

文章编号: 2095-2163(2023)04-0020-07

中图分类号: TP391.4

文献标志码: A

基于粗糙集和集成剪枝的人脸表情识别方法

唐玉梅, 李丹杨, 吴亚婷, 黄仕松, 陈星, 吴义青

(贵州大学 大数据与信息工程学院, 贵阳 550025)

摘要: 人脸表情识别在疲劳驾驶监测等场景有着广泛的运用。针对人脸表情识别难度大, 单一分类器泛化能力较弱的缺点, 基于集成学习理论, 提出一种基于粗糙集和集成剪枝的人脸表情识别方法。首先, 更改卷积神经网络参数得到若干基分类器; 其次, 结合粗糙集理论, 根据基分类器的预测结果构建信息决策表, 将分类器选择转化为知识约简过程, 剔除系统中弱分类器或冗余分类器, 选出分类器子集; 最后, 用大多数投票法将选择出来的分类器子集组合。和多个集成剪枝算法对比, 本文集成剪枝算法在表情数据集上具有较高的识别准确率。

关键词: 表情识别; 卷积神经网络; 集成剪枝; 粗糙集理论

A facial expression recognition method based on rough set and integrated pruning

TANG Yumei, LI Danyang, WU Yating, HUANG Shisong, CHEN Xing, WU Yiqing

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] Facial expression recognition is widely used in fatigue driving monitoring and other scenes. Aiming at the disadvantages of difficult face expression recognition and weak generalization ability of single classifier, a face expression recognition method based on rough set and integrated pruning is proposed based on ensemble learning theory. Firstly, the convolutional neural network parameters are changed to obtain some base classifiers. Secondly, combined with rough set theory, an information decision table is constructed according to the prediction results of base classifier, and classifier selection is transformed into a knowledge reduction process. Weak or redundant classifiers are removed from the system, and classifier subsets are selected. Finally, the subsets of selected classifiers are combined using majority voting methods. Compared with multiple integrated pruning algorithms, the integrated pruning algorithm in this paper has higher recognition accuracy in facial expression data sets.

[Key words] facial expression recognition; convolutional neural network; integrated pruning; rough set theory

0 引言

面部表情识别是近年来模式识别与人工智能领域的研究热点, 在疲劳驾驶监测、刑事案件侦破、课堂教育、医疗诊断等场景有着广泛的应用^[1-2]。尽管已有学者对表情识别技术展开一系列的研究^[3-5], 但由于光照、年龄、性别、拍照姿势等综合因素的影响, 要想获得较高的表情识别准确率在机器学习和模式识别领域仍是一个复杂且富有挑战的问题。

深度学习成为表情识别领域的研究热点^[6-9]。但是, 由于识别过程通常采用单一模型进行分类预

测, 对未知数据的预测能力有限, 集成学习能通过组合若干分类器提高模型的泛化能力, 有效弥补单一模型在表情识别领域的不足^[10-11]。

随着集成分类器数目的增多, 分类器之间的多样性降低, 产生冗余的分类器, 不但不能提升预测系统的准确率, 却还会增加计算量和存储空间, 降低预测的时效性。从所有基分类器中选择少而好的分类器集成, 有效弥补集成学习存在的上述缺陷, 由此产生“选择性集成”, 达到了以较小的集成的规模, 在保证分类准确率稳中有升的同时减少计算负载的效果^[12-14]。Li等学者^[15]提出基于图的动态剪枝算法, 将图论的思想运用到分类器选择过程, 证明选择

作者简介: 唐玉梅(1997-), 女, 硕士研究生, 主要研究方向: 机器学习、选择性集成; 李丹杨(1991-), 女, 博士, 副教授, 主要研究方向: 机器学习、大数据信息融合、选择性集成、人脸表情识别; 吴亚婷(1979-), 女, 讲师, 主要研究方向: 电路分析及传感器网络应用; 黄仕松(1998-), 男, 硕士研究生, 主要研究方向: 机器学习、选择性集成; 陈星(1998-), 男, 硕士研究生, 主要研究方向: 机器学习、选择性集成; 吴义青(1999-), 男, 硕士研究生, 主要研究方向: 机器学习、选择性集成。

通讯作者: 李丹杨 Email: danyangcl@163.com

收稿日期: 2022-05-10

性集成是一种有效提高人脸表情识别准确率的方法。贾澎涛等学者^[16]为了提高人脸表情识别的性能,提出一种基于集成学习理论的二次优化选择性集成分类。

分类器选择性集成在表情识别取得一定的效果,但现有相关研究的文献较少,而且,如何选出“好而不同”的分类器没有统一的衡量标准,主流的选择思想是通过准确率或多样性选择分类器子集。然而,如何有效地度量并利用分类器之间的差异性的问题也没有得到很好的解决。基于粗糙集理论的知识约简无需先验知识,在不损失分类或学习信息的前提下降低数据维度,简化学习过程,在理论和应用上都取得了丰硕的研究成果^[17-20]。粗糙集理论也被运用于人脸表情识别领域。冯林等学者^[21]提出一种基于邻域粗糙集与量子遗传算法的人脸表情特征选择。段丽等学者^[22]提出一种基于粗糙集的表情特征选择。但这些方法目的是为了降低人脸表情特征的维度,找到一个特征子集,提升分类器的性能。

