

基于睡眠心电数据的人工智能抑郁评估系统

肖碧波¹, 唐健豪², 曾加玉³, 王肖文², 魏晓滨¹

1. 佛山市第三人民医院设备科, 广东 佛山 528041; 2. 佛山市第一人民医院设备科, 广东 佛山 528000; 3. 南部战区空军医院医学工程科, 广东 广州 510062

【摘要】睡眠障碍是抑郁症的评估因素之一, 可从心电数据中提取特征值用于抑郁评估, 辅助临床诊断。设计开发一种基于睡眠心电数据的人工智能抑郁评估系统, 其中包含心电数据的分析、心电滤波和心电R波的定位。其间并进行数图转换, 生成心电RR间期图。利用TensorFlow进行图像的增强处理后, 构建卷积神经网络, 最终得出睡眠心电数据的人工智能抑郁评估模型。

【关键词】睡眠; 心电数据; 人工智能; 抑郁

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2022)12-1538-06

Artificial intelligence system for depression assessment based on sleep ECG data

XIAO Bibo¹, TANG Jianhao², ZENG Jiayu³, WANG Xiaowen², WEI Xiaobin¹

1. Department of Equipment, the Third People's Hospital of Foshan, Foshan 528041, China; 2. Department of Equipment, the First People's Hospital of Foshan, Foshan 528000, China; 3. Department of Medical Engineering, Air Force Hospital of Southern Theatre Command, Guangzhou 510062, China

Abstract: Sleep disorder is one of the evaluation factors of depression, and the characteristic values can be extracted from ECG data for depression assessment and clinical diagnosis. An artificial intelligence system using sleep ECG data to carry out depression assessment is designed and developed, with the functions of ECG data analysis, ECG filtering and ECG R wave localization. The ECG RR interval graph is generated through the conversion from data to graph. After image enhancement with TensorFlow, a convolution neural network is constructed, and finally an artificial intelligence model for depression assessment based on sleep ECG data is obtained.

Keywords: sleep; ECG data; artificial intelligence; depression

前言

最近有证据表明睡眠障碍先于抑郁, 伴有睡眠障碍的抑郁症患者可能表现出更严重的症状和治疗困难, 甚至有研究认为改善睡眠就是改善抑郁^[1]。基于睡眠心电数据的抑郁症研究显示, 睡眠质量与抑郁症及其严重程度和预后存在明显的相关性^[2-4]。抑郁症的典型症状主要为情感低落、思维迟缓以及言语动作减少^[5]。抑郁症患者给家庭和社会带来沉重的负担, 约15%的抑郁症患者死于自杀^[6-7]。世界卫

生组织、世界银行和哈佛大学的一项联合研究表明, 抑郁症已经成为中国疾病负担的第2大病^[8]。失眠在抑郁症患者中非常普遍, 数据显示至少2/3的抑郁症患者存在失眠问题^[9]。有研究发现改善抑郁症患者的睡眠质量能够减轻抑郁症的严重程度^[10]。因此, 评估抑郁症患者的睡眠质量十分重要^[11]。在评估抑郁症的方法上, 睡眠的心电分析能有效对抑郁症与正常人进行区分^[12]。基于抑郁症患者睡眠心电的差异能使用人工智能(AI)学习的方式对患者症状进行评估。本研究设计了一种通过分析睡眠心电数据并进行AI机器学习, 进而对抑郁症作出有效评估的方法。

1 系统架构

系统基于Python3.7设计, 由数图转换和图像训练两大部分组成, 见图1。

【收稿日期】2022-06-14

【基金项目】广东省自然科学基金(2021A1515011735); 佛山市医学类科技攻关项目(2020001005781)

