A Deep Learning Algorithm for Detecting AMI

Xiantong Xiang

Shandong University

Thursday, April 4, 2024

EuroIntervention. 2021 Oct 20;17(9):765-773. doi: 10.4244/EIJ-D-20-01155. PMID: 33840640; PMCID: PMC9724911.

Contents

Background - Receiver Operating Characteristic

- Consider a two-class prediction problem (binary classification), in which the outcomes are labeled either as positive (p) or negative (n).

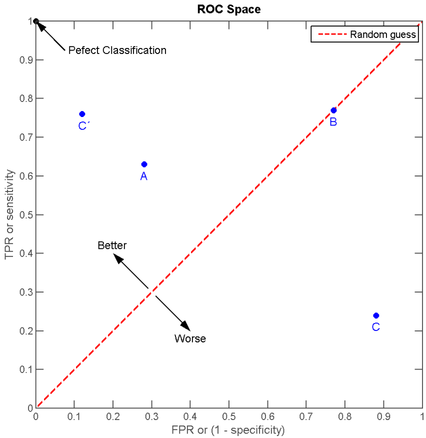
Total population $= P + N$	**Predicted Positive (PP)**	**Predicted Negative (PN)**
Positive (P) [a]	**True positive** (TP), hit[b]	**False negative** (FN), miss, underestimation
Negative (N)[d]	**False positive** (FP), false alarm, overestimation	**True negative** (TN), correct rejection[e]

Background - Receiver Operating Characteristic

- A ROC space is defined by FPR and TPR as x and y axes, respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs).

$$TPR = \frac{TP}{P}, \quad FPR = \frac{FP}{N}$$

- The experimenter can adjust the threshold , which will in turn change the false positive rate.

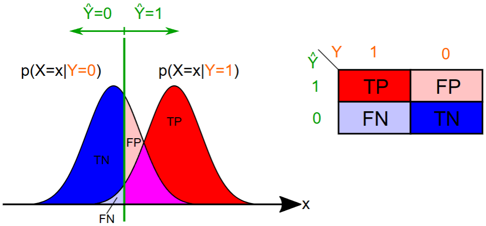


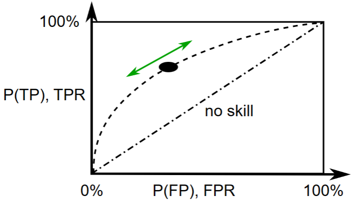
Background - Receiver Operating Characteristic

- Given a threshold parameter T, the instance is classified as "positive" if $X > T$, and "negative" otherwise. X follows a probability density $\boldsymbol{f}_1(\boldsymbol{x})$ if the instance actually belongs to class "positive", and $\boldsymbol{f}_0(\boldsymbol{x})$ if otherwise. Therefore, the true positive rate is given by

$$TPR(T) = \int_T^{\infty} \boldsymbol{f}_1(\boldsymbol{x})\mathrm{d}x$$

$$FPR(T) = \int_T^{\infty} \boldsymbol{f}_0(\boldsymbol{x})\mathrm{d}x$$





Background - Precision-Recall ROC

- In information retrieval, precision is a measure of result relevancy, while recall is a measure of how many truly relevant results are returned.
- Precision (P) is defined as the number of true positives (TP) over the number of true positives plus the number of false positives (FP)

$$P = \frac{TP}{TP + FP}$$

- Recall (R) is defined as the number of true positives (TP) over the number of true positives plus the number of false negatives (FN)

$$R = \frac{TP}{TP + FN}$$

- These quantities are also related to the F_1 score, which is the harmonic mean of precision and recall.

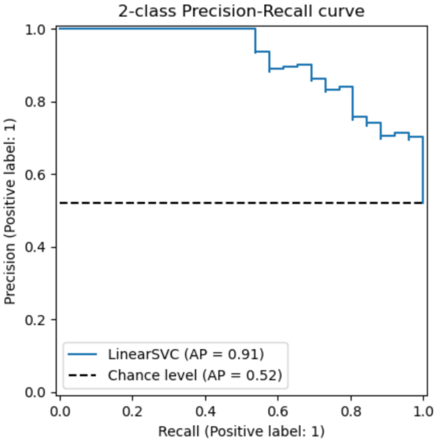
$$F_1 = \frac{2TP}{2TP + FP + FN}$$

Background - Precision-Recall ROC

- The relationship between recall and precision can be observed in the stairstep area of the plot - at the edges of these steps a small change in the threshold considerably reduces precision, with only a minor gain in recall.
- Average precision (AP)

$$AP = \sum (R_n - R_{n-1}) P_n$$

where P_n and R_n are the precision and recall at the nth threshold.



Background - Cohen's kappa

- Cohen's kappa measures the agreement between two raters who each classify N items into C mutually exclusive categories. The definition of κ is

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e}$$

where p_o is the relative observed agreement among raters, and p_e is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly seeing each category.

- For k categories, N observations to categorize and n_{ki} the number of times rater i predicted category k:

$$p_e = \sum_k \widehat{p_{k12}} \overset{lnd}{=} \sum_k \widehat{p_{k1}}\widehat{p_{k2}} = \sum_k \frac{n_{k1}}{N}\frac{n_{k2}}{N} = \frac{1}{N^2}\sum_k n_{k1}n_{k2}$$

where $\widehat{p_{k12}}$ is the estimated probability that both rater 1 and rater 2 will classify the same item as k, while p_{k1} is the estimated probability that rater 1 will classify an item as k (the rating of the two raters are independent).

