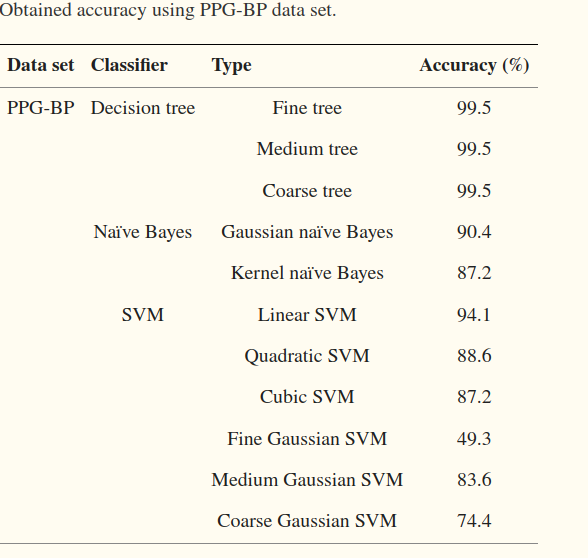
Review of Literature

Purpose: The purpose of this literature review is to collate previous research and understanding of PPG signals and Machine Learning Techniques.

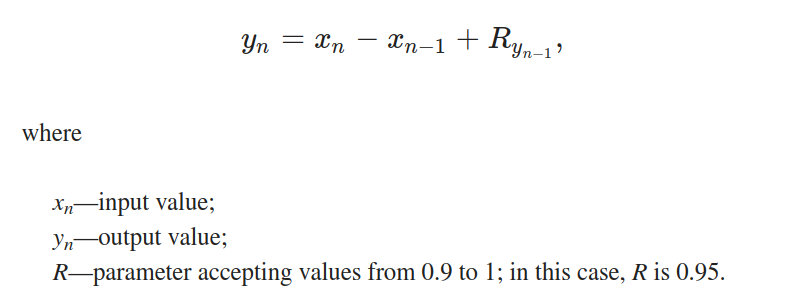
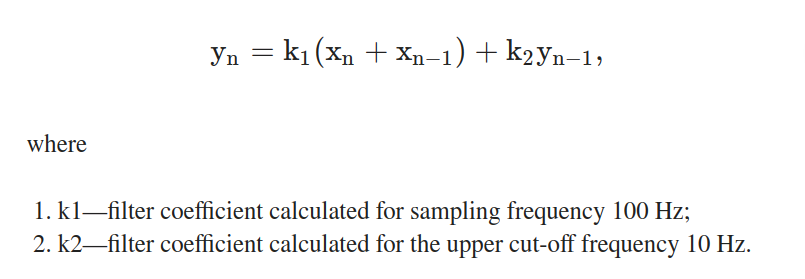
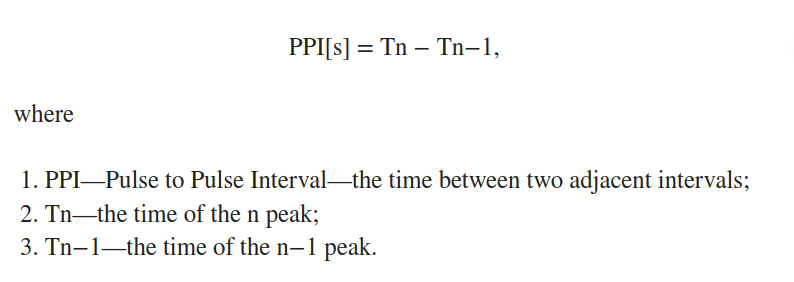
Sadad, T., Bukhari, S. A. C., Munir, A., Ghani, A., El-Sherbeeny, A. M., & Rauf, H. T. (2022). Detection of Cardiovascular Disease Based on PPG Signals Using Machine Learning with Cloud Computing. Computational intelligence and neuroscience, 2022, 1672677. <https://doi.org/10.1155/2022/1672677>

Notes – **PPG and AI**

* WPM (Wearable Patient Monitoring) transmit medical information to mobile and web applications but face **accuracy, precision and reliability issues**
* PPG (Photoplethysmography) eliminates many issues such as invasive procedures for continuous monitoring and uncomfortable non-continuous cuff-based reading
  + PPG uses human skin vessels to find changes in light transmitted or reflected through the photoelectric sensors providing information on:
    - Variation of human blood volume
    - Differentiates the systolic and diastolic processes of the heart associated with BP
  + The light emitting diode and photodiode are used to evaluate the variation in the reflected light
  + PPG can be used for distinct aspects of cardiovascular surveillance including
    - Blood oxygen saturation
    - Heart Rate
    - BP estimation
    - Cardiac Output
    - Respiration
    - Arterial Aging
    - Endothelial Control
    - Microvascular Blood Flow
    - Autonomic Function
* Evaluation of PPG: Periodograms, Spectrograms and Wavelets are used to evaluate the biosignals
* **\*\*\* If I need more data points – Article states there is a comprehensive clinical and anonymous database of patients admitted to Guilin People’s Hospital, China that can be used** 
  + Ensure data has complete heartbeat, less noise, cycles, and motion artifacts
    - To do: I need to remove the noise from the dataset
  + PPG-BP dataset is evaluated through conventional method of Machine Learning
    - Supervised machine learning classifiers: Decision Tree (DT), naïve Bayes (NB), support vector machine (SVM)
      * Decision Tree – classifies instances via top-down approach using (Fine, Medium, Coarse with split criteria of misclassification
        + Uses an ensemble classifier involving boosting and bagging to achieve better performance
  + PPG-DaLiA is evaluated through **1D CNN-LSTM (Long short-term memory)**
    - Used for Heart Rate Detection
    - 1D CNN can be used to extract robust features() of 1D Time-series sequence data using convolutional operations through several filters
      * 2 – Convolutional layers (conv1D)
      * 1- LSTM layers
      * 2- Fully connected Layers
      * Soft-Max Output lLayer
    - Performance was evaluated using accuracy, confusion matrix and ROC curve
      * ROC – evaluate classification performance through graphical portrayal
      * Confusion matrix = TP, TN, FP and FN parameters (True, Negative, False, Positive)
    - Study obtained 95%+ based on these methods
    - 