【作者简介】肖碧波, 硕士, 高级工程师, 研究方向: 生物医学工程, E-mail: xiaobibo@163.com

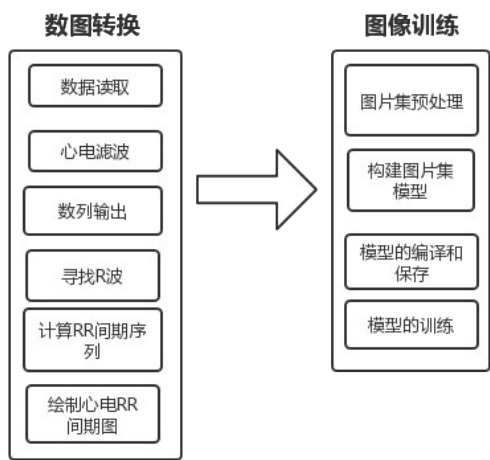


图1 系统总结构

Figure 1 General structure of the system

数图转换主要实现分析采集的心电信号、读取心电数据。对心电数据进行滤波得出R波序列,通过对R波序列的分析,构建RR间期的庞加莱图,并将RR间期的庞加莱图进行保存。

图像训练是利用数图转换后的RR间期庞加莱图进行图像分类识别,设计的前提是需要由医生判定被采集心电数据的人员是否属于抑郁症患者。通过获得足量的抑郁症患者心电数据和正常人的心电数据,利

用抑郁症患者和正常人的数图转换后的RR间期庞加莱图进行机器的二分类训练,由系统对抑郁症和正常人的图像进行识别学习,最终得出图像训练的模型。

2 数图转换

通过Python读取心电数据,并将心电数据进行滤波,目的是寻找心电数据中的R波位置。通过寻找R波位置,计算得出RR序列,并绘制心电RR序列的散点图。通过AI训练使机器分辨出正常人与抑郁证患者的心电RR间期不同并实现匹配。

2.1 数据读取

心电采集使用SEU-SLEEP睡眠检测仪,检测仪采集使用者的睡眠心电数据,并以dat文件格式对使用者的心电数据进行存储。使用Hexeditor查看,数据格式为16位十六进制数据。其中数据包括电极脱落标志、呼吸数据、运动数据和心电数据。

在对使用者的抑郁症评估中,系统设计需要采用使用者的有效心电数据。系统设计的第一步是读取dat格式文件的数据,并将心电数据进行分离处理。在使用Python设计程序中,需引用Python的os库对dat格式进行读取。使用os.open对dat文件以二进制的打开方式进行读取。见图2。

Address	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
00000000	c0	00	00	ff	44	60	fd	94	1c	27	00	a4	ff	2f	00	c0
00000010	00	00	ff	a0	20	fd	4c	f6	27	00	a4	ff	2f	00	c0	00
00000020	00	ff	c9	94	fd	15	eb	27	00	a4	ff	2f	00	c0	00	00
00000030	ff	de	f5	fc	e9	7c	27	00	a4	ff	2f	00	c0	00	00	ff
00000040	eb	4b	fc	c2	29	27	00	a4	ff	2f	00	c0	00	00	ff	f2
00000050	f8	fc	9f	72	27	00	a4	ff	2f	00	c0	00	00	ff	f7	b7
00000060	fc	80	98	27	00	a4	ff	2f	00	c0	00	00	ff	fa	c6	fc
00000070	64	61	27	00	a4	ff	2f	00	c0	00	00	ff	fc	c1	fc	49
00000080	c0	27	00	a4	ff	2f	00	c0	00	00	ff	fd	fb	fc	30	68
00000090	27	00	a4	ff	2f	00	c0	00	00	ff	fe	ca	fc	18	73	27
000000a0	00	a4	ff	2f	00	c0	00	00	ff	ff	64	fc	02	25	27	00
000000b0	a4	ff	2f	00	c0	00	00	ff	ff	a6	fb	ed	1d	27	00	a4
000000c0	ff	2f	00	c0	00	00	ff	ff	d4	fb	d9	0b	27	00	a4	ff
000000d0	2f	00	c0	00	00	ff	ff	f5	fb	c6	34	27	00	a4	ff	2f
000000e0	00	c0	00	00	ff	ff	de	fb	b4	09	27	00	a4	ff	2f	00
000000f0	c0	00	00	ff	ff	fb	fb	a2	af	27	00	a4	ff	2f	00	c0
00000100	00	00	00	00	4e	fb	91	fd	27	00	a4	ff	2f	00	c0	00
00000110	00	00	00	2f	fb	82	a4	27	00	a4	ff	2f	00	c0	00	00
00000120	ff	ff	e1	fb	74	3a	27	00	a4	ff	2f	00	c0	00	00	ff
00000130	ff	d5	fb	65	bb	27	00	a4	ff	2f	00	c0	00	00	ff	ff

