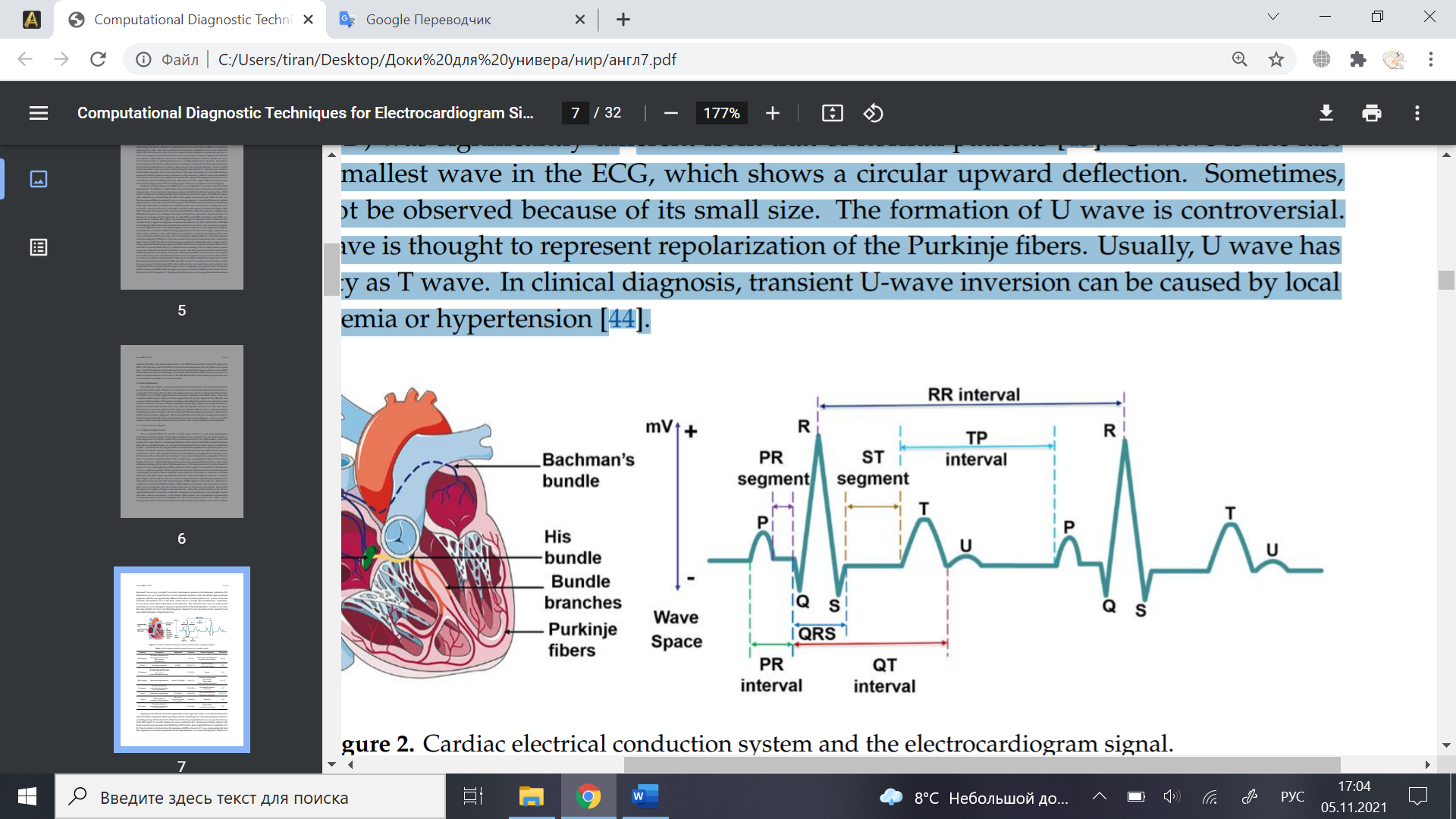
Computational Diagnostic Techniques for Electrocardiogram Signal Analysis Liping Xie \* , Zilong Li, Yihan Zhou, Yiliu He and Jiaxin Zhu College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China; longzilipro@gmail.com (Z.L.); 20185423@stu.neu.edu.cn (Y.Z.); 20185457@stu.neu.edu.cn (Y.H.); 20175563@stu.neu.edu.cn (J.Z.) \* Correspondence: xielp@bmie.neu.edu.cn Received: 8 September 2020; Accepted: 4 November 2020; Published: 5 November 2020

Вейвлет-преобразования широко используются для снижения шума ЭКГ благодаря их хорошим частотно-временным характеристикам, характеристикам сжатия коэффициентов, характеристикам разбавления шума и удалению избыточности. Вейвлет-преобразование не изменяет временные характеристики сигнала, а вейвлет-анализ обладает хорошими локальными характеристиками как во временной области, так и в частотной области, которая непрерывно разлагает сигнал на детальные коэффициенты высокой частоты и приблизительные коэффициенты низкой частоты. Поскольку шумовая составляющая обычно появляется в детальных коэффициентах, шумовой сигнал может быть отфильтрован посредством порогового квантования детального коэффициента. Наконец, сигнал ЭКГ может быть получен путем вейвлет-реконструкции высокочастотных и низкочастотных коэффициентов. Следовательно, выбор подходящих пороговых функций является ключом к достижению желаемого эффекта фильтрации шума. Эффективность стратегии пороговой обработки зависит от типа порогового метода и пороговых правил для конкретного приложения.

