

doi: 10.3969/j.issn.1674-1242.2024.02.002

基于 ERP 信号的目标检测研究综述

杨倩倩¹, 王伟²

(1. 上海理工大学, 上海 200093;

2. 中国人民解放军海军特色医学中心, 上海 200093)

【摘要】 目标检测是一个广泛应用于多个领域的研究课题。传统的计算机视觉方法存在复杂度高、泛化能力差、模型对象可迁移性差等缺点, 而人脑在目标检测方面具有较强的认知能力和感知能力, 仅需要很少的训练样本就可以达到较好的检测效果。基于事件相关电位 (Event-Related Potentials, ERP) 的脑机接口 (Brain-Computer Interface, BCI) 是一种全新的提升目标检测性能的方法, 它能够将脑内信息与外部环境进行交互。在此基础上, Gerson 提出了一种以脑电信号为基础的快速串行视觉呈现 (Rapid Serial Visual Presentation, RSVP) 范式, 包括显示方式、有效视觉范围、目标复杂度、刺激频率和持续时长等。该文还对 ERP 信号的预处理、特征提取及分类等方法进行了综述, 并对该领域的局限性和发展趋势进行了探讨。

【关键词】 事件相关电位; 脑机接口; 快速串行视觉呈现; 目标检测; 特征提取

【中图分类号】 TP391

【文献标志码】 A

文章编号: 1674-1242 (2024) 02-0111-12

A Review of Target Detection Based on ERP Signal

YANG Qianqian¹, WANG Wei²

(1. University of Shanghai for Science and Technology, Shanghai 200093, China;

2. PLA Naval Medical Center, Shanghai 200093, China)

【Abstract】 Target detection is a research topic which is widely used in many fields. The traditional methods of computer vision have some problems, such as high complexity, poor generalization ability and weak mobility of model objects. However, human brain has higher cognitive ability and perception ability in object detection, only a few training samples are needed to achieve better detection results. The Brain-Computer Interface (BCI) based on Event-Related Potentials (ERP) is a new method to improve the performance of target detection, which can interact with information in the brain and the external environment. On this basis, a new pattern of Rapid Serial Visual Presentation (RSVP) based on EEG signals is proposed by Gerson, it includes display mode, effective visual range, target complexity, stimulus frequency and duration. In addition, the methods of ERP preprocessing, feature extraction and classification are reviewed. Finally, the limitation and development trend of this field are discussed.

【Key words】 Event-Related Potentials(ERP); Brain-Computer Interface(BCI); Rapid Serial Visual Presentation(RSVP); Target Detection; Feature Extraction

收稿日期: 2023-08-27。

基金项目: 后勤科研重点项目 (BHJ22J024, BHJ14C017), 国防科技创新特区项目 (19-163-12-ZT-002-002-01, 21-163-12-ZT-006-002-14)。

作者简介: 杨倩倩 (2000—), 女, 江苏省淮安市人, 硕士研究生, 从事脑机接口研究。

通信作者: 王伟, 男, 研究员, 硕士生导师, 电话 (Tel.): 19945657097, 邮箱 (E-mail): wwang_fd@fudan.edu.cn。

0 引言

目标检测是一项在各领域都具有深入研究与广泛应用的技术,其任务是在大幅面复杂场景图像中,对微小实例物体进行预定义,旨在准确地识别和定位感兴趣的目标物体。随着人工智能的迅速发展,计算机视觉自提出后逐渐进入各行业领域,并在很多行业领域取得了突破性进展,如图像识别与分类^[1]、人体动作捕捉^[2]、医学影像检测^[3]等。计算机视觉领域的目标检测是指利用计算机视觉技术和算法,通过对图像或视频进行分析和处理,识别并定位图像中的目标物体。虽然基于计算机视觉的目标检测已被广泛研究,但是目前依然存在一些困难及检测障碍。其局限性主要有目标遮挡、样本数量少及无法满足在线需求等。

因此,完全依赖计算机视觉难以适应复杂场景的需求。基于快速串行视觉呈现(Rapid Serial Visual Presentation, RSVP)的目标检测是指当受试者看到目标图像时,就会被诱发出相应的特异性脑电图(Electroencephalogram, EEG)波形,通过检测事件相关电位(Event-Related Potential, ERP)信号中受刺激影响的P300等电位,即可实现对目标图像的检测与识别。由于人脑具有记忆力、注意力、概念抽象等能力,对目标检测任务有高敏感性,能够在几百毫秒内对一幅图片或一个场景中的重点敏感信息产生神经响应,可以克服计算机目标视觉检测所需时间长、无法满足在线需求的问题。人脑只需要进行短期少量样本的学习,即可进行目标检测,可以克服计算机视觉检测样本数量少的问题。同时,人脑具有知识迁移能力,能迅速适应环境,

可以克服计算机视觉检测领域易受目标遮挡、光线强弱等问题。在传统方式中,人需要手动直接对目标进行标记,该方式精准性高,速度较快,在小量任务中效果较好。但随着待检测样本数量的增多,人会出现视觉、精神、肢体等一系列生理疲劳现象,检测性能下降^[4]。因此,如何将计算机视觉强大的信息处理能力和人脑应对复杂场景的高级认知与快速感知能力相结合,逐渐成为当前目标检测领域的研究热点^[5]。

随着近年来脑机接口(Brain-Computer Interface, BCI)技术的发展,结合BCI系统的方式逐渐进入目标检测领域。BCI系统是一种通过解译人类大脑信息完成任务的人机交互系统^[6]。脑电图作为一种非侵入性、高时间分辨率、成本较低且具有多通道记录能力的脑电信号测量技术,是研究BCI最常见的方式。ERP是一种特殊的脑诱发电位,是外加某种特定刺激(如听觉、视觉或触觉)时所获得的脑电波。它作为BCI系统的重要分支之一,是反映了大脑高级思维活动的一种客观方法。

ERP有两个重要特性,即潜伏期恒定和波形恒定。由于ERP信号具有特定的波形和电位分布,且潜伏期与刺激之间存在严格的锁时关系,因此可以通过以RSVP为主的实验范式高效率地提取大脑在观看目标时产生的ERP信号,在计算机处理后即可完成目标检测。在ERP信号中,存在许多ERP成分与目标检测相关,最常使用的ERP成分是P300。P300出现在目标刺激后250~750ms^[7],典型的P300ERP波形如图1所示。