图2 dat数据文件16位十六进制数据

Figure 2 Hexadecimal data of dat data file

从原始数据分析,电极脱落标志位为第0至第2位,心电数据位于第6至第8位。设计是从完整数据列中按15位数据循环读取电极脱落标志位和心电数值。并将读取的数值赋值至对应的电极脱落标志序列和心电数据序列中。在电极脱落标志存在的情况下,心电信号并不作为有效的心电数据。在电极脱落时,电极脱落标志赋值为1。在电极脱落序列标志为0时,采集的

心电信号才符合抑郁评估使用标准。

2.2 心电滤波

通过心电数据的读取,获得心电数列。利用心电数据绘图得出如图3所示原始的心电数据波形图和电极脱落数据图。对图3进行分析,据此得出的原始心电数据比较杂乱,存在较大的上下幅度波动。在使用采集的心电数据时,需对心电数据进行基线滤波。

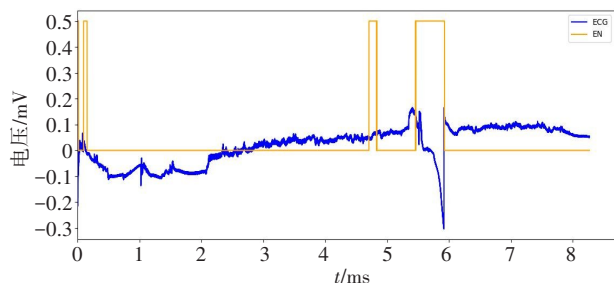


图3 心电读取数据(蓝色)和电极脱落数据(黄色)绘图
Figure 3 ECG reading data (blue) and electrode shedding data (yellow)

对原始心电数据进行基线滤波,程序设计使用低通滤波器。原始数据进行滤波得出原始数据波形的包络线,然后将原始数据减去该包络线,进而将该波形拉成水平。原始心电数据与包络线的数列相减后即可初步得到水平的心电数据,初步处理后的心电数据减去平均值可得出基线滤波后的数据,见图4。

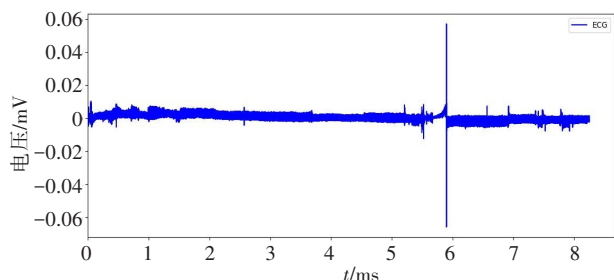


图4 心电信号基线滤波和高频滤波后图像
Figure 4 ECG baseline filtered and high-frequency filtered data

基线滤波后的心电数据还存在高频杂波,在处理基线滤波后的数据进行再设置低通滤波将高频的杂波去除,得到能初步使用的心电信号序列。核心代码如下:

```
#配置滤波器8表示滤波器的阶数
b, a = signal.butter(8, Hz, 'lowpass')
#ECG为要过滤的信号
ecg_outline = signal.filtfilt(b, a, ECG)
ecg_norm = ECG - ecg_outline
#求ECG数列的平均值avg
avg = sum(ECG)/len(ECG)
#ECG数列减去平均值avg得到水平基线的ECG
#数列
for i in range(len(ECG)):
    #配置滤波器8表示滤波器的阶数对ECG数据进行高频滤波
    ECG[i] = ECG[i] - avg
```

'lowpass')

#ECGfinal为滤波后的心电数据

ECGfinal = signal.filtfilt(b, a, ECG)

2.3 数列输出

经过处理的心电信号在进行寻找R波前进行了一次转换存储,将心电信号存储为csv文件。在心电处理的过程中,使用csv文件输出的数列通过使用heartpy.datautils库进行心电信号分析,同时使用csv文件存储使用普通的办公软件进行读取,增加心电数据的可读性,为寻找R波波峰提供读取数据的基础。

2.4 寻找R波

心电数据经数列输出后转存为csv文件。寻找R波需先设置读取心电csv文件,得到心电数列dataset。项目设计使用滑窗的形式寻找数列的峰值。程序设计使用pandas库的rolling_mean计算移动窗口的平均值。并将窗口的平均值储存在新增列hart_rollingmean中。