Врачи часто ставят диагноз, наблюдая за морфологическими характеристиками зубцов P-QRS-T, что во многом зависит от опыта врачей и обычно занимает много времени. Компьютерный анализ, основанный на проектировании признаков, значительно повысил эффективность и точность анализа ЭКГ. Для сигналов ЭКГ, состоящих из множества точек данных, можно выделить небольшое количество характеристик с точки зрения пиковой амплитуды, морфологии, распределения энергии и энтропии, частотного содержания, интервалов между событиями, которые могут представлять поведение сигнала ЭКГ. Эти функции, обычно применяемые для диагностики ЭКГ, включают функции P-QRS-T, статистические характеристики, морфологические особенности, характеристики в частотной области и другие более сложные параметры, которые обеспечивают эффективные инструменты для оценки врачей. Усовершенствованные алгоритмы извлекают функции в соответствии с потребностями задачи и автоматически выбирают конкретные функции для достижения точной диагностики. Как ядро ​​анализа ЭКГ, извлечение и выбор признаков играют решающую роль в работе алгоритма.

Зубец U - это последняя неустойчивая и самая маленькая волна на ЭКГ, которая показывает круговое отклонение вверх. Иногда зубец U может не наблюдаться из-за его небольшого размера. Формирование зубца U спорно. Обычно считается, что зубец U отражает реполяризацию волокон Пуркинье. Обычно зубец U имеет ту же полярность, что и зубец T. В клинической диагностике преходящая инверсия зубца U может быть вызвана локальной ишемией миокарда или гипертензией.



Изображение выглядит как текст

Автоматически созданное описание

**МЕТОД ФУРЬЕ**

Преобразование Фурье (FT) как хорошо известный инструмент для исследования сигнала в частотной области широко используется при обработке сигналов. FT преобразует сигнал из временной области в частотную, и можно легко найти амплитудный и фазовый спектр каждой частотной составляющей. Быстрое преобразование Фурье (БПФ) значительно сокращает время умножения, необходимое для вычисления дискретного преобразования Фурье, и значительно повышает скорость обработки. Анализ частотной области может быть использован для определения характерных точек для анализа сигнала ЭКГ. Комплекс QRS - самый яркая форма волны в сигнале ЭКГ, которая служит основой для автоматического определения частоты сердечных сокращений. Gothwal et al. идентифицировали пики в сигнале ЭКГ с помощью БПФ. Такие характеристики, как показатели интервала RR и частота сердечных сокращений, рассчитанная по этим точкам, были классифицированы нейронными сетями для выявления заболеваний. Хотя БПФ предоставляет подробную информацию о частоте, оно не может предоставить никакой информации относительно времени появления частотных компонентов для нестационарного входного сигнала. Эта проблема может быть решена с помощью кратковременного преобразования Фурье (STFT). STFT может использоваться для преобразования нестационарного сигнала из временной области в частотно-временную область с использованием оконной функции, которая также называется оконным преобразованием Фурье. Сигналы были разделены на серию квазистационарных частей в пределах периода окна. STFT был исследован спектральной информацией, которая предоставляет информацию о времени и частоте одновременно для сигнала, меняющегося во времени. Многие особенности частотной области в медицине связаны с определенными заболеваниями. При возникновении аритмии комплекс QRS, очевидно, изменится, что приведет к аномальному изменению высокочастотной составляющей. Minami et al. преобразовали каждый комплекс QRS в спектр Фурье, а затем рассчитали спектр мощности. Характеристики были переданы в алгоритм нейронной сети для классификации ритмов. Метод позволяет отличать наджелудочковые ритмы от желудочковых с высокой чувствительностью и специфичностью (≥0,98). Распределение энергии сигнала ЭКГ может быть вычислено с помощью STFT, затем характеристики извлекаются из распределения энергии и передаются в алгоритмы классификации для диагностики. Хуанг и др. использовали STFT и 2D-CNN для классификации аритмий ЭКГ. Сигналы ЭКГ во временной области были сначала преобразованы в частотно-временные спектрограммы с помощью STFT. Впоследствии спектрограммы на основе STFT были поданы в 2D-CNN, что привело к средней точности 99,00%. STFT извлекает частотно-временную информацию, касающуюся точного местоположения частотных компонентов сигнала ЭКГ, однако основным недостатком этого STFT является то, что существует компромисс между его временным и частотным разрешениями в STFT, главным образом из-за фиксированного размера окна. Длинное окно приводит к хорошему разрешению по частоте и плохому разрешению по времени, в то время как короткое окно приводит к хорошему разрешению по времени и плохому разрешению по частоте. Чтобы преодолеть этот недостаток, вейвлет-преобразование использует схему разрешения шкалы времени для анализа сигнала ЭКГ.