综上,本文提出一种面向人脸表情识别的基于知识约简的分类器选择方法,首次将粗糙集运用于分类器的选择过程,综合知识约简、邻域粗糙集和一致性指标的思想,提出了基于一致性阈值的邻域粗糙集模型,发掘分类器之间的依赖关系,对集成系统进行剪枝,依次选出对集成系统贡献最大的“最优”分类器子集集成,并运用于人脸表情的识别领域,取得比多个集成剪枝算法更好的效果。

1 基于粗糙集和分类器选择的人脸表情识别

为了减少人脸表情识别分类器的集成规模,本文提出一种基于粗糙集和集成剪枝的人脸表情识别方法,实验流程如图 1 所示。首先,训练分类器,选择卷积神经网络算法对人脸表情数据加以训练,得到若干基分类器并对验证数据集进行预测;其次,进行分类器选择,即根据预测结果构建分类器选择的信息决策表,综合邻域粗糙集^[23-24]和一致性指标的思想,提出了基于一致性阈值的邻域粗糙集模型,利用粗糙集对信息决策表进行知识约简,从而选出分类器子集;最后,集成分类器,即使用大多数投票法将选出的分类器子集进行集成。

1.1 基分类器训练

卷积神经网络(CNN)可以更好地展现图像的深层次信息,对大型图像处理表现十分出色。因此,本文选择卷积神经网络作为基分类器,拟合表情数

据。

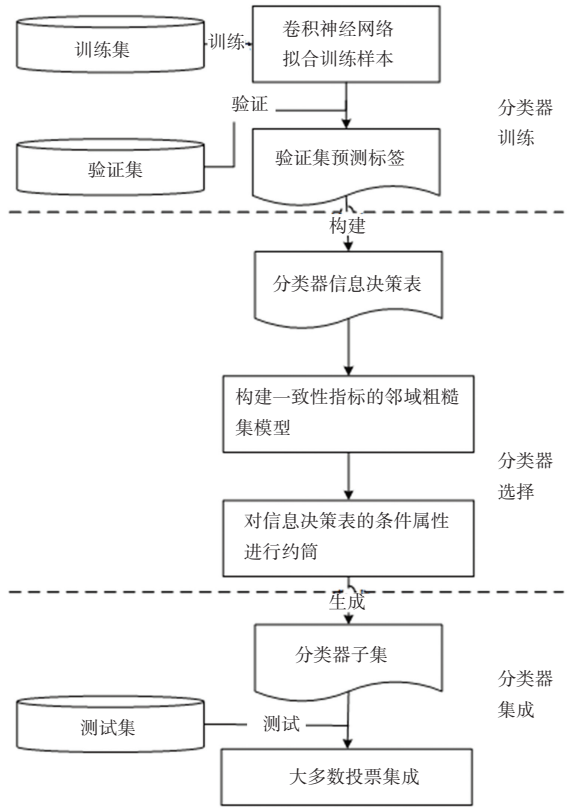


图 1 基于粗糙集集成剪枝的人脸表情识别模型
Fig. 1 Facial expression recognition model based on rough set and integrated pruning

卷积神经网络模型如图 2 所示,由图 2 可知,卷积神经网络基本结构包括输入层、卷积层、池化层、全连接层和输出层。本文共构造了 31 种不同 CNN 模型,各模型的不同之处包括网络层数、卷积核大小、填充方式、步长、激活函数等,从而增加分类器间的多样性,具体参数设置将在实验部分讨论。

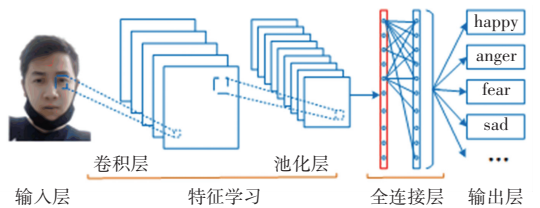


图 2 卷积神经网络模型
Fig. 2 Convolutional neural network model

1.2 基于邻域粗糙集的分类器选择模型

粗糙集不需要任何先验知识或附加信息,处理不相容、不精确和不完整数据的问题,可用于知识约简。因此,本文拟基于粗糙集属性重要度指标,衡量分类器的重要程度,依次选出重要度较大的分类器子集集成。