近年来,随着算法性能的提升、BCI硬件设施

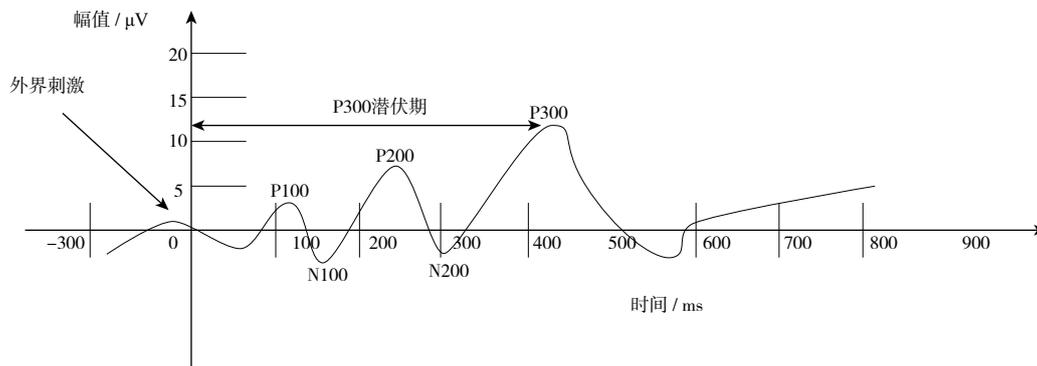


图1 典型的P300ERP波形

Fig.1 Typical P300 ERP waveform diagram

的升级，基于 ERP 的目标检测技术在具体场景中的应用越来越多。相比传统计算机视觉的目标检测，它更能适应复杂场景的需求，集中在人脸身份检测、军事作战及医疗辅助等方面，这些场景需要准确快速地识别当前接收到的图像，且运用了计算机技术强大的信息处理能力，能够及时抵抗复杂环境的影响。军事侦察是目前基于 ERP 的目标检测应用和研究最多的一个领域。其中最具代表性的研究是 2012 年美国加利福尼亚州圣地亚哥大学开发的单兵军事威胁探测装置——认知技术威胁警报系统。另外，在进行大幅面遥感图像中目标图像的识别时，BCI

系统可以凭借人脑的高级感知能力进行快速识别并及时发出警报。而计算机视觉样本训练有限且不能满足在线需求，完全靠计算机技术不能应付此类复杂场景的需求。基于 ERP 信号的目标检测的应用如图 2 所示。

基于 ERP 信号的目标检测的基本流程如图 3 所示，主要包括图片刺激、人脑、脑电信号数据采集、脑电信号分析、图像目标检测等步骤。

本综述由以下几部分组成：①梳理了以 RSVP 为主的实验范式，包括其设计因素与研究局限等，旨在为 RSVP 范式的发展现状和未来前景提供有价值

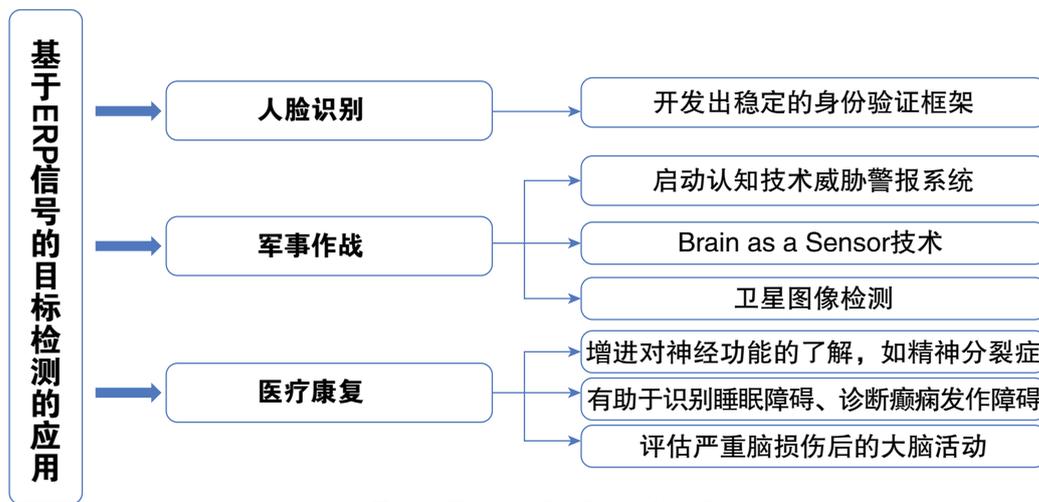


图 2 基于 ERP 信号的目标检测的应用
Fig.2 Applications of target detection based on ERP signal



图 3 基于 ERP 信号的目标检测的基本流程
Fig.3 Basic process of target detection based on ERP signal

的参考。②总结了目前脑电信号分析中的预处理方法、特征提取与特征分类算法。③指出了目前基于 ERP 信号的目标检测存在的问题和挑战并展望了未来。

1 实验范式

Polich 和 Donchin^[8] 于 1988 提出的 Oddball 范式可以诱发 P300 等 ERP 成分。该范式规定受试者必须将一系列刺激分为两类：目标和非目标。目标比非目标出现的频率更低（通常占总刺激的 5%~10%）。Gerson 等^[7] 基于 Oddball 范式于 2005 年提出了一种用于目标检测的范式——RSVP 范式。RSVP 范式是指将图像按序列呈现给受试者，每张

图像呈现给受试者的时间为 50~500ms，图像序列中通常包含目标图像与非目标图像。当受试者看到目标图像时，就会被诱发出相应的特异性 EEG 波形，通过检测 ERP 信号中受刺激影响的 P300 等电位，即可实现对目标图像的检测与识别。

1.1 RSVP 设计因素

如何设计实验范式以诱发更明显的 ERP 信号并提高目标检测效率，一直是相关领域的研究热点。RSVP 范式是由多因素设计的，每个因素都可能造成分类结果的偏差。RSVP 范式的设计因素如图 4 所示。下面主要介绍显示模式、有效视觉范围、刺

