

文章编号: 2095-2163(2020)12-0064-05

中图分类号: U471.1

文献标志码: A

# 基于深度学习的自动驾驶车辆模型迁移路径规划研究

欧阳可可, 杨国平

(上海工程技术大学 机械与汽车工程学院, 上海 201620)

**摘要:**为解决传统路径规划方法中无车辆动力学约束和消除车辆模型跟踪误差的问题,本文提出一种基于深度学习的自动驾驶车辆模型迁移路径规划方法。首先,根据真实环境建立虚拟行车环境模型,该模型应用经深度学习训练后的最优自动驾驶策略;其次,通过 MATLAB 自动驾驶场景设计器将实际场景问题迁移至虚拟抽象模型中;最后,经过仿真实验验证该方法。实验结果表明,所提方法能够减少横向跟踪误差和提高模型的泛化性能,减小过度依赖问题。  
**关键词:**自动驾驶车辆;深度学习;模型迁移;路径规划

## Research on Migration Path Planning of Autonomous Vehicle Model Based on Deep Learning

OUYANG Keke, YANG Guoping

(College of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**[Abstract]** In order to solve the problem of no vehicle dynamics constraints in the traditional path planning method and eliminate the tracking error of the vehicle model, a deep learning-based autonomous vehicle model migration path planning method is proposed. First establish a virtual driving environment model based on the real environment, which applies the optimal autonomous driving strategy after deep learning training; then transfer the actual scene problem to the virtual abstract model through the MATLAB automatic driving scene designer, and finally verify it by simulation experiment this method. Experimental results show that the proposed method can reduce the lateral tracking error, improve the generalization performance of the model, and reduce the problem of excessive dependence.

**[Key words]** Self-driving vehicle; Deep learning; Model migration; Route plan

### 0 引言

自动驾驶技术飞速发展,但目前的自动驾驶系统还不能绝对的安全。对行进中的车辆动态外观识别与分类,追踪其行驶路径轨迹具有重要研究和现实意义。因此,车辆跟踪技术已经成为了智能交通等诸多应用领域中至关重要的技术之一<sup>[1]</sup>。随着深度学习的快速发展,其作为机器学习领域中的一个课题开始受到了广泛的关注,轨迹规划在基于视觉控制方面取得了很大的成就。PAXTON C 等通过神经网络学习和深度 Q 网络控制策略的方法,可以简单的实现复杂的运动规划任务<sup>[2]</sup>;Hotz 等人用基于深度学习的方法,完成了轨迹跟踪和自动驾驶车辆潜在空间解码图像的任务<sup>[3]</sup>;LIU W 等提出基于深度学习的方法来建立控制序列与对应轨迹之间映射关系,为实现自动驾驶车辆的路径规划<sup>[4]</sup>。

Q 学习能够很好的解决离散空间的低维度问题,而深度 Q 网络(deep Q network, DQN)是针对高维状态空间并且能有效的处理,但是基本不能解决高维连续动作空间的问题。基于深度学习的深度确

定性策略梯度(DDPG)算法,优势在于能够处理连续动作空间,利用 Actor-Critic 框架,融合 DQN 优势。同时,针对传统轨迹规划方法无法消除模型误差和对跟踪控制依赖程度高的问题,利用 DDPG 算法降低轨迹规划方法对样本数据的依赖。本文所提方法与传统轨迹规划方法相比,自动驾驶车辆行驶时能减小横向误差,并能提高模型的泛化性能。

### 1 行车环境描述

#### 1.1 虚拟行车环境模型设计

根据自动驾驶车辆行为决策中给定的车辆驾驶场景描述,设计如图 1 所示行车环境模型。

图 1 为具有三车道的圆形地图,其中白色区域是不可跨越的障碍物,两条蓝线区域是供车辆行驶的三车道,橘黄色线条表示期望路径  $pat h_d = (x_d, Y_d, \varphi_d)$ 。自动驾驶车辆  $Car = (x_c, y_c, \varphi_c, v)$  为在圆形地图中行驶自动驾驶的策略,  $x_c, y_c, \varphi_c$  为自动驾驶车辆当前位姿信息。环境模型的关键回报函数 Reward 为式(1):