然而,经典粗糙集局限于论域上严格的等价关

系,完全按照集合间的包含关系定义下近似集,当分类器预测值与样本实际类别不符时,难以求解下近似,从而无法区分不同分类器的重要程度。邻域粗糙集支持离散型和连续型数据处理,借助给定的半径来约束样本之间的相似性,进而实现邻域信息粒化,一致性指标进一步弥补粗糙集对噪声敏感的不足。邻域粗糙集通过分类器对样本的预测结果计算样本间的距离,得到各样本的邻域,并根据邻域信息判断样本是否能够划分到下近似中,对错误预测的样本有较好的容忍性。

1.2.1 构建信息决策表

根据粗糙集模型,利用基分类器对验证集数据的预测结果构建信息决策表 $S = \langle U, C \cup D \rangle$ 。其中, $U = \{x_1, x_2, \dots, x_n\}$ 为论域,由所有验证集样本组成, x_i 为第 i 个验证样本; $C = \{c_1, c_2, \dots, c_m\}$ 为条件属性, C 表示分类器池,属性值为第 k 个分类器 c_k 对样本 x_i 的预测结果 pre_{ik} ; $D = \{d_i\}$ 是决策属性,由验证集样本真实标签组成,所有属性值均为离散数值。可以得到的信息决策表见表 1。

表 1 基于基分类器对验证集预测结果构建信息决策表

Tab. 1 Based on the base classifier, an information decision table is constructed for the predicted results of validation set

样本 (U)	条件属性 (C)				决策属性 (D)
	c_1	c_2	\dots	c_m	
x_1	pre_{11}	pre_{12}	\dots	pre_{1m}	d_1
\dots	\dots	\dots	\dots	\dots	\dots
x_n	pre_{n1}	pre_{n2}	\dots	pre_{nm}	d_n

1.2.2 构建邻域一致性粗糙集模型

针对表 1 分类器池构建的信息决策表,可定义样本间存在邻域关系如下:

定义 1 若对论域 U 中的任意样本 x_i, x_j, x_n , 都存在唯一确定的实函数 Δ 与之对应,且 Δ 满足:

(1) $\Delta(x_i, x_j) \geq 0$, 当且仅当 $x_i = x_j, \Delta(x_i, x_j) = 0$;

(2) $\Delta(x_i, x_j) = \Delta(x_j, x_i)$;

(3) $\Delta(x_i, x_n) \leq \Delta(x_i, x_j) + \Delta(x_j, x_n)$ 。

则 $\langle U, \Delta \rangle$ 为距离空间或度量空间, Δ 为 U 上的距离函数。本次研究使用汉明距离作为距离函数,“ \oplus ”为异或运算符,则对任意 $x_i, x_j \in U$, $x_i = (pre_{i1}, \dots, pre_{im})$ 和 $x_j = (pre_{j1}, \dots, pre_{jm})$ 的距离表示为:

$$\Delta(x_i, x_j) = \sum_{k=1}^m pre_{ik} \oplus pre_{jk} \quad (1)$$

对分类器池 C , 若属性子集 $B \subseteq C$, B 中包含 m

个分类器,则第 i 个验证样本 x_i 在 B 上的 δ_1 邻域可以表示为:

$$\delta_B(x_i) = \{x_j \mid \frac{\Delta(x_i, x_j)}{m} \leq \delta_1, x_j \in U\} \quad (2)$$

其中, δ_1 为邻域半径。通过式(1)、式(2)计算各样本邻域,得到分类器子集 $B \subseteq C$ 上的邻域关系 NR_B 。

定义 2 在邻域决策系统 $\langle U, NR_B, D \rangle$ 中, NR_B 为在分类器子集 $B \subseteq C$ 上的邻域关系,决策属性 D 将论域 U 中的所有样本划分为 P 个等价类: X_1, X_2, \dots, X_P 。则决策属性 D 关于分类器子集 B 的邻域下近似为:

$$\underline{NR_B}D = \{\underline{NR_B}X_1, \underline{NR_B}X_2, \dots, \underline{NR_B}X_P\} \quad (3)$$

其中, $\underline{NR_B}X_i$ 为邻域被 X_i 完全包含的对象的集合,表示为:

$$\underline{NR_B}X_i = \{x_i \mid \delta(x_i) \subseteq X_i, x_i \in U\} \quad (4)$$

图 3 为针对二分类问题的邻域粗糙集上下近似划分示意图,设 δ 为邻域半径,由式(4)分析可知,左侧 x_1 邻域内所有样本属于同一类,则 x_1 被列入 O 类下近似中。然而,右侧 x_2 邻域内的样本既有 O 类,也有 Δ 类, x_2 不能被划分到任何等价类的下近似。

