is a schematic diagram of a computer architecture in accordance with the embodiments of the present disclosure.

DETAILED DESCRIPTION OF EMBODIMENTS

The following detailed description illustrates various embodiments of the present disclosure. Those skilled in the art would recognize that other embodiments for carrying out or practicing the present disclosure are also possible.

The measures, values, shapes and geometric references (such as perpendicularity and parallelism), when associated with words such as “approximately” or other similar terms such as “almost” or “substantially”, are to be understood as less of measurement errors or inaccuracies due to production and/or manufacturing errors and, above all, unless there is a slight deviation from the value, measure, shape or geometric reference to which it is associated. For example, these terms, if associated with a value, preferably indicate a deviation of no more than 10% of the value itself.

Furthermore, when used, terms such as “first”, “second”, “superior”, “inferior”, “principal” and “secondary” do not necessarily identify an order, relationship priority or relative position, but can simply be used to more clearly distinguish between different components.

Unless otherwise specified, as apparent from the following discussions, terms such as “processing”, “computing”, “determination”, or the like are understood to refer to the action and/or processes of a computer or similar electronic calculation that manipulates and/or transforms represented data as physical, such as electronic quantities of computer system registers and/or memories in, other data similarly represented as physical quantities within computer systems, registers or other information storage, transmission or display devices. The measurements and data reported in this text are to be considered, unless otherwise indicated, as performed in an ICAO International Standard Atmosphere (ISO 2533:1975).

The diagnostic procedure is based on the detection of acoustic waves deriving from physiological processes which mainly, but not exclusively, affect the cardiovascular

system and which, in addition to the signals relating to the regular functioning of the organs, can include further less common signals associated with the presence of pathologies in the patient tested.

According to a first aspect, the present disclosure preferably provides a system for
5 detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds.

The system preferably comprises a laser device preferably configured to generate a first laser signal. Said laser signal preferably comprising a wavelength between 400 nm and 2500 nm and preferably having a power of at least 0.1 mW to at most 5 mW. The first
10 laser signal is preferably directed to a region comprising neck area and/or chest area of the subject to obtain a second laser signal by the reflection of the first laser signal;

The system, preferably, also comprises a camera unit configured to (i) acquire the second reflected laser signal periodically at a signal acquisition frequency which is preferably of at least 600 Hz, and preferably (ii) generate a plurality of image frame data
15 comprising spots. Each of the acquisitions preferably generates an image comprising spots;

The system, preferably, also comprises a control unit configured to perform real-time calibration of the laser device and the camera unit parameters.

The system, preferably, also comprises a server communicatively connected to
20 the laser device, the camera unit, and the control unit, wherein the server comprises a processor and a memory that stores a set of machine-readable instructions operable, when executed by the processor, to:

preferably, receive the plurality of image frame data comprising spots from
the camera unit,

preferably process the plurality of image frames to extract time-series heart sound data by (i) preferably detecting variation of the position of each of the spots between consecutive image frames, (ii) preferably obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal preferably comprises the overall variation of the image frames collected consecutively in a time interval, and (iii) preferably converting the raw signal into amplitude to obtain the time-series heart sound data,

preferably automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs and preferably by analysing the time series heart sound data more preferably using a statistical model and/or a machine learning model, and

preferably predict the cardiovascular conditions of the subject using the machine learning model on the heart sound data.

The advantage of the embodiments according to the present disclosure is that the embodiments enable healthcare providers to optimize the detection, prevention, diagnosis, and treatment of cardiovascular diseases by considering the unique characteristics of each subject. The system preferably identifies and preferably extracts heart sounds from regions with optimal signal quality by segmenting the second laser signal and compares amongst each other to assess the quality of the decoded data and if there is variation in Signal-to-Noise Ratio. The system helps in the classification between subjects with cardiovascular disease and healthy subjects. The system preferably further assesses the severity of the pathology preferably by providing a probability score, where the higher probabilities indicate higher severity of the illness which can be further stratified into mild, moderate, and severe cardiovascular conditions. The system preferably provides detection and subtyping labels for cardiology conditions

such as valvular heart disease, arrhythmia, coronary artery disease, etc. It preferably further quantifies the corresponding cardiology comorbidities present in the subject into mild, moderate, and severe, enabling better triaging and care.

The laser device may point the light at the base of the subject's neck region and/or chest area to extract the heart sound data. The laser device may be a diode laser device. The diode laser device is advantageous since the device can have small and compact dimensions and allows to make the system easier to install. The first laser signal is a laser beam with a diameter preferably between 2 mm and 6 mm. The first laser signal may include a power of at least 2 mW. The first laser signal may preferably include wavelengths between 500 nm and 560 nm, more preferably between 510 nm and 550 nm, even more preferably between 520 nm and 540 nm. These wavelength values have the advantage of allowing the camera unit to improve signal acquisition.