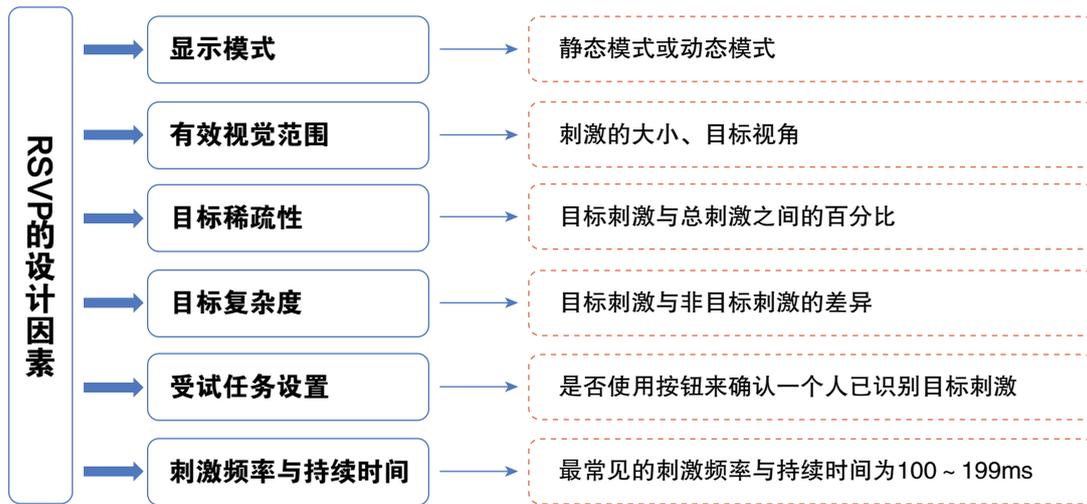


图 4 RSVP 范式的设计因素
Fig.4 Design factors of RSVP paradigm

激频率与持续时间这 3 个设计因素。

1.1.1 显示模式

RSVP 范式的呈现模式可分为静态模式和动态模式^[9]，分别指根据特定频率显示的静态目标图像和运动目标动图或视频。动态模式比静态模式更复杂，因为受试者必须观看整个场景，而不是特定的图像。

1. 静态模式

在静态模式下，显示的图像具有相同的入口和出口——图像在屏幕上被短暂地显示（通常为 50~500ms）。静态模式的图像占据了显示屏的大部分，因此，即使它们只是被短暂地呈现，受试者也有可能识别目标。为达到实验效果，研究人员会向

受试者提出各种实验要求，如下所述。

(1) 在演示之前，向受试者展示目标图像，并要求受试者在一系列图像中识别该图像。在该过程中，受试者正确识别目标图像和非目标图像的效率最高可以达到 10 张 /s^[10]。

(2) 受试者被要求确定一种目标类型，如图像集合中的动物、飞机、书籍等。在这种模式下，图像呈现的速度应该减慢（4 张 /s）。

(3) 在展示图像序列后，受试者可能会立即看到一张图像，并被问道：“这个图像是否出现在你刚刚看到的序列中？”

2. 动态模式

在动态模式下，受试者需要观看一系列较短的

视频片段或动图，并在该过程中识别一个或多个目标。这些目标需要分散以避免 P300 的抑制。动态模式常用于视频监控领域。在动态模式下，研究人员会向受试者提出各种实验要求，如下所述。

(1) 在演示之前，会给受试者一个目标，如要求他们识别一个移动场景中的“人”或“车辆”^[11]。

(2) 受试者可以被要求确定一个目标事件。在这种情况下，目标是跨越空间和时间的。受试者需要综合运动特征来决定该事件是否为目标事件，如将目标定义为“在火车站留下可疑包裹的人”^[12]。

1.1.2 有效视觉范围

在 RSVP 范式中，较小的目标和处于图像边沿的目标比较难以被识别。为了充分利用受试者的有效视觉范围，需要设置一个合理的图像尺寸、视距（受试者与显示屏之间的距离）和视角（受试者与显示屏的相对位置）。

视频监控领域最常用的图像尺寸是 500 × 500 像素。所有使用这种图像尺寸的监控研究都取得了很高的准确性，高达 80%。其他领域一般使用较小的图像尺寸，如 360 × 360 像素，也可以实现较高的准确性。因此，可以得出结论：并不存在一种通用的适用于所有场景的最佳图像尺寸。不同的应用场景和任务需要有针对性地选择合适的图像尺寸来实现最佳效果。

视角是图像在眼睛中呈现的角度，人眼最敏锐的视力范围为 1°~15°^[13]，当目标视角在 8° 以内时，人眼的聚焦效果最佳。Dias 和 Parra^[14] 的一项研究表明，当目标刺激位于受试者的视场中心时，受试者的目标检测率最佳；当目标刺激偏离视场中心 4° 时，所有受试者的检测准确率都持续下降；

当目标刺激被放置在偏离视场中心 8° 的视角时，这种下降进一步加剧。具体实验结果如表 1 所示。

褚凯轩等^[15]进一步分析了目标视角对 RSVP 目标检测系统的影响。研究结果表明，当目标视角在 8° 左右时，受试者的平均目标检测率最高，达到 82% 以上；当目标视角小于 0.5° 或大于 15° 时，所有受试者的平均目标检测率均出现了一定程度的衰减。

1.1.3 刺激频率与持续时间

RSVP 范式的另一个设计因素是刺激频率与持续时间，因为该范式的重点是快速呈现数据，以便在短时间内分析大型数据集。刺激频率与持续时间通常为 50~500ms^[16]。为了防止出现扫视盲点^[17]，即一个人在两个位置之间扫视前后 50ms 内发生目标丢失，目标刺激频率与持续时间需要大于 50ms。在 RSVP 范式中，刺激频率与持续时间的上限在文献中定义不清。然而，我们发现每张图像 500ms 是所有 RSVP 研究中使用的最长刺激频率与持续时间。

刺激频率与持续时间在不同的应用中通常是不同的，Sajda 等^[16] 在一项研究中要求两名受试者识别有人物的场景：在每场实验中，刺激频率与持续时间从 200ms 减少到 100ms，再减少到 50ms。研究结果表明，当刺激频率与持续时间较短(50ms)时，两名受试者的表现都有所下降。这表明 RSVP 范式中合适的刺激频率与持续时间是 100~199ms，这个时间段可以在目标检测的准确性和速度之间取得平衡。对相关文献的调研可证明以上观点，如表 2 所示。

目标识别难度和刺激频率与持续时间之间可能存在显著的相关性。例如，给定刺激集的最佳刺激

表 1 不同目标视角下的目标检测准确率
Tab.1 Target detection accuracy under different view angles

视角偏离视场中心的角度		<0.5°	4°	8°
目标检测准确率 / %	受试者 1	81	75	66
	受试者 2	86	78	71
	受试者 3	88	80	70
	受试者 4	79	69	64
	受试者 5	77	65	60
	平均	82.2	73.4	66.2

表 2 RSVP 范式相关文献调研
Tab.2 RSVP literature survey

刺激持续时间 /ms	文献数量 / 篇	正确率范围 / %
<100	13	66~93
100~199	29	70~94
200~299	26	67~98
300~399	10	77~91
400~499	8	81~94
>500	17	72~90