$$Reward = \begin{cases} -1, \\ R_{action} + R_{reward}. \end{cases} \quad (1)$$

作者简介: 欧阳可可(1995-),男,硕士研究生,主要研究方向:智能驾驶与智能车辆的路径规划。

收稿日期: 2020-09-16

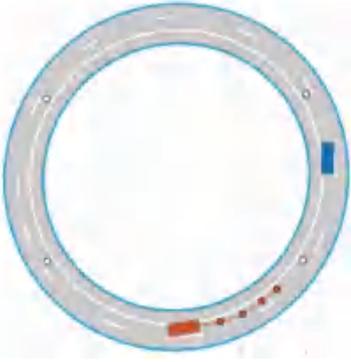


图 1 虚拟行车环境模型

Fig. 1 Virtual driving environment model

其中,  $R_{action}$  表示自动驾驶车辆前轮转角  $\delta_m$ ;  $R_{reward}$  表示自动驾驶车辆获得的奖励;  $\delta$  间的差额惩罚;  $\lambda$  表示正惩罚系数;  $get\_reward$  表示自动驾驶车辆在期望路径上行驶所获得的奖励, 式(4):

$$R_{action} = -\lambda * \|\delta_m - \delta\|^2, \quad (2)$$

$$R_{reward} = \begin{cases} 0, & get\_reward = 0, \\ 0.1, & get\_reward = 1. \end{cases} \quad (3)$$

$$get\_reward = \begin{cases} 0, & \text{其它;} \\ 1, & |\Delta x| \leq \varepsilon_1 \& |\Delta y| \leq \varepsilon_2 \& |\Delta \varphi| \leq \varepsilon_3. \end{cases} \quad (4)$$

$(\Delta x, \Delta y, \Delta \varphi)$  表示自动驾驶车辆当前位姿  $x_c, y_c$ ;  $\varphi_c$  与期望路径之间的差值;  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  为容误差。

自动驾驶车辆的最优驾驶策略为  $\pi$ , 给定策略为  $\pi_{reset}$ , 式(5)

$$\begin{cases} x_0 \in [X_{min}, X_{max}], \\ y_0 \in [Y_{min}, Y_{max}], \\ \varphi_0 \in [\varphi_{min}, \varphi_{max}]. \end{cases} \quad (5)$$

## 1.2 车辆动力学模型

基于图 2 的简单单轨模型车辆的运动学方程可表达为式(6):

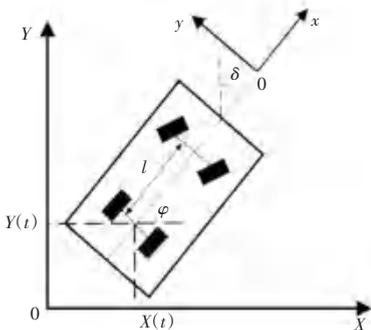


图 2 简单单轨模型

Fig. 2 Simplified monorail model

$$\begin{cases} \dot{X} = v \cos(\varphi + \beta), \\ \dot{Y} = v \sin(\varphi + \beta), \\ \dot{\varphi} = \frac{v}{l} (\tan \delta_f - \tan \delta_r) \cos \beta. \end{cases} \quad (6)$$

其中,  $X$  和  $Y$  分别为车辆在大地坐标系下的横坐标和纵坐标;  $\varphi$  为车辆横摆角;  $\beta$  为质心侧偏角;  $v$  为车速;  $\delta_f$  和  $\delta_r$  分别为单轨模型前后车轮的转角。

考虑到车辆的质心侧偏角较小, 故  $\beta \approx 0$ 、 $\delta_r = 0$ , 所以式(6)可简化为式(7):

$$\begin{cases} \dot{X} = v \cos \varphi, \\ \dot{Y} = v \sin \varphi, \\ \dot{\varphi} = \frac{v}{l} \tan \delta_f. \end{cases} \quad (7)$$

