

# 基于灰度共生矩阵和精度高斯支持向量机的 中国手语手指语识别

蒋贤维<sup>1</sup> 张妙娴<sup>2</sup> 朱兆松<sup>1</sup>

(南京特殊教育师范学院数学与信息科学学院 南京 210038)<sup>1</sup>

(南开大学周恩来政府管理学院 天津 300071)<sup>2</sup>

**摘要** 手语识别是打破聋人和健听人之间交流障碍的有效途径。中国手语一般可以分为手势语和手指语,手势语因为地区性和个体差异性导致种类和变化繁多,识别相对困难,所以需要不断学习和训练;手指语通过拼音字母的表现形式给出结果,表达具有确定性,尤其在姓名、特殊含义、抽象表达方面效果明显。手语识别中,大部分的研究主要聚焦于某种手势,围绕手形、方向、位置和运动轨迹等关键特征,并结合某些学习算法来提升识别的准确率,然而最基本可靠的手指语识别却往往被忽略。为此,文中提出了一种基于灰度共生矩阵(GLCM)和精度高斯支持向量机(FGSVM)的方法来更准确有效地识别中国手语手指语。首先构建手指语数据集,即通过数码相机直接获取手指语图像或者从视频中选取关键帧作为手语图像素材,然后将手形从图像背景中分割出来,把每个图像调整为  $N \times N$  的特定尺寸并转换为灰度图像;其次是提取特征,即对灰度图像中强度值的数量进行降维,同时创建对应的灰度共生矩阵,通过调整像素间的距离和角度等参数来获取增强的数据特征;最后,将提取的图像的特征数据提交到精度高斯支持向量机分类器中,进行 10 倍交叉验证和分类测试。对 30 种类别的 510 个中国手语手指语图像样本的实验结果表明,基于 GLCM-FGSVM 的分类准确率最高可达到 92.7%,可以认为该方法在中国手语手指语分类方面卓有成效。

**关键词** 手语识别,中国手语,手指语,灰度共生矩阵,精度高斯支持向量机

**中图分类号** TP391.41 **文献标识码** A

## Recognition of Chinese Finger Sign Language Based on Gray Level Co-occurrence Matrix and Fine Gaussian Support Vector Machine

JIANG Xian-wei<sup>1</sup> ZHANG Miao-xian<sup>2</sup> ZHU Zhao-song<sup>1</sup>

(School of Mathematics and Information Science, Nanjing Normal University of Special Education, Nanjing 210038, China)<sup>1</sup>

(Zhou Enlai School of Government, Nankai University, Tianjin 300071, China)<sup>2</sup>

**Abstract** Sign language recognition is an effective way to break the barriers between communication between deaf and hearing people. Generally, Chinese sign language can be divided into gesture language and finger language. Regional and individual differences lead to a wide variety, therefore gesture language recognition is relatively difficult, which requires constant learning and training. The finger language gives the result through the expression of the Chinese pinyin letters, which is deterministic, especially in terms of name, special meaning, and abstract expression. Most of the researches in sign language recognition concentrate on a certain gesture, focusing on key features such as hand shape, direction, position and motion trajectory, and combine some learning algorithms to improve the recognition accuracy, but neglect the most basic and reliable finger recognition. To this end, an effective method using gray level co-occurrence matrix (GLCM) and fine Gaussian support vector machine (FGSVM) was proposed to solve the problem of identifying Chinese finger sign language more accurately and effectively. The research method is as follows. Firstly, the finger sign language data set was constructed. The finger language image was directly obtained by the digital camera or got from the key frame of the video, meanwhile the hand shape was segmented from the image, and each image was adjusted to  $N \times N$  specific size and converted to grayscale images. Secondly, feature extraction was performed to reduce the dimension of the intensity values in the grayscale image, and at the same time, the corresponding gray level co-occurrence matrix was created, and the enhanced data features were obtained by adjusting the parameters of inter-pixel distance and angle. Finally, the extracted image feature data were submitted to the fine Gaussian support vector machine classifier based on the 10-fold cross-validation classification. Experiments on 510 Chinese finger sign language image samples from 30 categories show that the classification accuracy based on GLCM-FGSVM is up to 92.7%, and this method can be considered

as effective approach in Chinese finger sign language classification.

