Article Electrocardiogram Classification Based on Faster Regions with Convolutional Neural Network Yinsheng Ji 1,2, Sen Zhang 1,2 and Wendong Xiao 1,2,\* 1 School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China; s20180575@xs.ustb.edu.cn (Y.J.); zhangsen@ustb.edu.cn (S.Z.) 2 Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China \* Correspondence: wdxiao@ustb.edu.cn; Tel.: +86-158-0165-7835 Received: 31 March 2019; Accepted: 22 May 2019; Published: 5 June 2019

Большинство методов классификации ЭКГ в основном основаны на одномерных данных ЭКГ. Эти методы обычно должны извлекать характеристики формы волны, интервал соседней волны, а также амплитуду и период каждой волны в качестве входных данных. Основное различие между ними - выбор классификатора.

Ранее использовали нейронную сеть АД для классификации сердечных сокращений ЭКГ с точностью классификации 93,9%. Были предложены коэффициенты дискретного косинусного преобразования (DCT) из сегментированных импульсов ЭКГ, которые затем подвергали анализу главных компонентов для уменьшения размерности, и вероятностную нейронную сеть (PNN) для автоматической классификации.

Использовалась искусственная нейронная сеть, основанная на многократной прямой связи, чтобы классифицировать сердцебиение ЭКГ на шесть категорий и, наконец, была получена степень точности классификации 90,6%. Osowski разработал классификатор, который каскадирует нечеткий самоорганизующийся слой и многослойный персептрон, и реализовал семь классификаций сердечных сокращений ЭКГ с точностью классификации 96%.

**\*много примеров исследований\***

В этой статье мы использовали одномерные данные ЭКГ из записей базы данных MIT-BIH и пациента. Сначала мы предварительно обработали данные ЭКГ пациента, у которого были довольно серьезные шумовые помехи. Поскольку модель Faster R-CNN обрабатывает двумерные изображения в качестве входных данных, мы преобразовали сигналы ЭКГ в изображения ЭКГ. Наконец, мы использовали Faster R-CNN, чтобы разделить сердечные сокращения ЭКГ на пять категорий. Общая процедура показана на рисунке 1.

**\*описан способ избавления от шумов\***

Перед преобразованием одномерного сигнала ЭКГ в двумерное изображение необходимо определить положение зубца R сигнала ЭКГ. В этой статье для нахождения зубца R. было принято дискретное вейвлет-преобразование (DWT).

Вейвлет-преобразование дало хорошие результаты в улучшении защиты от помех и точности группового обнаружения QRS. Согласно теории вейвлет-преобразования, точка пика зубца R соответствует пересечению нуля максимумов модуля. Положение пика зубца R определяется путем определения положение максимумов модуля зубца R, а затем начальная и конечная точки зубца QRS ищутся вперед и назад в соответствии с положением пика R. В эксперименте использовались метод DWT и метод шумоподавления с адаптивным порогом для обнаружения зубца R сигнала ЭКГ. Максимум модуля и пересечение нуля были обнаружены, чтобы определить положение группы QRS. Затем был использован метод адаптивного шумового порога, чтобы судить, было ли обнаруженное положение пика зубцом R или глитчем.

В этой статье для извлечения биений ЭКГ использовался метод поиска со скользящим окном. Текущая точка зубца R использовалась в качестве ориентира для поиска слева в течение 150 мс. Если точка существовала, координаты левой конечной точки записывались; в противном случае поиск был остановлен. Затем в течение 150 мс выполнялся поиск справа, используя текущую точку зубца R. Если точка местоположения существовала, координаты правой конечной точки записывались; в противном случае поиск был остановлен. Наконец, мы разрезаем график слева направо от конечных точек в качестве входной выборки сети глубокого обучения в эксперименте.

