

文章编号 1004-924X(2014)12-3427-08

采用多模糊支持向量机决策融合的积雨云检测

金 炜^{*}, 符冉迪, 范亚会, 王文龙, 田文哲
(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要:采用决策融合策略, 提出了一种基于多模糊支持向量机(FSVM)的积雨云检测方法以解决添加更多的特征可增加云分类识别的准确率而特征维数过高又会造成过拟合现象的矛盾。该方法首先从训练云图提取光谱特征、通道亮温差特征、一阶灰度直方图纹理特征、灰度共生矩阵纹理特征以及 Gabor 小波特征, 并组成包含 5 类特征的训练样本集; 然后针对每类特征, 训练 5 个 FSVM 子分类器。最后对各子分类器的结果在输出空间进行加权决策融合, 以提高积雨云检测的准确率。实验结果表明, 本文方法不仅较好地解决了积雨云检测中由于特征维数过高而造成的过拟合现象, 而且能自适应地确定不同特征的权重, 检测准确率优于各 FSVM 子分类器和包含所有输入特征的单 FSVM 分类器, 有望在卫星云图分析中得到应用。

关键词:多模糊支持向量机; 加权决策融合; 卫星云图; 积雨云检测

中图分类号: TP751; P412.15 **文献标识码:** A **doi:** 10.3788/OPE.20142212.3427

Cumulonimbus detection by using decision fusion of multi-FSVM

JIN Wei^{*}, Fu Ran-di, FAN Ya-hui, Wang Wen-long, Tian Wen-zhe

(College of Information Science and Engineering, Ningbo University, Ningbo 315211, China)

^{*} Corresponding author, E-mail: xyjw1969@126.com

Abstract: A cumulonimbus detection approach was proposed based on Multi Fuzzy Support Vector Machine(FSVM) by using a decision fusion strategy to solve the contradiction that adding more features will increase the accuracy of cloud classification while cause over fitting phenomenon due to high feature dimensions. Firstly, spectral features, the brightness temperature difference of multi-channels, first order histogram texture features, gray level co-occurrence matrix texture features and Gabor wavelet features were extracted from training cloud images to form a training sample set which contains 5 kinds of features. Then, five FSVM sub-classifiers were trained respect to each kind of feature. Finally, the output of each sub-classifier was fused by weighted decision in the output space to improve the detection accuracy of the cumulonimbus. Experimental results show that the proposed approach solves the over fitting phenomenon in cumulonimbus detection caused by the too high feature dimensions and can determine the weight of different features adaptively. The results also demonstrate that the accuracy is not only superior to each FSVM sub-classifier but also to the FSVM classifier trained by all the input features at once. The proposed approach is expected to be applied in the analy-

收稿日期: 2014-08-25; **修订日期:** 2014-10-16.

基金项目: 国家自然科学基金资助项目(No. 61271399, No. 61373068); 宁波市自然科学基金资助项目(No. 2011A610192, No. 2013A610055); 宁波市科技创新团队研究计划资助项目(No. 2011B81002); 宁波大学研究生教育改革研究重点项目(No. JGZDI201202); 浙江省信息与通信工程重中之重学科项目(No. XKXL1425, No. XKXL1306)

sis of satellite cloud images.

Key words: Multi Fuzzy Support Vector Machine (Multi-FSVM); weighted decision fusion; satellite cloud image; cumulonimbus detection

1 引言

雷暴、冰雹、龙卷风、局部强降雨等强对流天气与积雨云密切相关,它们严重威胁着人类的生命财产安全,因此准确检测以及预警积雨云至关重要^[1-2]。卫星云图分析是检测积雨云发生发展过程的重要手段,然而早期建立的基于气象专家人工标注云类的方法却随着大型气象卫星的发展、超高分辨率云图数据的海量增长显得捉襟见肘,而基于计算机的云类自动识别应运而生,并在强对流天气的监测中发挥越来越重要的作用^[3]。