Background - Binary classification confusion matrix

- In the traditional 2×2 confusion matrix employed in machine learning and statistics to evaluate binary classifications, the Cohen's Kappa formula can be written as :

$$\kappa = \frac{2 \times (TP \times TN - FN \times FP)}{(TP + FP) \times (FP + TN) + (TP + FN) \times (FN + TN)}$$

where TP are the true positives, FP are the false positives, TN are the true negatives, and FN are the false negatives.

Method - Implementation details of the DLM

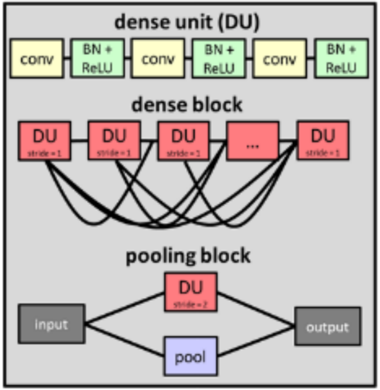
- Supposing that a standard 12-lead ECG signal comprised 12 sequences of N numbers, the ECG signal sequence

$$X = [x_{1,1}, x_{1,2}, \cdots, x_{1,N}; x_{2,1}, x_{2,2}, \cdots, x_{2,N}; \cdots; x_{12,1}, x_{12,2}, \cdots, x_{12,N}]$$

  was used as the input, and the output was a one-hot encoder of AMI categories (STEMI, NSTEMI, and non-AMI) and the IRA of STEMI (STEMI-LMCA, STEMI-LAD, STEMI-LCx, and STEMIRCA).
- For example, a label of STEMI is encoded as [1,0,0], and a label of NSTEMI is encoded as [0,1,0]. Each output label corresponded to a segment of the input. Because the ECG information was mostly provided by morphologic changes with shift invariance, convolutional layers with weight sharing were used to adapt to this situation and reduce the hazard of overfitting.

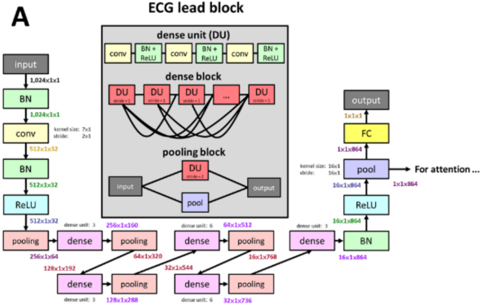
Method - Dense Unit

- We defined a “dense unit” as a neural combination as follows:
- (1) a batch normalisation layer to normalisation.
- (2) a rectified linear unit (ReLU) layer for non-linearisation.
- (3) a 1×1 (3×1) convolution layer with 4K filters to reduce the dimensions of the data (extract features).



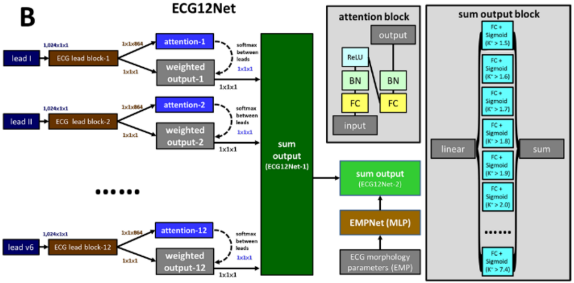
Method - ECG Lead Block

- Dense blocks cannot be concatenated when the size of the feature maps changes. Thus, a pooling block was used to concatenate each dense block for downsampling in our architecture.Each dense block was concatenated by the pooling block to integrate the features of the previous blocks.



Method - ECG12Net

- ECG12Net comprised 12 ECG lead blocks corresponding to lead sequences. They designed an attention mechanism based on a hierarchical attention network to concatenate these blocks, increasing the interpretive power of ECG12Net.



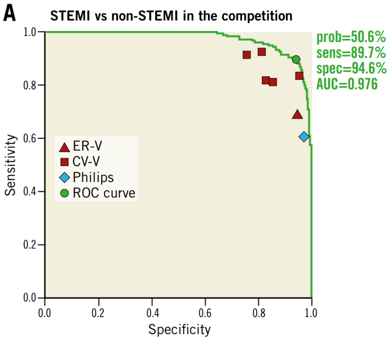
-

$$\text{MultiLogLoss} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} y_{i,j}\log(p_{i,j})$$

Method - TATISTICAL ANALYSIS

- A significance level of $p<0.05$ was used throughout the analysis.
- The receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) were applied to evaluate the competition results.
- They also used precision-recall ROC (PRROC) to evaluate the model performance in hypothetical real-world situations.
- Because the proportions of STEMI, NSTEMI, and non-AMI were distorted in the competition set, we re-weighted the samples based on the incidences in the real world (0.1%, 0.2%, and 99.7% of STEMI, NSTEMI, and non-AMI cases, respectively).

Experiment - ROC



- DLM:

$$AUC = 0.976$$

$$\text{sensitivity} = 89.7\%$$

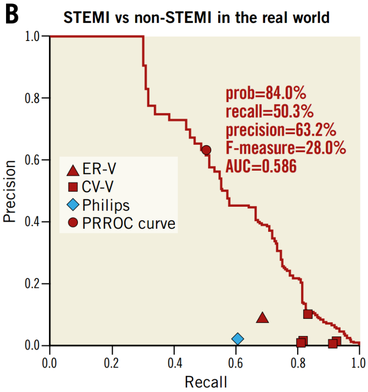
$$\text{specificity} = 94.6\%$$

- Gphysicians and the Philips:

$$\text{sensitivity} : 60.5 - 92.6\%$$

$$\text{specificity} : 76.0 - 97.5\%$$

Experiment - PRROC



- DLM(appropriate cut-off point):

$$\mathrm{AUC} = 0.586$$

$$\text{sprecision} = 63.2\%$$

$$\text{recall} = 50.3\%$$

Experiment - Performance Rankings and Consistency Analysis