В этой статье мы использовали Faster R-CNN на основе сети ZF для классификации ЭКГ. Faster R-CNN состоит из сети ZF, сети предложений региона (RPN) и сети Fast R-CNN. Среди них сеть ZF представляет собой архитектуру CNN, которая используется для извлечения карты характеристик ЭКГ изображений. Затем сеть RPN запускается на карте функций и генерирует примерно 20 000 прямоугольных блоков, которые сортируются по количеству баллов от большого к меньшему. Затем первые 300 прямоугольных ящиков принимаются в качестве входных данных для сети Fast R-CNN, которая обеспечивает более высокую точность при одновременном сокращении временной сложности. Наконец, сеть Fast R-CNN выводит вероятность категории и матрицу координат (содержащую четыре значения координат).

Сеть предложения региона использует нейронную сеть и объединяет три процесса: генерирование блоков кандидатов, извлечение функций и их классификацию в сетевой модели. Наконец, RPN реализует сквозное обучение и обнаружение. RPN принимает изображение любого размера в качестве входных данных и выводит набор блоков-кандидатов, где каждый блок имеет оценку для оценки сходства между блоком и целью.

Небольшая сеть скользит по карте характеристик свертки последнего выходного слоя общего сверточного слоя. Каждое скользящее окно сопоставляется с вектором низкой размерности, который выводится на два полностью связанных слоя одного и того же уровня - слой регрессии прямоугольного кадра и слой классификации прямоугольного кадра.

Якорный механизм предложен Faster R-CNN. В каждой позиции скользящего окна одновременно прогнозируются k предложенных областей, поэтому слой регрессии прямоугольного кадра имеет 4k выходных данных, то есть значения координат k блоков (x\_1, y\_1, x\_2, y\_2 ). Уровень классификации прямоугольных кадров имеет 2k выходов. Центр якорей k является центром текущего скользящего окна. Faster R-CNN использует три шкалы и три соотношения сторон экрана; таким образом, в каждом положении скольжения имеется k = 9 анкеров. Всего имеется якорей W × H × k для сверточной карты признаков размером W × H.

Каждой привязке присваивается метка, которая является «целевой» (положительная метка) или «нецелевым» (отрицательная метка). Сеть Faster R-CNN наследует механизм потери многозадачности сети Fast R-CNN. Для якоря функция потерь определяется как

**Изображение выглядит как текст

Автоматически созданное описание**

**\*формулы пошли\***

Существует много типов сердечных сокращений ЭКГ, и многие сердечные сокращения аналогичной формы должны полагаться на врачей, обладающих специальным опытом, чтобы их точно идентифицировать. Таблица 2 показывает принятый в настоящее время стандарт классификации сердечных сокращений ЭКГ.

Изображение выглядит как текст

Автоматически созданное описание

Эффективность классификации измеряется по четырем критериям: специфичности (Spe), чувствительности (Sen), положительной оценке (Pre) и точности (Acc), как указано ниже:

(1) Специфичность (Spe): доля нормальных сердечных сокращений ЭКГ, которые правильно классифицированы, что представляет собой правильную частоту не пациентов.

(2) Чувствительность (Sen): доля аномальных сокращений ЭКГ, которые правильно классифицированы, что представляет собой правильное соотношение пациентов.

(3) Положительная частота (Pre): количество аномально классифицированных сокращений ЭКГ правильно классифицируется как доля аномальных сокращений ЭКГ.

(4) Точность (Acc): пропорция правильной классификации сердечных сокращений ЭКГ ко всем сердечным сокращениям ЭКГ.

Таблица 3 показывает определение четырех результатов классификации. Согласно определениям специфичности, чувствительности и точности формулы следующие:

Изображение выглядит как текст

Автоматически созданное описание

Структура Faster R-CNN, используемая в этой статье, имеет три аспекта для классификации сердечных сокращений: блок целевой позиции, целевая классификация и оценка. Целевая позиция включает координаты верхней левой вершины и нижней правой вершины прямоугольного блока. Целевая классификация включает результат суждения модели о категории изображения в этой позиции, которая в этом эксперименте имела только пять дискретных значений. Оценка - это значение вероятности, показывающее, насколько вероятно, что сердцебиение ЭКГ должно быть в этой категории, что также называется коэффициентом достоверности.