设定纵向速度恒定, 建立动力学模型, 式(8):

$$\begin{cases} m \ddot{y} = -m x \dot{\varphi} + 2 F_{yf} + 2 F_{yr}, \\ I_z \ddot{\varphi} = 2a F_{yf} - 2b F_{yr}, \\ \dot{X} = x \cos \varphi - y \sin \varphi, \\ \dot{Y} = x \sin \varphi + y \cos \varphi, \\ \dot{\varphi} = \varphi. \end{cases} \quad (8)$$

其中:  $F_{yf}$  和  $F_{yr}$  分别表示车辆在纵向方向上前面和后面所受的力;  $\varphi$  为车辆横摆角;  $I_z$  为车辆绕  $z$  轴的转动惯量;  $a, b$  分别为质心到前后轴的距离;  $X, Y$  分别为惯性坐标系下的横纵坐标;  $x, y$  分别为车身坐标系下的横纵坐标;  $m$  为车身质量。

## 2 模型迁移轨迹规划方法

### 2.1 DDPG 实现框架与算法流程

在自动驾驶场景设计器中, 虚拟行车环境模型采用 DDPG 算法训练的深度学习模型。根据国内外学者研究表明, 单个 Q 神经网络算法, 在很大程度上会导致学习过程不稳定, Q 神经网络算法的缺点是神经网络的不同参数更新频率会导致不能有效的促进学习。因此, DDPG 的两个神经网络拷贝, 分别为 online 网络和 target 网络, 分别对应于策略网络、Q 网络。DDPG 算法的优点是比较稳定, 两种网络参数基本不会变化。

DDPG 采用确定性策略, 先初始化在线策略神经网络参数  $\theta^\mu$  和  $\theta^Q$ , 然后将在线策略网络的参数复制给对应的目标策略网络参数。DDPG 算法流程如图 3 所示。输入状态为  $(d_1, \dots, d_9, v, \delta_m)$ , 输出动作为  $\delta$ 。此算法的优点是需采样的数据少, 算法

效率高。

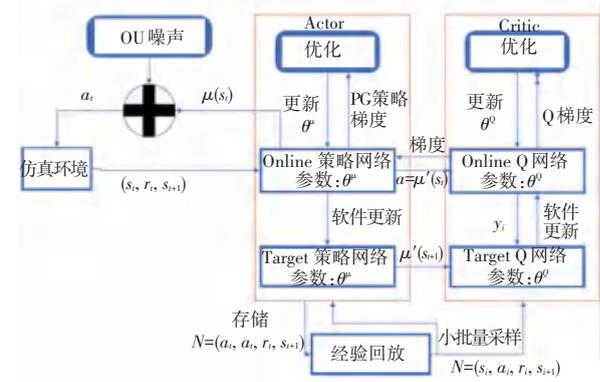


图3 DDPG 算法流程图

Fig. 3 DDPG algorithm flow chart

## 2.2 模型迁移策略

本文根据真实环境建立虚拟行车环境模型,将实际场景问题迁移至虚拟抽象模型,式(9):

$$\begin{cases}
 \dot{\theta} = \varphi_{tar} - \varphi_{ref}, \\
 \dot{x}' = x_{clM} \cos \theta - y_{clM} \sin \theta, \\
 \dot{\varphi}' = \varphi_{clM} - \theta, \\
 \dot{y}' = x_{clM} \sin \theta - y_{clM} \cos \theta, \\
 \dot{x}'_{tar} = x_{tar} \cos \theta - y_{tar} \sin \theta, \\
 \dot{y}'_{tar} = x_{tar} \sin \theta + y_{tar} \cos \theta, \\
 \dot{\varphi}'_{tar} = \varphi_{ref}, \\
 (\Delta x, \Delta y, \Delta \varphi) = (x' - x'_{tar}, y' - y'_{tar}, \theta), \\
 (x_c, y_c, \varphi_c) = (x_{ref} + \Delta x, y_{ref} + \Delta y, \varphi_{ref} + \Delta \varphi), \\
 \delta = \{\delta_1, \delta_2, \dots, \delta_t\}, \\
 \zeta = \{p_1, p_2, \dots, p_t\}
 \end{cases} \quad (9)$$