The camera unit is positioned at a distance between 0.05 m and 15 m. More preferably it is placed at a distance between 0.25 m and 1.0m, even more preferably between 0.5 m and 0.75 m. In detail, the camera unit is preferably placed at a greater distance from the focal distance by a value between 10 cm and 30 cm, more preferably between 15 cm and 25 cm. This distance has the advantage of facilitating the visualization of the second laser signal on the screen of the camera unit. In some embodiments, the camera unit is configured to acquire the second reflected laser signal periodically at a signal acquisition frequency of at least 0.8 kHz, more preferably of at least 1.0 kHz, even more preferably of at least 1.2 kHz.

The camera unit may be provided with other supplementary parts such as lenses with filters to capture the reflected light. A secondary camera may be provided with or without edge-compute models for automated detection of the base of the subject's neck region and/or chest area. The generated image frame data from the camera unit is

communicated to the server through a wired data-transfer module and/or Bluetooth module and/or Wifi module. The image frame data may be captured by the camera unit in MP4 format files or other relevant formats including NumPy arrays, etc. and communicated to the server. The server preferably comprises additional modules to
5 provide data collection feedback, data anonymization, pre-and/or post-processing for converting the second laser signal to the heart sound data. The image frame data may be recorded and stored in an encrypted and anonymized format on the local [device/clinic/hospital/ institute] storage and/or in a cloud database.

The time series heart sound data may be divided into equal or variable length segments,
10 typically of 5-10 seconds duration for further analysis. The backend of the machine learning model preferably uses shell scripting, Python with its corresponding signal processing and ML/AI libraries, Docker, data protection & anonymization protocol systems, cloud computing with real-time APIs for ease of data transfer and analysis, and signal processing libraries. Finally, the inferences from the analysis preferably are put
15 into reports/outputs that are automatically generated to help general practitioners and/or automated patient triaging systems and/or cardiologists and/or other clinicians to better manage patients and provide personalised patient care.

In some embodiments, a motion description algorithm is used to quantify image motion between the consecutive image frames. The motion is estimated between consecutive
20 frames and then quantified to convert the image into time series heart sound data. Accordingly, a first algorithm reconstructs the maps of the displacements of the spots in each pixel of the surface of the camera unit. In this way, the first algorithm calculates the vectorial displacement along the surface of each of the spots. From the displacement of each spot of the same image acquired, the first algorithm calculates the mean vector
25 displacement of the entire image. The first algorithm can also show graphic maps in

which the spots and the vectorial displacements associated with the spots and the entire images are reported. The first algorithm generates a raw signal. It is a function obtained from the interpolation of the movements recorded for each acquisition instant of the images. Therefore, the raw signal includes the overall variation of images collected
5 consecutively in a time interval. The raw signal contains both information of diagnostic interest and disturbing elements.

In a filtration sub-step, a second algorithm is used. It analyzes the raw signal and transforms it into a frequency spectrum. Therefore, it comprises components, each having a unique frequency. The components therefore have different origins and it is
10 necessary to eliminate the components of interest from those that are not necessary. Therefore, the second algorithm carries out an operation of elimination of part of the components obtaining the residual components. The residual components are therefore at least the components having the frequencies associated with the phenomena of interest. In particular, the components having frequencies having values from 0 Hz to a
15 value between 15 Hz and 25 Hz, more preferably between 17 Hz and 23 Hz, even more preferably between 18 Hz and 22 Hz are preferably removed from the spectrum. In fact, in these frequency ranges no signals associated with known phenomena of diagnostic interest are detected. Therefore, the components that fall into this frequency range are removed from the spectrum. The residual components constitute the signals. The latter
20 are the signals of diagnostic interest. The signals separated in frequency are correlated to the presence of pathologies. It is known in the literature that the main sound coming from the heart corresponds to a first signal which can vary from 10 Hz to 140 Hz and is due to the closure of the mitral and tricuspid valves. This signal is preferably distinguishable by a secondary sound coming from the heart, separated by the systolic
25 pause from the first. This signal can vary from 10 Hz to 400 Hz and is caused by the

closure of the aortic and pulmonary valves. This type of signal is obviously present in every patient. Other signals that can be detected reach higher frequencies. In fact, they can reach frequencies ranging from 20 Hz up to 1000 Hz. In particular, these signals are indicators of the presence of pathologies. In particular, the frequency range of the signals

5 between a value of at least 15 Hz and a value between 400 Hz and 1000 Hz is preferably analysed, preferably between a value of 650 Hz and 750 Hz, even more preferably between a value of 675 Hz and 725 Hz.

The heart sound data preferably include the prominent S1 and S2 sounds caused by the heart valves opening and closing, the S3 and S4 sounds which might or might not

10 indicate pathology, the corresponding murmurs as blood flows, and/or the turbulent moment of the blood as it interacts with plaques and cholesterol in the arteries. The server preferably automatically quantifies heart sounds & murmurs, and labels them into S1, S2, S3, S4, and/or murmurs using a statistical model and/or a machine learning model on the heart sound data.