В таблице 9 показаны результаты классификации Faster R-CNN. Точность классификации каждого класса составила более 99%. Средняя точность классификации составила 99,21%.

Изображение выглядит как текст, снимок экрана, монитор

Автоматически созданное описание

Преимущества Faster R-CNN следующие:

i. Извлечение функций вручную не требуется.

II. Частоту дискретизации исходного сигнала ЭКГ не нужно учитывать.

iii. Подход нечувствителен к качеству сигнала ЭКГ.

iv. Высокая точность классификации.

Недостатки Faster R-CNN следующие:

i. Образцы обучающей выборки необходимо пометить вручную.

II. Требуются долгие часы обучения и специализированное оборудование для эффективного обучения наборов данных (графический процессор (GPU)).

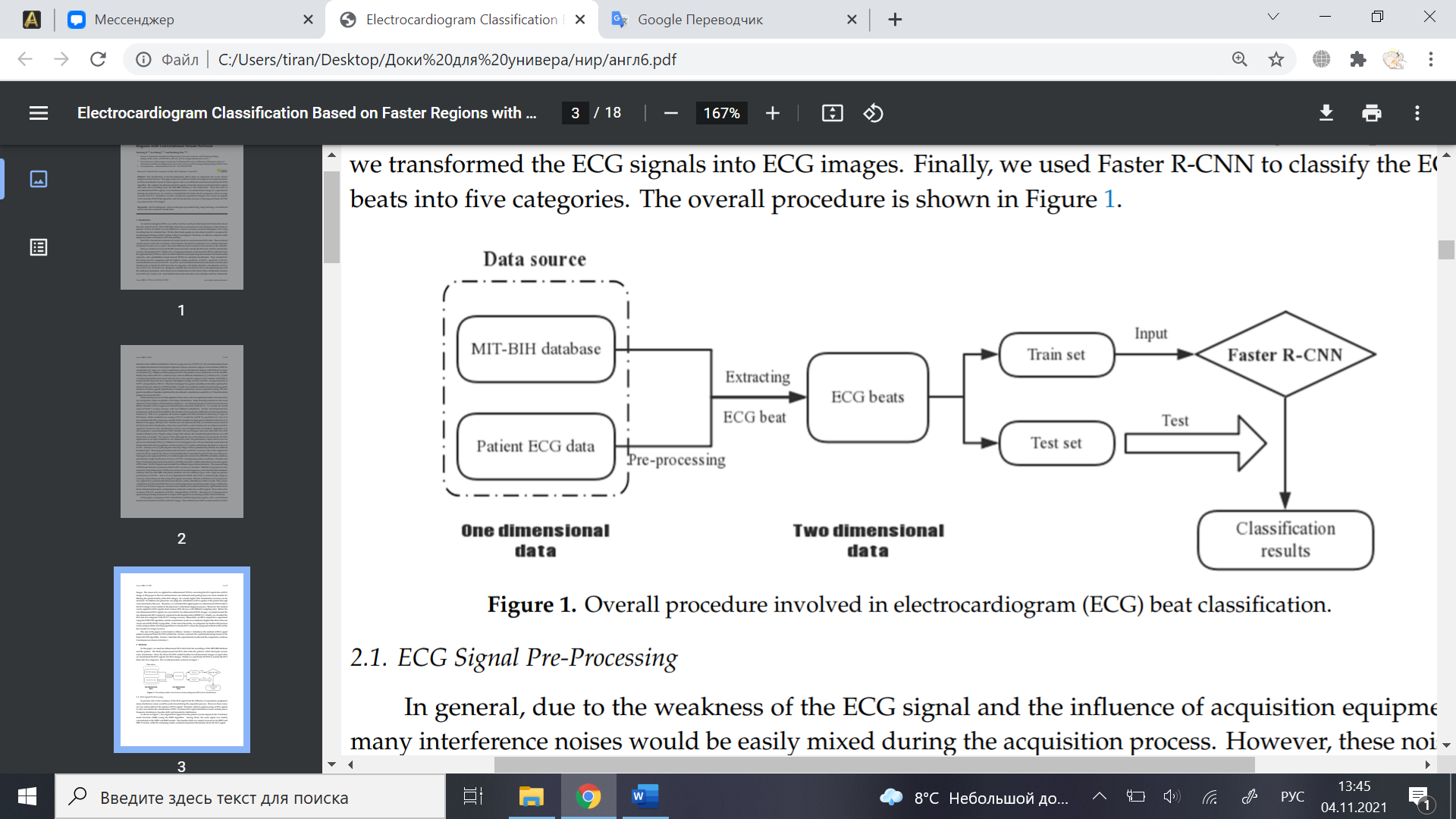
Однако, как только обучение сигналов ЭКГ завершено, классификация сигналов сердцебиения ЭКГ выполняется быстро. Классификация тестовой выборки ЭКГ занимает около 0,025 с.