顺序读取心电数据dataset中hart数列的每一个数据,比较窗口平均值与hart数列的整体平均值大小。将二者中较大值重新赋值与rollingmean。然后比较rollingmean与心电数列的datapoint的大小。当datapoint高于rollingmean时,将datapoint在dataset中的位置记录,并将所标识的位置序号添加至peaklist中。标记的R波见图5,核心代码如下:

```
#计算滑动平均移动窗口的平均值
mov_avg = dataset['hart'].rolling(int(hrw*fs)).mean()
#表格中增加一列'hart_rollingmean'存储滑动平均值
dataset['hart_rollingmean'] = mov_avg
window = []
peaklist = []
#使用listpos变量在不同的列移动
listpos = 0
for datapoint in dataset.hart:
    #得到局部均值
    rollingmean = dataset.hart_rollingmean[listpos]
    if rollingmean > avg:
        rollingmean = rollingmean
    else:
        rollingmean = avg
    #检测不到R波不作记录
    if (datapoint < rollingmean) and (len(window) < 1):
        listpos += 1
```

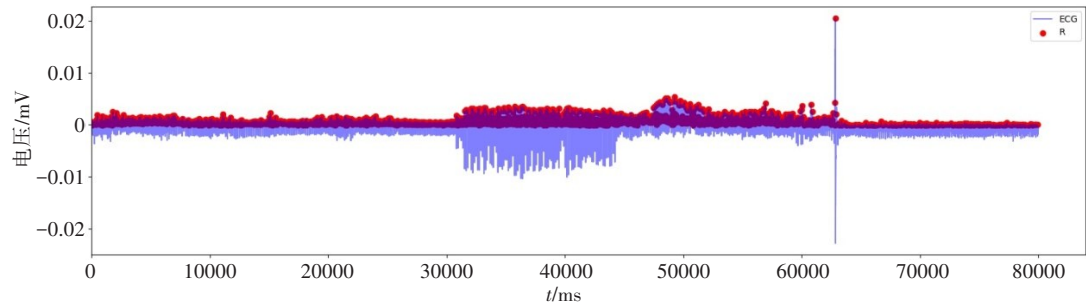



图5 心电数列R波标记图
Figure 5 R wave marked graph of ECG series

```
#信号高于平均值,进行标注
elif (datapoint > rollingmean):
    window.append(datapoint)
    listpos += 1
#信号低于局部均值确定最高点
else:
    maximum = max(window)
    #确定R波的位置
    beatposition = listpos - len(window) + (window.
index(max(window)))
    #存储R波位置
    peaklist.append(beatposition)
    #清空标注
    window = []
    listpos += 1
```

2.5 计算RR间期序列

在获取了R波位置数列 peaklist 后,需统计RR间期。并将RR间期的数值赋值于RR_list数组中。程序设计通过读取peaklist数列,将peaklist的后一个数值减去前一个数值得出相邻的两个R波的间隔。相邻的两个RR_i间隔值赋值于数列xp,将RR_{i+1}的数值赋值于数列yp。

2.6 绘制心电RR间期图

数图转换的最后一个环节是将心电RR间期数列转换为图片,并将图片保存。Python的数据绘图使用Matplotlib库,通过利用Matplotlib库统一绘制尺寸和标尺相同的散点图,见图6。其中x轴为心电RR_i间隔值,y轴为心电RR_{i+1}间隔值,图片绘制后进行保存。

3 图像模型训练

项目在进行数图转换后,需进行图像的模型训练。项目使用TesensoFlow1.15.0进行设计。
项目共使用32份抑郁症患者的睡眠心电数据和22份正常心电数据。通过使用32份心电数据经过数