15 The statistical model employed to analyse the heart sound data extracts simple heart sound time domain statistical features like mean, median, variance, standard deviation, skewness, kurtosis, etc. Additionally, the other possible combined features such as Peak-Peak Mean, Mean Square Value, Hjorth Parameter Activity, Hjorth Parameter Mobility, Hjorth Parameter Complexity, Maximum Power Spectral Frequency, Maximum

20 Power Spectral Density, Power Sum, etc are derived. The non-linear entropy features such as Shannon Entropy, Singular Entropy, Kolmogorov Entropy, Approximate Entropy, C0 Complexity, Correlation Dimension, Lyapunov Exponent, Permutation Entropy, Spectral Entropy, etc. are also derived. When required, ICA and other wavelet-based features are obtained to derive more potent features that are prevalent in cardiology

25 illness and co-morbid groups. All these help in the better interpretable analysis of the

signals and in identifying the signatures of the pathology. Features are selected based on the information on the illness groups. Low variance filters, high correlation filters, random forests, forward feature selection, etc. are used for automated feature selection. Gender, age, and BMI are usual covariates that affect the features selection process.

- 5 Based on the illness pathology suspected and the ML models used, different features are automatically weighed and selected for the subject. Multiple features are also merged/combined to arrive at a new derived feature that has more information on the illness pathology, which can have better predictive value in the classification and prognosis. The system may use dimensionality reduction algorithms such as Wavelets,
- 10 PCA, ICA, factor analysis, etc. to reduce the dimensionality of the feature set. This is used to improve the accuracy of the models with minimal data and speed of further analysis using minimal features while retaining the information around data distribution. In some embodiments, normative modeling is used to analyse the heart sound data that uses a wide spectrum of cardiac data collected from a diverse and extensive population.
- 15 The normative modelling sets 'normal' benchmarks for various cardiac parameters. When individual cardiac health data is compared to these benchmarks, any deviations from these norms can be effectively identified. The comparison can serve as an early warning system for potential heart conditions. The normative modelling provides a comprehensive, context-rich perspective for individual cardiac health assessments,
- 20 enabling personalised cardiac care and longitudinal modelling.

In some embodiments, synthetic heart sound data is generated to account for class imbalance via synthetic minority oversampling technique (SMOTE), Neural-Based Generative Models, etc. This allows for generating more data from the same distribution as the classes under investigation and further develop more robust Machine learning

25 models for the detection and prognosis of cardiology illnesses.

In some embodiments, biometric identification of the subject is performed. The heart sound data is anonymised and/or is compliant with GDPR/HIPAA. The subject is mapped to the existing ID by searching in the database or a new ID is created for the subject. The subject can be distinctly identified using their unique heart sound which is quantified using the second laser signal captured from the neck area and/or chest area of the subject. This is quantified as an anonymised ID of the subject and matched to their Patient ID on the EHR records or other systems to keep track of the distinct subject.

In some embodiments, a summarised data of the subject is generated using captured data such as heart rate (beats per minute), heart rhythm, intensity of heart sounds, duration of heart sounds, presence of additional sounds or murmurs, splitting of heart sounds, respiratory variations, etc. Heart sounds such as S1, S2, S3, S4, murmurs, etc. are automatically labelled for better data assessment and patient management. Automated reports with the summarised data of the subject are provided to the clinicians. The clinicians are given the normative, detection/illness score and/or prognostic scores, and/or the subjects are given the normative and feedback scores. The report also includes the subjective data collected using EHR, questionnaire scores, symptom scores, etc. This ensures a holistic approach to patient and clinical reporting.

The server may be an on-edge PCB/microcontroller and/or a cloud server. In some embodiments, processing of the plurality of image frames to extract the time-series heart sound data is performed in the on-edge PCB/microcontroller and quantification of heart sounds, and prediction of the cardiovascular conditions using the statistical model and/or the machine learning model are performed in the cloud server. The time-series heart sound data may be transferred from the on-edge PCB/microcontroller using Wi-Fi and/or a Bluetooth module to the cloud server or any other external devices for the quantification

of the heart sounds, and prediction of cardiovascular conditions using the statistical model and/or the machine learning model.

Optionally, the machine learning model predicts the cardiovascular conditions in the subject based on normative percentile scores, longitudinal scores and prognostic scores
5 of the heart sound data.

The system may use dimensionality reduction algorithms such as Wavelets, PCA, ICA, factor analysis, etc. to reduce the dimensionality of the feature set. This is used to improve the accuracy of the models with minimal data and speed of further analysis using minimal features while retaining the information around data distribution. Shallow
10 learning algorithms such as logistic regression, support vector machine, random forest, decision trees, naive Bayes, etc., and/or deep learning-based algorithms such as Convolutional Neural Networks (CNNs), Transformer-based Models, Recursive Neural networks (RNNs), Long Short-Term Memory (LSTM), etc may be used. Unsupervised clustering such as k-means, t-SNE, etc. may be used on heart sound feature data to
15 arrive at data-driven illness clusters and comorbid groups. This allows for better prognostic predictability. Further, Bayesian, Monte Carlo Simulation, etc. approaches are used to derive probabilistic illness stage and/or progress. As more data points are obtained, machine learning based supervised and/or unsupervised learning approaches are used to learn across different subjects to further improve the accuracy of
20 cardiovascular disorder diagnostics. Ensemble features and/or models are further used to improve the predictability of the model.