**Keywords** Sign language recognition, Chinese sign language, Finger sign language, Gray level co-occurrence matrix, Fine Gaussian support vector machine

## 1 引言

当前全世界约有数亿人患有听力障碍,根据第二次全国残疾人抽样调查,在中国约有 2780 万听障人士,这是个数量庞大的群体。由于没有更好、更方便的工具,他们中的大多数人总是无法清楚地向社会表达诉求。例如,在一起案件调查中,他们往往很难向警察传达明确和真实的意思,这严重影响了他们的权益。可见,聋人群体与社会健康人群之间的确存在沟通障碍。手语识别被认为是一种破冰这种障碍的创新,它可以通过将手语翻译成文本或语音等来弥合聋人和健听人之间的间隙,这有利于促进社会和谐发展。作为当下研究的一个热点,手语识别被视为一个涵盖了图像处理、模式识别和人工智能等多学科的研究领域<sup>[1]</sup>,受到了越来越多的关注和重视。

手语指的是“人们在聋人环境中使用手的指式、动作、位置和朝向,配合面部表情,按照一定的语法规则来表达特定意思的交际工具”<sup>[2]</sup>。中国手语一般分为手势语和手指语两大类,前者强调表达的过程,往往同时使用双手,辅以面部表情,强调情景,相对比较复杂;后者主要强调 30 个基本指语,通过它们的组合来表达拼音或一些特殊的意义,相对比较单一和确定。因此,手语识别中常用的特征表示有手形、方向、位置和运动轨迹<sup>[3]</sup>等,大部分的研究和实验也是主要围绕这些关键特征展开。如 Pan 等<sup>[4]</sup>基于手部分割的自适应肤色建模方法来更准确地导出手部轮廓。司阳等<sup>[5]</sup>基于彩色-深度图像的手语识别算法来解决传统手语识别方法中普遍存在的背景干扰、特征提取不完整及识别率低等问题。许天然等<sup>[6]</sup>采用 Canny 算子提取图像边缘并进行图形填充,基于模板匹配算法检测白背景下的 26 种手势。林水强等<sup>[7]</sup>的分割算法结合多序列背景模型、肤色跟踪和几何形状,并基于决策树分类识别来自动态帧方法获取的多个几何特征。郑津津等<sup>[8]</sup>选用结构特征进行手语识别等。同时,识别过程中,许多学习方法和分类算法被广泛应用,如决策树和随机森林<sup>[9]</sup>,隐马尔可夫模型(HMM),动态时间扭曲算法<sup>[10]</sup>,机器学习方法包括 k-最近邻和支持向量机<sup>[11-12]</sup>,卷积神经网络(CNNs)<sup>[13-14]</sup>,反向传播神经网络算法<sup>[15]</sup>,极限学习机(ELM)<sup>[16]</sup>,递归神经网络(RNN)<sup>[17]</sup>等。

手语识别方法研究的最终目的是高效、快捷地实现自动手语识别,打破聋人和健听人之间交流的障碍。本文主要聚焦中国手语手指语的识别工作。手指语虽然只有 30 个手指字母分类(包含 26 个基本拼音字母和 4 个翘舌发音),但其意义重大。因为中国手语地区性差异较大,同一个意思在不同地区有各种表示方式,所以手势语经常会出现同一个意思有不同的表示。即使我们国家已经出版了《中国手语》上、下册标准手语,但在实际使用过程中聋人并不广泛接受和使用。刘艳虹、顾定倩、程黎等《我国手语使用状况的调查研究》<sup>[18]</sup>显示,59.6%的成年聋人和 33%的聋哑学生选择使用本地手

语,77.3%的教师认为本省内各特殊学校的教师之间使用的手语“有些差异”或“差异明显”,76.9%的聋生和成年聋人认为学校老师和工作人员手语不一致。因此,作为唯一的确定性表达,手指语的优势突显出来。手指语通过拼音离散成字母的表现形式给出结果,首先在定义其他手势含义的过程中起到决定作用,其次在姓名、特殊含义、抽象表达方面效果尤其明显。本文的主要工作是获得图像中必要的手部区域和手形特征(直接获取中国手语图像或截取视频关键帧),分析深度图像中手区域的属性并进行分类和识别,这可以被视为我们之前语料库建设<sup>[19]</sup>工作的延伸和丰富。文中首先运用灰度共生矩阵(GLCM)<sup>[20]</sup>的方法从中国手语手指语图像中获取特征,它可以大大减小提取特征之间的相关性,并降低图像识别过程中的难度,然后使用精度高斯支持向量机(FGS-VM)从 510 个中国手语手指语样本图片中分类识别手指语图像。实验表明,该方法在识别中国手语手指语方面展现出良好的预期。