可以简单地将基于计算机的卫星云图分析识别概括为三个步骤,即目标图像预处理、特征提取及最终的分类识别,它所具有的快速、客观、准确的云类识别能力是传统人工标注法所无法比拟的。Barnali 等将特征提取与神经网络相结合,提出了一种积雨云识别方法,但由于神经网络方法建立在经验风险最小化的基础上,准确分类需要大量样本,这在实际应用中往往很难满足。支持向量机(Support Vector Machine, SVM)在小样本、非线性及高维模式识别中有很大优势^[4-5],Xu 等利用 SVM 对台风云系中的积雨云进行分割,取得了较好的效果^[6]。然而,由于卫星云图的观测范围广,且云随时处于动态的变化过程中,一幅目标云图可能涵盖了某类云的产生、成熟、消亡的全过程,对其的自动分类识别很难提取到“类内高凝聚,类间高差异”的特征。因此,在对卫星云图进行特征提取时,往往将更多特征纳入特征集以增加识别的准确率,但这造成了一个矛盾:一方面希望通过参考更多的特征以反映云类间的本质区别,从而增加云分类识别准确率,而另一方面却由于特征维数过高造成“过拟合”现象。而且,不同特征的识别性能差异甚大,如果添加与识别相悖的无关特征反而会影响最终分类结果的准确性。因此,如何平衡目标识别准确率与特征选择两者间的关系成为云分类识别中亟需解决的问题,也是实现积雨云准确检测的关键^[6-8]。由于传统 SVM 不具有模糊处理能力,难于应对云图中不同

云类之间所表现出的模糊和不确定性,模糊支持向量机(Fuzzy Support Vector Machine, FSVM)通过引入模糊隶属度实现对模糊信息的处理,有望在积雨云检测中发挥作用。

针对以上问题,本文提出了一种采用多 FSVM 决策融合的积雨云检测方法。该方法从卫星云图中提取包含 5 类特征的训练样本集,并针对每类特征,训练 5 个 FSVM 子分类器,通过对每个子分类器输出结果的加权决策融合,实现积雨云检测。

2 积雨云及其特征提取

积雨云根据其发展状况可分为秃积雨云和鬃积雨云,其中秃积雨云处于积雨云的初始阶段,而鬃积雨云则处于其发展极盛阶段。当太阳辐射强烈,下垫面温度升高,或冷空气南下遇到温暖空气时,就会造成温暖空气急速上升,这些蕴含丰富水分的温暖空气上升到一定高度遇冷达到冻点就开始冰晶化,形成积雨云。积雨云浓厚,垂直尺度极大,云底约在 400~1 000 m,而云顶能达到 8 000~10 000 m 左右。它的出现往往带来强降水、雷电、冰雹、强烈的外旋气流、下击暴流等强对流天气,给人们的日常生活带来严重影响,因此,对其进行准确检测意义重大^[3]。

本文所提出的采用多 FSVM 决策融合的积雨云检测方法的训练过程包括 4 步:(1)多特征的提取:提取目标云图的光谱特征、通道亮温差特征、一阶灰度直方图纹理特征、灰度共生矩阵纹理特征以及 Gabor 小波特征,组成包含 5 类特征的训练样本集;(2)多 FSVM 的训练:针对 5 类特征,训练 5 个 FSVM 子分类器;(3)估算权重系数:对 5 个 FSVM 的输出结果进行加权,使不同特征识别结果具有不同的“重要性”;(4)决策融合:对 5 个 FSVM 的加权输出结果进行决策融合,建立最终的积雨云检测系统,用于日常卫星云图的分析处理。

2.1 多特性提取

气象卫星云图覆盖范围广,以我国 FY2-D 卫

星为例,其所成云图的经纬度跨度分别为 $50^{\circ}\text{E}\sim 145^{\circ}\text{E}$ 、 $-5^{\circ}\text{S}\sim 60^{\circ}\text{N}$,其中所包含的地理、海洋、大气信息庞大而繁杂,提取能反映积雨云与背景云图本质区别的特征,实现积雨云的精确检测难度较大。本文分别提取待测云图的多通道光谱特征、通道间亮温差特征、一阶灰度直方图纹理特征、基于灰度共生矩阵(Gray Level Co-occurrence Matrix, GLCM)的纹理特征以及 Gabor 小波高频信息特征,构成 5 个独立的样本特征集,用于后续 FSVM 的训练与测试。