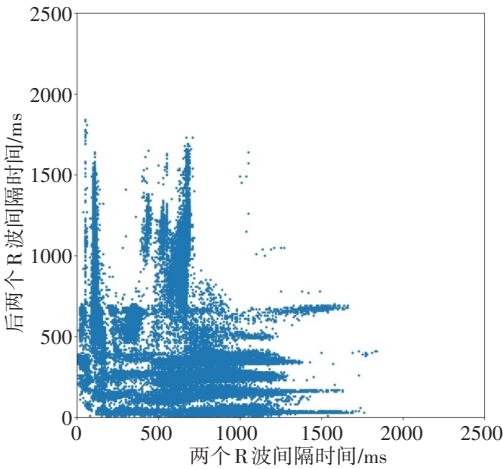


图6 心电RR间期图
Figure 6 ECG RR interval graph

图转换后的心电RR间期图和22份正常心电RR间期图进行分类训练。系统设计使用二分类的方式,通过AI学习抑郁症患者和正常人心电RR间期图的区别作出二分类的识别。

3.1 图片集预处理

进行分类后的图片需进行分类处理。其中项目设置为二分类,即设计为抑郁症患者的心电RR间期图和正常人心电RR间期图两个类别。在两个类别中按照8:2的比例设计训练集和验证集的图片数量。

训练文件夹图像通过使用图像生成器,设置一次训练所选取的样本数 batch_size,读取训练的文件夹路径 directory,标签数组模式 class_mode 设置为 "binary"返回1D的二值标。训练集图像设定 shuffle 为打乱数据模式,并对训练图像进行增强。验证图片集设置与训练图片集的差异在于一次训练所选取的样本数 batch_size 较训练图片集数量少,且验证图片集并不需要对图片顺序进行打乱设定。核心代码如下:

```
#训练集路径
train_dir = image_path + "/train"
#验证集路径
```

```

validation_dir = image_path + "/validation"
#定义训练集图像生成器,并进行图像增强
train_image_generator=ImageDataGenerator(rescale
=1./255,
rotation_range=40,
width_shift_range=0.2,
height_shift_range=0.2,
shear_range=0.2,
zoom_range=0.2,
horizontal_flip=True,
fill_mode='nearest')
#使用图像生成器从文件夹 train_dir 中读取样本
train_data_gen=train_image_generator.
flow_from_directory(directory=train_dir,
batch_size=batch_size,
shuffle=True, target_size=(im_height, im_width),
class_mode='binary')
#训练集样本数
total_train = train_data_gen.n
#定义验证集图像生成器,并对图像进行预处理
validation_image_generator =
ImageDataGenerator(rescale=1./255)
#使用图像生成器从验证集 validation_dir 中读取
样本
val_data_gen=validation_image_generator.
flow_from_directory(directory=validation_dir,
batch_size=5,
shuffle=False,
target_size=(im_height, im_width),
class_mode='binary')
#验证集样本数
total_val = val_data_gen.n

```

3.2 构建图片集模型

Keras库是TensorFlow中进行神经网络模型的训练。项目设计使用Sequential神经网络框架,在Sequential使用中封装了神经网络结构。使用一组输入和一组输出,各层间按照顺序进行堆叠,通过多层堆叠,构建神经网络。

其中第一层使用keras.layers.Conv2D()进行卷积,定义卷积核filters,卷积核尺寸kernel_size,激活函数activation和输入形状input_shape。第二层池化层keras.layers.MaxPool2D设定池窗口的大小pool_size和步幅strides。使用keras.layers.Flatten降低输出维度。全连接层设置keras.layers.Dense函数的计划函数activation。设计代码如下:

```
#定义LeNet
```

```

model=tf.keras.Sequential([tf.keras.layers.
Conv2D(filters=6, kernel_size=(5, 5), activation='relu',
input_shape=(200, 200, 1)),
tf.keras.layers.MaxPool2D(pool_size=(2, 2),
strides=2),
tf.keras.layers.Conv2D(filters=16, kernel_size=(5,
5), activation='relu'),
tf.keras.layers.MaxPool2D(pool_size=(2, 2),
strides=2),
tf.keras.layers.Flatten(),
tf.keras.layers.Dense(120, activation='relu'),
tf.keras.layers.Dense(84, activation='relu'),
tf.keras.layers.Dense(10, activation='softmax')
])
#模型打印
model.summary()