The change in the subject's acoustics of the heart over time is estimated by the features across the same subject to obtain functioning and health of the heart and the circulatory system. This allows assessing the changes in illness pathology. One way to achieve this
25 is by assessing the normative scores of the subject over two different time points. This

can have implications in deriving the prognostic scores by determining if the changes in the score are towards or away from the healthy baseline.

The machine learning model uses interpretable models such as Grad-CAM, Saliency map, etc. to derive the acoustic (sound) markers leading to a cardiovascular pathology, which can, in turn, be modulated using pharmacological and/or lifestyle change and/or other therapies that can change the illness's (sound) signatures and nudge it towards the healthy baseline and/or make sure the illness does not deteriorate. This feedback loop is the key to enabling preventative medicine. The efficacy of the system and the accuracy of the outputs of each stage may be assessed using metrics such as percentile deviation, positive predictive value (PPV), confusion matrix, accuracy, precision, recall (sensitivity, True Positive Rate), specificity, F1 score, Precision-Recall (PR) curve, Receiver Operating Characteristics (ROC) curve, PR vs ROC curve, etc. As the model gets more data, it is iteratively trained to improve its performance.

Optionally, the time series heart sound data comprises a frequency having a value between 20 Hz and 1000 Hz.

Optionally, the server is further configured to automatically detect the neck area, and/or the chest areas of the subject and control the movement of the laser device to direct the first laser signal and the camera unit to capture the second laser signal through the control unit.

Optionally, the camera unit is a high-frequency CCD (charge-coupled device), CMOS (complementary metal oxide semiconductor), sCMOS (scientific Complementary Metal-Oxide-Semiconductor) image sensor, Raspberry Pi camera.

Optionally, the camera unit captures the second laser signal at a distance between 0.05 m and 15 m.

Optionally, the system comprises a low pass filter to filter frequencies less than 20 Hz.

Optionally, the second laser signal is captured at 400x400 FOV resolution by the camera unit and segmented into four 200x200 FOV frames and/or sixteen 100x100 FOV frames and/or sixty-four 50x50 FOV and/or two hundred and fifty-six 25x25 FOV, and processed separately to obtain corresponding heart sound signals.

- 5 Optionally, the server is configured to process the plurality of image frames to extract the time-series heart sound data using image motion description models or motion tracking models.

Optionally, the server is configured to compare the time-series sound data among each other to determine variation in signal-to-noise ratio.

- 10 Optionally, the server is configured to perform the real-time calibration of the laser device and the camera unit through the control unit based on the determined variation in Signal-to-Noise Ratio to acquire optimal signals.

Recording the heart sounds data at an optimal signal-to-noise ratio requires precise alignment and/or parameters of the laser device and the camera unit so that the reflected
15 light from the laser device shone on the subject is being captured by the camera unit. Additionally, the lens of the camera unit makes sure that the second laser signal being recorded by the camera unit is focused, enabling the capture of robust signals optimal for further analysis. The server provides feedback to the laser device that enables the real-time quick calibration of the laser device via setting the configuration & parameters.

- 20 The quality of the heart sound data is assessed using the quality analysis /quality control module. Depending on the quality of the heart sound data, the server provides feedback to the laser device to re-adjust the configuration and/or parameters to acquire optimal signals. The variables, including patient information such as skin tone, height, weight, Body mass index (BMI), past clinical data, etc obtained from the Electronic
25 Health/Medical Record (EHR/EMR) system are used to further optimise the parameters

of the laser device (i.e.) and laser point location to gather heart sounds data at optimal SNR & frequency for further analysis. A scoring on a scale of 1-10 on the quality of the signal captured may be provided, 1 being data of very low quality and 10 being data of high quality to quantify heart sounds. This is given as feedback to the laser device and
5 the camera unit to enable automated calibration.

Optionally, the machine learning model is trained using a multimodal dataset comprising (i) phonocardiogram (PCG) data of age and sex-matched subjects with labelled ID and/or cardiovascular diseases and/or heart conditions and/or heart murmurs and/or heart sound labels such as S1, S2, etc., and (ii) clinical/symptoms data and/or demography
10 data of the subject.

The Clinical data includes prior illness and relapse, dietary changes/restrictions, blood parameters, past/current medication and routine, different therapeutic regimes, etc. The demographic data includes patient skin tone, patient age, gender, marital status, family size, ethnicity, income range, education, etc. The Clinical data and the demographic data
15 may be anonymized and automated to maintain confidentiality. The patient and/or clinical general physicians, nurses, EHR database systems, etc may manually or automatically upload the required clinical and demographic data into the database.