2.1.1 多通道光谱特征

Hong G 等人^[6]利用红外云图光谱信息反演云顶亮温用于强对流云团的识别,取得初步效果。但此种单一利用红外通道云顶亮温的阈值分割法易受时空变化的影响,很难将目标云类从复杂的云图中分离出来。本文的研究对象主要为 FY2-D 卫星云图, FY2-D 卫星所携带的可见光红外自旋扫描辐射计有 5 个成像通道,包括可见光通道(VIS, $0.51\sim 0.90\ \mu\text{m}$)、两个红外分裂窗通道(IR1, $10.29\sim 11.45\ \mu\text{m}$; IR2, $11.59\sim 12.79\ \mu\text{m}$)、水汽通道(IR3, $6.32\sim 7.55\ \mu\text{m}$)以及中红外通道(IR4, $3.59\sim 4.09\ \mu\text{m}$)。不同通道对云顶亮温、反照率、水汽含量等大气物理信息的反映各有侧重,因此给云及大气水汽的遥感带来了极大的方便。由于云类的光谱特征在云图上最直接的反映是灰度明暗大小,同时为了削弱随机噪声的影响,本文提取每个样本点的 IR1、IR2、水汽、IR4 通道云图上对应位置的邻域灰度均值作为光谱特征, VIS 通道云图由于受光照、季节、纬度的影响较大,因此不纳入光谱特征集。

2.1.2 通道亮温差特征

Mecikalski J R 等人^[9-10]通过分析强对流天气期间 GOES 卫星的观测数据,发现光谱为 $6.5\ \mu\text{m}$ 与 $10.7\ \mu\text{m}$ 的通道亮温差可以反映强对流云团的云顶高度,当其差值为负时表明此时云顶高度已经超过了对流层,而当差值在 $-35\sim -5\ \text{K}$ 时极有可能发生强对流天气;同时,光谱为 $10.7\ \mu\text{m}$ 与 $13.3\ \mu\text{m}$ 通道亮温差则可以描述强对流云发展前期积状云的状态,当差值为 $-25\sim -5\ \text{K}$ 时可以作为强对流云的判别依据。虽然 FY2-D 所带扫描辐射计与 GOES 卫星有所不同,但通过对照 FY2-D 卫星与 GOES 卫星成像通道的光谱范围,发现 GOES 卫星的 $6.5\ \mu\text{m}$ 与 $10.7\ \mu\text{m}$ 的

成像通道恰好被 FY2-D 卫星的水汽及 IR1 通道覆盖,而 FY2-D 卫星 IR2 通道的光谱范围接近 GOES 卫星的 $13.3\ \mu\text{m}$ 通道,因此本文提取 FY2-D 卫星云图 IR1 通道与水汽通道、IR1 通道与 IR2 通道两种通道亮温差特征用于积雨云检测。

2.1.3 灰度直方图纹理特征

纹理作为一种重要的信息特征,可以提供图像区域的光滑度、稀疏性和规则性信息^[11],卫星云图纹理细节丰富,如层云纹理光滑均匀,积云纹理褶皱有斑点,卷云纹理呈纤维状,提取有效的纹理特征,可以大大提高云图分析的准确性。由于图像的灰度分布往往反映了局部结构的空间重复周期,周期大的纹理比周期小的纹理看上去要粗糙,而基于灰度直方图的统计特征能够表征纹理的粗细程度,因此在云图特征提取时,通过相应邻域的一阶灰度直方图,计算均值、方差、平滑度、三阶距、一致性、熵等 6 种统计量作为云图的灰度直方图纹理特征,用于积雨云检测(为了控制特征维数,纹理特征提取时仅针对 IR1 通道云图进行),具体公式可见参考文献^[12],本文不再赘述。