Most ECG classification methods are mainly based on one-dimensional ECG data. These methods usually need to extract the waveform’s characteristics, the interval of adjacent wave, and the amplitude and period of each wave as input. The main difference between them is the selection of the classifier

Early on, Yuzhen et al. used the BP neural network to classify the ECG beat, with the classification accuracy rate reaching 93.9%. Martis et al. proposed discrete cosine transform (DCT) coefficients from the segmented beats of ECGs, which were then subjected to principal component analysis for dimensionality reduction, and a probabilistic neural network (PNN) for automatic classification.

Luo et al. used an artificial neural network, based on multi-order feedforward, to classify the ECG beat into six categories, and finally obtained a classification accuracy rate of 90.6%. Osowski et al. designed a classifier that cascades the fuzzy self-organizing layer and the multi-layer perceptron, and realized seven classifications for ECG beats with a classification accuracy rate of 96%.

In this paper, we used one-dimensional ECG data from the recordings of the MIT-BIH database and the patient. We firstly preprocessed the ECG data from the patient, which had quite serious noise interference. Since the Faster R-CNN model handles two-dimensional images as input data, we transformed the ECG signals into ECG images. Finally, we used Faster R-CNN to classify the ECG beats into five categories. The overall procedure is shown in Figure 1.



Before transforming the one-dimensional ECG signal into a two-dimensional image, the R wave position of the ECG signal needs to be found. In this paper, discrete wavelet transform (DWT) was adopted to find the R wave.

Wavelet transform achieved good results in improving the anti-interference and accuracy of the QRS group detection. According to the wavelet transform theory, the R wave peak point corresponds to the zero crossing of the modulus maxima. The R wave peak position is located by detecting the position of the modulus maxima of the R wave, and then the start and end points of the QRS wave are searched forward and backward according to the R peak position. In the experiment, the DWT method and the adaptive threshold denoising method were used to detect the R wave of the ECG signal. The modulus maximum and the zero crossing were detected to find the position of the QRS group. Then, the adaptive noise threshold method was used to judge whether the detected peak position was an R wave or a glitch.

In this paper, the sliding window search method was used to extract the ECG beat. The current R wave point was used as a reference to search the left for 150 ms. If the point existed, the coordinates of the left endpoint were recorded; otherwise, the search was stopped. Then, the right was searched for 150 ms with the current R wave point as reference. If the location point existed, the coordinates of the right endpoint were recorded; otherwise, the search was stopped. Finally, we cut the graph from the left to the right endpoints as the input sample of the deep learning network in the experiment.