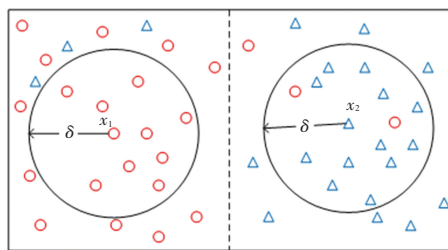


图 3 邻域粗糙集近似划分示意图

Fig. 3 Approximate partition of neighborhood rough sets

图 3 说明邻域粗糙集的近似划分方法对被错误标识的样本敏感度较高。因此,本文引入一致性阈值 δ_2 解决上述问题。假设 $P(L_j \mid \delta_B(x_i))$ 为 x_i 域内 $label_j$ 类的概率, n 是邻域内样本的数量, n_j 是第 j 类样本的数量。给定分类器子集 $B \subseteq C$, $x_i \in U$, $\delta_B(x_i)$ 是样本 x_i 在属性子集 B 上的邻域,则 x_i 的邻域决策函数可以定义为:

$$ND(x_i) = L_l$$

$$P(L_l \mid \delta_B(x_i)) = \max_j P(L_j \mid \delta_B(x_i)) = \max_j \frac{n_j}{n} \quad (5)$$

即 L_l 为 x_i 域内概率最大的样本类,当 $\frac{n_j}{n} \geq \delta_2$,

且 L_i 与样本 x_i 的决策属性 d_i 一致时,将 x_i 加入决策属性 d_i 所在等价类 X_i 的下近似,从而有效降低邻域粗糙集对错误表示样本干扰的敏感性。此时式(4)可重新定义为:

$$\underline{NR}_B X_i = \{x_i \mid ND(x_i) = d_i \text{ 且 } \frac{n_j}{n} \geq \delta_2, x_i \in U\} \quad (6)$$

1.2.3 对信息决策表的属性约简

不同分类器对集成预测系统具有不同影响,删除不重要的分类器能够降低数据维度,从而减小分类器的集成规模。设 red 为约简集, $c_k \in \{C - red\}$ 表示预选入约简集的分类器, D 表示样本的真实预测值。由式(6)得出属性子集 $\{red \cup c_k\}$ 约束下的下近似集,从而计算出下近似集中样本的总数 $|POS_{red \cup c_k}(D)|$, 则决策属性对选入的分类器序列的依赖度可表示为:

$$\gamma_{red \cup c_k}(D) = \frac{|POS_{red \cup c_k}(D)|}{|U|} \quad (7)$$

依赖度反映论域中能够被确切分类的样本比例。 $|U|$ 表示样本的总数,由粗糙集依赖度的性质可知,依赖度越大,则 D 对 $\{red \cup c_k\}$ 的依赖程度越强,也就是选出的分类器对预测结果影响越大。

分类器的重要度可由添加该分类器所引起的信息量的变化来度量,由此得出分类器 $c_k \in \{C - red\}$ 在已有分类器集合 red 的基础上相对决策属性 D 的重要度为:

$$\sigma(c_k, red, D) = \gamma_{red \cup c_k}(D) - \gamma_{red}(D) \quad (8)$$

将分类器的重要度排序,选出 $\sigma(c_k, red, D)$ 最大时对应的第 k 个分类器 c_k , 如果 $\max(\sigma(c_k, red, D)) > 0$, 则将 c_k 加入到约简集合 red , 参与最终的分类器集成。

1.2.4 基于邻域粗糙集的分类器选择算法

本文分类器选择算法,由空属性集合 $red = \emptyset$ 开始,选出分类器池 C 中属性重要度最大,即对当前系统最重要的分类器 c_k 添加到 red 中。其后的每次迭代,将原属性集合剩下的属性 $\{C - red\}$ 中最重要分类器添加到 red , 当属性的重要度 $\max(\sigma(c_k, red, D)) = 0$ 时,停止迭代。对表1信息决策表 S 中的条件属性约简,结果 red 即为最终选择的分类器。分类器选择过程如算法1所示。

算法1 基于邻域粗糙集的分类器选择算法

输入 $S = (U, C \cup D)$, δ_2 ; 这里, U 为所有验证集,条件属性 C 为分类器集合, D 为验证集的真实标签, δ_2 为一致性阈值