**ВЕЙВЛЕТ МЕТОД**

Вейвлет-анализ как мощный метод обработки сигналов для анализа временных рядов с множеством различных временных масштабов может разложить сигнал во временной шкале на сдвинутые и масштабированные версии базового вейвлета. По сравнению с преобразованием Фурье, вейвлет-преобразование предоставляет переменное «частотно-временное окно», которое позволяет нам динамически изменять его масштаб. Размер окна будет увеличиваться в области низких частот и сужаться в области высоких частот, поэтому это подходящий инструмент для всех диапазонов частот, что приводит к повышению точности как во временной, так и в частотной области. Вейвлет предоставляет метод сжатия или улучшения функций. Благодаря высокому разрешению по времени и частоте, вейвлет-преобразование может распознавать абстрактные и скрытые особенности сигналов ЭКГ. Существуют различные базы вейвлетов, такие как Морле, Мексиканская шляпа, Мейер, Добеши, Симлет и Койфлет, Хаар и биортогональные вейвлеты. Жизненно важно выбрать подходящую основу вейвлета и соответствующее количество уровней разложения для интересующего сигнала с использованием вейвлет-преобразования. Различные базы вейвлетов имеют свои характеристики и подходят для конкретного приложения, например, вейвлет Хаара прост и быстр, обеспечивая при этом эффективность памяти, и он может разделять сигнал и шум без значительной потери информации о сигнале. Вейвлет Мейера наделен лучшими характеристиками локализации, в то время как биортогональный вейвлет полезен для реконструкции сигнала, поскольку он имеет линейные блоки фазовых фильтров с симметричным свойством. Вейвлет Добеши обладает симметрией с энергетическим спектром, в основном распределенным вокруг низких частот, что, как обнаружено, более подходит для обнаружения R-пика. На практике соответствующую вейвлет-функцию следует выбирать в соответствии с такими параметрами, как поддерживающая длина, симметрия, исчезающий момент, регулярность и подобие. При разумном выборе баз вейвлетов следует должным образом учитывать эти свойства для конкретного приложения. В семействе вейвлетов с дискретным вейвлет-преобразованием (DWT) легко работать на компьютере, а эффективность вычислений относительно выше по сравнению с другими типами вейвлетов. DWT также наиболее широко используется при анализе ЭКГ. Jayachandran et al. использовали свойства множественного разрешения вейвлет-преобразования для выявления незначительных изменений в сигнале ЭКГ для выявления инфаркта миокарда (ИМ). Они использовали DWT для разложения ЭКГ на различные уровни разрешения. Путем сравнения энергоэнтропийных характеристик сигнала ЭКГ от 2282 нормальных и 718 ИМ в вейвлет-области, предложенный метод позволил различить нормальные и ИМ ЭКГ с точностью более чем 95%. Hong et al. применили дискретное вейвлет-пакетное преобразование (DWPT) для выделения признаков на ЭКГ в 12 отведениях с высокой размерностью и множественными сигналами. Было обнаружено, что многолинейный анализ главных компонент (MPCA) на основе тензора вейвлетов может извлекать признаки более эффективно, чем PCA на основе векторов. До сих пор существует множество методов вейвлет-анализа, которые оказались очень эффективными при анализе ЭКГ. Mohit Kumar et al. применили метод гибкого аналитического вейвлет-преобразования (FAWT) для выделения признаков и вычислили параметр перекрестного информационного потенциала (CIP) из реальных значений коэффициентов FAWT для захвата скрытой информации из сокращений ЭКГ. Высокая точность классификации обычно достигается за счет оптимизации уровня декомпозиции. Однако точность метода DWT часто зависит от предположения, что форма базовой функции аналогична форме сигналов ЭКГ. Трудно предсказать форму базовой функции, особенно для нестационарных сигналов ЭКГ.

**МЕТОД СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК**

Статистические характеристики, такие как среднее значение, дисперсия, отклонение, энтропия Шеннона, также оказались эффективными при анализе ЭКГ. Статистические характеристики невысоки и не зависят от реперных точек сигналов ЭКГ. Таким образом, на них обычно не влияют ложные реперные точки сигнала ЭКГ. Кроме того, статистические функции также используются при обнаружении шумовых артефактов для сбора сигналов для повышения точности. Ли и др. разработал метод в реальном времени для обнаружения артефактов движения и шума (MN). Они разделили высокочастотные компоненты с помощью функции внутренней моды первого порядка (F-IMF). Затем энтропия, среднее значение и дисперсия Шеннона, рассчитанные на основе временного ряда F-IMF, были выбраны в качестве характеристик артефактов MN для обнаружения присутствия артефактов MN с чувствительностью 96,63%. Морфологические особенности сигналов ЭКГ относятся к показателям, полученным ЭКГ после математической морфологической обработки, которые обычно достигаются путем морфологической фильтрации. Как разновидность нелинейного преобразования, морфологический фильтр может локально изменять геометрические характеристики сигнала. Одним из важнейших понятий морфологической фильтрации является структурный элемент, называемый зондом. Связь между частями сигнала проверяется перемещением щупа. По сравнению с методом частотно-временного преобразования соответствующие структурные элементы могут хорошо сохранять сигнал. Однако большинство статистических признаков и морфологических признаков не являются исчерпывающими и заметными, и большинство исследований не используют только эти признаки, но часто используют их в сочетании с другими видами признаков или используют статистические методы для дальнейшего извлечения существующих признаков. Raj et al. объединил метод декомпозиции словаря со статистическими характеристиками и предложил метод извлечения статистических признаков на основе сверхполного словаря Габора. Сигналы ЭКГ были разложены на элементарные волны переполненным словарем Габора. Из элементарных волн были извлечены четыре статистических характеристики, включая временную задержку, частоту, параметр ширины и квадрат коэффициента расширения. Эти характеристики были переданы в модели классификаторов и оптимизированную SVM-двойник по методу наименьших квадратов, достигнув высокой точности классификации 99,11%. Этот метод, сочетающий словарную декомпозицию и статистические характеристики, значительно повысил скорость обнаружения ЭКГ, сократил объем вычислений и повысил точность. По сравнению со статистикой первого и второго порядка (например, среднее значение и дисперсия) статистика более высокого порядка (HOS) привлекла большое внимание в области анализа ЭКГ. Доказано, что статистика более высокого порядка подходит для оценки асимметрии и эксцесса и не учитывает аддитивный гауссов шум. Мариньо и др. использовали статистические методы, включая HOS и матрицу структурной совместимости для выделения признаков, а также классификаторы (машина опорных векторов, многослойный персептрон, байесовский метод и лес оптимальных путей) для распознавания аритмий. Набор функций, основанный на HOS, получил самую высокую точность по сравнению с другими функциями. Комбинация HOS и байесовского классификатора позволила достичь наивысшей точности 94,3%. Результаты показали, что использование HOS для описания типов аритмии клинически надежно.