Georgieva-Tsaneva, G., Gospodinova, E., & Cheshmedzhiev, K. (2022). Cardiodiagnostics Based on Photoplethysmographic Signals. Diagnostics, 12(2), 412. <https://doi.org/10.3390/diagnostics12020412>

Notes – **PPG Mathematical Evaluation**

* Diagnosis of cardiovascular disease can be assisted by mathematical methods for heart rate variability
  + Decreased heart rate variability (HRV) is considered a deterioration in the health of the individual and a reason to look for cardiovascular disorders and vascular disease
* PPG makes it possible to measure various parameters and noninvasive
  + The shape of the recorded signals indicates fluctuations in blood volume and reflect heart rate
    - The light read has a variable component(AC component) and quasi-DC component
      * Changes in AC component are due to changes in arterial blood volume
      * Changes in DC component are due to the following factors:
        + Average Blood Volume
        + Optical Properties of the Skin
        + Respiration
        + Vasomotor activity
* Mathematically based analyses of the form of PPG signals provide useful clinical information on several medical issues concerning general health including:
  + Blood Pressure, Respiratory Processes, SNS (Sympathetic Nervous System). Nervous System Activity, HRV, Detection of irregular heartbeat, infectious diseases, diabetes and others
* HRV can be determined by:
  + Variation of Time series formed by the PP interval corresponds to the variation of RR intervals in ECG (electrocardiographic signals)
    - Detected P peaks in PPG signal
    - PP intervals time series formed
* Past studies referenced:
  + PPG signals to track sleeping patterns
  + PPG signals recorded from multiple sites studied breathing type (normal or deep) on the shape of the pulse wave and characteristics
    - Changes in signals recorded during deep breathing could not be used to extract characteristics
  + Multi-Wavelength PPG was used to monitor blood sugar levels and diagnose diabetes
  + PPG optical sensor was used to examine the cardiovascular system in different conditions. Found limitation when operating in a magnetic field
  + PPG and various sources of noise including variation, physiological processes, and external perturbations. \***Be a good source to look at on eliminating noise from our PPG signals**
  + PPG signals for cardiovascular monitoring of those with cardiovascular disease and healthy controls. Mathematically based analysis of data in time and frequency domain with nonlinear methods were used. \***Good source to read**
  + Placement of PPG signals is important as the signal quality and shape of pulse wave may change based on where it is placed.
    - Optical sensors are influenced by ambient temperature
    - Wrist sensors have decreased efficiency of recorded signals because there are many bones, tendons and muscles
* PPG contains noise components that must be optimally reduced before applying mathematical analysis of the studied data
  + Source – individual characteristics of the patient – skin thickness, age, sex, obesity etc
  + Source – physiological Processes – respiration, temperature, location of sensor, and venous pulsation
  + Source – external disturbance – movement, environment, and outdoor light
  + \*\* It is imperative to remove as much noise as possible
* Algorithm for Preprocessing PPG Signals: INCOMING SIGNALS RATHER THAN COLLECTED DATA
  + They used C ++ language but should be reproducible with Python
    - An averaging filter is applied to eliminate small, short-term disturbances in the signals
      * Verification that the number of each value has reached present number. If the present number is reached the total number is divided by the number of reading values and the arithmetic mean value is obtained.
    - The signal received from the sensors contain two components: constant and variable
      * Variable part is the P vertices
      * A digital filter was used to remove the DC part
      * 
* Eliminate large deviation in the input data by applying a median filter
  + The last n values of signals are sorted by size
    - If n is odd the middle element is selected
    - If n is even the arithmetic mean of the two element is calculated
* Remove high frequency components of the input signal by applying a low pass filter
  + If maximum HR is 220 bpm, then the heart is about 3.7 Hz. Outliers are eliminated
  + 
* Determine the slope of the signal curve. If the current value is greater than previous value, then maximum value is updated
* Determine the P peak as follows:
  + If a condition is met and the current value is less than previous signal value, this curve has a downward direction and this is the P peak,
  + Otherwise read next value
* A second check is made to observe if the maximum is not correctly determined by searching the data for the last 0.2s to determine the exact position of P peak.
* Determine The distance of two consecutive peak as the time postion for each signal peak is known
* 
* Save data as variable for mathematical analysis to diagnose cardiovascular disease
* Results:
  + Mean Squared Error was used to evaluate the math and differentiate between healthy and individuals with syncope
  + Relative error calculated the number of QRS complexes/pulse waver per 1 hour and mean PP intervals per millisecond
  + They used linear methods to compare ECG and PPG signals to see which was more effective
    - Value parameters SDNN, SDANN and RMSSD were significant as they lower in patients with syncope compared to healthy individuals
  + Geometric shape of the distribution of PP intervals was analysed using histograms and scattergrams
    - Normal RR/PP intervals and its triangular approximation of HRV triangular index and TINN
    - Scattergram of two adjacent RR/PP intervals