2.1.4 灰度共生矩阵纹理特征

GLCM 与一阶灰度直方图不同,它反映的是图像中两像素之间的联合分布。假设二维图像用 $f(x, y)$ 表示,灰度级为 L ,那么定义 GLCM 为 $[P(i, j, d, \theta)]$,它表示从灰度为 i 的像素点出发,在间隔为 $d=(d_x, d_y)$ 的点出现灰度为 j 的像素的个数。其中 $i, j=0, 1, \dots, L-1, \theta$ 一般取为 0° 、 45° 、 90° 、 135° ,分别对应 $d=(d_x=1, d_y=0)$ 、 $d=(d_x=1, d_y=-1)$ 、 $d=(d_x=0, d_y=-1)$ 、 $d=(d_x=-1, d_y=-1)$ 。基于 GLCM 的二阶纹理特征,即是首先对图像进行一次统计得到其对应的 GLCM $[P(i, j, d, \theta)]$,然后在此基础上计算其二次统计量作为图像的纹理特征。Haralick 最早提出 28 种 GLCM 的二次统计量对纹理进行定性描述^[13]。本文针对卫星云图的特性,仅挑选了 0° 方向上的二阶距、对比度、相关性、熵 4 种二次统计量,形成 4 维的 GLCM 纹理特征向量,用于积雨云的检测。

2.1.5 Gabor 小波高频特征

基于小波变换的特征提取法^[14],旨在通过对原始空间图像进行频域变换来挖掘图像有用的高低频信息。Gabor 小波即是采用 Gabor 函数作为母函数的小波变换法。假设二维图像用 $f(x, y)$

表示,则其离散 Gabor 小波变换表示为:

$$\mathbf{I}(x, y) = \sum_s \sum_t f(x-s, y-t) g_{m,n}^*(s, t), \quad (1)$$

式中: $g^*(s, t)$ 为 Gabor 小波函数的共轭, s 和 t 为滤波掩模尺寸变量, m 和 n 则为小波变换尺度与方向, 而 $\mathbf{I}(x, y)$ 为 Gabor 小波变换的输出系数。

本文利用 Gabor 小波变换对云图进行 3 尺度 3 方向的分解, 利用得到的 9 个高频信息直接进行单独重构, 将重构得到的高频图像对应点作为特征, 构成 9 维特征向量用于积雨云检测。

3 决策融合

如上所述, 本文通过 5 种不同方法提取了目标云图的 5 类共 25 个特征用于积雨云检测, 在实际云图分析时, 可将这 25 个特征规范化后共同组成特征向量来训练学习机器, 虽然从直观上讲由于参考了更多的特征可以更好地反映云类间的本质区别, 从而增加云分类识别准确率, 但实验结果表明, 由于引入的特征维数过高, 造成了“过拟合”现象, 在卫星云图的积雨云检测中并没有取得预期的效果。因此, 本文依次利用 5 类特征训练 5 个 FSVM, 通过加权系数决策融合法对 5 个 FSVM 输出进行融合, 最终“集思广益”得到积雨云检测结果。

3.1 模糊支持向量机 (FSVM) 及其输出模糊概率的拟合

本文选用 FSVM 作为积雨云检测分类器, 针对 5 类特征中的每类特征, 各训练一个 FSVM 用于后续的检测。基于统计学习 VC (Vapnik-Chervonenkis Dimension) 维理论的 SVM 具有结构风险最小化的特性, 在解决非线性、小样本、高维数的模式识别问题中表现突出, 应用广泛。但使用传统 SVM 时, 训练集中各个样本的重要性均一样, 表明其不具备处理模糊信息的能力, 不善于表达云图所具有的模糊、不均匀、云型复杂多变等特点。FSVM 通过引入模糊隶属度 μ_i , 可对不同的样本 x_i 采用不同的权重系数, 可以减少甚至忽略非重要样本和噪声对分类面的影响, 实现对模糊信息的处理。虽然 FSVM 较好地解决了输入样本的不确定性问题, 然而对于输出类别的判定却仍然采用传统 SVM 的 sign 函数, 即样本属于