В клинической практике все изменения параметров ЭКГ выявляются путем визуальной оценки и ручной интерпретации для выявления сердечно-сосудистых заболеваний. Однако из-за нестационарной и нелинейной природы сигналов ЭКГ индикаторы сердечно-сосудистых заболеваний могут появляться на временной шкале случайным образом. Нелинейные характеристики ЭКГ, такие как энергия (Ee), энтропия (Ez), фрактальная размерность (FD) и относительный вейвлет (RWz), могут быть извлечены, чтобы показать некоторые диагностические детали, которые нельзя просто обнаружить с помощью визуальной оценки. Нелинейные методы могут лучше работать в сложных нелинейных отношениях между функциями. Локально линейное встраивание (LLE) - это метод уменьшения нелинейной размерности без учителя, который пытается сохранить структуру данных с помощью нелинейного метода в соответствии с локальными особенностями набора данных.

Он может работать намного быстрее, чем другие подходы. Однако LLE чрезвычайно чувствителен к шумам, и LLE не будет хорошо сохранять локальную геометрию наборов данных в пространстве внедрения, если в данных есть выбросы. Метод на основе ядра стал одним из самых популярных подходов к извлечению сложной нелинейной информации, встроенной в набор данных ЭКГ. Ядро LLE (KLLE) может восстанавливать нелинейные данные как линейную комбинацию своих соседей. Ли и др. сопоставил 12-мерные особенности сегментов ЭКГ типа одиночных сокращений в 7-мерное пространство вложения, описываемое двумя координатами ядра LLE. Результаты показали, что соотношение истинно / ложно для предложенного метода превосходит другие выбранные методы. Пространство KLLE показывает лучшую производительность при потере диагностической информации в сигналах ЭКГ. Анализ основных компонентов ядра (KPCA) - это популярное нелинейное обобщение PCA, которое подходит для обработки линейно несепарабельных наборов данных. Основная идея KPCA состоит в том, чтобы отобразить исходные данные в многомерное пространство с помощью функции ядра (Φ), а затем применить к нему стандартный алгоритм PCA, который может извлечь более полное нелинейное представление основных компонентов. Однако KPCA требует больше времени, чем PCA.

Модель классификации отвечает за вынесение суждений на основе ряда входных данных, что в конечном итоге позволяет достичь цели диагностики заболевания. Ранние классификаторы, состоящие из серии пороговых значений и оценочных предложений, могут имитировать человеческую логику для выполнения некоторых простых задач. Ранние модели легко интерпретируемы, но они не могут эффективно анализировать сложные задачи. В настоящее время модели машинного обучения становятся зрелыми и могут поддерживать точность обнаружения аритмии более 90%. Однако в процессе анализа разработка функций по-прежнему является трудоемкой и требует чрезмерных шагов обработки, которые значительно ослабляют эффективную информацию о данных. Модели глубокого обучения - это мощные аналитические модели, появившиеся в последние годы. Хотя этот тип модели требует больших вычислительных ресурсов, он является мощным и может значительно сократить использование искусственных свойств. Сквозной анализ, основанный на моделях глубокого обучения, может напрямую обрабатывать необработанные данные без ручных методов, что обеспечивает возможность полностью автоматического анализа ЭКГ, особенно при классификации сложных проблем.

**МЕТОД МАШИННОГО ОБУЧЕНИЯ**

Polat и Güne¸s применили анализ главных компонентов (PCA), чтобы уменьшить 279-мерные особенности до 15 измерений, и использовали машину опорных векторов наименьших квадратов (LS-SVM) для обнаружения аритмий ЭКГ. Полученная точность классификации составляет 100% с 70–30% набора данных обучающего теста или 80–20% набора данных обучающего теста.

Производительность модели можно улучшить с помощью нескольких классификаторов машинного обучения для оптимизации решений. Osowski et al. предложила высоконадежную экспертную систему на основе двух SVM для распознавания сердцебиения. Система применила статистику высшего порядка (HOS) и характеристику комплекса QRS по Эрмиту для генерации признаков, и два подмножества признаков были введены в классификатор SVM, образуя два нейронных классификатора. Метод наименьших квадратов был применен для оптимизации весов схемы взвешенного голосования для объединения двух классификаторов. Частота ошибок 4,09% после слияния была значительно ниже, чем у двух классификаторов, которые составляли 5,43% и 6,28% соответственно. Ye et al. предложил метод слияния с двумя отведениями, основанный на оценке вероятности для автоматической классификации сердечных сокращений. Процедура была независимо применена к данным от двух отведений ЭКГ, и два решения были объединены для окончательного решения по классификации. Метод достиг высокой точности 99,3% на основе MIT-BIH и эффективно снизил частоту ложного обнаружения.