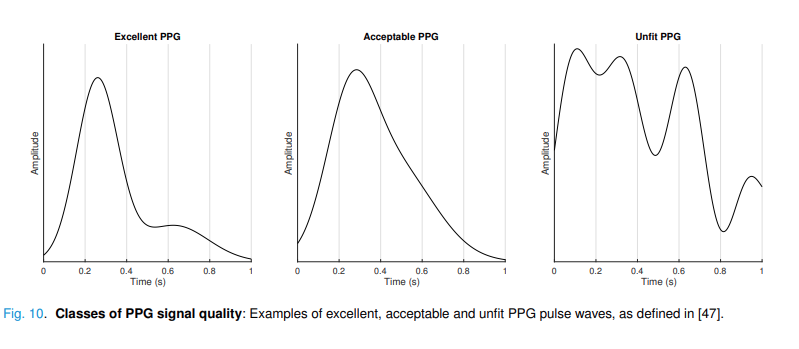
Priyadarshini, R & M., Kalimuthu & Nikesh, S & Manjunathan, Bhuvaneshwari. (2021). Review of PPG signal using Machine Learning Algorithms for Blood Pressure and Glucose Estimation. IOP Conference Series: Materials Science and Engineering. 1084. 012031. 10.1088/1757-899X/1084/1/012031.

Notes – **PPG and AI**

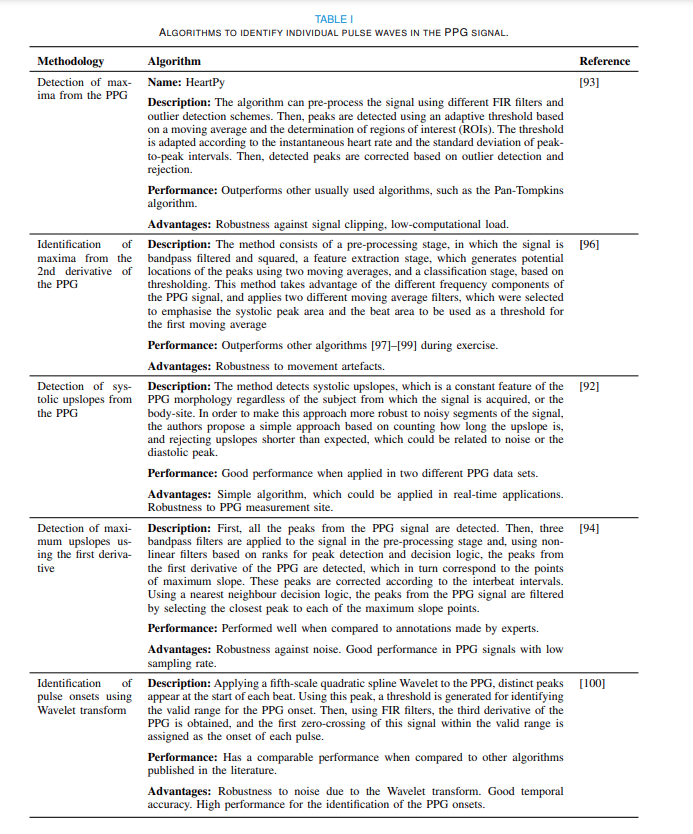
* PPG signal is recorded for three minutes from individuals and stored as BGL values
  + The sample of the signals are corrupted by baseline drift, noise and discontinuity. They were ignored in this training dataset
* Parameters like frequency, amplitude and slope were included in the PPG curve with its waveform. These were the features in the training dataset.
* The pulse area under the PPG curve is the total area calculated at the minimum and maximum of the cardiac cycle by the height of the pulse difference.
  + The estimation of the pulse area depends on the Trapezoidal law of integration
* The AD module has some common techniques which was utilized to detect voice in speech technique but is used to reject the corrupted parameters
* The Finite State Automation(FSA) is used for the segmentation of PPG signals and it involves three states namely S1 (spurious/lack of signal), S2(PPG –in) and S3(PPG-out)
* The paper mainly focuses on the model for estimating the BGL and BP values and training these values are achieved by the implementation of four non linear algorithms of regression.
  + 1) Decision Tree Regression: final outcome should be a number
  + 2)Support Vector Machine: support vector regression works my minimising the absolute magnitude of errors.
  + 3)Random Forest Regression – supervised learning method that uses classifcation as well as regression
* The shape of the photoplethysmograph signals often changes so it makes it difficult to extract any useful signal characteristics. For the whole feature, the algorithm is suggested, and the extraction is based on the raw signal values, but it doesn't depend on the shape.