输出 S 的属性约简 red , red 中的集合即为选

出的分类器子集

初始化 $\emptyset \rightarrow red$

for each $c_k \in \{C - red\}$

临时决策表 $S' = \langle U, \{red \cup c_k\}, D \rangle$

for each $x_i \in U$

根据算法2计算样本 x_i 在属性子集 $\{red \cup c_k\}$ 的邻域 $\delta(x_i)$

由式(5)找出 $\delta(x_i)$ 中样本出现最多的类别 $label_l$,

由式(6), if $L_i = d_i$, 且 $\frac{n_j}{n} \geq \delta_2$

$POS_{red \cup c_k}(D) \cup x_i \rightarrow POS_{red \cup c_k}(D)$

end if

end for

由式(7)计算选入的分类器依赖度 $\gamma_{red \cup c_k}(D)$

由式(8)计算分类器 c_k 的重要度 $\sigma(c_k, red, D)$

找出 $\max(\sigma(c_k, red, D))$ 对应的分类器 c_k

if $\max(\sigma(c_k, red, D)) > 0$

$\{red \cup c_k\} \rightarrow red$

else

return red

end if

end for

算法2 样本邻域计算算法

输入 决策系统 $S = (U, C \cup D)$, 邻域半径 δ_1

输出 样本 x_i 的邻域 $\delta(x_i)$

for each $x_i \in U$

令 $\delta(x_i) = \emptyset$;

$\forall x_j \in U$, 由式(1)计算2个样本之间的距离 $dis(x_i, x_j)$;

由式(2), 若 $dis(x_i, x_j) \leq \delta_1$, 则 $\delta(x_i) = \delta(x_i) \cup \{x_j\}$

end for

1.3 分类器集成

模型求解得到分类器子集 $red = \{c_1, c_2, \dots, c_m\}$, m 为分类器子集数量,使用大多数投票法将选出的分类器子集进行决策层融合,获得最大投票数的标签作为最终的预测类别 $mv(x)$ 。输入测试样本 x , 大多数投票公式如下:

$$mv(x) = \underset{l \in L}{\operatorname{argmax}} \sum_{i=1}^m I(c_i(x) = l) \quad (9)$$

其中, $I(\cdot)$ 是一个判别函数; L 表示所有样本类别; $c_i(x)$ 是第 i 个分类器对给出的观测样本 x 的预测标签。

2 实验结果及分析

2.1 实验准备

2.1.1 数据集与实验环境

实验采用 FER2013^[25]、JAFPE^[26]、CK+^[27] 三个经典人脸表情数据集进行实验。其中, FER2013 均为灰度图像, 包含 28 709 张测试集图像, 公共验证集和私有验证集图像各有 3 589 张。JAFPE 选取 10 名日本女学生在实验环境下根据指示做出的 213 张表情图像, 每个人包括愤怒等 7 种基本表情。CK+ 数据集由 123 个类别、593 个图形序列组成, 其中 326 个图形序列带有表情标签, 本实验随机选取 58 个带标记的序列的前 3 张图片组成 CK+ 的中性表情图片。

实验过程将所有表情统一成大小为 48 * 48 的单通道图, 并将数据集划分为训练集、测试集和验证集三个部分。FER2013 训练集数据量较大, 实验中作为训练集, 用于训练分类器, 获得模型参数。公共验证集作为验证集 (fer_valid), 用于验证分类器效果, 进行模型参数微调。测试集包括 FER2013 中的私有验证集 (FER)、CK+ 和 JAFPE, 用于测试整个识别系统的性能。图 4 为 3 个数据集的部分表情示例。图 4 (a) ~ (c) 中, 从左至右依次为: anger, disgust, fear, happy, sadness, surprise, neutral。

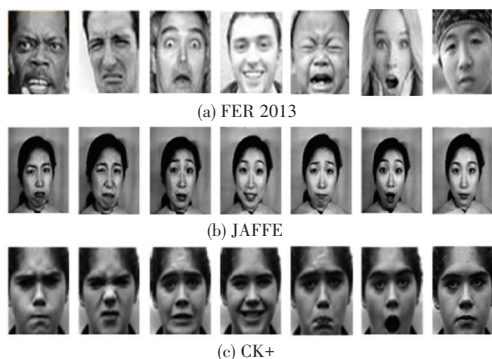


图 4 所有数据集图片示例

Fig. 4 Sample images of all datasets

2.1.2 实验评价标准

准确率 (Accuracy, Acc) 常作为衡量分类效果的标准。设 TP 和 FP 分别表示被正确分类和被错误分类的样本数量, Total 为总的样本数量, 则 Acc 计算公式如下:

$$Acc = \frac{TP}{TP + FP} = \frac{TP}{Total} \quad (10)$$

2.1.3 训练细节

实验基于 TensorFlow 框架, 主要通过调整超参数、

网络结构共生成 31 个卷积神经网络分类器形成分类器池。所更改的神经网络参数类型及范围见表 2。

表 2 卷积神经网络参数设置

Tab. 2 Parameter settings of convolutional neural network

参数类型	参数范围	参数类型	参数范围
网络层数	3~8	学习率	0.000 01~0.01
卷积核	[3,3]到[7,7]	激活函数	Relu, leaky_relu, Sigmoid
池化层核	[2,2],[3,3]	特征图数量	32, 60, 64, 100, 300, 128
池化步幅	2,3	边缘填充	SAME, VALID