本文第 2 节主要介绍了数据预处理和特征提取的过程,着重介绍了基于 GLCM 的特征提取、约简以及增强优化;第 3 节阐述了实验方法,介绍了 10 倍分层交叉验证技术和精度高斯支持向量机分类器,并给出了设计方法的整体流程图;第 4 节给出了实验环境和结果,并展开相应的讨论;最后总结全文并展望未来。

## 2 预处理及特征提取

### 2.1 预处理

鉴于现代手机的高分辨率和实验场景的方便快捷,也为了更逼近真实情景,我们通过手机拍摄选取了 17 个不同样本的 510 张中国手指语图片(1080×1080 大小),每个样本对应包含 26 个基本字母和 4 个翘舌音共计 30 个类别(见图 1)。然后使用 Adobe Photoshop CS 软件手动分割每个中国手语手指语图片,将图中手形区域剥离出来,并将图像大小调整设置为 256×256,同时归一化设置背景色值为 RGB(0,0,0),最后保存为基本不损失图像信息的 TIF 压缩格式(见图 2)。



图 1 一个样本的 30 个分类源图



图 2 剥离的手指语手形图

通过以上操作,经过预处理的图像具有相同的大小和背

景,且不影响测试结果。为了后续创建灰度共生矩阵,我们通过工具软件 Matlab 程序将彩色图像统一转换成灰度图像,其伪代码如下:

```
Begin
Creatmaindir;%创建实验根目录
Get subdir;%获取子目录
for j=1:length(subdir)%遍历每个子目录
    if(有效子目录) then
        Get subdirpath;%获取当前文件夹下所有 tif 格式图片的路径
    endif
endfor
for i=1:length(subdirpath)%遍历所有 tif 图片文件
    Getimgpath;%获取当前图片的路径
    Rgb2gray;%转换为灰度图像
    imshow();%显示图像,检验效果
endfor
End
```

选取两个已经转换为灰度图像的手语示例,如图 3 所示。其中,(a)图像表示字母 m,(b)图像表示字母 x。

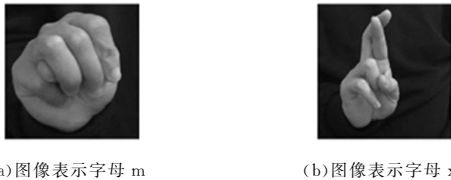


图 3 转换为灰度图像的手语手形图

## 2.2 特征提取

灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)是一种考量像素空间关系的有效统计方法,也被称为灰度级空间依赖性矩阵,是一种把图像转化为矩阵数据的常用方法。本文方法使用 Matlab 中的 graycomatrix 函数,通过计算具有灰度级强度值  $i$  和  $j$  的两个像素基于某个特定角度  $A$  (可以使用“偏移”参数  $offset$  来指定当前像素和其他像素的空间关系)及相邻距离  $D$  (同样可以通过参数  $offset$  来设置)出现的频率,来创建灰度共生矩阵,从而实现图像特征的数值化。例如,当前角度为  $0$  度时,灰度共生矩阵中的每个元素  $(i, j)$  表示指定具有值  $i$  的像素水平地与具有值  $j$  的像素相邻的次数;一般默认距离为 1,根据需要可以通过修改参数  $D$  的值来调节两个像素间的距离。