In this paper, we used Faster R-CNN based on the ZF net to classify the ECG. As we can see in Figure 8, Faster R-CNN is composed of the ZF net, region proposal network (RPN) net, and Fast R-CNN net. Among them, the ZF net is a CNN architecture, which is used to extract the feature map of ECG

images. Then, the RPN net runs on the feature map and generates approximately 20,000 rectangular boxes, which are sorted according to scores from large to small. Then, the first 300 rectangular boxes are taken as inputs for the Fast R-CNN net, which maintains higher accuracy while reducing time complexity. Finally, the Fast R-CNN net outputs the probability of a category and a coordinate matrix (containing four coordinate values).

The region proposal network adopts the neural network, and integrates the three processes of generating candidate boxes, extracting features, and classifying them into a network model. Finally, the RPN realizes end-to-end training and detection. The RPN takes an image of any size as input and outputs a set of candidate boxes, where each box has a score for evaluating the similarity between the box and the target.

The small network slides on the convolution feature map of the last layer output of the shared convolutional layer. Each sliding window is mapped to a low-dimensional vector, which is output to two fully connected layers of the same level—the rectangular frame regression layer and the rectangular frame classification layer.

An anchor mechanism is proposed by Faster R-CNN. At each position of the sliding window, k area suggestions are predicted at the same time, so the rectangular frame regression layer has 4k outputs, that is, coordinate values of k boxes (x\_1, y\_1, x\_2, y\_2). The rectangular frame classification layer has 2k outputs. The center of the k anchors is the center of the current sliding window. Faster R-CNN uses three scales and three aspect ratios; thus, there are k = 9 anchors at each sliding position. There is a total of W × H × k anchors for a convolutional feature map of size W × H.

Each anchor is assigned a label that is a “target” (positive label) or “non-target” (negative label). The Faster R-CNN network inherits the multitasking loss mechanism of the Fast R-CNN network. For an anchor, its loss function is defined as

Изображение выглядит как текст

Автоматически созданное описание

**\*формулы пошли\***

There are many types of ECG beats, and many similarly shaped beats must rely on physicians with specialized experience to be able to accurately identify them. Table 2 shows the currently accepted classification standard for ECG beats.

Изображение выглядит как текст

Автоматически созданное описание

The Faster R-CNN framework used in this paper has three aspects for the classification of ECG beats: target position box, target classification, and score. The target position involves the coordinates of the upper left vertex and the lower right vertex of a rectangular box. The target classification involves the result of the model’s judgment on the image category at that position, which had only five discrete values in this experiment. The score is a probability value to show how likely the ECG beat is to be in this category, which is also called the confidence coefficient.

The classification performance is measured by four criteria, specificity (Spe), sensitivity (Sen), positive rate (Pre), and accuracy (Acc), as follows:

(1) Specificity (Spe): The proportion of normal ECG beats that are correctly classified, which represents the correct rate of non-patients.

(2) Sensitivity (Sen): The proportion of abnormal ECG beats that are correctly classified, which represents the correct ratio of patients.

(3) Positive rate (Pre): The number of abnormally classifiedEУCG beats is correctly classified as the proportion of abnormal ECG beats.

(4) Accuracy (Acc): Proportion of correct classification of ECG beats to all ECG beats

Table 3 shows the definition of the four classification results. According to the definitions of specificity, sensitivity, and accuracy, the formulas are as follows:

Изображение выглядит как текст

Автоматически созданное описание

Table 9 shows the classification results of Faster R-CNN. The classification accuracy of each class was more than 99%. The average classification accuracy rate was 99.21%.

Изображение выглядит как текст, снимок экрана, монитор

Автоматически созданное описание

The benefits of Faster R-CNN are as follows:

i. No manual feature extraction is required.

ii. The sampling rate of the original ECG signal does not need to be considered.

iii. The approach is insensitive to the ECG signal quality.

iv. High classification accuracy.

The drawbacks of Faster R-CNN are as follows:

i. Training set samples need to be manually labeled.

ii. Requires long training hours, and specialized hardware to efficiently train datasets (graphics processing unit (GPU)).

However, once the training of the ECG signals is completed, the classification of ECG heartbeat signals is fast. It takes about 0.025 s to classify a test sample of ECG beats.