To enable robust predictive modeling using minimal data collected using the camera unit, a baseline model is trained using phonocardiogram (PCG) and a transfer learning
20 approach is used to further optimise the weights of the models using data collected for the camera unit. This enables generalisable diagnostic & prognostic models for clinical utility. The stages of this is as follows:

Phase 1: Building the Initial Model on PCG Data

Phonocardiogram (PCG) datasets are preferably used to train, test, and validate the
25 initial baseline model. The dataset includes the relevant subjects with heart murmurs

and/or heart valve defects and/or coronary artery disease (labels for the model). The dataset to train the baseline model is taken from a representative population to account for Age, Sex and other covariates

The initial model is preferably trained using PCG data of age and sex-matched
5 individuals with mild, moderate, and severe valvular heart conditions and further accounted for other covariates such as BMI, skin tone, other clinical comorbidities, etc. The model is preferably also tested and validated using the PCG data from the same distribution. The model uses multimodal data to train (i.e.) in addition to PCG data, age, sex, and other clinical indications are used as input to the model. At the end of this step,
10 a robust baseline model that accounts for covariates is formed. This model can be used for transfer learning.

Phase 2: Transfer Learning

The data collected from the laser device and the camera unit with better SNR & higher frequency is preferably used to further train and optimise the weights of the baseline
15 model. The subject data distribution in the PCG dataset is matched with the data collected from subjects using the laser device and the camera unit but doesn't necessarily need to since the final layers of the model are optimised using a multimodal dataset. The newly trained model is preferably compared against an out-of-sample validation set of data collected from the laser device and the camera unit to obtain the
20 accuracy of the newly optimised model. The data distribution in both training and validation will be similar, accounting for covariates such as age, sex, and other clinical indications.

Optionally, the machine learning model is trained and/or modelled on a labelled dataset using the extracted features as input and corresponding labels as target outputs.

Optionally, the server is connected to an Electronic Health/Medical Record (EHR/EMR) and/or clinical/hospital/local backend server system to obtain the demography data comprising height, weight, Body mass index (BMI) and clinical/symptoms data of the subject.

- 5 Optionally, the machine learning model is trained on a labelled subject ID dataset for Biometric authentication using the subject's heart sound data.

Optionally, total energy distribution in the recorded second laser signal is integrated across the image frame. The integration is performed separately for each of the plurality of image frames to convert into the time series heart sound data.

- 10 Optionally, the normative percentile score is determined based on a derived heart sound metrics of the subject against a general population's heart sound metric considering factors such as sex, age, and other relevant statistical variables.

Optionally, the longitudinal score is determined based on the change in baseline of the derived heart sound metrics of the subject from one visit to the next and/or after

- 15 therapeutic intervention.

Changes in the baseline heart sound data of a subject from one visit to the next and/or after therapeutic intervention compare the new activity state to normative modeling, e.g. is the heart sound signatures getting better or worse, is the therapeutic intervention working for the individual, etc. This further helps to optimise a better treatment plan.

- 20 Optionally, the server is configured to display the heart sounds in real-time on an interactive user interface or output the heart sounds via speaker/headphones in real-time.

Optionally, the server is configured to classify the heart murmurs into Systolic murmur, Diastolic murmur, and/or Continuous murmur using the machine learning model on the

- 25 heart sound data.

According to a second aspect of the present disclosure, a method for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds, wherein the method comprises the steps of:

preferably generating a first laser signal comprising a wavelength between 400
5 nm and 2500 nm preferably with a power of at least 0.1 mW to at most 5 mW by a laser device, wherein the first laser signal is directed to a region comprising neck area and/or chest area of the subject to obtain a second laser signal by the reflection of the first laser signal;

preferably acquiring the second reflected laser signal periodically at a signal
10 acquisition frequency of at least 600 Hz by a camera unit, wherein parameters of the laser device and the camera unit are preferably calibrated in real-time through a control unit;

preferably generate a plurality of image frame data comprising spots, wherein each of the acquisitions preferably generates an image comprising spots;

15 r preferably receiving the plurality of image frame data comprising the spots from the camera unit by a server, wherein the server is communicatively connected to the laser device, the camera unit and the control unit, wherein the server preferably comprises a processor and a memory that stores a set of machine-readable instructions operable, when executed by the processor, to:

20 preferably process the plurality of image frames to extract time-series heart sound data by (i) preferably detecting variation of the position of each of the spots between consecutive image frames, (ii) preferably obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal comprises the overall variation of the image frames collected

consecutively in a time interval, (iii) preferably converting the raw signal into amplitude to obtain the time-series heart sound data,

preferably automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by preferably analysing the time series

5 heart sound data using a statistical model and/or a machine learning model, and

preferably predict the cardiovascular conditions of the subject preferably using the machine learning model on the heart sound data.

To obtain the heart sound data, the laser device and the camera unit are preferably arranged in a mutually determined position and the subject is brought into a position and
10 maintained stably until the next acquisition phase. The subject may be in a sitting position, standing position or a lying position that causes a portion of the subject's skin to be invested by a first laser signal preferably coming from the laser device. The skin portion of the subject preferably corresponds to the base of the neck of the subject. This area of the body, being highly vascularized, makes it easier to detect the vibrations
15 associated with physiological processes and the functioning of internal organs. When abnormal heart sounds associated with anomalies are detected, this procedure makes it possible to carry out a preliminary analysis which can refer to further more in-depth checks.

Optionally, the machine learning model predicts the cardiovascular conditions in the
20 subject based on normative percentile scores, longitudinal scores and prognostic scores of the heart sound data.