Huang, J.-D., Wang, J., Ramsey, E., Leavey, G., Chico, T. J. A., & Condell, J. (2022). Applying Artificial Intelligence to Wearable Sensor Data to Diagnose and Predict Cardiovascular Disease: A Review. Sensors, 22(20), 8002. <https://doi.org/10.3390/s22208002>

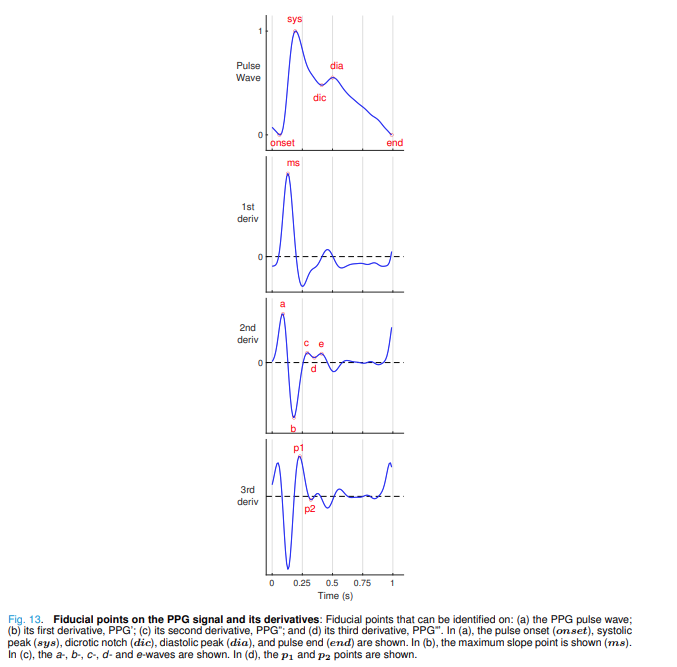
Notes – **Overview of PPG signal processing and synthesis**