图 5 为生成的分类器对不同数据集的预测结果的箱线叠加散点图, 用于显示基分类器对不同数据集预测准确率分布情况。箱线图由上至下各横向依次表示基分类器预测准确率的最大值、上四分位数、中位数、下四分位数和最小值。箱线图中处于最大值和最小值之外的数据视为异常值, 说明该分类器预测效果不具有代表性。fer_valid 和 FER 均来自于同一数据库的不同划分, 且样本总体数量一致, 由此推测会出现类似的预测效果。从图 5 可得到证明, 分类器对 2 个数据集的预测结果的准确率均主要分布在 0.45~0.6 之间。CK+ 上测试结果相对较差, 准确率主要分布在 0.4~0.53 之间, 而 JAFPE 数据库采集的亚洲人脸表情, 和训练数据集相比由于种族不同, 面部解析的方式有所不同, 从而预测效果不佳, 准确率主要分布在 0.3~0.45 之间。

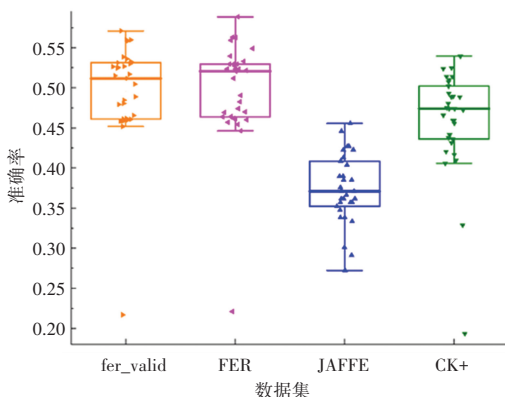


图 5 基分类器对各数据集预测结果的箱线叠加散点图

Fig. 5 Box line superposition scatter diagram of predicted results of each data by base classifier

2.2 实验结果分析

根据式 (10) 计算测试集的预测准确率作为集成算法性能评估的标准。利用邻域关系构建的粗糙集, 下近似元素的个数可以衡量相对重要度, 依据属性重要度准则进行知识约简, 采用前向贪心搜索策略依次选出分类器。邻域半径过大, 会导致大量不同类别的样本划分到同一邻域; 一致性阈值过小, 会失去对邻

域样本的约束作用,反之,会输出空的特征子集。

为了说明所提出的基于邻域粗糙集分类器选择的表情识别模型的有效性,设置 9 个不同的邻域半径 δ_1 , 分别为 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5 及 6 个不同的一致性阈值 δ_2 , 分别为 0.55, 0.6, 0.65, 0.7, 0.75, 0.8。

邻域半径 δ_1 随着不同一致阈值 δ_2 变化,在各数据集上的集成效果如图 6 所示。

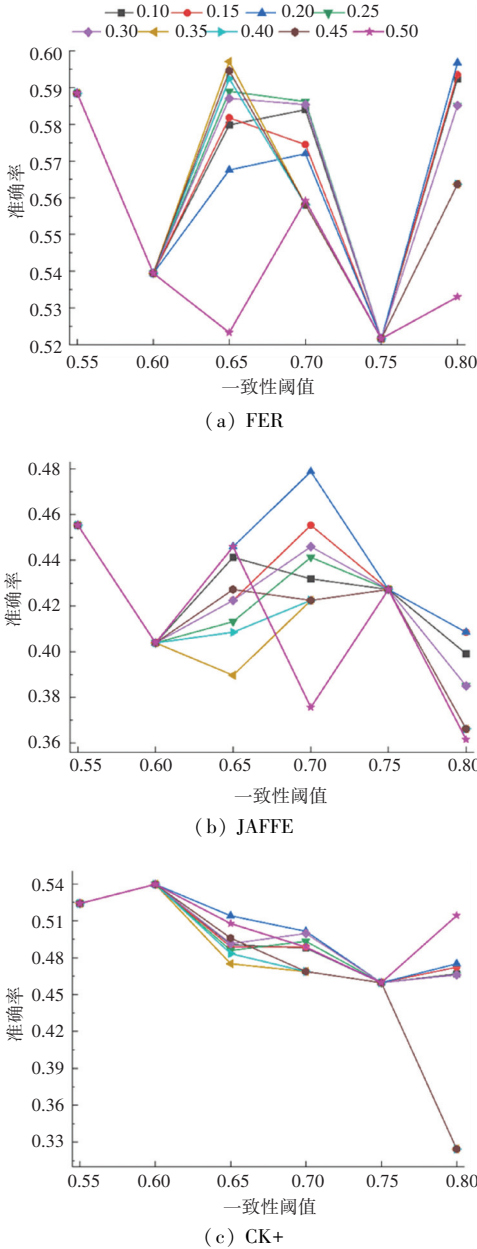


图 6 邻域半径 δ_1 随一致性阈值 δ_2 变化的选择集成效果

Fig. 6 Influence of domain radius δ_1 with consistency threshold δ_2 on the selected integration effect of each data set