本实验中每个预处理图像在转换为灰度图像后的矢量空间具有  $256 \times 256 \times 1 = 65526$  个维度,相比原图已经实现了降维。在此基础上,我们仍然可以通过 graycomatrix 函数的缩放功能将强度范围的数量从 256 降到 8 来减小图像尺寸,即现在每个图像的矢量空间被降到了  $8 \times 8 \times 1 = 64$  维,并产生了一个对应的灰度共生矩阵。由此可见,灰度共生矩阵的大小与灰度级的数量息息相关。最后,汇总所有图像对应的灰度共生矩阵,针对 510 个样本图像,我们将得到一个大小为  $510 \times 65$  (含特征响应值,见图 4)的特征矩阵  $g lcm510$ 。主要代码如下:

```
...
glcms=graycomatrix(I,'offset',[0 0 0 0]);%I 为当前待处理图像,offset=[0 D;-D D;-D 0;-D -D]
```

```
temp=reshape(glcms,[1,64]);%转换每个  $8 \times 8$  的灰度共生矩阵为 1 行 64 列向量
```

```
temp=horzcat(temp,i);%基于分类增加特征响应值
```

```
g lcm510=[g lcm510;temp];%生成一个最后的特征矩阵
```

```
...
```

510x65 double		55	56	57	58	59	60	61	62	63	64	65	
495	1812	0	0	0	0	0	0	0	0	0	26	32	16
496	2001	23	0	0	0	0	0	0	0	0	81	282	17
497	3914	77	0	0	0	0	0	0	0	0	81	282	17
498	2845	179	0	0	6	10	7	31	267	2846	18		
499	3171	64	0	0	0	1	1	5	71	1059	19		
500	2692	94	0	2	0	1	0	3	123	1236	20		
501	4266	127	1	1	2	1	3	5	134	468	21		
502	3448	65	0	0	0	0	0	0	67	95	22		
503	2647	291	0	1	2	3	1	5	359	4729	23		
504	3314	229	0	1	3	2	9	11	252	1312	24		
505	2112	30	0	0	0	0	2	9	19	44	25		
506	2214	166	0	0	0	4	13	16	158	850	26		
507	2545	30	0	0	0	1	6	1	27	43	27		
508	1901	178	0	0	3	12	7	12	175	952	28		
509	3488	114	0	0	1	1	2	4	111	788	29		
510	3971	148	0	0	2	1	2	1	295	1737	30		

图 4 510 个图像对应的灰度共生矩阵  $g lcm510$

在默认情况下生成的图像灰度共生矩阵还不能满足高精度识别的需求,为了获得更丰富的图像特征,我们可以基于不同角度和不同距离取值和计算,并选取多个特征值的均值或范围矩阵作为增强的图像特征,以提高分类识别率。在函数 graycomatrix( $I$ , 'offset',  $value$ ) 中,参数  $offset$  决定了当前选择的感兴趣像素点和它邻近像素的距离和角度,  $offset$  的值对应矩阵  $[0 D; -D D; -D 0; -D -D]$ , 其中  $D$  表示距离,常用值为  $1 \sim 4$ 。  $[0 D]$  表示水平方向,  $[-D D]$  表示  $45^\circ$  方向,  $[-D 0]$  表示角度为  $90^\circ$ ,  $[-D -D]$  表示角度为  $135^\circ$ 。因此,我们可以用以下公式来表示当前图像对应的 4 个角度的特征值:

$$Feature_a = glcms(\theta=0^\circ, D)$$

$$Feature_b = glcms(\theta=45^\circ, D)$$

$$Feature_c = glcms(\theta=90^\circ, D)$$

$$Feature_d = glcms(\theta=135^\circ, D)$$

在此基础上,我们还可以用以下两个公式来求像素间距离  $D$  的灰度共生矩阵的平均值矩阵( $g lcm_m$ )和灰度共生矩阵的范围矩阵( $g lcm_r$ ):

$$g lcm_m = ((Feature_a + Feature_b + Feature_c + Feature_d) / 4)$$

$$g lcm_r = \text{Max}(Feature_a, Feature_b, Feature_c, Feature_d) - \text{Min}(Feature_a, Feature_b, Feature_c, Feature_d)$$

经过增强处理后的特征矩阵数据更合理,实验效果准确。尽管灰度共生矩阵提供了关于图像的灰度方向、间隔和变化的信息,但是它不能直接提供差异纹理的属性。因此,如有必要,可以计算在灰度共生矩阵的基础上定量描述纹理特征的统计特征。通常使用纹理特征统计属性的方差和相关性,以下是它们对应的计算公式:

$$Variance = \sum_{i=0}^{quant_i} \sum_{j=0}^{quant_j} p(i, j) * (i - Mean)^2$$

$$Correlation = \frac{\sum_{i=0}^{quant_i} \sum_{j=0}^{quant_j} (i - Mean) * (j - Mean) * p(i, j)^2}{Variance}$$