其中:  $\varphi_{tar}$  为目标航向角;  $\varphi_{ref}$  为参考点航向角;  $\theta$  表示两个航向角之差。

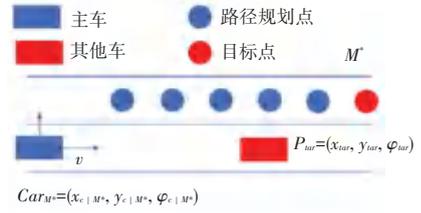
真实环境场景  $M^*$  如图 4(a) 所示, 车道数量、车道宽度等是不变的路况信息, 障碍物信息由自动驾驶车辆测距光束反馈, 是变化的路况信息。根据此时的路况信息可以看出, 前方存在速度小于本车的车辆, 由测距光束反馈得出自动驾驶车辆需要换道行驶的最优驾驶策略。自动驾驶车辆  $Car_{M^*}$ ,  $P_{tar}$  经模型迁移后的场景如图 4(b) 所示。

## 3 自动驾驶车辆轨迹规划仿真测试

### 3.1 深度学习模型训练

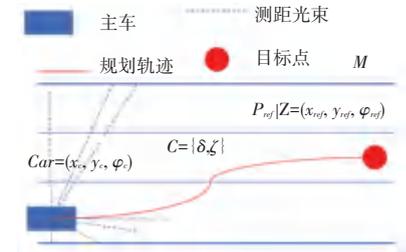
虚拟仿真环境 Env 中圆形地图车道宽 3.75 m, 自动驾驶车辆大小 2 m × 4 m, 速度恒定  $v = 36$  km/h, 测距光束维  $n = 7$ , 且最远测距  $d_{max} = 20$  m, 转角最大输出  $\delta_{max} = 0.2$  rad, 期望路径  $path_d$  为中间车道中心线, 误差范围为  $\varepsilon_1 = \varepsilon_2 = 0.2, \varepsilon_3 = 0.3$ , 连续动作

惩罚系数  $\lambda_1 = 0.01$ , 随机初始位置  $x_0 \in [0, 1000]$ ,  $y_0 \in [0, 60], \varphi_0 \in [-\pi/2, \pi/2]$ , 步长 0.1 s。同时, 利用深度学习框架 Tensorflow。



(a) 真实环境场景

(a) Real environment scene



(b) 模型迁移后场景

(b) Scene after model migration

图4 模型迁移示意图

Fig. 4 Schematic diagram of model migration

深度学习模型  $M$  中的参数对实验的影响: 随着折扣因子越增大, 更重视以往经验; 而随着折扣因子减小, 更重视当前回报; 随着学习速率的增大, 保留之前训练的效果就随之减少。通过多次试验尝试, 选择了以下面参数为标准的深度学习模型  $M_\theta$ : 初始方差  $\text{var}_{max} = 2$ , 最小方差  $\text{var}_{min} = 0.01$ , 衰减率为  $10^{-4}$ , 折扣因子  $\gamma = 0.6$ 。

由图 5(a) 可知, 应用深度学习的轨迹规划方法中平均值  $Q$  在迭代次数 0-10 000 范围内波动较大, 然后逐渐平缓增加, 平均值  $Q$  在 62 000 步时出现了抖动, 最后趋于稳定。由图 5(b) 可知, 策略梯度选择因为随机选择因数出现波动, 从而导致策略选择在中间范围出现了轻微的波动。根据图 5(c) 可以看出噪声在迭代次数 0-6 000 范围内逐渐降低, 然后趋于稳定。由图 5(d) 可知, 损失在迭代次数 0-8 000 范围内缓慢降低, 然后逐渐趋于 0 附近, 说明评价网络有效。由此说明模型训练的结果是最优策略必然导致的。

由图 6(d) 和图 7(d) 车辆的前轮转角依然在合理的范围之内变化, 但是在 0.8 s 处也出现了较大的转角增量。根据图 7(e) 和图 7(f) 可以看出初始航向角不同时, 横向误差都在合理的范围之内, 说明该方法具有良好的轨迹跟踪能力, 误差的极值出现在到达目标点时的转弯处。