* PPG signal reflects the variation in attenuation of incident optical radiation by different tissue components within the tissue volume
  + High frequency variation (the AC Part) = changes in arterial blood volume with each heartbeat
  + Low frequency variation (the DC Part) = changes in components such as venous and capillary blood, and bloodless tissue
  + Exhibits a quasi-periodic pattern consisting of an arterial pulse wave for each heartbeat resembling an arterial blood pressure pule wave
* Preprocessing – Digital FIltering: COMPREHENSIVE
  + Removal of noise from PPG signals is a crucial step in extracting valuable information from it.
  + Digital filtering consists of convolving the input signal with filter coefficients to produce a filter signal
    - Need to find the b[k] and a[l] coefficient that give the desired response of the filter
    - Two types of filters:
      * Finite Impulse Response
      * Infinite Impulse Response
    - Filters have an order, cut-off frequency and are of a particular type: low-pass, high-pass, band-pass, or band-stop
      * The order determines the behavior of the slop of transitions between pass bands and reject bands
      * Higher order filters have steeper transitions but also require a longer duration of input signal and generate higher delay
      * Cut-off frequency determines the frequency at which gain of resulting signal is around 70% of original amplitude
      * High and low pass have ONE transition and one cut-off frequency
      * Band-pass and band-stop have TWO transition bands and TWO distinct cut-off frequencies
    - **The most common used FIR filter in the field of biomedical processing, including preprocessing the PPG signal, is the MOVING AVERAGE FILTER**
      * Other common filters include the median filter which instead of using mean value of n samples, it uses the median value
    - The other class of filter is known are IIR
      * The most common type of these filters are Butterworth, type I and II Chebyshev filters and Elliptic filters
    - FIR filters are preferred but has longer delays suitable for real-time data and are linear.
* Preprocessing – Motion Artefact Removal: COMPREHENSIVE
  + PPG probe movements may result in episodes of significant alterations to the PPG waveform.
    - Robust signal processing algorithms are required to reduce the impact of artefacts on PPG measurements and to cover a range of clinical applications
    - Removal is tricky because the noise appears quite similar to physiological responses.
    - Denoising techniques:
      * Manual identification and labelling for exclusion of noisy pulses
      * Cluster techniques approaches to extract the most consistent presentations of pulse shapes in a recording
      * Independent component analysis by exploiting the quasi-periodicity of the PPG signals
        + Combining independent component analysis and block interleaving with low pass filtering under condition of dual-wavelength measurement
      * Signal decomposition and reconstruction
      * General framework TROIKA (signal decompisiTION for denoising, sparse signal RecOnstructioin for hIgh resolution for spectrum estimation and spectral peaK tracking with verification)
      * Periodic moving average filtering
      * Wavelet denoising as a pre=processing stage for robust HR detection
  + Signal quality is expected to be higher in measurements acquired at rest but reduces with movement
    - Noise hinders the extraction of reliable information from PPG signal in these activities
    - Identification of noise-corrupted signal segments are eliminated from analyses
  + There are three classes of PPG pulse waves
    - 1. excellent PPG featuring salient systolic and diastolic waves
    - 2. acceptable PPG in which systolic and diastolic waves cannot be distinguished and heart rate can be estimated
    - 3. Unfit PPG in which both waves cannot be distinguished, and heart rate cannot be estimated.
    - 
  + For morphological analysis, a minimum of 2 minutes was recommended with 86-95% of PPG being diagnostic quality
    - If recordings in which 36-50% of beats exhibit an artifact, then it should be rejected
  + Most methods begin with the identification of clearly poor-quality segments using checks of physiological plausability then apply more sophisticated techniques to identify the remaining less evident poor-quality segments
    - Extraction of morphological, spectral and trend-based characteristics from a single PPG pulse wave and then a pulse wave is classified as high or low quality based on empirically determined or machine learning derived thresholds
      * These characteristics are usually extracted in the time domain such as amplitude or timing characteristic
        + Timing characteristics = systolic phase duraction, ratio of systolic to diastolic phase duration, interbeat intervals, average pulse rate
        + Shape characteristics = # of diastolic peaks, # of times the signal changes from positive to negative or vice versa, signal to noise ratio and the comparison of the accuracy of different systolic waves
        + Higher order of statistics give an indication of distribution of data
        + Shannon entropy has been proposed to measure the presenetce of disorder in PPG pulse wave
    - Algorithms to detect similarities between successive pulse waves such as:
      * Variations in pulse wave duration, systolic phase duration, and pulse wave amplitude
      * Template-matching techniques are popular measures of regularity in PPG segments as they are thought to have similar morphologies
        + Comparison between each pulse wave and average representation of these pulse waves measuring Euclidean distance and ratio of amplitudes
        + Obtaining mean value of correlation coefficient between each cycle and extracted template
        + Modifying the width of each pulse when comparing it to the template using Dynamic Time warping
* Preprocessing: Calculating Derivatives
  + The simplest approach to calculate a derivative is to use the single-sided difference quotient, in which the derivative, x 0 , of a signal, x, is calculated from the difference between adjacent points (x 0 [n] = (x[n + 1] − x[n])/T, where x[n] is the signal, and T is the time interval).
    - Highly susceptible to high frequency noise because it only uses two points on the signal.
  + Symmetric difference quotient (x 0 [n] = (x[n + 1] − x[n − 1])/T) can provide a more accurate estimate for the derivative, it is similarly susceptible to noise.
  + Two approaches to reducing susceptibility to noise: low-pass filtering the signal (such as below 7 Hz [63]) prior to differentiation, or using more signal points in the calculation.
    - More signal points can be included by using differential quadrature methods, in which the derivative is calculated as the weighted sum of multiple points on the signal. One example is Savitzky-Golay filters [64], [65], which incorporate smoothing, thus reducing the impact of high frequency noise on calculated derivatives
* **TIME DOMAIN ANALYSIS**
  + Individual pulse waves can be characterized by several features including: amplitude, timing, and shape as well as variability of these features.
  + After preprocessing – one can use time domain to

1. **Identify individual pulse wave for analysis**

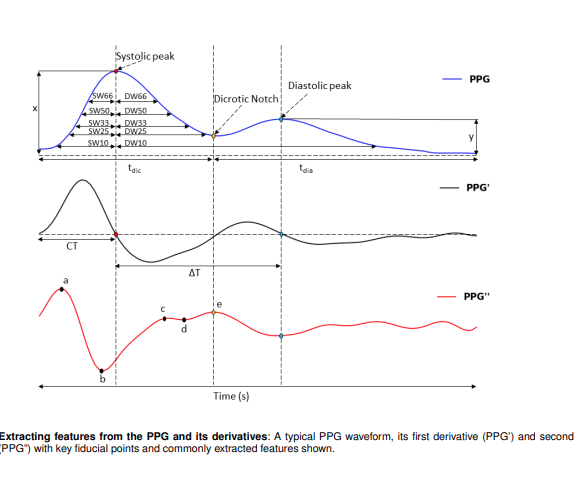
* Individual pulse waves can exhibit two distinct peaks esp. In young patients
* Noise and low frequency variations can create disturbances
* Most methods are based on the identification of the systolic peak
  + First, filter the signal to emphasize the desired components of the PPG
    - Detection of peaks using thresholds
    - Identification of peaks as zero-crossing points on first derivative with adaptive thresholding
    - Identifying maxima and minima points of PPG
    - Identifying points on first and second derivative of PPG
    - Identifying upslopes in PPG
    - Wavelet transforms
    - Local maxima scalogram
  + Second, extraction of candidate pulses
  + Third, identifying peaks or onsets
  + Fourth, correction of these peaks and onsets
  + 

1. **Identify fiducial points** on each pulse wave and its derivatives from which pulse wave features can be calculated