## 3 实验方法

### 3.1 分层交叉验证

交叉验证方法通常被用于评估分类器的统计相关性,其中  $K$  倍交叉验证由于简单易用而被广泛应用。它同时使用所有的数据进行训练和验证,其工作机制如下:创建整个数据集的  $K$  倍分区,选择  $K-1$  倍数据进行训练,使用剩余的进行

验证,重复  $K$  次,最后平均  $K$  次实验的错误率。分层  $K$  倍交叉验证可以解决某些折叠可能与其他折叠有完全不同分布情况的问题。

另一个要考虑的重要问题是确定折叠次数即  $K$  值。如果  $K$  值设置得过大,真实误差率和估计值的偏差将很小,但估计量的方差将很大,且计算十分耗时;反之, $K$  值过小将导致估计量偏差很大。在本实验中,我们选择 10 倍交叉分类验证(见图 5)。这是最常用的交叉验证模式<sup>[21]</sup>,它既可以避免过度拟合,又可以实现样本外估算。所有程序均在 MATLAB R2018b 平台和 Classifier Learner 应用程序上完成测试。

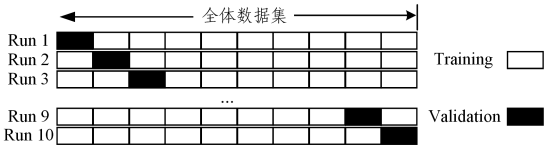


图 5 基于 10 倍交叉验证的流程图

### 3.2 FGSVM 分类识别

对于计算量不是很大且参数训练对硬件设备要求不苛刻的任务,支持向量机(Support Vector Machine, SVM)这样的浅层网络往往能取得优异成绩<sup>[22]</sup>,本文中我们选用了支持向量机作为分类器。它是一种监督学习算法,既可用于分类,又可以解决回归问题。在 SVM 算法中,我们基于每个特征数的值为特定坐标的值来绘制  $n$  维空间的数据点,最终目标是寻找能将数据分为两类的最佳超平面。因此,简单来说,支持向量指的是  $n$  维空间中观察样本的坐标, SVM 即是将样本分为两类的最佳超平面。最佳超平面被用来分离基于类成员的数据集。训练分类器的目标就是找到最大化类边界和训练模式之间边界的值。当无法找到合适的最佳超平面时,我们还可以通过核函数把输入数据集转换为更高维度的空间来尝试解决。本实验中我们通过支持向量机分类取得了良好的效果,其中尤其出色的是精度高斯支持向量机(FGSVM),其分类准确率达到 92.7%。

相较于线性支持向量机以及决策树分类,精度高斯支持向量机效果极佳的原因是,一些二维空间中线性不可分的问题如果映射到四维空间后将变得线性可分,因此这也给我们提供了原始的解决线性不可分割问题的基本思路,即尝试转换为高维空间并使其线性可分。转换的关键部分是找到合适的映射方法或函数,一般称之为核函数。核函数的基本功能是接受两个低维空间中的向量,并在一定变换后计算高维空间中向量的内积值。高斯核函数就是一类常用函数,可以用于解决类似问题。这就解释了该实验中精度高斯支持向量机获得了最好结果的原因。

### 3.3 方法流程图

本文方法主要分为以下 3 个步骤,其流程图如图 6 所示。

**步骤 1 构建数据集。**获取 510 幅彩色样本图像,每幅图像的初始大小为  $1080 \times 1080$ ,颜色通道值为 3,其意义为代表 30 个基本手语中的某一个。图像预处理,从背景中分割出基本手形,同时设置背景颜色值为零,调整图像大小为  $256 \times 256$ ,保留颜色通道不变。

**步骤 2 特征提取。**将预处理后的彩色图像转为灰度图

像,基于 graycomatrix 函数降维并创建灰度共生矩阵。增强特征,在不同角度和距离获取均值和范围矩阵。

**步骤 3 训练及分类识别。**将所有样本数据在 10 倍交叉验证下提交精度高斯支持向量机训练及分类识别,分析其相关性能和参数特点。

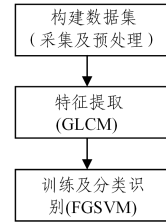


图 6 方法流程图

## 4 实验结果

### 4.1 实验环境

基于 MATLAB R2018(b)和分类学习者应用程序,按方法流程图步骤进行实验。所有的训练测试和计算是在个人计算机上进行的,其基本配置是 CPU:2.5GHz Core i7,内存 16 GB,安装了 Windows 10 操作系统。实验数据集为自定义的 510 幅预处理的中国手语手指语图片,包含 30 个分类(26 个基本拼音字母和 4 个翘舌发音),由随机选取的 17 位志愿者提供,通过高分辨率像素的手机摄像头采集。