频率与持续时间高度依赖识别该集合中目标的难度^[18]。低杂波、高对比度、无遮挡和大目标尺寸的图像集可能适合更短的刺激频率与持续时间，而高杂波、低对比度、高遮挡、小目标尺寸的图像集则适合较长的刺激频率与持续时间^[20]。

1.2 RSVP 的研究局限和发展前景

近年来，越来越多的人投入到 RSVP 的研究中，RSVP 成为研究 BCI 的主要范式之一，推进了 RSVP 实际应用的进展。但在研究过程中仍然存在一些难点，有望从认知能力融合、数据集公开、受试者群体多样化、刺激干扰等方面继续拓展 RSVP 的研究深度和广度，为 RSVP 在科学和实际研究中提供更广阔的应用前景。

1.2.1 认知能力融合

RSVP 目前还没有将较大的数据量、多天多测、跨受试者验证等问题融入一个全面的研究中，导致目前的实验结果难以得到真正意义上的泛化性和鲁棒性。

1.2.2 数据集公开

与其他神经认知研究不同，目前 RSVP 没有一个公开的大型数据库能为跨受试者、跨天等问题的研究提供统一的数据基础。因此，未来还需要在数据公开、采集合理、受试者数量等方面发展一个可

供更多研究者使用的数据集。

1.2.3 受试者群体多样化

目前绝大多数研究的受试者都是正常人，即认知能力没有明显障碍或未患疾病的人。而 RSVP 最初旨在通过信号解码为眼球运动障碍患者提供帮助^[20]。虽然正常人的研究可用于军事警务、辅助医生诊断等，但是未来研究有必要针对那些患有记忆障碍或视觉障碍的人专门设计编码范式，为疾病患者提供帮助，使 RSVP 具有更广阔的应用场景。

1.2.4 刺激干扰

RSVP 脑机接口未来可能会在目标刺激中使用具有大量干扰物刺激的现实生活图像。这是一项困难的任务，但它将增强 RSVP 脑机接口与现实应用的相关性，将 RSVP 脑机接口与其他脑机接口范式相结合，因此开始受到越来越多的关注。

2 EEG 信号分析

2.1 预处理方法

由于设备或传输线产生的干扰，原始的 EEG 信号中含有许多影响分类的噪声。为了获得相对干净的 EEG 信号并提高分类性能，需要对原始信号进行合理的预处理，去除 EEG 信号中的伪迹^[21]。EEG 信号预处理流程如图 5 所示。

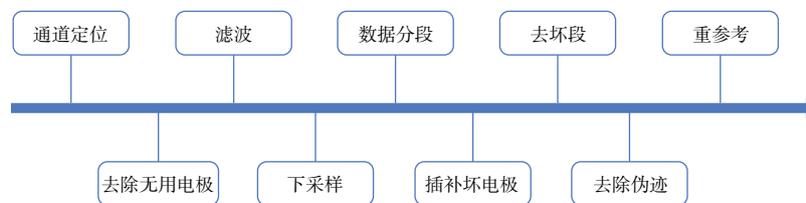


图 5 EEG 信号预处理流程
Fig.5 EEG signal preprocessing process

为了确定电极在头皮上的位置,以便后续的信号处理和分析,首先需要将采集到的单个或多个 EEG 通道数据进行电极定位,电极定位后可以将无用的电极信号从数据中剔除。为了消除特定频率范围内的干扰和噪声,可以利用滤波衰减过频明显不具备目标特征的频率信号。独立成分分析 (Independent Component Analysis, ICA) 和主成分分析 (Principal Component Analysis, PCA) 是目前常见的 EEG 信号伪迹处理方法。ICA 是一种高效的盲源信号分离方法,根据输入信号源的统计特性,将信号源有效地分离成独立的成分^[22]。PCA 是一种通用的降维工具,通过计算特征的协方差矩阵,将原始特征旋转至新的特征空间,形成所有时间点数据共同的表征信息^[23]。此外,还有一些去除伪迹的方法,如共同平均参考值 (Common Average Reference, CAR)、共同空间模式 (Common Spatial Pattern, CSP) 及盲源分离和小波变换相结合等。去除伪迹后需要对数据进行重参考,防止数据受到单个电极的较大影响,常见的重参考方法有平均参考、双侧乳突参考等。

2.2 特征提取

在通过 EEG 信号进行目标检测的过程中,特征提取是关键部分,特征提取的质量将直接影响分类的准确性。目标检测技术一般用来提取 ERP 时域特征。本节将介绍常见的 ERP 特征提取方法,重点从两个方面进行提取:有监督的空间滤波、无监督的空间滤波。

2.2.1 有监督的空间滤波

EEG 数据具有低信噪比和高空间维度的缺点。空间滤波是解决这些问题的一种有效方法。在脑机接口研究领域,CSP^[24]、xDAWN^[25]和 LDA 波束形成器^[26]被广泛用于生成有监督的空间滤波。对于 CSP,目标是最大化判别活动与共同活动之间的比值,从而使两种刺激 EEG 信号的判别方差最优;对于 xDAWN,目标是最大化信噪比;LDA 波束形成器被用于源信号重建,使信噪比最大化。

设 $\mathbf{X} \in \mathbf{R}^{C \times T}$ 为每个刺激对应的 EEG 时段,其中 C 为通道数, T 为时间长度。空间滤波是通过找到一组投影向量 $\mathbf{w} \in \mathbf{R}^{C \times n}$ (n 为分量的个数) 将 \mathbf{X}

投影到一个子空间。

$$\mathbf{X}_{\text{sub}} = \mathbf{w}'\mathbf{X} \quad (1)$$

1. CSP

CSP 生成通道权重集,可用于将多通道 EEG 数据投影到低维子空间中,该变换可以最大化两类信号矩阵的方差。设 $\mathbf{X}_+(i) \in \mathbf{R}^{C \times T}$ 、 $\mathbf{X}_-(i) \in \mathbf{R}^{C \times T}$ 为两种不同实验条件下的第 i 个事件对应的 EEG 时段 (C 为通道数, T 为时间长度),两种情况下的协方差矩阵可估计为

$$\sum_c = \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{X}_c(i)\mathbf{X}_c'(i)}{\text{trace}(\mathbf{X}_c(i)\mathbf{X}_c'(i))} \quad (2)$$