Сквозная модель предоставляет возможность для полного интеллектуального анализа ЭКГ, который по своей сути объединяет извлечение и классификацию признаков в обучающий объект и выводит на них процесс принятия решений. Основываясь на 91 432 ЭКГ в одном отведении от 53 549 пациентов, команда Ханнана разработала DNN для различения 12 различных сердечных ритмов. Они сравнили классификационную чувствительность модели DNN со средней специфичностью, достигнутой кардиологами. Результаты показали, что площадь под кривой рабочей характеристики приемника (ROC) модели составила 0,97. Значение F1 модели составило 0,837, что выше среднего показателя кардиолога 0,780. Эти результаты продемонстрировали, что предлагаемый подход сквозного глубокого обучения может классифицировать широкий спектр различных аритмий с высокой диагностической эффективностью, аналогичной показателям кардиологов. Улучшение может быть отнесено за счет способности DNN к обучению признакам, которая может реализовывать функцию выделения и классификации признаков. CNN, состоящие из иерархических нейронных сетей, используются как для извлечения признаков, так и для классификации. Его сверточные слои чередуются со слоями субдискретизации, которые могут научиться извлекать индивидуальные особенности пациента. Полностью связанные слои, следующие за сверточными слоями, идентичны MLP. Эти слои выполняют задачу классификации для принятия окончательного решения. Kiranyaz et al. предложила адаптивную 1D CNN с тремя слоями CNN и двумя слоями MLP для классификации ЭКГ в реальном времени отдельного пациента, что позволяет избежать необходимости извлечения каких-либо характеристик вручную и постобработки, поэтому показывает большой потенциал для работы в режиме реального времени. реализация для мониторинга сердца из-за относительной вычислительной сложности. Zhai et al. предложил 2D CNN-сеть для классификации аритмий. По сравнению с 1D-CNN, предложенной Kiranyaz et al., Эта модель улучшает чувствительность суправентрикулярных эктопических сокращений (SVEB) более чем на 12,2%. Матрица фиксирует морфологию одиночного сердцебиения и временные отношения в соседних сердцебиениях путем преобразования трех соседних сердечных сокращений в ЭКГ в двумерную матрицу связи. Следовательно, по сравнению с 1D CNN, этот метод может более эффективно исследовать взаимосвязь между различными компонентами ЭКГ в соседних сердцебиениях, тем самым улучшая производительность классификатора. В дополнение к CNN, рекуррентная нейронная сеть (RNN) может полностью использовать информацию во временной области, поэтому она широко используется при изучении естественного языка и анализе сигналов. Долговременная краткосрочная память (LSTM) как популярный тип RNN, который сохраняет информацию во временной области, которая играет важную роль при обработке последовательных данных. Tan et al. применили CNN для извлечения признаков ЭКГ и LSTM в качестве классификатора для автоматической диагностики сердечно-сосудистых заболеваний. За счет использования трехслойной LTSM вместо CNN в качестве выходного уровня модель имеет лучшую диагностическую производительность, чем 11-слойная модель CNN, предложенная Acharya et al. Йилдирим предложил модель с хорошей эффективностью и низкими вычислительными затратами для распознавания сигналов аритмии ЭКГ. Модель сжатия на основе глубокого сверточного автокодировщика (CAE) в сочетании с сетями LSTM для распознавания ритмов ЭКГ, что значительно сократило время обучения с 4,5 до 0,6 часа. По сравнению с классической технологией машинного обучения обучение с подкреплением обеспечивает более быстрый механизм обучения и более адаптируется к изменениям в окружающей среде. Андерсен и др. предоставила надежную сквозную модель в реальном времени, включая CNN и LSTM, для автоматического обнаружения фибрилляции предсердий в длительных записях ЭКГ.

Сквозные нейронные сети также имеют некоторые недостатки. Во-первых, из-за отсутствия предварительных знаний для поддержки сквозные нейронные сети полагаются на высококачественные наборы данных по сравнению с классическими методами машинного обучения. Более крупный и полный набор данных поможет нейронной сети хорошо адаптироваться. Кроме того, нейронные сети с глубокой архитектурой, включающие большое количество параметров, легко переигрывают обучающие данные. Следовательно, в тренировочном процессе требуются определенные тренировочные навыки. В-третьих, модель становится черным ящиком, что снижает интерпретируемость нейронной сети. Дизайн модели может быть изменен только структурой модели и методами обучения для получения лучших результатов. Однако, благодаря своим мощным функциям, сквозные нейронные сети по-прежнему будут одним из основных направлений исследования анализа ЭКГ.

Wavelet transforms are widely used in ECG noise reduction due to their good time-frequency characteristics, coefficient compression characteristics, noise dilution characteristics and removal of redundancy. The wavelet transform does not change the signal’s time characteristics, and the wavelet analysis possesses good local characteristics in both time domain and frequency domain, which continuously decomposes the signal into the detail coefficients of high frequency and the approximate coefficients of low frequency. Since the noise component usually appears in the detail coefficients, the noise signal can be filtered out by threshold quantization of the detail coefficient. Finally, the ECG signal can be obtained by wavelet reconstruction of the high-frequency and low-frequency coefficients. Therefore, the selection of appropriate threshold functions is the key to achieve the desired effect of the noise-filtering effect. The performance of the thresholding strategy depends on the type of threshold method and the threshold rules for a specific application.