Individualised optimal therapeutic strategy can be prognostically predicted amongst pharmacological medications, lifestyle change recommendations, etc. for the subject to return to health control baseline and help in illness alleviation. The heart and circulatory
25 system health of the subjects is monitored over time and the optimal therapeutic

interventions change the acoustic/sound states are assessed and further modulates the intervention when required based on the subject's response to the therapy, leading to personalised treatment plan.

Embodiments of the present disclosure substantially eliminate or at least partially address the aforementioned technical drawbacks in existing technologies in providing a system and method for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject by retrieving and analyzing heart sounds using statistical heuristics-based approaches and/or machine learning approaches and/or deep learning based modelling.

10 DETAILED DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic preferred illustration of a system **100** for detecting, screening, diagnosing, and predicting cardiovascular conditions in subjects by retrieving and analyzing heart sounds according to an embodiment of the present disclosure. The system **100** preferably includes a laser device **102** preferably configured to generate a first laser signal comprising a wavelength between 400 nm and 2500 nm preferably with a power of at least 0.1 mW to at most 5 mW. The first laser signal is preferably directed to a region comprising neck area and/or chest area of the subject **106** to preferably obtain a second laser signal by the reflection of the first laser signal. A camera unit **104** is preferably configured to (i) preferably acquire the second reflected laser signal periodically at a signal acquisition frequency of at least 600 Hz, and (ii) preferably generate a plurality of image frame data comprising spots. Each of the acquisitions preferably generates an image comprising spots. A control unit **108** is preferably configured to perform real-time calibration of parameters of the laser device **102** and the camera unit **104**. A server **110** is communicatively preferably connected to one or more of the laser device **102**, the camera unit **104**, and the control unit **108**. The server **110** is

preferably configured to receive the plurality of image frame data comprising the spots from the camera unit **104**. The server **110** is preferably configured to process the plurality of image frames to extract time-series heart sound data by (i) preferably detecting variation of the position of each of the spots between consecutive image frames, (ii) preferably obtaining the variation aggregate of the consecutive image frames to generate a raw signal including the overall variation of the image frames collected consecutively in a time interval, and (iii) preferably converting the raw signal into amplitude to obtain the time-series heart sound data. The server **110** is preferably configured to automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time-series heart sound data using a statistical model and/or a machine learning model. The server **110** is preferably configured to predict the cardiovascular conditions of the subject using the machine learning model on the heart sound data.

FIG. 2 is a schematic preferred illustration of the server **110** of **FIG.1** including various modules according to an embodiment of the present disclosure. The server **110** preferably includes one or more of database **202**, data receiving module **204**, time-series heart sound data extraction module **206**, heart sound quantification module **208**, cardiovascular condition prediction module **210**. The data receiving module **204** is preferably configured to receive the plurality of image frame data including the spots from the camera unit **104**. The time-series heart sound data extraction module **206** is preferably configured to process the plurality of image frames to extract time-series heart sound data by (i) preferably detecting variation of the position of each of the spots between consecutive image frames, (ii) preferably obtaining the variation aggregate of the consecutive image frames to generate a raw signal including the overall variation of the image frames collected consecutively in a time interval, and (iii) preferably converting the raw signal into amplitude to obtain the time-series heart sound data. The heart sound

quantification module **208** is preferably configured to automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time-series heart sound data using a statistical model and/or a machine learning model. The cardiovascular condition prediction module **210** is preferably configured to
5 predict the cardiovascular conditions of the subject using the machine learning model on the heart sound data. The server **110** preferably includes one or more of a calibration module to calibrate the parameters of the laser device **102** and the camera unit **104** in real-time through a control unit. The server **110** preferably includes a quality analysis / quality control module configured to ensure the quality and consistency of the time-series
10 heart sound data.

FIGS. 3A-3C is a flowchart illustrating a preferred method for predicting cardiovascular conditions in subjects by retrieving and analyzing heart sounds according to an embodiment of the present disclosure. At a step **302**, the method preferably includes generating a first laser signal comprising a wavelength between 400 nm and 2500 nm
15 with a preferred power of at least 0.1 mW to at most 5 mW by a laser device. The first laser signal is preferably directed to a region comprising neck area and/or chest area of the subject to obtain a second laser signal by the reflection of the first laser signal. At a step **304**, the method preferably includes acquiring the second reflected laser signal periodically at a signal acquisition frequency of at least 600 Hz by a camera unit.
20 Parameters of the laser device and the camera unit are calibrated in real-time through a control unit. At a step **306**, the method preferably includes generating a plurality of image frame data comprising spots by the camera unit. Each of the acquisitions preferably generates an image comprising spots. At a step **308**, the method preferably includes receiving the plurality of image frame data comprising the spots from the camera unit by
25 a server. At a step **310**, the method preferably includes processing by the server, the

plurality of image frames to extract time-series heart sound data by (i) preferably detecting variation of the position of each of the spots between consecutive image frames, (ii) preferably obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal comprises the overall variation of the image frames collected consecutively in a time interval, and (iii) preferably converting the raw signal into amplitude to obtain the time-series heart sound data. At a step **312**, the method preferably includes automatically determining by the server, normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time-series heart sound data using a statistical model and/or a machine learning model. At a step **314**, the method preferably includes predicting by the server, the cardiovascular conditions of the subject using the machine learning model on the heart sound data.