CSP 优化问题可表述为

$$\{\max, \min\}_{\mathbf{w} \in \mathbf{R}^C} \frac{\mathbf{w}'\sum_+ \mathbf{w}}{\mathbf{w}'\sum_- \mathbf{w}} \quad (3)$$

该优化方法是由两个协方差矩阵的同时数字化给出的,这可以通过求解广义特征值来实现。

$$\sum_+ \mathbf{w} = \lambda \sum_- \mathbf{w} \quad (4)$$

2. xDAWN

xDAWN 的基本目标是增强目标刺激对应的单信噪比。设记录的信号为 $\mathbf{X} \in \mathbf{R}^{N_t \times N_s}$, 其中 N_t 为记录的 EEG 信号的时间长度, N_s 为通道数。

它考虑了以下模型

$$\mathbf{X} = \mathbf{D}\mathbf{A} + \mathbf{H} \quad (5)$$

式中, $\mathbf{D} \in \mathbf{R}^{N_t \times N_t}$ (N_t 为目标刺激对应的 ERP 时间样本数) 为托普利茨矩阵; \mathbf{A} 为 ERP 对目标的响应, \mathbf{D} 的第一列元素除与目标刺激相对应的元素外均为零; \mathbf{H} 为正在进行的 EEG 活动。

xDAWN 的问题变成了如何生成空间滤波器,从而通过空间滤波增强同步响应的问题。

$$\mathbf{X}\mathbf{U} = \mathbf{D}\mathbf{A}\mathbf{U} + \mathbf{H}\mathbf{U} \quad (6)$$

式中, $\mathbf{U} \in \mathbf{R}^{N_t \times N_f}$ (N_f 为空间滤波器个数)。优化解如下。

$$\hat{\mathbf{U}} = \arg \max_{\mathbf{U}} \frac{\text{Trace}(\mathbf{U}'\hat{\mathbf{A}}'\mathbf{D}'\mathbf{D}\mathbf{A}\mathbf{U})}{\text{Trace}(\mathbf{U}'\mathbf{X}'\mathbf{X}\mathbf{U})} \quad (7)$$

式中, $\hat{\mathbf{A}}'$ 为响应的最小二乘估计,计算方法详见文献 [14]。

3. LDA 波束形成器

LDA 波束形成器提出了一种对相关源具有鲁棒性的方式，可以最大化 EEG 信号的信噪比。利用 LDA 波束形成器生成有监督的空间滤波分为 3 个步骤：①空间方向图估计；②协方差矩阵估计；③空间滤波优化。设列向量 $P_1 \in R^{C \times 1}$ 、 $P_2 \in R^{C \times 1}$ 为特定组件在两种不同实验条件下的空间格局，其中 C 为通道数。将差异模式记为 $P = P_1 - P_2$ ，协方差矩阵 $\sum \in R^{C \times C}$ 。LDA 波束形成器的优化问题可以表述为

$$\begin{cases} \text{minimize } w' \sum w \\ \text{s.t. } w' p = 1 \end{cases} \quad (8)$$

可确定最优解为

$$w = C^{-1} p (p' C^{-1} p)^{-1} \quad (9)$$

与前两种方法不同，LDA 波束形成器生成一个最优空间滤波器（只有一个空间滤波器），使信噪比最大。

到目前为止，我们已经介绍了 3 种有监督的空间滤波。采用空间滤波方法的总体效果是提高了信噪比，降低了计算成本，对许多存在高维特征向量问题的分类算法（特别是当可用的训练样例很少时）来说，可以取得相对较好的结果。

2.2.2 无监督的空间滤波

1. ICA

脑电图源活动是指脑电图源内产生的时变远场电位，其容积传导至头皮电极，它也可以被定义为盲源分离（Blind Source Separation, BSS）方法。如果存在 n 个线性混合物 x_1, x_2, \dots, x_n ，那么向量 x 为

$$x = AS \quad (10)$$

式中， A 表示大小为 $n \times n$ 的混合矩阵； S 表示独立分量。ICA 的目标是找到一个矩阵 W ，它将是矩阵 A 的逆，以逆混合效应。给出一个独立分量，如方程（11）所示。

$$y = Wx \quad (11)$$

ICA 依赖信号的线性，可以提高感兴趣信号的信噪比，已广泛应用于脑电信号领域的去噪和去伪^[27]。

2. PCA

PCA 是一种统计技术，它使用特征值分解将一组相关变量转换为一组线性不相关变量，其中每个结果变量称为主成分^[28]。对于多变量数据集，特别是高维空间中的数据，PCA 降维特别有效。对于定义的 $n \times n$ 维数据集，计算平均向量 i 和协方差矩阵（大小为 $z \times z$ ）求和。然后，找到特征值和特征向量，并朝着特征值递减的方向进行排序。在基于 RSVP 的脑机接口文献中，PCA 迄今仅用于特征降维^[29]。

2.3 特征分类

在完成 EEG 信号的特征提取后，需要选择合适的分类方法对 EEG 信号特征进行有效的识别和分类。基于 ERP 的目标检测分辨率低、模糊性强，且在提取特征过程中可用特征少^[30]，从而增大了模型的分类难度。目前的 EEG 信号特征分类方法主要有空间滤波、机器学习、深度学习 3 类，具体如图 6 所示。

2.3.1 空间滤波

基于空间滤波的分类方法将多通道信号组合成

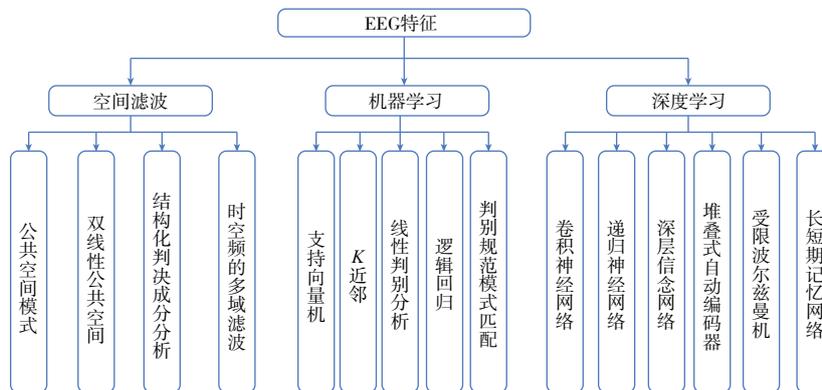


图 6 EEG 信号特征分类方法

Fig.6 EEG sigal feature classification method

一个新的时间序列，对其提取特征并选取后再进行分类。现有方法包括公共空间模式及其衍生的公共空间-频谱模式、公共稀疏空间-频谱模式、公共空间-时间模式^[31]、双线性公共空间^[32]等。