Doctors often make diagnoses by observing the morphological characteristics of P-QRS-T waves, which largely depends on doctors’ experience and usually takes a long time. Computer-aided analysis based on feature engineering has greatly improved the efficiency and accuracy in ECG analysis. ECG signals consisting of abundant data points can be extracted a small number of features in terms of its peak amplitude, morphology, energy and entropy distribution, frequency content, intervals between events, which can represent the behavior of the ECG signal. These features commonly applied to ECG diagnosis include P-QRS-T features, statistical features, morphological features, frequency-domain features and other more complex parameters, which provide effective tools for doctors’ judgment. Advanced algorithms extract features according to the needs of the task and automatically selects specific features to achieve precise diagnosis. As the core of ECG analysis, feature extraction and selection play a decisive role in the performance of the algorithm.

U-wave is the last unsteady and smallest wave in the ECG, which shows a circular upward deflection. Sometimes, U-wave may not be observed because of its small size. The formation of U wave is controversial. Generally, U-wave is thought to represent repolarization of the Purkinje fibers. Usually, U wave has the same polarity as T wave. In clinical diagnosis, transient U-wave inversion can be caused by local myocardial ischemia or hypertension.

Изображение выглядит как текст

Автоматически созданное описание

Fourier transform (FT) as a well-known tool for investigating a signal in frequency domain, is widely used in signal processing. FT converts a signal from the time domain into the frequency domain, and the amplitude spectrum and phase spectrum of each frequency component can be easily found. Fast Fourier Transform (FFT) greatly reduces the times of multiplication required for computing the discrete Fourier transform and significantly improves the processing speed. Frequency domain analysis can be used to locate feature points for ECG signal analysis. The QRS complex is the most striking waveform within the ECG signal, which serves as the basis for the automated determination of the heart rate. Gothwal et al. identified the peaks in the ECG signal using FFT. Feature as RR interval metrics and heart rate calculated from these feature points were classified by neural networks to identify the diseases. Although FFT provides detailed frequency information, it fails to provide any information regarding the time of occurrence of the frequency components for a non-stationary input signal. This problem can be solved by short-time Fourier transform (STFT). The STFT can be used to convert non-stationary signal from time domain to time-frequency domain using a window function, which is also called windowed Fourier transform. The signals were separated into a series of quasi-stationary parts within the window period. STFT has been explored spectral information that provides time and frequency information simultaneously for a signal varying with time. Many frequency-domain features in medicine are associated with certain diseases. When arrhythmia occurs, QRS complex will change obviously, resulting in abnormal change of high-frequency component. Minami et al. converted each QRS complex to a Fourier spectrum, and then the power spectrum was calculated. The features were fed to neural network algorithm for rhythm classification. The method could discriminate supraventricular rhythms from ventricular ones with high sensitivity and specificity (≥0.98). Energy distribution of the ECG signal can be computed by STFT, the features are then extracted from the energy distribution and fed to classification algorithms for diagnosis. Huang et al. utilized STFT and 2D-CNN for ECG arrhythmia classification. The time-domain signals of ECG were first transformed into time-frequency spectrograms by STFT. Subsequently, the STFT-based spectrograms were fed to a 2D-CNN, resulting in an averaged accuracy of 99.00%. STFT extracts time-frequency information regarding the exact location of frequency components of ECG signal, however, the major drawback of this STFT is that there is a tradeoff between its time and frequency resolutions in STFT mainly due to the fixed window size. A long window results in good frequency resolution and poor time resolution, while short window results in good time resolution and poor frequency resolution. To overcome this drawback, the wavelet transform employs a time-scale resolution scheme for ECG signal analysis.

Wavelet analysis as a powerful signal processing technique for analyzing time series with many different timescales can decompose the signal on a time-scale domain into shifted and scaled versions of the base wavelet. Compared with Fourier transform, wavelet transform provides a variable “time-frequency window”, which allows us to change its scale dynamically. The window size will wide at low frequencies region and narrow at high frequencies region, therefore, it is a suitable tool for all frequency ranges, resulting in an improvement of accuracy both in time and frequency domain. Wavelet provides a method for compression or enhancement features. Because of its high resolution of time and frequency, wavelet transform can recognize the abstract and hidden features of ECG signals. There are various wavelet bases, such as Morlet, Mexican hat, Meyer, Daubechies, Symlets and Coiflets, Haar and Biorthogonal wavelets. It is vital to choose a suitable wavelet base and appropriate number of decomposition levels for the signal of interest using the wavelet transform. Different wavelet bases have their characteristics and are suitable to a specific application, for example, the Haar wavelet is simple and fast while esulting in memory efficiency, and it can separate signal and noise without significant loss of the signal information. Meyer wavelet is endowed with better localization characteristics, while biorthogonal wavelet is useful for signal reconstruction because it has linear phase filter banks with symmetric property. Daubechies wavelet possesses symmetry with the energy spectrum mainly distributed around low frequencies, which is found more suitable for the detection of R-peak. In practice, the appropriate wavelet function should be selected according to the parameters such as supporting length, symmetry, vanishing moment, regularity and similarity. A reasonable selection of wavelet bases should pay proper consideration to these properties for a specific application.