### 4.2 实验结果

基于 10 倍交叉验证,我们对比了精度高斯支持向量机和决策树分类、线性支持向量机以及  $k$ -最近邻分类算法的 4 种情况,如表 1 所列。可以看出,基于灰度共生矩阵和精度高斯支持向量机的方式的准确率为 92.7%,明显领先于决策树的 38.2%;相对于线性支持向量机的 86.7% 和  $k$ -最近邻算法的 91.1%,本文方法的准确率依然很有优势。

表 1 基于 10 倍交叉验证的不同方法下的识别准确率

方法	准确率/%
灰度共生矩阵+决策树	38.20
灰度共生矩阵+线性支持向量机	86.70
灰度共生矩阵+k-最近邻	91.10
灰度共生矩阵+精度高斯支持向量机(本文)	92.70

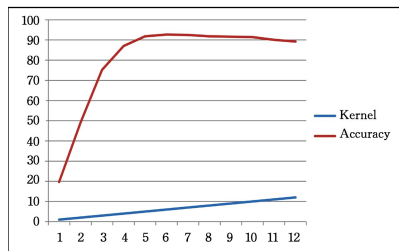
在灰度共生矩阵和精度高斯支持向量机的方式下,我们又分别测试了基于同一角度、不同像素距离下的灰度共生矩阵对应的准确率和同一像素距离、不同角度参数生成的灰度共生矩阵的对应准确率(见表 2),最后发现在距离值  $D=3$  和角度值为  $135^\circ$  参数上取得最好的效果 92.7%,并获得均值矩阵对应的准确率 91% 和范围矩阵对应的准确率 92.2%。

表 2 不同灰度共生矩阵对应的识别准确率

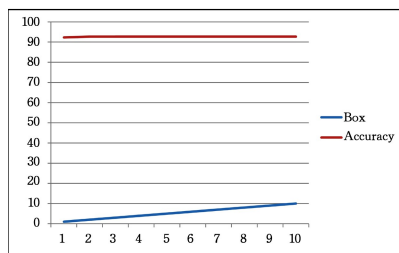
DC(距离值)	准确率/%	A( $^\circ$ )	准确率/%
1	87.8	0	91.6
2	91.2	45	89.9
3	92.7	90	90.2
4	92.7	135	92.7

在运用精度高斯支持向量机分类器时,我们发现精度高斯支持向量机在相同框式约束级别(box constraint level)情况下准确率随核规模(manual kernel scale)线性上升变换,但到某一极值后开始下降(见图 7 左图);但在相同的手动内核规模下,精度不会随着不同的框式约束级别而明显改变(见

图 7 右图)。实验找到的最佳框式约束级别为 2,手动内核比例模式下内核比例参数为 6,同时针对多类分类实施选择一对一的方法。



(a)



(b)

图 7 精度高斯支持向量机分类器的参数-准确率关联图

具有 92.7% 的最佳分类精度高斯支持向量机的混淆矩阵如图 8 所示。其中,矩阵的每一列对应目标类中的实例(真实类),矩阵的每行表示输出类中的实例(预测类),红色标出的是所有错误分类情况。

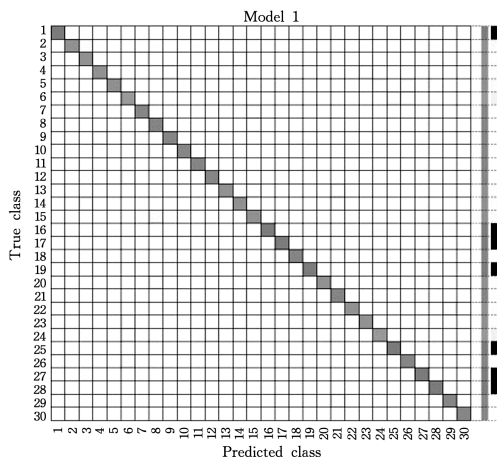


图 8 FGSVM 混淆矩阵(电子版为彩色)