此外，Sajda等^[16]采用结构化判决成分分析（Hierarchical Discriminant Component Analysis, HDCA）方法，其核心思想是找到一组最佳加权系数，对RSVP实验中的脑电数据进行先空间、后时间的线性加权，得到对应图像的评分，使目标类图像的评分值尽量高于非目标类图像的评分值。如图7所示为HDCA方法的训练过程。Sajda等还开发了皮层耦合计算机视觉系统（C3Vision），该系统将人类视觉与计算机视觉进行了结合。张宁宁^[33]提出了一种时-空-频的多域滤波方法（MDF-HDCA），该方法结合不同角度的信号特点进行加权分析，从而提高了HDCA算法的性能。

2.3.2 机器学习

Mathan等^[34]采用支持向量机（Support Vector Machine, SVM）方法将某受试者训练得出的分类器应用于其他受试者，证明RSVP系统具有在不同

受试者之间的泛化能力。Vareka等训练了一个堆叠自动编码器（Stacked Auto-Encoder, SAE）并将其与线性判别分析（Linear Discriminant Analysis, LDA）和多层感知机（Multi-Layer Perceptron, MLP）的性能进行了对比，结果表明SAE方法对于处理高维的EEG数据优于LDA和MLP。Xiao等^[35]提出的判别规范模式匹配（Discriminative Canonical Pattern Matching, DCPM）方法结合判别模式空间滤波和典型模板匹配用于RSVP分类，实现了比IDCA更好的分类效果。

2.3.3 深度学习

P300检测方面的研究主要基于卷积神经网络（Convolutional Neural Networks, CNN）^[36]、长短期记忆（Long Short-Term Memory, LSTM）网络等。2018年，Lawhern等^[37]提出了用于EEG数据的经典网络——EEGNet。EEGNet一经发布，便在EEG各个范式中取得了较好的结果，之后大部分的深度学习研究方法研究都是基于EEGNet做的改进。EEGNet结构如图8所示。2019年，Chen等^[38]提出了一种CNN，探究不同频率下的RSVP信号，认为CNN

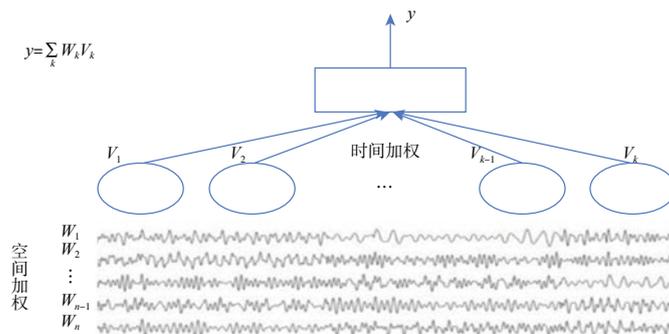


图7 HDCA方法的训练过程
Fig.7 Training process of HDCA

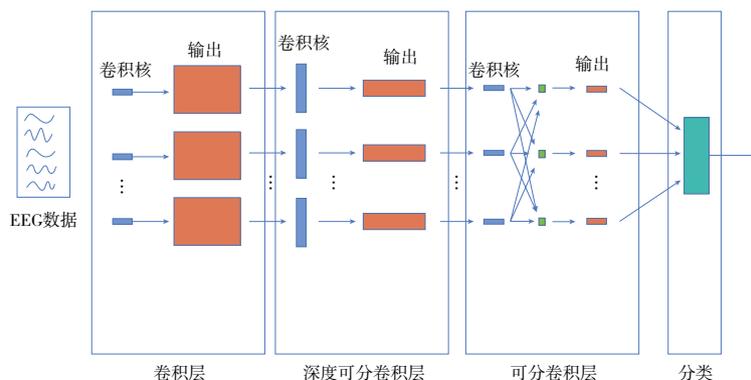


图8 EEGNet结构
Fig.8 Schematic diagram of EEGNet structure

对于不同信号的提取具有天然的优势。CNN 结构如图 9 所示。此外，还衍生出了 BN3、MACRO^[39] 等网络模型，这些模型通过特定的网络结构，利用大量的训练参数和训练数据集提取时空信息。

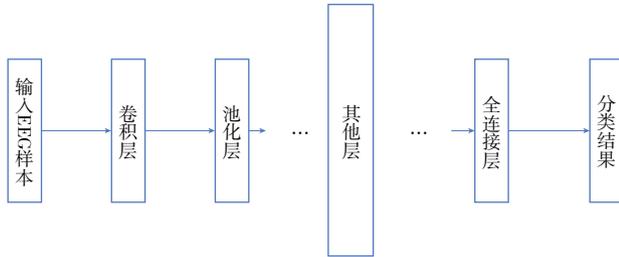


图 9 CNN 结构
Fig.9 Schematic diagram of CNN structure

3 总结

近年来，基于 ERP 的目标检测研究逐渐成为科学领域的研究热点，相关文章数量逐年增长。相关研究主要发展为基于人脑对目标的神经响应和基于计算机强大的信息处理能力两大分支^[40]。本节归纳了这些研究的机制、局限性、优势和适用场景，具体如表 3 所示。

同时，由于不同个体之间存在认知的多样性和复杂性，基于 ERP 的目标检测技术仍然面临许多

问题和挑战，主要有以下几个方面。

3.1 个体差异和泛化性能

不同个体之间的 ERP 信号表现出较大的差异，相关算法的泛化能力和鲁棒性仍然较弱，使建立具有良好泛化性能的目标检测模型变得困难。模型需要能够适应各种个体的信号特征，并在不同场景下保持稳定和可靠的检测性能。未来可考虑对实验范式进行创新以解决个体差异大的问题。

3.2 实时性要求

某些应用场景对目标检测具有实时性要求，而 ERP 信号的采集和处理需要一定的时间延迟。因此，在保持良好检测性能的同时，还需要考虑如何优化算法以实现实时目标检测。未来可考虑结合强化学习进行在线训练，通过用户与具体应用的交互获得最佳在线分类效果。

3.3 噪声和干扰

ERP 信号容易受到来自环境、生理、设备的噪声和干扰影响，如眼动、肌电、伪迹等。这些噪声和干扰可能会掩盖或削弱目标信号，对目标检测的准确性产生负面影响，因此未来需要考虑更有效的去噪和信号增强方法。