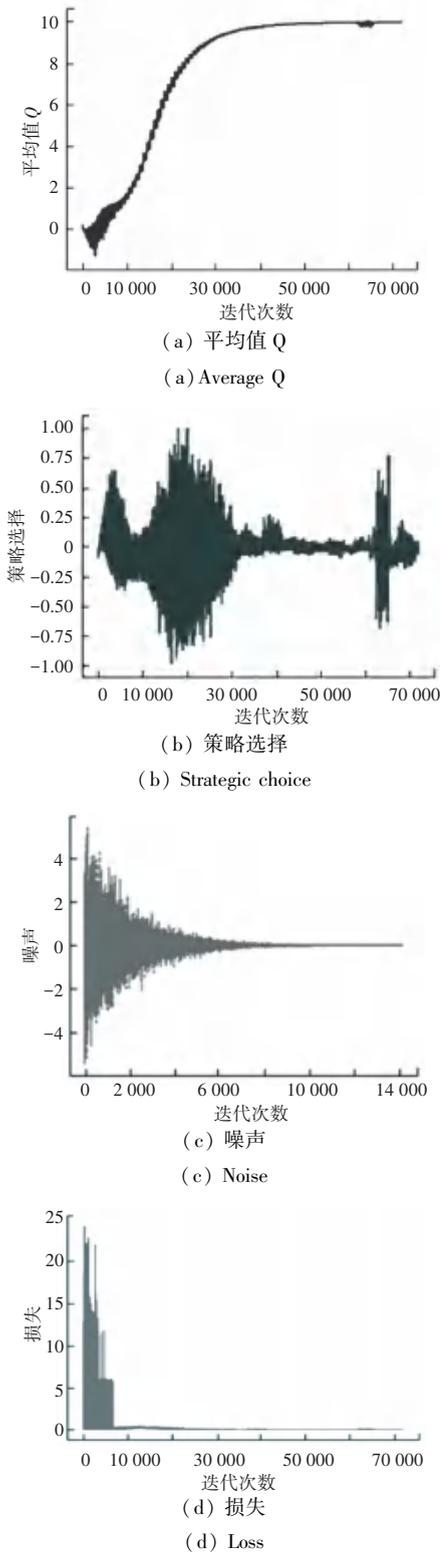


图 5 深度学习模型训练结果

Fig. 5 Deep learning model training results

### 3.2 自动驾驶车辆轨迹规划验证

仿真测试过程中,自动驾驶车辆分别以不同的偏航角出发,如图 6 和图 7 所示,都能顺利完成过弯的驾驶行为,在弯道行驶的控制过程中,车辆的侧向加速度基本上平稳,最大值未超过规定限制。