* Distinct points that can be identified on either pulse wave or derivatives
  + Pulse onset, the systolic and diastolic peaks
    - On original pulse: systolic peak = maximum, pulse onset = minimum
    - Diastolic peak can be determined as the maximum point after the location of the dicrotic notch
      * In older subjects, the first local maxima in the second derivative after the e wave
  + Maximum point of first derivative indicating the point of maximum slope in the original signal
  + Four distinct points of second derivatives: a, b,c, d, e waves.
    - The e wave is the dicrotic notch (esp. In young persons)
    - The a wave based on squaring of PPG followed by knowledge-based threshold approach
      * B wave can then be determined after removing a wave and so on.
  + Two points on the third derivative
    - P1 is the first maxima on the third derivative after occurrence of the b wave and p2 is the last local minimum of the third derivative before the d wave
    - 

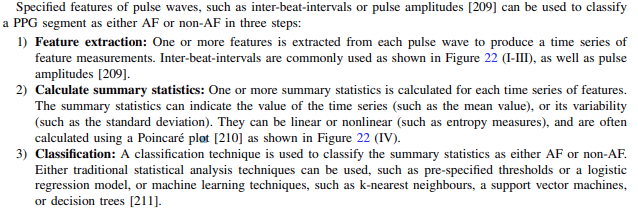
1. **Calculate pulse wave features** to quantify the amplitude, timing, and shape of each pulse wave

* On the original pulse wave, the systolic amplitude, that is, the height of the PPG from the baseline to its peak, has been correlated with stroke volume
* Potential indicators of total peripheral resistance include the width at half the amplitude of the PPG
* Diastolic time, CT, and pulse width from both systolic and diastolic portions of the pulse wave, and their ratios, have been proposed as indicators of blood pressure
* On the second derivative, the amplitude ratios of the b, c, d, and e-waves with respect to the a-wave are frequently used for pulse wave analysis.
  + Decreases in c/a, d/a, and e/a indices, and an increase in b/a, are thought to reflect increased arterial stiffness, and hence these indices decrease with age
  + The cross-correlation approach can also be used to study autonomic function and cardiovascular control.



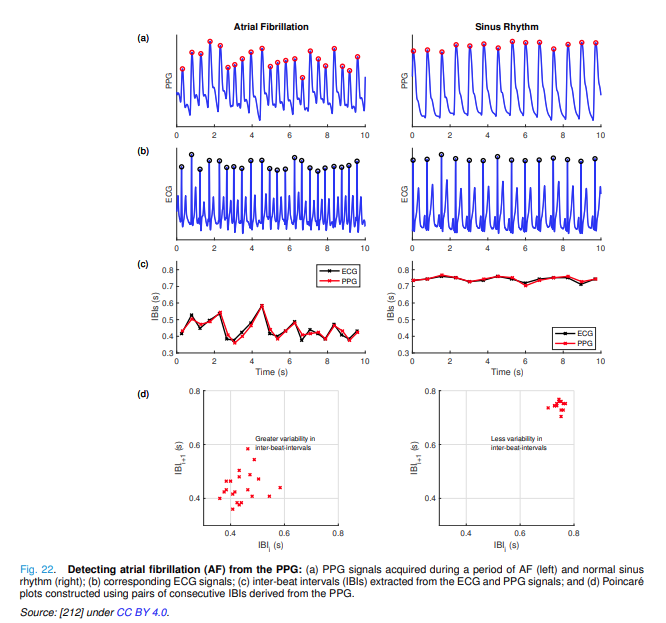
1. **Obtain representative feature measurements from** a recording by averaging features derived from each pulse wave
   * 1. variability, features derived from pulse waves should be averaged over multiple pulse waves to obtain representative measurements.
     2. variability is compounded in the presence of noise, such as during artefact due to movement between the sensor and tissue. A high proportion of noise in a recording can impact the validity of a measurement
     3. There is no global standardisation for the number of beats needed to obtain representative pulse wave features, although suggestions have been made, e.g., 60 high quality beats
     4. Simple statistical measures can be used to quantify PPG features and their variability over multiple pulse waves, for example using simple mean and standard deviation values calculated, or perhaps median and inter-quartile range values.
     5. A range of shape related measures can also be extracted from average pulse waves (an ‘ensemble average’ of multiple pulse waves)
2. **Identify abnormal pulse waves** to remove or replace feature measurements derived from them
   1. Consequently, features extracted from them are often removed (or replaced) prior to further analysis.
   2. Abnormal pulse waves are often identified through analysis of inter-beat-intervals. M