图 6(a) 展示了随着一致性阈值变化,在不同邻域半径下选择后集成的效果总体呈现相同的趋势,邻域半径 $\delta_1 = 0.5$ 时预测效果大幅下降后再上升,这是

由于邻域半径设置过大可能导致较多不同类别的样本划分到同一邻域内,而较小的一致性阈值不能很好地取起对邻域决策的约束作用;同理,由图 6(b)、(c) 可看出,邻域半径 $\delta_1 = 0.5$ 时预测效果与其它不同邻域半径的预测趋势不一致,说明邻域半径设置过大,会增加邻域样本的不一致性。综合图 6 在 3 个不同的数据集上的集成效果,总体上,一致性阈值 δ_2 取 0.65 和 0.7 时,对不同的邻域半径下集成效果影响较为突出,说明只有选取合适的一致性阈值,才能更好地对样本进行正确分类,有效度量不同分类器的重要程度。

2.3 与其它方法比较

为了验证本文集成剪枝方法的有效性,使用经典表情数据集,选用多个不同剪枝算法与本文所提模型进行对比。通过多次实验,选择每种算法在各个数据集上表现效果最好的一组数据,得到的结果见表 3。

表 3 不同剪枝算法在各个数据集上的准确率

Tab. 3 The accuracy of different pruning algorithms on each data set

算法	数据集		
	FER	JAFFE	CK+
Baseline	0.577	0.399	0.461
SEFSBRS ^[28]	0.577	0.409	0.472
互信息 ^[29]	0.583	0.484	0.473
Kappa ^[30]	0.572	0.432	0.465
OO ^[31]	0.608	0.455	0.524
Non-Consistent ^[32]	0.589	0.455	0.524
本文算法	0.598	0.479	0.540

由表 3 可知,将所有分类器进行大多数投票集成的结果作为基线,各集成剪枝算法的准确率在大部分情况下高于将所有分类器集成。和基线相比,本文的集成剪枝算法在 FER 数据集上准确率提升了 2.1%,在 JAFFE 数据集上提升了 8%,在 CK+ 上提升了 7.9%,且本文提出的集成剪枝算法在 JAFFE 和 CK+ 数据集上都获得比其它集成剪枝算法更高的准确率,在 FER 上准确率比文献[32]方法略低,原因是采用大多数投票法集成,当存在多个相同票数的类型时,随机选择其中一个类型,会出现误判的情况,进而导致准确率略微下降。

3 结束语

本文结合粗糙集和集成学习的思想,研究重点在于对多个卷积神经网络模型进行集成剪枝,拓宽粗糙集的运用领域,并将提出的集成剪枝算法运用到人脸表情识别领域,在提高人脸表情识别率的同时,减小分类器的集成规模,进而提升识别系统的稳

定性的效果。由于集成剪枝算法准确率一定程度上取决于基分类器预测效果,因此,后续工作将使用先进的深度学习网络构建分类器池,再进行集成剪枝操作,以提高人脸表情识别的准确率。