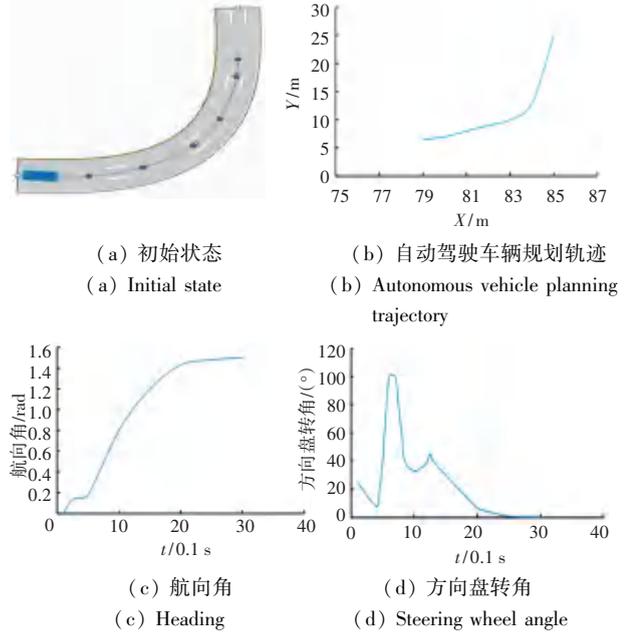


图 6 弯道行驶实验一

Fig. 6 Curve driving test one

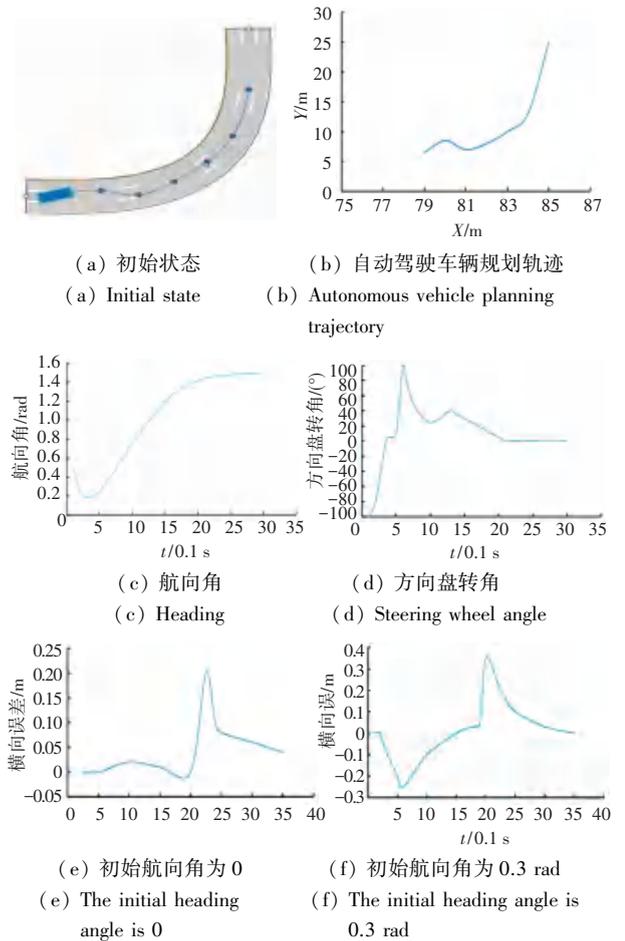


图 7 弯道行驶实验二

Fig. 7 Curve driving test two

### 4 结束语

仿真测试的结果表明,本文所提的轨迹规划方法能够较为精准的跟踪规划路径,且侧向加速度和

前轮转角变化均在合理的范围之内。通过在 MATLAB 自动驾驶场景设计器中,车辆以恒速行驶通过弯道的验证实验。实验结果表明,本文所提法能够减少横向跟踪误差和提高模型的泛化性能,减小过度依赖问题。

## 参考文献

[1] 赵海. 视频序列下的车辆轨迹异常行为识别[D]. 昆明: 昆明理工大学, 2014.

- [2] PAXTON C, RAMAN V, HAGER G D, et al. Combining neural networks and tree search for task and motion planning in challenging environments [C]// IEEE/RSJ International Conference on Intelligent Robots and Systems ( IROS ). Vancouver, BC: IEEE, 2017; 6059-6066.
- [3] SANTANA E, HOTZ G. Learning a driving simulator. arXiv 1608.01230[cs.LG]. 2016, <https://arxiv.org/abs/1608.01230v1>.
- [4] LIU W, LI Z, LI L, et al. Parking like a human: a direct trajectory planning solution[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(12): 3388-3397.

(上接第 63 页)

测试集进行测试。每次的交叉验证完成后,保留每次所得到训练数据和测试数据。

将每个算法得到的训练数据和测试数据分别相加并计算平均值,就得到了新的训练集和测试集,接着利用逻辑回归算法作为次级学习器进行第二层的训练,并在测试集上进行测试,得到最终决策结果,如图 2 所示。通过 5 折交叉验证,可以避免再次训练数据产生过拟合。