In wavelet family, discrete wavelet transform (DWT) is easy to work on a computer, and the calculation efficiency is relatively higher compared with other types of wavelets. DWT is also the most widely used in the analysis of ECG [66–70]. Jayachandran et al. took advantage of multiresolution properties of wavelet transformation to identify subtle сhanges in the ECG signal for detection of myocardial infarction (MI). They used DWT to decompose ECG into various resolution levels. By comparing energy–entropy characteristics of ECG signal from 2282 normal and 718 MI in the wavelet domain, the proposed method discriminated the normal and MI ECG beat with more than 95% accuracy. Hong et al. applied discrete wavelet packet transform (DWPT) to perform feature extraction on 12-lead ECG with high-dimensionality and multiple signals. It was found that wavelet tensor-based multilinear principal component analysis (MPCA) could extract features more efficiently than vector-based PCA. There are still many wavelet analysis methods that have been proved to be very effective in ECG analysis. Mohit Kumar et al. applied flexible analytic wavelet transform (FAWT) technique to feature extraction, and computed cross information potential (CIP) parameter from the real values of the coefficients of FAWT to capture the hidden information from ECG beats. High classification accuracy was usually achieved by optimizing decomposition level. However, the accuracy of DWT method often depends on the assumption that the shape of the basis function is similar to that of the ECG signals. It is difficult to predict the shape of the basis function, especially for non-stationary ECG signals.

Statistical features, such as mean value, variance, deviation, Shannon entropy, have also been proved to be effective in ECG analysis. The statistical features are low complexity and independent of fiducial points of the ECG signals. Thus, they are usually unaffected by fake fiducial points of the ECG signal. Besides, the statistical features are also used in detection of noise artifacts to pick up signals for improving accuracy. Lee et al. developed a real-time method for detection of motion and noise (MN) artifacts. They separated high-frequency components using first-order intrinsic mode function(F-IMF). Then, the Shannon entropy, mean, and variance calculated from the F-IMF time series were selected as features of MN artifacts to detect the presence of MN artifacts with sensitivity of 96.63%. Morphological features of ECG signals refer to the indicators obtained by the ECG after mathematical morphological processing, which are usually achieved by morphological filtering. As a kind of nonlinear transformation, the morphological filter can locally modify geometric characteristics of a signal. One of the most important concepts in morphological filtering is a structural element, called probe. The relationship between the parts of the signal is checked by moving the probe. Compared with the time-frequency transform method, appropriate structural elements can preserve the signal well. However, most of the statistical features and morphological features are not comprehensive and prominent, and most of the studies do not use these features alone, but often use them in combination with other kinds of features, or use statistical methods to further extract the existing features. Raj et al. combined dictionary decomposition method with statistical characteristics, and proposed an overcomplete Gabor dictionary-based statistical feature extraction method. ECG signals were decomposed into elementary waves by the overcomplete Gabor dictionary. Four statistical features, including time delay, frequency, width parameter, and square of expansion coefficient were extracted from the elementary waves. These features were fed to classifier models, and optimized least-square twin SVM, achieving a high classification accuracy of 99.11%. This method combining dictionary decomposition and statistical characteristics, significantly improved the ECG detection speed, reduced computation, and improved accuracy. Compared with first-order and second-order statistics (e.g., mean and variance), higher-order statistics (HOS) have attracted extensive attention in the field of ECG analysis. Higher-order statistics are proved to be suitable for estimation of skewness and kurtosis, and are blind to additive Gaussian noise. Marinho et al. used statistical methods including HOS and structural-co-occurrence matrix for features extraction, and used classifiers (support vector machine, multi-layer perceptron, bayesian, and optimum-path forest) to discriminate arrhythmias. The feature set based on HOS obtained the highest accuracy compared with other features. The combination of HOS and Bayesian classifier achieved the highest accuracy of 94.3%. The results demonstrated that it is clinically reliable to use HOS for describing types of arrhythmia.

In clinical practice, all the changes of ECG parameters are detected by visual evaluation and manual interpretation to detect the presence of cardiovascular disease. However, due to the nonstationary and nonlinear nature of ECG signals, cardiovascular disease indicators may appear randomly on the time scale. Non-linear features of ECG, such as energy (Ee), entropy (Ez), fractal dimension (FD), and relative wavelet (RWz) can be extracted to show some diagnostic details that can’t be simply detected by visual evaluation. Nonlinear methods can perform better in complex nonlinear relationships among the features [110]. Locally linear embedding (LLE) is an unsupervised nonlinear dimensionality reduction method, which tries to preserve the data structure by a non-linear method according to local features of the dataset.

It can perform much faster than the other approaches. However, LLE is extremely susceptive to noises, and LLE will not preserve well the local geometry of the data sets in the embedding space if there are outliers in the data. Kernel-based method has become one of the most popular approaches to extract the complicated nonlinear information embedded on an ECG dataset. Kernel LLE (KLLE) can reconstruct nonlinear data as a linear combination of its neighbors. Li et al. mapped 12-dimensional features of ECG segments of single beat type into 7-dimensional embedding space described by two coordinates of kernel LLE. The results showed that the true/false ratio for the proposed method outperformed other selected methods. The KLLE space shows better performance in loss of diagnostic information in ECG signals. Kernel principal component analysis (KPCA) is a popular nonlinear generalization of PCA, which is suitable for processing linearly nonseparable data sets. The basic idea of KPCA is to map the original data into a high dimensional space via a kernel function (Φ), and then to apply the standard PCA algorithm to it, which can extract a more complete nonlinear representation of the principal components. However, KPCA takes more time than PCA.