* **FREQUENCY DOMAIN ANALYSIS**
  + Provides valuable isnight into cardiovascular dynamics and allows for the extraction of features for diagnosis and monitoring of diseases
  + Two groups:
    - Classical method based on Fourier Transform (Spectral Analysis)
      * The Fourier Transform uses sinusoidal waves to map the contents of the signal to frequencies
        + Can be expressed in terms of amplitude, frequency, phase
        + Sampling frequency should be more than twice the maximum frequency of interest.
        + Duration of the recording influences the resolution of the resulting frequency spectrum: the resolution is a function of fs and the number of data points used to calculate the spectrum.
        + The power spectral density (PSD, or power spectrum) and Welch’s periodogram method are commonly used.
        + Some methods are based on autoregressive models (AR), moving average models (MA) or a combination of both (ARMA).
        + Different approaches have been proposed to estimate the power spectrum using AR models, such as the Yule-Walker, Burg, covariance, and modified covariance methods.
        + Other non-parametric modern approaches for spectral analysis are based on eigenanalysis frequency estimation,
    - Modern methods based on model of the source and signal
* Time-Frequency Analysis
  + Study of a signal simultaneously in both time and frequency domains, can be performed using several methods. Time-frequency analysis allows one to investigate how spectral features evolve over time, even in signals of short duration
  + Wavelet transforms (WT) have a number of advantages over traditional methods, including the tailoring of both time and frequency resolutions without one being traded off against the other, and they can identify transients in non-stationary signals.
    - Can offer time-localised filtering and help determine discontinuities and other events
    - Provides sparse representation of data and useful for denoising or compressing data while preserving important features
      * The Continuous Wavelet Transform (CWT) is commonly used to analyse 1-D data such as a single PPG signal.
      * Wavelet analysis was performed using MATLAB (The MathWorks Inc.), using the cwt function with the Morse wavelet selected
* Machine Learning
  + PPG-based machine learning algorithms have been developed to classify patients according to whether they have coronary artery disease and atrial fibrillation. Machine learning has also been used to estimate parameters from a PPG signal, including blood pressure, haemoglobin levels, blood glucose, blood oxygen saturation, stress levels, and respiratory rate.
  + **Linear Regression**
    - This makes the model perform well on smaller datasets. However, its performance degrades as the complexity and nonlinearity of the input and target variables increases. Linear regression has been applied in several PPG analysis studies
  + **Tree-Based Algorithms**
    - a first branch, a PPG signal could be placed into one of two groups according to whether the derived heart rate is above a threshold. At a second branch, the signal could be placed into one of two subgroups according to the level of heart rate variability. This splitting process continues until reaching the end of the branching structure, at which point either a class label is provided (such as atrial fibrillation or not)
    - Decision trees are used for classification, and regression trees are used for parameter estimation.
      * Random Forest is the most popular tree-based learning algorithm
        + Multiple trees are involved in the decision-making process
        + Each tree returns a categorical value and final category is selected based on voting criterion
      * Random forests can perform better than linear models and can handle large datasets with input vectors of high dimensions. They can also model highly nonlinear relationships and achieve a comparable performance with a multilayer perceptron neural network
      * random forests are prone to overfitting and are more computationally complex than single tree-based algorithms
        + FOR ATRIAL FIBRILLATION S. Fallet et al., “Can one detect atrial fibrillation using a wrist-type photoplethysmographic device?” Medical and Biological Engineering and Computing, vol. 57, no. 2, pp. 477–487, 2019. https://doi.org/10.1007/s11517-018-1886-0 [163] S
  + **Support Vector Machines**
    - a popular non-parametric machine learning algorithm that can be used for classification and regression.
    - The objective of an SVM is to find the decision boundary that maximizes the distance between the data points and the hyperplane.
    - The boundaries are set by finding the soft margins that maximize the margin between the hyperplane and the closest points to it, known as support vectors
    - The disadvantages of the SVM are that it does not work well on very large datasets, is prone to noise in the data, and the kernel application increases computational complexity. Nonetheless, it is still widely used to analyse PPG signals
  + **Neural Networks and Deep Learning**
    - Different architectures of neural networks are designed to handle different tasks, the most popular ones are Feedforward, Convolutional, ResNet, Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Encoder-Decoder, and Autoencoders.
    - Neural Networks can scale better on larger datasets but this comes at the cost of added computational complexity
    - FOR ATRIAL FIBRILLATION: S. P. Shashikumar et al., “A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology,” in 2017 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2017. IEEE, 2017, pp. 141–144. <https://doi.org/10.1109/BHI.2017.7897225>
      * I. Gotlibovych et al., “End-to-end Deep Learning from Raw Sensor Data: Atrial Fibrillation Detection using Wearables,” arXiv preprint, p. arXiv:1807.10707, 2018. <http://arxiv.org/abs/1807.10707>
      * S. Kwon et al., “Deep Learning Approaches to Detect Atrial Fibrillation Using Photoplethysmographic Signals: Algorithms Development Study,” JMIR mHealth and uHealth, vol. 7, no. 6, p. e12770, 2019. <https://doi.org/10.2196/12770>
  + **Nonlinear analysis in phase space**
    - PPG signal can be transformed into phase space using time-delay embedding
      * is transformed into phase space using Takens’ Embedding Theorem. This process consists of transforming the PPG signal into m new co-ordinates of (x(t), x(t − 1τ ), ..., x(t − (m − 1)τ ), where x(t) is the PPG signal, m is the embedding dimension, and τ is the time delay. The resulting representation of the signal is termed an ‘attractor’. The choice of embedding dimension affects the resulting attractor, with dimensions used previously.
      * The attractor can be used immediately for analysis [192], or it can be transformed further into a 2D attractor using symmetric projection attractor reconstruction
      * This has the advantage of reducing the influence of baseline drift, and facilitating analysis in two dimensions. Some analyses make use of a density plot of the 2D attractor to assess variability in pulse wave timing and morphology, with potential applications in identifying illnesses such as sepsis which affect autonomic nervous system functionality
    - Allows for visualisation and analysis of pulse wave morphology and variability
      * Does not require the identification of individual pulse waves
  + **Estimating Physiological Parameters**
    - PPG pulse wave contains far more information in its shape, height, and timing than just pulse rate.
      * HEART RATE:
        + heart rate or pulse rate can be estimated by simply identifying peaks corresponding to heartbeats
        + counts the peaks in a fixed time window (usually > 10 seconds), and then extrapolates to calculate the heart rate in beats per minute.
        + Filtering is crucial if the PPG are noisy
        + frequency-domain methods are used, particularly when PPG signals are affected by motion artefact
        + Spectral analysis techniques (such as Welch’s averaged periodogram) can be used to identify the heart rate as the frequency corresponding to the maximum spectral power
      * INTERBEAT INTERVALS
        + the intervals between consecutive heartbeat
        + Inter-beat intervals (IBIs) are extracted by measuring the time delay between occurrences of a particular fiducial point on consecutive pulse waves.
        + their reliability has been found to be greater when measured using the pulse onset rather than the systolic peak.
        + The best-performing algorithm to identify pulse onsets has been found to be based on identifying the intersection point of tangent lines to the apex point of the first derivative of the PPG signal, and to the foot of the PPG pulse
      * DETECTING ATRIAL FIBRILLATION
        + Atrial fibrillation (AF) is the most common arrhythmia, causing a fivefold increase in stroke risk [207]. Fortunately, the risk of stroke can be reduced through medication if it is diagnosed.
        + Typically diagnosed during a short period of 30 sec
        + Techniques to identify AF from a PPG segment typically either analyse specified features of PPG pulse waves, or use deep learning without the need for any specified features