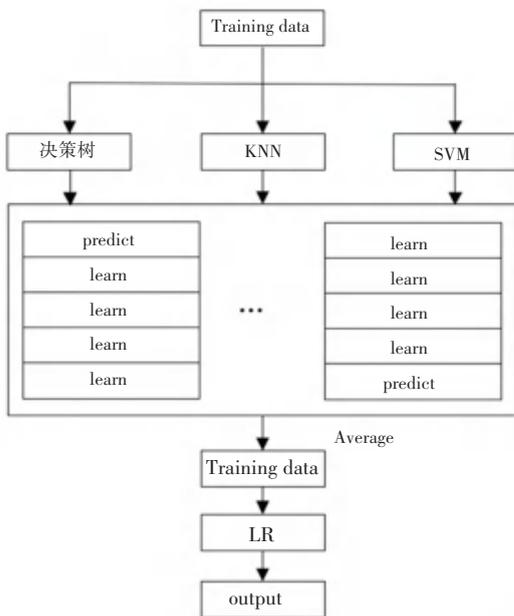


图 2 stacking 训练框架图

Fig. 2 Training based stacking frame diagram

## 2 实验结果及分析

本文首先对两种模态的 MRI 影像进行图像预处理、特征提取以及配准之后,利用 SVM、KNN 和决策树 3 种分类算法,通过基于学习法的决策级的融合算法得到最终分类决策结果。在实验中,对比单模态 DCE-MRI 和 DWI 的实验结果,如图 3 所示。DCE 影像 AUC 值为 0.827, DWI 影像的 AUC 值为 0.793,而基于学习法的决策级融合算法,结合两种模态影像的 AUC 值为 0.854。

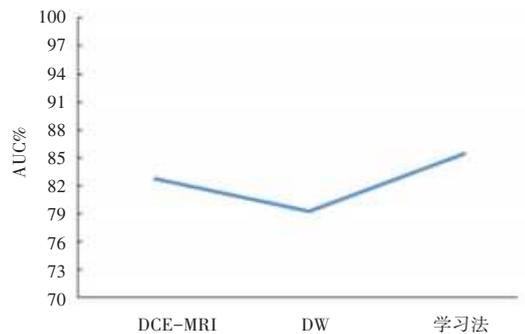


图 3 单模态与多模态融合 AUC 值对比图

Fig. 3 Comparison chart of AUC value of single and multi-fusion

可以看出相比较单模态影像数据的分类结果,多模态影像数据的分类效果比其他两种单模态影像的效果都要更好,结合多模态影像数据能够辅助医生提高诊断的准确率。

## 3 结束语

本文对乳腺 MRI 的 DWI 和 DCE-MRI 二种模态的影像进行预处理、配准以及融合操作,获取二种模态下的更详细的病灶信息,提高乳腺癌良恶性分类地准确度,也更利于医生的辅助诊断。

## 参考文献

- [1] ORGANIZATION W H. Last global cancer data[OL]. <https://www.iarc.fr/featured-news/latest-global-cancer-data-cancer-burden-rises-to-18-1-million-new-cases-and-9-6-million-cancer-deaths-in-2018/>.
- [2] ANDERSON B O, BRAUN S, LIM S, et al. Early detection of breast cancer in countries with limited resources[J]. Breast J, 2003, 9(2): 1-9.
- [3] ANDERSON B O, BRAUN S, CARLSON R W, et al. Overview of breast health care guidelines for countries with limited resources[J]. Breast J, 2003, 9(2): 42-50.
- [4] 刘欣婷, 时伟, 武林松. 彩色多普勒超声与钼靶 X 线诊断乳腺癌的对比分析[J]. 安徽医学, 2015, 36(4): 460-462.
- [5] 师红莉, 许秋霞. 多模态核磁共振成像技术对乳腺癌的诊断价值[J]. 中国实用医药, 2016, 11(26): 121-122.
- [6] 陈婷, 鲁珊珊, 张晶, 等. MRI 动态增强和弥散加权成像诊断子宫内腺癌肌层浸润和术前分期的对照研究[J]. 南京医科大学学报(自然科学版), 2014, 34(12): 1753-1756.
- [7] 李航. 统计学习方法[M]. 清华大学出版社, 2012.