Classification model is responsible for making judgments based on a series of input data, thus ultimately achieving a goal of disease diagnosis. The early classifiers consisting of a series of thresholds and judgment sentences can simulate the human logic to perform some simple tasks. The early models are highly interpretable, but they cannot effectively analyze complex tasks. Machine learning models currently become mature and can maintain more than 90% accuracy in arrhythmia detection. However, in the analysis process, feature engineering is still labor-intensive, and needs excessive processing steps which greatly weaken the effective information of the data. Deep learning models are powerful analytical models that have emerged in recent years. Although this type of model is computationally expensive, it is powerful and can greatly reduce the use of artificial features. The End-to-End analysis based on deep learning models can directly process raw data without hand-crafted techniques, which provides the possibility for fully automatic analysis of ECG, especially in the classification of complex problems.

Polat and Güne¸s applied principal component analysis (PCA) to reduce 279-dimensional features to 15 dimensions, and used least square support vector machine (LS-SVM) for detection of ECG arrhythmias. The obtained classification accuracy is 100% with 70–30% of training-test dataset or 80–20% of training-test dataset.

Performance of the model can be improved by multiple machine learning classifiers for decision optimization. Osowski et al. proposed a highly reliable expert system based on two SVMs for heartbeat recognition. The system applied higher order statistics (HOS) and Hermite characterization of QRS complex for generation of features, and the two subsets of features were fed into SVM classifier, forming two neural classifiers. Least mean square method was applied to optimize the weights of the weighted voting integrating scheme for the combination of the two classifiers. The error rate of 4.09% after fusion was significantly lower than that of the two classifiers, which were 5.43% and 6.28% respectively. Ye et al. proposed a two-lead fusion method based on probability estimation for automatic heartbeat classification. The procedure was independently applied to the data from two ECG leads and the two decisions were fused for the final classification decision. The method achieved a high accuracy rate of 99.3% based on MIT-BIH, and effectively reduced the false detection rate.

The end-to-end model provides the possibility for complete intelligence of ECG analysis, which inherently fuses feature extraction and classification into the learning body and outputs decision-making. Based on 91,432 single-lead ECGs from 53,549 patients, Hannun’s team developed a DNN to distinguish 12 different heart rhythms. They compared the classification sensitivity of the DNN model with the average specificity achieved by cardiologists. The results showed that the area under the receiver operating characteristic curve (ROC) of the model was 0.97. The F1 value of the model was 0.837, which was higher than the cardiologist’s average of 0.780. These findings demonstrated that the proposed end-to-end deep learning approach could classify a broad range of distinct arrhythmias with high diagnostic performance similar to that of cardiologists. The improvement may be attributed to the feature learning capability of DNN, which can realize the function of feature extraction and classification. CNNs consisting of hierarchical neural networks are used for both feature extraction and classification. Its convolutional layers alternate with subsampling layers, which can learn to extract patient-specific features. The fully connected layers following the convolutional layers are identical to MLP. These layers perform the classification task to produce the final decision. Kiranyaz et al. proposed an adaptive 1D CNN with three CNN layers and two MLP layers for real-time ECG classification of an individual patient, which avoids the need for any manual feature extraction and postprocessing, therefore, it shows great potential for a real-time implementation for heart monitoring because of its relative computational complexity. Zhai et al. proposed a 2D CNN network for arrhythmia classification. Compared with the 1D-CNN proposed by Kiranyaz et al., this model improves the sensitivity of supraventricular ectopic beats (SVEB) by more than 12.2%. The matrix captures the morphology of a single heartbeat and the time relationship in the adjacent heartbeats by converting the three adjacent heartbeats in the ECG into a 2D coupling matrix. Hence, compared with 1D CNN, this method can more effectively investigate the relationship between various ECG components in adjacent heartbeats, thereby improving the performance of the classifier. In addition to CNN, recurrent neural network (RNN) can take full use of time-domain information, so it is widely used in natural language learning and signal analysis. Long short-term memory (LSTM) as a popular type of RNN that retains time domain information, which is prominent in the processing of sequential data. Tan et al. applied CNN to extract ECG features, and LSTM as a classifier to automatically diagnose CVD. By using three-layer LTSM instead of CNN as the output layer, the model has better diagnostic performance than the 11-layer CNN model proposed by Acharya et al. Yildirim proposed a model with good effectivity and low computational cost to recognize ECG arrhythmia signals. A deep convolutional auto-encoder (CAE) based compression model combined with LSTM networks to recognize ECG beats, which significantly reduced the training from 4.5 h to 0.6 h. Compared with the classic ML technology, reinforcement learning provides a faster learning mechanism and is more adaptable to changes in the environment. Andersen et al. provided a robust real-time End-to-end model including CNN and LSTM for automatic detection of atrial fibrillation in long-time ECG recordings.

End-to-end neural networks also have some disadvantages. First, because there is no prior knowledge to support, end-to-end neural networks rely on high-quality data sets compared to classic machine learning methods. A larger and more comprehensive data set will help the neural network to fit well. Besides, neural networks with deep architectures, which involve a large number of parameters, easily overfit the training data. Hence, certain training skills are required in the training process. Third, the model becomes a black box, which reduces the interpretability of the neural network. The model design can only be modified by model structure and training methods to obtain better results. However, because of its powerful functions, end-to-end neural networks will still be one of the main research directions of ECG analysis.