```

3.3 模型的编译和保存

在配置训练方法时,需告知训练时使用的优化器optimizer、损失函数loss和准确率的评测标准metrics。此外还需启用callbacks.ModelCheckpoint,训练好深度学习模型,callbacks.ModelCheckpoint是模型的权重,作为持续训练的基础。设定filepath,保存模型的路径。核心代码如下:

```

#模型编译
model.compile(optimizer=tf.keras.optimizers.
Adam(learning_rate=0.0001),
loss='binary_crossentropy',
metrics=['accuracy'])
#模型保存
cp_callback = tf.keras.callbacks.ModelCheckpoint
(filepath=checkpoint_save_path,
save_weights_only=True,
save_best_only=True)

```

3.4 模型的训练

模型的训练使用model.fit,函数model.fit返回一个History对象,记录loss和其他指标的数值随epoch的变化。设定输入数据x,梯度下降的样本数。训练时一个batch的样本会被计算一次梯度下降,使目标函数优化一步。训练终止时的epoch值,训练将在达到该epoch值时停止。训练集和验证集的损失和准确度图见图7。验证集validation_data和callbacks调用回调函数cp_callback。核心代码如下:

```

#模型训练
history = model.fit(x=train_data_gen,
steps_per_epoch=total_train // batch_size, epochs=
epochs,

```

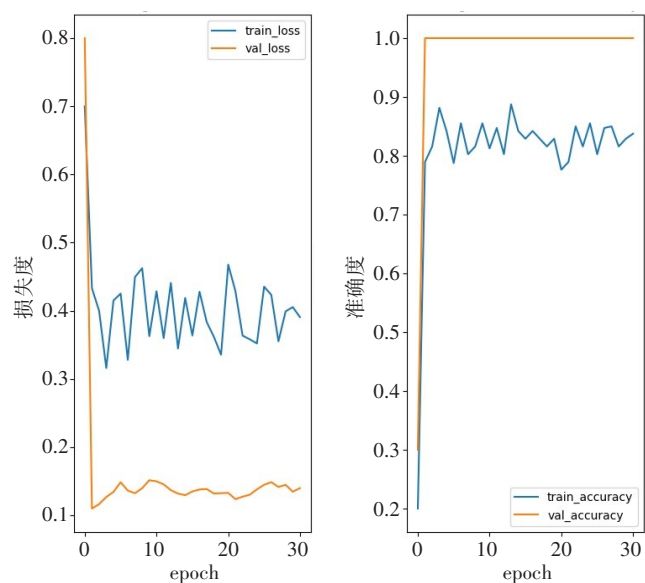


图7 训练集和验证集的损失和准确度图

Figure 7 Loss and accuracy on training set and validation set

```
validation_data=val_data_gen,
validation_steps=total_val // batch_size,
callbacks=[cp_callback])
#记录训练集和验证集的准确率和损失值
history_dict = history.history
train_loss = history_dict["loss"]
train_accuracy = history_dict["acc"]
val_loss = history_dict["val_loss"]
val_accuracy = history_dict["val_acc"]
```

3.5 模型的调用

系统利用输入需要预测的患者心电数据,滤波处理后生成心电RR间期图,加载完成训练后的模型。使用model.predict输入需要进行预测分类的心电RR间期图,并输出预测结果。系统测试输入已诊断为抑郁症患者的睡眠心电数据,并完成该患者的诊断预测分类,见图8。

```
Loading Patient ECG Dat...
ECG Filter...
ECG Data Csv Saving ...
RR Image Saving ...
Loading Image ...
Loading Model ...
This Patient most likely belongs to Depression (87.5%)
>>>
```

图8 AI系统分类预测图

Figure 8 Prediction by AI system

4 结 语

本项目设计提取患者的心电数据,进行抑郁症患者与正常人的区分。将原始数据视像化,并使用

TensorFlow 进行图像的学习分类,简化了整个AI系统的学习结构流程。系统实现了使抽象的数据转变为实例图像,更形象具体地体现了抑郁症患者与正常人睡眠心电的区别,实现了基于变值测量模型的心电信号可视化展示和研究^[13]。

【参考文献】

[1] McCall WV, Blocker JN, D'Agostino R, et al. Treatment of insomnia in depressed insomniacs: effects on health-related quality of life, objective and self-reported sleep, and depression[J]. J Clin Sleep Med, 2010, 6(4): 322-329.

[2] Müller MJ, Olschinski C, Kundermann B, et al. Subjective sleep quality and sleep duration of patients in a psychiatric hospital[J]. Sleep Sci, 2016, 9(3): 202-206.