表 3 目标检测相关文献研究
Tab.3 Research on target detection related literatures

参考文献	机制	局限性	优势	适用场景	年份
[44]	基于迭代最小距离平方误差原理的异步检测框架	训练样本有限，分类精度较低，视频诱导的 EEG 波动性差	克服目标起始时间和检测延迟的时间抖动，降低虚警率，跨任务通用模型	样本数据量大，适用于提高各项任务分类精度、虚警率要求低的场景	2020
[41]	基于判别典型模式匹配分类方法抑制脑电噪声，识别 ERP 的典型模式	计算复杂度较高	可识别各类 ERP 成分，分类效果好	需要综合考虑多个 ERP 成分特征的处理场景	2020
[42]	基于公共空间模式算法创新，对所有维度进行滤波并定义每种信息模式的正则化项	计算复杂度较高	相比公共空间模式，减少了噪声的影响和过拟合问题，在样本数据量较小的情况下可达到较高的分类精度	对耗时要求宽松，样本数据量小，适用于脑电电极阻抗大的处理场景	2021
[43]	基于三重 RSVP 范式的深度学习方法	提取的特征与刺激图像有特定的联系，但联系程度和形式未知	建立 EEG 信号和刺激图像之间的相关性，可肉眼观察 EEG 解码后的图像，为 EEG 研究提供新思路	设计更有效的特征分类算法	2021
[45]	基于 EEG 信号的听觉目标识别方法	提出的 EEG 分类算法需要进一步验证在线系统的检测准确率、虚警率和响应时间，受试者较少	在分类性能上，基于 CNN 的分类器在训练集和测试集上的最优准确率高于基于 SVM 的传统分类器	提出的方法可以作为在目标检测系统中加入人工智能感知环境的理论基础	2022
[46]	基于脑机接口系统增强 ERP 组件的物联网 UAV 目标检测	联合决策方法可能不适用于动态目标，且单样本的检出率较低	噪声时空均衡器成功地增强了 ERP 判别分量，使提出的算法优于其他基线算法	把重点放在无人机图形信息的安全性上。确保图片信息安全对监控、救援和搜索有影响	2023

3.4 样本需求量大

以深度学习为主的大部分模型结构简单,可实现端到端分类,但需要大量训练样本才能达到较好的检测效果。而脑电领域进行实验的成本较高,对于特定任务难以获得大量数据。迁移学习方法通过将大量先验知识迁移到特定任务中,可以与深度学习方法相结合解决样本需求量大的问题。

虽然目标检测技术面临许多挑战,但研究者们正在采取各种方法来应对这些挑战。这些方法包括迁移学习、胶囊网络等技术,以及组合模型、多模态特征和实验范式的创新等。随着与虚拟现实、增强现实、眼动和计算机等技术的结合,目标检测技术将具有广泛的应用前景。

参考文献

- [1] RUSSAKOVSKY O, DENG J, SU H, *et al.* Imagenet large scale visual recognition challenge[J]. **International Journal of Computer Vision**, 2015,115(3):211-252.
- [2] SILVER D, HUANG A, MADDISON C J, *et al.* Mastering the game of go with deep neural networks and tree search[J]. **Nature**, 2016, 529(7585):484-489.
- [3] 徐晓. 计算机视觉中物体识别综述 [J]. **电脑与信息技术**, 2013 (5) : 4-6.
XU Xiao. Review of object recognition in computer vision[J]. **Computer and Information Technology**, 2013(5):4-6.
- [4] 曲良承. 针对遥感图像目标检测的人机对比分析研究 [D]. 西安: 西安电子科技大学, 2020.
QU Liangcheng. A research on man-machine contrast analysis of object detection in remote sensing image[D]. Xi'an: Xidian University, 2020.
- [5] 李子杏, 闫超, 周晗, 等. 基于ERP信号的目标检测研究综述 [J]. **计算机工程与应用**, 2021, 57 (23) : 37-49.
LI Zixing, YAN Chao, ZHOU Han, *et al.* A review of target detection based on ERP signal[J]. **Computer Engineering and Applications**, 2021, 57(23):37-49.
- [6] RAMADAN R A, VASILAKOS A V. Brain computer interface: control signals review[J]. **Neurocomputing**, 2017 (223): 26-44.
- [7] GERSON A D, PARRA L C, SAJDA P. Cortical origins of response time variability during rapid discrimination of visual objects[J]. **Neuroimage**, 2005, 28(2): 342-353.
- [8] POLICH J, DONCHIN E. P300 and the word frequency effect[J]. **Electroencephalography and Clinical Neurophysiology**, 1988, 70(1): 33-45.
- [9] SAJDA P, POHLMAYER E, WANG J, *et al.* In a blink of an eye and a switch of a transistor: cortically coupled computer vision[J]. **Proceedings of the IEEE**, 2010, 98(3):462-478.
- [10] CECOTTI H, ECKSTEIN M P, GIESBRECHT B. Effects of performing two visual tasks on single-trial detection of event-related potentials[C]//2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Montevideo, Uruguay: IEEE, 2012: 1723-1726.
- [11] WEIDEN M, KHOSLA D, KEEGAN M. Electroencephalographic detection of visual saliency of motion towards a practical brain-computer interface for video analysis[C]//Proceedings of the 14th ACM International Conference on Multimodal Interaction, 2012: 601-606.
- [12] ROSENTHAL D, DEGUZMAN P, PARRA L C, *et al.* Evoked neural responses to events in video[J]. **IEEE Journal of Selected Topics in Signal Processing**, 2014, 8(3): 358-365.
- [13] 周前祥, 姜国华, 曲战胜, 等. 视场角设置对虚拟环境中目标判断的影响 [J]. **航天医学与医学工程**, 2003 (4) : 292-295.
ZHOU Qianxiang, JIANG Guohua, QU Zhansheng, *et al.* Effects of field viewing angles on object judgement in virtual environment[J]. **Space Medicine & Medical Engineering**, 2003(4):292-295.
- [14] DIAS J C, PARRA L C. NO EEG evidence for subconscious detection during rapid serial visual presentation[C]//2011 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE, 2011: 1-4.
- [15] 褚凯轩, 常天庆, 郭理彬, 等. 目标视场角对P300-RSVP目标检测系统的影响分析 [J]. **计算机工程与应用**, 2018, 54 (23) : 115-119.
CHU Kaixuan, CHANG Tianqing, GUO Libin, *et al.* Analysis of influence of target perspective on P300-RSVP target detection system[J]. **Computer Engineering and Applications**, 2018, 54(23):115-119.
- [16] SAJDA P, GERSON A, PARRA L. High-throughput image search via single-trial event detection in a rapid serial visual presentation task[C]//First International IEEE EMBS Conference on Neural Engineering. IEEE, 2003: 7-10.
- [17] GONSALVEZ C L, POLICH J. P300 amplitude is determined by target-to-target interval[J]. **Psychophysiology**, 2010, 39(3):388-396.
- [18] WARD R, DUNCAN J, SHAPIRO K. Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention[J]. **Perception & Psychophysics**, 1997 (59): 593-600.
- [19] ROUSSELET G A, THORPE S J, FABRE-THORPE M. How parallel is visual processing in the ventral pathway?[J]. **Trends in Cognitive Sciences**, 2004, 8(8): 363-370.
- [20] SHAPIRO K L, RAYMOND J, ARNELL K. The attentional blink[J]. **Trends in Cognitive Sciences**, 1997, 1(8) : 291-296.
- [21] PEDRONI A, BAHREINI A, LANGER N. Automagic: standardized preprocessing of big EEG data[J]. **NeuroImage**, 2019 (200):460-473.
- [22] HESSE C W, JAMES C J. The fast ICA algorithm with spatial constraints[J]. **IEEE Signal Processing Letters**, 2005, 12(11): 792-795.
- [23] 张柏雯. 基于结构磁共振和深度学习模型的AD病程分类研究 [D].