图 9 给出了对手指语字母 t 分类识别的 ROC 图。类似地,通过对比筛选 ROC 值较低的分类(ROC 值为 0.76~0.82),我们发现其对应的手指语字母也是实际基于图像特征不容易被识别的或因相似性而容易混淆的情况,比如 t, m, sh, ng 等。这也再次证明了我们的基于灰度共生矩阵和精度高斯支持向量机分类方式的准确性和实用性。

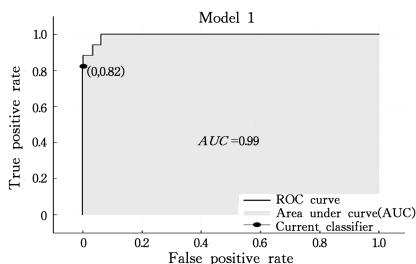


图 9 分类 t 的 ROC 图

**结束语** 本文介绍了中国手语手指语和手势语的区别,突出了手指语在研究和应用过程中的唯一确定性和重要性。在本研究中,我们提出了一种灰度共生矩阵和精度高斯支持向量机的中国手语手指语识别方法,并在像素点的不同角度和距离创建灰度矩阵测试分类效果,获得均值矩阵对应的准确率 91% 和范围矩阵对应的准确率 92.2%,并在距离值  $D=3$  和角度值为  $135^\circ$  参数上取得最好的分类识别效果 92.7%,实验结果显示这一方法是准确有效的。本方法的缺点是图像预处理部分未能实现自动化,训练和识别的速度还不够快。

因此,我们未来的研究工作将集中在以下几个方面:(1)通过计算机程序实现中国手语图像中关键区域的自动分割;(2)寻找并增加更多附加特征值,以提高分类准确度;(3)选择测试和对比其他高级分类器,如极端学习机、参数优化的核支持向量机<sup>[23]</sup>和卷积神经网络 CNN<sup>[24]</sup>等;(4)将我们的方法应用于其他相似的特殊教育和康复工程领域。

## 参考文献

- [1] ZHAN T. Pathological brain detection by artificial intelligence in magnetic resonance imaging scanning[J]. Progress in Electromagnetics Research, 2016, 156: 105-133.
- [2] 杨军辉. 中国手语和汉语双语教育初探[J]. 中国特殊教育, 2002(1): 33.
- [3] DIMITRIS METAXAS M D, CAROL N. Scalable ASL sign recognition using model-based machine learning and linguistically annotated corpora [C]// 8th Workshop on the Representation & Processing of Sign Languages; Involving the Language Community, Language Resources and Evaluation Conference 2018. Miyazaki, Japan, 2018: 1-5.
- [4] PAN T Y, YL L, YEH C W, et al. Sign language recognition in complex background scene based on adaptive skin colour modeling and support vector machine [J]. International Journal of Big Data Intelligence, 2018(5): 1-2.
- [5] 司阳,任松,肖秦琨,等. 基于彩色-深度图像的手语识别算法[J]. 科学技术与工程, 2018, 18(11): 104-109.
- [6] 许天然,吴焱,苏红旗. 基于移动终端的汉语手语识别技术研究[J]. 科技资讯, 2012, 19(1): 24-26.
- [7] 林水强,吴亚东,陈永辉. 基于几何特征的手势识别方法[J]. 计算机工程与设计, 2014, 35(2): 637-639.
- [8] 郑津津,徐士海. 基于结构特征提取的手语识别系统研究[D]. 合肥:中国科学技术大学, 2015.
- [9] SU R L, X C, CAO S, et al. Random Forest-Based Recognition of Isolated Sign Language Subwords Using Data from Accelerometers and Surface Electromyographic Sensors [J]. Journals Sensors, 2016, 16(1): 100.
- [10] WASHEF AHMED K C, SOMA M. Vision based Hand Gesture Recognition using Dynamic Time Warping for Indian Sign Language [C]// International Conference on Information Science (ICIS), 2016: 1-6.
- [11] BISWAS N, CHAKRABORTY S, MULLICK S S. A parameter independent fuzzy weighted k-Nearest neighbor classifier [J]. Pattern Recognition Letters, 2018, 101: 80-87.
- [12] HAO Z Y, ALIFU K, LI X H, et al. Chinese finger language recognition use CapsNet [J]. Application Research of Computers, 2018, 36: 216-228.