参考文献

- [1] LIU Sijia, MA Ruihua, LUO Yang, et al. Facial expression recognition and ReHo analysis in major depressive disorder[J]. *Frontiers in Psychology*, 2021, 12.
- [2] YAN Lingyu, SHENG Menghan, WANG Chunzhi, et al. Hybrid neural networks based facial expression recognition for smart city[J]. *Multimedia Tools and Applications*, 2021.
- [3] ZHANG Xi, ZHANG Feifei, XU Changsheng. Joint expression synthesis and representation learning for facial expression recognition[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022(3): 1681-1695.
- [4] GUO Yifei, HUANG Jian, XIONG Mingfu, et al. Facial expressions recognition with multi-region divided attention networks for smart education cloud applications [J]. *Neurocomputing*, 2022, 493: 119-128.
- [5] SUN Z, CHIONG R, HU Z, et al. A dynamic constraint representation approach based on cross-domain dictionary learning for expression recognition[J]. *Journal of Visual Communication and Image Representation*, 2022, 85: 103458.
- [6] LI Jing, JIN Kan, ZHOU Dalin, et al. Attention mechanism-based CNN for facial expression recognition [J]. *Neurocomputing*, 2020, 411: 340-350.
- [7] ZHU F, CAI J, HE M, et al. Channel and band attention embedded 3D CNN for model development of hyperspectral image in object-scale analysis [J]. *Chemometrics and Intelligent Laboratory Systems*, 2022, 224: 104537.
- [8] GONÇALVES C B, SOUZA J R, FERNANDES H. CNN architecture optimization using bio-inspired algorithms for breast cancer detection in infrared images[J]. *Computers in Biology and Medicine*, 2022, 142: 105205.
- [9] GE Huilin, ZHU Zhiyu, DAI Yuewei, et al. Facial expression recognition based on deep learning[J]. *Computer Methods and Programs in Biomedicine*, 2022, 215: 106621.
- [10] RENDA A, BARSACCHI M, BECHINI A, et al. Comparing ensemble strategies for deep learning: An application to facial expression recognition [J]. *Expert Systems with Applications*, 2019, 136: 1-11.
- [11] TSAI K, TSAI Y, LEE Y, et al. Frontalization and adaptive exponential ensemble rule for deep-learning-based facial expression recognition system [J]. *Signal Processing: Image Communication*, 2021, 96: 116321.
- [12] ASAFUDDOULA M, VERMA B, ZHANG M. A divide-and-conquer-based ensemble classifier learning by means of many-objective optimization [J]. *IEEE Transactions on Evolutionary Computation*, 2018, 22(5): 762-777.
- [13] WANG Yuchen, Yang Yang, LIU Yunxia, et al. A recursive ensemble learning approach with noisy labels or unlabeled data [J]. *IEEE Access*, 2019, 7: 36459-36470.
- [14] JIANG Weili, CHEN Zhenhua, XIANG Yan, et al. SSEM: A novel self-adaptive stacking ensemble model for classification[J]. *IEEE Access*, 2019, 7: 120337-120349.
- [15] LI Danyang, WEN Guihua, LI Xu, et al. Graph-based dynamic ensemble pruning for facial expression recognition [J]. *Applied Intelligence*, 2019, 49(9): 3188-3206.
- [16] 贾澎涛, 李阳. 基于选择性集成分类器的面部表情识别研究 [J]. *计算机应用研究*, 2017, 34(12): 3825-3827.
- [17] CHELLY D Z, ZARGES C. A detailed study of the distributed rough set based locality sensitive hashing feature selection technique[J]. *Fundamenta Informaticae*. 2021, 182(2): 111-179.
- [18] YAO Jingtao, MEDINA J, ZHANG Yan, et al. Formal concept analysis, rough sets, and three-way decisions [J]. *International Journal of Approximate Reasoning*, 2022, 140: 1-6.
- [19] ZABIHIMAYVAN M, DORAN D. Fuzzy rough set feature selection to enhance phishing attack detection [C]//2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). New Orleans, LA, USA: IEEE, 2019: 1-6.
- [20] GAO Wei, WEN Jingxin, JIANG Nan, et al. A study of data fusion based on combining rough set with BP neural network [C]// 2009 Ninth International Conference on Hybrid Intelligent Systems. Shenyang, China: IEEE, 2009.
- [21] 冯林, 李聪, 沈莉. 基于邻域粗糙集与量子遗传算法的人脸表情特征选择方法[J]. *合肥工业大学学报(自然科学版)*, 2013, 36(01): 39-42.
- [22] 段丽, 张建明. 基于粗糙集的表情特征选择[J]. *计算机工程与应用*, 2010, 46(32): 177-179.
- [23] HU Meng, TSANG E C C, GUO Yanting, et al. Attribute reduction based on overlap degree and k-nearest-neighbor rough sets in decision information systems [J]. *Information Sciences*, 2022, 584: 301-324.
- [24] BAI Juncheng, SUN Bingzhen, CHU Xiaoli, et al. Neighborhood rough set-based multi-attribute prediction approach and its application of gout patients [J]. *Applied Soft Computing*, 2022, 114: 108127.
- [25] GOODFELLOW I J, ERHAN D, LUC C P, et al. Challenges in representation learning: A report on three machine learning contests [J]. *Neural Networks*, 2015, 64: 59-63.
- [26] LYONS M, AKAMATSU S, KAMACHI M, et al. Coding facial expressions with Gabor wavelets [C]//Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. Nara, Japan: IEEE, 1998: 200-205.
- [27] COHN L P, KANADE J F, SARAGIH T, et al. The extended Cohn-Kanade ataset (CK+): A complete dataset for action unit and emotion-specified expression [C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops. San Francisco, CA: IEEE, 2010: 94-101.
- [28] 杨勇. 基于粗糙集理论的表情识别研究[D]. 成都: 西南交通大学, 2009.
- [29] 杨素妨, 曾红春. 融合多特征与互信息选择集成多核极限学习机的影像分类方法[J]. *遥感信息*, 2021, 36(01): 56-60.
- [30] 孙涛. 集成学习结构多样性研究[D]. 南京: 南京大学, 2018.
- [31] MARTINEZ-MUNOZ G, SUÁREZ A. Pruning in ordered bagging ensembles [C]//Proceedings of the 23rd international conference on Machine learning (ICML '06). New York: ACM, 2006: 609-616.
- [32] LUDMILA I K, CHRISTOPHER J W. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy [J]. *Machine Learning*, 2003, 51: 181-207.