Deep learning in the ECG realm

What is a CNN?

* The benefit of deep neural network is that they can be modified to more specialised tasks. CNN are convolutional neural networks that contain convolutional layers within them. These layers are able to function more efficiently than normal dense layers as they are able to capture a spatial relationship in the data which in the case of ECGs is a temporal relationship being captured .
* Usually, the convolutional layers are 2D and therefore a specialised for computer vision tasks like identifying what time of day a photo is taken in in which a photo has two dimensions height and width. However, when we deal with ECGs, especially when they are digitally recorded, we then are dealing with a single dimension across time.
* Previously it was required to transform the data into two dimensions when to fit 2D Convolutional layers however this created issues in requiring a large data set, the model dealing with overfitting and being less efficient overall. 1D convolutional networks served to purpose of requiring no such preprocessing you can just put your raw 1d signal and have feature automatically extracted for the purpose of things like classification. Instead of 2d matrices that are created by 2d CONV layers 1D Conv layers produce an output a 1d convolutional sequence that represent a weighted sum two 1D arrays. This allows us to have a more efficient back propagation process and therefore have an efficient way of extracting and analysing features from a raw 1d signal like an ECG (Kiranyaz *et al.*, 2019).
* Using 1D CNN has shown a lot of promise with detecting atrial fibrillation, there have been CNN models that have been tested on MIT-BIH atrial fibrillation database and have achieved over 97% in sensitivity and specificity when classifying signals into normal. Atrial fibrillation , atrial flutter and AV junctional rhythm(Petmezas *et al.*, 2021).

What is an autoencoder?

* An autoencoder still utilizes the CNN architecture but in a different way. The model architecture consists of an encoder, a decoder and in the middle as a latent space. The encoder, aided by convolutional layers can act to dilute the input into lower dimensionality which makes up the latent space. Essentially it is dimensionality reduction into a latent space, however with this method unlike principle component analysis (PCA) we can use non-linear transformation (due to using CNN architecture) to capture more complex and intricate relationships between the data as we reduce to a smaller dimensions. The decoder then works to reconstruct this signal and the loss function is based on the models ability to reconstruct the signal as accurately as possible like mean absolute error in which backpropagation is also used the adjust the networks parameter to provide the most accurate reconstructed signal that is possible(Bank, Koenigstein and Giryes, 2021).
* The promise of this model is that it can provide a more intricate set of features that can be clustered on or used for classification purposes in which researchers have regressed on these features to output classification based on the high-level features extracted from the autoencoder rather than the input itself. Researcher have shown promising success in classifying types arrhythmias, in which researcher have achieved 97% accuracy classifying 6 version of arrhythmia including : normal sinus beat, atrial fibrillation , ventricular bigeminy , pacing beat , atrial flutter , sinus brady cardia(Ramkumar *et al.*, 2022)