- [13] PAN C. Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU [J]. *Journal of Computational Science*, 2018, 28: 1-10.
- [14] YANG W B, YANG H C. Gesture recognition method based on convolutional neural network [J]. *Journal of Anhui Polytechnic University*, 2018, 33: 41-46.
- [15] WU J. Fruit classification by biogeography-based optimization and feedforward neural network [J]. *Expert Systems*, 2016, 33(3): 239-253.
- [16] LU S. Pathological Brain Detection in Magnetic Resonance Imaging Using Combined Features and Improved Extreme Learning Machines [J]. *Journal of Medical Imaging and Health Informatics*, 2018, 8: 1486-1490.
- [17] 刘艳虹, 顾定倩, 程黎, 等. 我国手语使用状况的调查研究 [J]. *语言文字应用*, 2013, 5(2): 35-41.
- [19] 徐鑫鑫, 黄元元, 胡作进. 连续复杂手语中关键动作的提取算法 [J]. *计算机科学*, 2018, 45(S2): 189-193.
- [20] MELLISA P A, JEKLIN H, SAKKA N. Mammograms Classification Using Gray-level Co-occurrence Matrix and Radial Basis Function Neural Network [J]. *Procedia Computer Science*, 2015, 59: 83-91.
- [21] LU H M. Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation [J]. *IEEE Access*, 2016, 4: 8375-8385.
- [22] 毛思晨. 基于卷积网络和长短时记忆网络的中国手语词识别方法研究 [D]. 合肥: 中国科技大学, 2018.
- [23] ZHOU X X, SHENG H. Combination of stationary wavelet transform and kernel support vector machines for pathological brain detection [J]. *Simulation*, 2016, 92(9): 827-837.
- [24] ZHANG H Y, YUAN J Z. Survey on New Methods of Vision-based Hand Gesture Recognition [J]. *Journal of Computational Science*, 2017, 44: 1-6.

(上接第 276 页)

- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet-classification with deep convolutional neural networks [C] // *Advances in Neural Information Processing Systems*. 2012: 1097-1105.
- [3] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time series [M] // *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1998.
- [4] ABDEL-HAMID O, DENG L, YU D. Exploring convolutional neural network structures and optimization techniques for speech recognition [C] // *INTERSPEECH 2013*. Lyon, 2013.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C] // *Advances in Neural Information Processing Systems*. 2012: 1097-1105.
- [6] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C] // *International Conference on International Conference on Machine Learning*. JMLR. org, 2015.
- [7] IOFFE S. Batch renormalization: Towards reducing minibatch dependence in batch-normalized models [C] // *Advances in Neural Information Processing Systems*. 2017: 1945-1953.
- [8] BA J L, KIROS J R, HINTON G E. Layer normalization [J]. *arXiv:1607.06450*, 2016.
- [9] WU Y, HE K. Group normalization [C] // *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018: 3-19.
- [10] SALIMANS T, KINGMA D P. Weight normalization: A simple reparameterization to accelerate training of deep neural networks [C] // *Advances in Neural Information Processing Systems*. 2016: 901-909.
- [11] REN M, LIAO R, URTASUN R, et al. Normalizing the normalizers: Comparing and extending network normalization schemes [C] // *ICLR*. 2017.
- [12] LIAO Q, KAWAGUCHI K, POGGIO T. Streaming Normalization: Towards Simpler and More Biologically-plausible Normalizations for Online and Recurrent Learning [J]. *arXiv:1610.06160v1*, 2016.
- [13] SPRINGENBERG J T, DOSOVITSKIY A, BROX T, et al. Striving for simplicity: The all convolutional net [C] // *ICLR*. 2015.
- [14] LIN M, CHEN Q, YAN S. Network in network [J]. *arXiv:1312.4400*, 2013.
- [15] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images; Technical Report; TR-2009 [R]. University of Toronto, 2009.
- [16] CLEVERT D A, UNTERTHINER T, HOCHREITER S. Fast and accurate deep network learning by exponential linear units (elus) [C] // *ICLR*. 2016.