Machine Learning Steps:

1. To assess the quality of the segment to determine whether it is of sufficient quality to determine whether it exhibits AF.
2. The second step is to use a deep learning model to classify the segment. The PPG signal can be used directly as an input to the model, and/or features extracted from it Additional signals or features can also be used, such as accelerometry data. The amount of labelled data required to train a deep learning model can be reduced by using transfer learning to fine-tune a pre-trained deep learning mode



Extra Reading – ATRIAL FIBRILLATION:

L. M. Eerikainen ¨ et al., “Atrial fibrillation monitoring with wrist-worn photoplethysmography-based wearables: State-of-the-art review,” Cardiovascular Digital Health Journal, vol. 1, no. 1, pp. 45–51, 2020. https://doi.org/10.1016/j.cvdhj.2020.03.001

S.-C. Tang et al., “Identification of atrial fibrillation by quantitative analyses of fingertip photoplethysmogram,” Scientific Reports, vol. 7, no. 1, p. 45644, 2017. <https://doi.org/10.1038/srep45644>

T. Pereira et al., “Photoplethysmography based atrial fibrillation detection: a review,” npj Digital Medicine, vol. 3, no. 1, p. 3, 2020. https://doi.org/10.1038/s41746-019-0207-9

P. H. Charlton, “Detecting atrial fibrillation from the photoplethysmogram,” 2021. https://commons.wikimedia.org/wiki/File:Detecting atrial fibrillation (AF) from the photoplethysmogram (PPG).svg

M. V. Perez et al., “Large-scale assessment of a smartwatch to identify atrial fibrillation,” New England Journal of Medicine, vol. 381, no. 20, pp. 1909–1917, 2019. https://doi.org/10.1056/NEJMoa1901183

Y. Guo et al., “Mobile photoplethysmographic technology to detect atrial fibrillation,” Journal of the American College of Cardiology, vol. 74, no. 19, pp. 2365–2375, 2019. https://doi.org/10.1016/j.jacc.2019.08.019

HEART RATE VARIABILITY

* Pulse rate variability (PRV) metrics assess the variability in inter-beat intervals (IBIs) derived from the PPG
* HRV has been largely used for the assessment of cardiac autonomic activity. Both branches of the autonomic nervous system (ANS), sympathetic and parasympathetic nervous systems, control the firing rate of the sinus node in the heart, and hence their behaviour determines heart rate
* PRV and HRV should not be considered equivalent
* As with HRV, the minimum recommended duration of PPG recordings for PRV analysis is 5 minutes, for short-term changes, while recordings with duration of up to 24 hours can be used to analyse long-term dynamics of pulse rate
* The most commonly used indices are extracted in the time and frequency domains
  + When calculating PRV indices, it is important to identify and correct for the presence of outliers and ectopic beats. Most outlier removal methods are based on physiologically-based feasibility checks, mainly comparing consecutive IBIs and ensuring that their duration does not vary too greatly allowing only ‘normal-to-normal’ (NN) intervals to be retained.