FIG. 4 is a schematic diagram of a computer architecture in accordance with the embodiments of the present disclosure. A representative hardware environment for practicing the embodiments herein is depicted in **FIG. 4**, with reference to **FIGS. 1** through **3**. This schematic drawing preferably illustrates a hardware configuration of a server **110** /computer system in accordance with the embodiments herein. The server **110** /computer preferably includes at least one processing device **10** and preferably a cryptographic processor **11**. The special-purpose CPU **10** and the cryptographic processor (CP) **11** may be interconnected via system bus **14** to various devices such as a random access memory (RAM) **15**, read-only memory (ROM) **16**, and an input/output (I/O) adapter **17**. The I/O adapter **17** can connect to peripheral devices, such as disk units **12** and tape drives **13**, or other program storage devices that are readable by the system. The server **110**/ computer can read the inventive instructions on the program storage devices and follow these instructions to execute the methodology of the embodiments herein. The server **110**/computer system preferably further includes a user

interface adapter **20** that connects preferably one or more of a keyboard **18**, mouse **19**, speaker **25**, microphone **23**, and/or other user interface devices such as a touch screen device (not shown) to the bus **14** to gather user input. Additionally, a communication adapter **21** connects the bus **14** to a data processing network **26**, and a display adapter **22** connects the bus **14** to a display device **24**, which provides a graphical user interface (GUI) **30** of the output data in accordance with the embodiments herein, or which may be embodied as an output device such as a monitor, printer, or transmitter, for example. Further, a transceiver **27**, a signal comparator **28**, and a signal converter **29** may be connected with the bus **14** for processing, transmission, receipt, comparison, and conversion of electric or electronic signals.

Modifications to embodiments of the present disclosure described in the foregoing are possible without departing from the scope of the present disclosure as defined by the accompanying claims. Expressions such as “including”, “comprising”, “incorporating”, “have”, “is” used to describe, and claim the present disclosure are intended to be construed in a non-exclusive manner, namely allowing for items, components or elements not explicitly described also to be present. Reference to the singular is also to be construed to relate to the plural.

LIST OF REFERENCE NUMERALS

- 100 – system
- 102 – laser device
- 104 - camera unit
- 106 – subject
- 108 – control unit
- 110 – server
- 202 – database

204 - data receiving module

206 - time-series heart sound data extraction module

208 - heart sound quantification module

210 - cardiovascular condition prediction module

5

CLAIMS

1. A system (100) for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject (106) by retrieving and analyzing heart sounds, wherein the system (100) comprises:

5 a server (110) communicatively connected to a laser device (102), a camera unit (104), and a control unit (108), wherein the laser device (102) configured to generate a first laser signal comprising a wavelength between 400 nm and 2500 nm with a power of at least 0.1 mW to at most 5 mW, wherein the first laser signal is directed to a region comprising neck area and/or chest area of the subject (106) to obtain a second laser
10 signal by the reflection of the first laser signal, wherein the camera unit (104) configured to (i) acquire the second reflected laser signal periodically at a signal acquisition frequency of at least 600 Hz, and (ii) generate a plurality of image frame data comprising spots, wherein each of the acquisitions generates an image comprising spots, wherein the control unit (108) configured to perform real-time calibration of the laser device (102)
15 and the camera unit (104) parameters, wherein the server (110) comprises a processor and a memory that stores a set of machine-readable instructions operable, when executed by the processor, to:

receive the plurality of image frame data comprising the spots from the camera unit (104),

20 process the plurality of image frames to extract time-series heart sound data by (i) detecting variation of the position of each of the spots between consecutive image frames, (ii) obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal comprises the overall variation of the image frames collected consecutively in a

time interval, (iii) converting the raw signal into amplitude to obtain the time-series heart sound data,

automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time-series heart sound data using a statistical model and/or a machine learning model, and

predict the cardiovascular conditions of the subject (106) using the machine learning model on the heart sound data.

2. The system (100) as claimed in claim 1, wherein the machine learning model predicts the cardiovascular conditions in the subject (100) based on normative percentile scores, longitudinal scores and prognostic scores of the heart sound data.

3. The system (100) as claimed in claim 1, wherein the time series heart sound data comprises a frequency having a value between 20 Hz and 2000 Hz.

4. The system (100) as claimed in claim 1, wherein the server (110) is further configured to automatically detect the neck area and/or the chest areas of the subject (106) and control the movement of the laser device (102) to direct the first laser signal and the camera unit (104) to capture the second laser signal through the control unit (108).

5. The system (100) as claimed in claim 1, wherein the camera unit (104) is a high-frequency CCD (charge-coupled device), CMOS (complementary metal oxide semiconductor), sCMOS (scientific Complementary Metal–Oxide–Semiconductor) image sensor, Raspberry Pi camera.

6. The system (100) as claimed in claim 1, wherein the camera unit (104) captures the second laser signal at a distance between 0.05 m and 15 m.