[3] 张关娅, 王平, 丁莉, 等. 睡眠及睡眠剥夺与抑郁症的相关性探讨[J]. 中华中医药杂志, 2018, 33(8): 3529-3531.

Zhang GY, Wang P, Ding L, et al. Study on the correlation between sleep, sleep deprivation and depression [J]. China Journal of Traditional Chinese Medicine and Pharmacy, 2018, 33(8): 3529-3531.

[4] Murphy MJ, Peterson MJ. Sleep disturbances in depression[J]. Sleep Med Clin, 2015, 10(1): 17-23.

[5] 胡静, 滕晶. 滕晶教授基于中医五神辨治抑郁症经验[J]. 中医药导报, 2021, 27(11): 175-178.

Hu J, Teng J. Professor Teng Jing's experience in treating depression based on the five spirits of traditional Chinese medicine[J]. Guiding Journal of Traditional Chinese Medicine and Pharmacology, 2021, 27(11): 175-178.

[6] 戚玲瑜, 谢春明. 抑郁症自杀相关的神经影像学研究[J]. 中华行为医学与脑科学杂志, 2021, 30(4): 373-378.

Qi LY, Xie CM. Neuroimaging study on suicide in depression[J]. Chinese Journal of Behavioral Medicine and Brain Science, 2021, 30(4): 373-378.

[7] 王晨, 许冬梅, 邵静, 等. 精神科住院抑郁患者自杀预防及护理干预措施专家共识[J]. 中华护理杂志, 2022, 57(18): 2181-2185.

Wang C, Xu DM, Shao J, et al. Expert consensus on suicide prevention and nursing intervention measures for inpatients with depression in psychiatric department[J]. Chinese Journal of Nursing, 2022, 57(18): 2181-2185.

[8] 李茜瑶, 周莹, 黄辉, 等. 疾病负担研究进展[J]. 中国公共卫生, 2018, 34(5): 777-780.

Li QY, Zhou Y, Huang H, et al. Research progress of disease burden [J]. Chinese Journal of Public Health, 2018, 34(5): 777-780.

[9] 孟繁强, 杨丽娟, 张居易, 等. 伴失眠的抑郁症患者睡眠认知特征及其对睡眠质量的影响[J]. 四川精神卫生, 2022, 35(4): 319-324.

Meng FQ, Yang LJ, Zhang JY, et al. Cognitive characteristics of sleep in depressive patients with insomnia and its influence on sleep quality [J]. Sichuan Mental Health, 2022, 35(4): 319-324.

[10] 马欣荣, 宋广荣, 马陈雯, 等. 睡眠质量对中老年抑郁患者认知功能的损害及影响因素研究[J]. 宁夏医学杂志, 2019, 41(10): 906-908.

Ma XR, Song GR, Ma CW, et al. Study on the cognitive impairment and influencing factors of sleep quality in middle-aged and elderly patients with depression [J]. Ningxia Medical Journal, 2019, 41(10): 906-908.

[11] 王肖文, 肖碧波, 陈旭创. 基于便携式睡眠心电采集及深度学习的抑郁症患者睡眠质量评估系统[J]. 医疗装备, 2022, 35(1): 28-29.

Wang XW, Xiao BB, Chen XC. Sleep quality assessment system for depressive patients based on portable sleep eeg acquisition and deep learning[J]. Medical Equipment, 2022, 35(1): 28-29.

[12] 杨炳新, 郭艳蓉, 郝世杰, 等. 基于数据增广和模型集成策略的图神经网络在抑郁症识别上的应用[J]. 计算机科学, 2022, 49(7): 57-63.

Yang BX, Guo YR, Hao SJ, et al. Application of graph neural network based on data augmentation and model integration strategy in depression recognition[J]. Computer Science, 2022, 49(7): 57-63.

[13] 吉艳, 郑智捷, 山路. 基于变值测量模型的心电信号可视化研究[J]. 图像与信号处理, 2016, 5(1): 25-32.

Ji Y, Zheng ZJ, Shan L. Research on ECG visualization based on variable measurement model [J]. Journal of Image and Signal Processing, 2016, 5(1): 25-32.

(编辑:黄开颜)