- 北京: 北京工业大学, 2018.
- ZHANG Baiwen. Research on classification of AD stage with structural MRI and deep learning model[D]. Beijing: Beijing University of Technology, 2018.
- [24] BLANKERTZ B, TOMIOKA R, LEMM S, *et al.* Optimizing spatial filters for robust EEG single-trial analysis[J]. **IEEE Signal Processing Magazine**, 2007, 25(1): 41-56.
- [25] RIVET B, SOULOUMIAC A, ATTINA V, *et al.* xDAWN algorithm to enhance evoked potentials: application to brain-computer interface[J]. **IEEE Transactions on Biomedical Engineering**, 2009, 56(8): 2035-2043.
- [26] JONMOHAMADI Y, POUDEL G, Innes C, *et al.* Comparison of beamformers for EEG source signal reconstruction[J]. **Biomedical Signal Processing and Control**, 2014 (14): 175-188.
- [27] JUNG T P, MAKEIG S, HUMPHRIES C, *et al.* Removing electroencephalographic artifacts by blind source separation[J]. **Psychophysiology**, 2000, 37(2): 163-178.
- [28] BIGDELY-SHAMLO N, VANKOV A, RAMIREZ R R, *et al.* Brain activity-based image classification from rapid serial visual presentation[J]. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, 2008, 16(5): 432-441.
- [29] ALPERT G F, MANOR R, SPANIER A B, *et al.* Spatiotemporal representations of rapid visual target detection: a single-trial EEG classification algorithm[J]. **IEEE Transactions on Biomedical Engineering**, 2014, 61(8):2290-2303.
- [30] HARIHARAN B, ARBELÁEZ P, GIRSHICK R, *et al.* Hypercolumns for object segmentation and fine-grained localization[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 447-456.
- [31] YU K, SHEN K, SHAO S, *et al.* Common spatio-temporal pattern for single-trial detection of event-related potential in rapid serial visual presentation triage[J]. **IEEE Transactions on Biomedical Engineering**, 2011, 58(9): 2513-2520.
- [32] YU K, SHEN K, SHAO S, *et al.* Bilinear common spatial pattern for single-trial ERP-based rapid serial visual presentation triage[J]. **Journal of Neural Engineering**, 2012, 9(4): 046013.
- [33] 张宁宁. 基于多通道脑电的图像识别方法研究 [D]. 秦皇岛: 燕山大学, 2020.
- ZHANG Ningning. Research on image recognition method based on multi-channel EEG signal[D]. Qinhuangdao: Yanshan University, 2020.
- [34] MATHAN S, WHITLOW S, MAZAEVA N. Sensor-based cognitive state assessment in a mobile environment[C]//Proceedings of the 11th International Conference on Human-Computer Interaction. 2005: 110-119.
- [35] XIAO X, XU M, JIN J, *et al.* Discriminative canonical pattern matching for single-trial classification of ERP components[J]. **IEEE Transactions on Biomedical Engineering**, 2019, 67(8): 2266-2275.
- [36] CECOTTI H, GRASER A. Convolutional neural networks for P300 detection with application to brain-computer interfaces[J]. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2010, 33(3): 433-445.
- [37] LAWHERN V J, SOLON A J, WAYTOWICH N R, *et al.* EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces[J]. **Journal of Neural Engineering**, 2018, 15(5): 056013.
- [38] CHEN J X, MAO Z J, ZHENG R, *et al.* Feature selection of deep learning models for EEG-based RSVP target detection[J]. **IEICE Transactions on Information and Systems**, 2019, 102(4): 836-844.
- [39] LAN Z, YAN C, LI Z, *et al.* MACRO: multi-attention convolutional recurrent model for subject-independent ERP detection[J]. **IEEE Signal Processing Letters**, 2021 (28): 1505-1509.
- [40] ZOU Z, CHEN K, SHI Z, *et al.* Object detection in 20 years: a survey[J]. **Proceedings of the IEEE**, 2023, 111(3): 257-276.
- [41] XIAO X L, XU M P, JIN J, *et al.* Discriminative canonical pattern-matching for single-trial classification of ERP components[J]. **IEEE Transactions on Biomedical Engineering**, 2019, 67(8): 2266-2275.
- [42] JALILPOUR S, SARDOUIE S H. RCTP: regularized common tensor pattern for rapid serial visual presentation spellers[J]. **Biomedical Signal Processing and Control**, 2021 (70): 102930.
- [43] QU L, CHEN D, YIN K. Research on EEG feature decoding based on stimulus image[C]//2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). IEEE, 2021, 4: 467-470.
- [44] XIAO X, XU M, JIN J, *et al.* Discriminative canonical pattern matching for single-trial classification of ERP components[J]. **IEEE Transactions on Biomedical Engineering**, 2020, 67(8): 2266-2275.
- [45] WANG Y, LUO Z, ZHAO S, *et al.* Spatial localization in target detection based on decoding N2pc component[J]. **Journal of Neuroscience Methods**, 2022 (369): 109440.
- [46] ZHANG Y, ZHANG H, GAO X, *et al.* UAV target detection for IoT via enhancing ERP component by brain computer interface system[J]. **IEEE Internet of Things Journal**, 2023, 10(19): 17243-17253.