7. The system (100) as claimed in claim 1, wherein the system (100) comprises a low pass filter to filter frequencies less than 20 Hz.

8. The system (100) as claimed in claim 1, wherein the second laser signal is captured at 400x400 FOV resolution by the camera unit (104) and segmented into four 200x200 FOV frames and/or sixteen 100x100 FOV frames and/or sixty-four 50x50 FOV and/or two hundred and fifty-six 25x25 FOV, and processed separately to obtain
5 corresponding heart sound signals.
9. The system (100) as claimed in claim 1, wherein the server (110) is configured to process the plurality of image frames to extract the time-series heart sound data using image motion description models or motion tracking models.
10. The system (100) as claimed in claims 8 and 9, wherein the server (110) is
10 configured to compare the time-series sound data among each other to determine variation in signal-to-noise ratio.
11. The system (100) as claimed in claims 8 and 9, wherein the server (110) is configured to perform the real-time calibration of the laser device (102) and the camera unit (104) through the control unit (108) based on the determined variation in Signal-to-
15 Noise Ratio to acquire optimal signals.
12. The system (100) as claimed in claim 1, wherein the machine learning model is trained and/or modeled using a multimodal dataset comprising (i) phonocardiogram (PCG) data of age and sex-matched subjects with labelled ID and/or cardiovascular diseases and/or heart conditions and/or heart murmurs and/or heart sound labels such
20 as S1, S2, etc., and (ii) clinical/symptoms data and/or demography data of the subject (106).
13. The system (100) as claimed in claim 12, wherein the machine learning model is trained and/or modeled on a labelled dataset using the extracted features as input and corresponding labels as target outputs.

14. The system (100) as claimed in claim 12, wherein the server (110) is connected to an Electronic Health/Medical Record (EHR/EMR) and/or clinical/hospital/local backend server system to obtain the demography data comprising height, weight, Body mass index (BMI) and clinical/symptoms data of the subject (106).
- 5 15. The system (100) as claimed in claim 12, wherein the machine learning model is trained on a labelled subject ID dataset for Biometric authentication using the subject's (106) heart sound data.
16. The system (100) as claimed in claim 1, wherein total energy distribution in the recorded second laser signal is integrated across the image frame, wherein the
10 integration is performed separately for each of the plurality of image frames to convert into the time series heart sound data.
17. The system (100) as claimed in claim 2, wherein the normative percentile score is determined based on a derived heart sound metrics of the subject (106) against a general population's heart sound metric considering factors such as sex, age, and other
15 relevant statistical variables.
18. The system (100) as claimed in claim 2 and 17, wherein the longitudinal score is determined based on the change in baseline of the derived heart sound metrics of the subject (106) from one visit to the next and/or after therapeutic intervention.
19. The system (100) as claimed in claim 1, wherein the server (110) is configured to
20 display the heart sounds in real-time on an interactive user interface or output the heart sounds via speaker/headphones in real-time.
20. The system (100) as claimed in claim 1, wherein the server (110) is configured to classify the heart murmurs into Systolic murmur, Diastolic murmur, and/or Continuous murmur using the machine learning modelling on the heart sound data.

21. A method for detection, screening, diagnosis, and prediction of cardiovascular conditions in a subject (106) by retrieving and analyzing heart sounds, wherein the method comprises the steps of:

generating a first laser signal comprising a wavelength between 400 nm and 2500
5 nm with a power of at least 0.1 mW to at most 5 mW by a laser device (102), wherein the first laser signal is directed to a region comprising neck area and/or chest area of the subject (106) to obtain a second laser signal by the reflection of the first laser signal;

acquiring the second reflected laser signal periodically at a signal acquisition frequency of at least 600 Hz by a camera unit (104),

10 generating a plurality of image frame data comprising spots, wherein each of the acquisitions generates an image comprising spots, wherein parameters of the laser device (102) and the camera unit (104) are calibrated in real-time through a control unit (108);

receiving the plurality of image frame data comprising the spots from the camera
15 unit (104) by a server (110), wherein the server (110) is communicatively connected to the laser device (102), the camera unit (104) and the control unit (108), wherein the server (110) comprises a processor and a memory that stores a set of machine-readable instructions operable, when executed by the processor, to:

process the plurality of image frames to extract time-series heart sound
20 data by (i) detecting variation of the position of each of the spots between consecutive image frames, (ii) obtaining the variation aggregate of the consecutive image frames to generate a raw signal, wherein the raw signal comprises the overall variation of the image frames collected consecutively in a time interval, and (iii) converting the raw signal into amplitude to obtain the time-
25 series heart sound data,

automatically determine normal/abnormal heart sounds by quantifying heart sounds and murmurs by analysing the time-series heart sound data using a statistical model and/or a machine learning model, and

5 predict the cardiovascular conditions of the subject (106) using the machine learning model on the heart sound data.

22. The method as claimed in claim 21, wherein the machine learning model predicts the cardiovascular conditions in the subject (106) based on normative percentile scores, longitudinal scores and prognostic scores of the heart sound data.