或者不属于的情况。而云图往往表现出模糊不确定性, 这就导致不能准确判断一些样本的类别, 同时, 只考虑两种极端情况也不适合于后续的决策融合。因此, 本文引入输出模糊, 将传统 SVM 的输出变换为概率输出形式。SVM 的输出为:

$$y = \text{sgn}(f(x)) = \text{sgn}\{(\omega^T \cdot x) + b\}. \quad (2)$$

对 ω 和 b 进行归一化, 即利用 $\omega/\|\omega\|$ 与 $b/\|\omega\|$ 代替原 ω 和 b , 此时任意样本到分类超平面的距离就可以写成:

$$d_x = \frac{f(x)}{\|\omega\|}, \quad (3)$$

而支持向量到分类超平面的距离为:

$$d_{sv} = \frac{1}{\|\omega\|}. \quad (4)$$

由式(3)、(4)可知, $f(x)$ 恰为 d_x 与 d_{sv} 的比, 因而可以对 $f(x)$ 进行模糊概率拟合, 将其输出变换为概率输出形式。本文利用 Sigmoid 函数对 SVM 输出进行直接拟合, 将输出形式变为:

$$P(y=1|f(x)) = \frac{1}{1+e^{A(f(x)+B)}}, \quad (5)$$

$$P(y=-1|f(x)) = 1 - P(y=1|f(x)), \quad (6)$$

其中: A 和 B 为 Sigmoid 函数的形态参数, 实验仿真时采用网格寻优法确定其具体数值。后验概率 $P(y=1|f(x))$ 表示某样本属于正类的概率, 这样在对样本进行类别测试时更接近于实际情况。

3.2 加权系数决策融合

决策融合可以简单地理解为利用多个分类器对同一样本进行判别, 然后根据多个判别结果进行统计决策, 以获得最终判别结果。由于传统的“最大票数”法只适用于分类器个数为奇数的情形, 因此本文引入后验概率并结合权重系数, 构造出一种适用于任意个分类器的决策融合法, 对多个 FSVM 子分类器的输出进行融合, 得到最终的积雨云检测结果。Louisa Lam 等人^[15-16]已证明: 当单个子分类器结果的准确率都大于 0.5 时, 随着子分类器个数的逐渐增加, 决策融合后的准确率将趋近于 1; 当单个子分类器结果准确率都小于 0.5 时, 随着子分类器个数的增加, 决策融合后的准确率将趋近于 0; 当单个子分类器结果准确率等于 0.5 时, 决策融合结果约等于 0.5。因此, 只要设计合理并保证每个 FSVM 子分类器正确率均在 0.5 以上, 最终决策融合的积雨云检测结果将有进一步的提高。

由于不同特征在分类识别中表现出的识别

率,也即反映目标本质的程度不尽相同。不同的特征可能对特定的目标具有高识别率,但对其它目标的识别能力可能不如人意。比如在积雨云检测中,光谱特征因为直接反映了云顶亮温因而具有较强的识别率,虽然积雨云蕴含的纹理特征与背景云类有所不同,但其识别结果的可信度相较光谱特征仍有一定的差距。如果将基于不同类特征所训练的 FSVM 的输出结果同等看待,从而决定最终的识别结果,就没有发挥决策融合优势。基于此,本文对 5 类特征所训练的 5 组 FSVM 子分类器的输出结果,在决策融合中赋予不同的“话语权”。上节已经对 5 组 FSVM 的输出结果进行了后验概率拟合,假设有 n 个训练样本 $\{x_1, x_2, \dots, x_n\}$,其对应类标为 $\{y_1, y_2, \dots, y_n\}$,子分类器个数为 k (本文 $k=5$),具体的决策加权融合算法如下:

Step 1: 每组 FSVM 输出权重系数的初始化,

$$\omega_s = \frac{1}{k}, s=1, \dots, k, \quad (7)$$

Step2: 对于每个样本 $x_j (j=1, 2, \dots, n)$, 根据每个子分类器输出的对应类别的后验概率, 计算出加权累加值:

$$u_i = \sum_{s=1}^k \omega_s P_{si}, \quad (8)$$

$$U_j = \{u_i | i=1, 2, \dots, c\}, \quad (9)$$

其中: P_{si} 表示第 s 个子分类器输出结果属于第 i 类的后验概率的大小, c 为类别个数(在积雨云检测时, $c=2$)。

Step3: 确定具有最大概率的 U 所对应的类别, 作为融合判别结果:

$$Y_j = \max(U_j), \quad (10)$$

Step4: 自适应迭代更新权重系数 ω_s 。将决策融合判别结果与每个子分类器的判别结果进行比

较,对判别错误的 l 个子分类器对应的权重系数自减标量 ϵ , 同时将每个子分类器输出的加权后验概率 P_{si} 按大小重新排列, 对前 l 个子分类器对应权重系数自增 ϵ , 以保证 $\sum \omega_s = 1$ 。返回 Step2, 直到决策融合输出与期望输出类标的平方误差累加值小于 θ , 或者所有样本均遍历完成, 此时就可将训练好的各组 FSVM 子分类器及决策加权系数用于实测云图的分析处理。

4 实验及结果分析

实验数据取自 FY2-D 卫星 2012 年 5 月到 9 月的云图, 通过参考气象专家的意见, 选取若干积雨云及背景样本, 并提取样本的多通道光谱特征、通道亮温差特征、一阶灰度直方图纹理特征、GLCM 纹理特征以及 Gabor 小波高频特征共 5 类, 如表 1 所示。

表 1 样本个数及每类特征维数

Tab. 1 Number of samples and each type of feature dimensions

类别	样本个数	光谱特征	亮温差	一阶纹理	GLCM	Gabor 小波特征
积雨云	600	4	2	6	4	9
背景	1 200	4	2	6	4	9

从中选取积雨云样本 200 个, 背景样本 400 个, 选用高斯径向基核函数, 分别利用 5 类特征训练 5 个 FSVM 子分类器。利用训练好的 5 个 FSVM, 另外随机选取 200 个积雨云, 400 个背景样本组成测试样本集进行测试, 其正确率及错分样本个数如表 2 所示。

表 2 采用不同特征 FSVM 子分类器测试样本的错分个数及分类正确率

Tab. 2 Misclassification numbers and correctoin rate of test samples for FSVM sub-classifiers using different input features

	光谱特征	亮温差	一阶纹理	GLCM	小波特征
错分个数(积雨云错分为背景/ 背景错分为积雨云)	40 (21/19)	44 (32/12)	56 (22/34)	150 (74/76)	202 (98/104)
正确率	93.33%	92.67%	90.67%	75.00%	66.33%

由上表可知, 对于光谱特征及亮温差特征, 错分样本中积雨云错分为背景的数目大于背景错分

为积雨云的数目, 而对于一阶纹理、GLCM 及小波特征则相反, 且 5 组 FSVM 子分类器对测试样

本的分类正确率均在 50% 以上,符合 Louisa Lam 等人的论证。因此,对 5 组 FSVM 子分类器的结果进行加权决策融合将有助于提高积雨云检测的准确率。本文对每个 FSVM 的输出进行模糊概率的拟合,其中 Sigmoid 函数形态参数 A 和 B 采

用网格寻优确定为 $A = -1.2, B = 0.2$,对每组 FSVM 的输出权重系数按 3.2 节的方法进行迭代确定,自减标量 ϵ 取为 0.002,针对包含 200 个积雨云,400 个背景的训练样本集,最终的预测结果如表 3 所示:

表 3 各子分类器的权重及最终决策融合分类正确率

Tab. 3 Weight of each sub-classifier and the final classification accuracy of decision fusion

	光谱特征	亮温差	一阶纹理	GLCM	小波特征
权重系数	0.218 0	0.262 0	0.310 0	0.126 0	0.084 0
错分个数(积雨云错分为背景/背景错分为积雨云)			20 (11/9)		
正确率			96.67%		

可以看出,通过迭代,不同的子分类器在决定最后输出结果时有不同的权重,反映了不同类别的特征对积雨云不同的检测能力。通过决策融合,错分样本数仅有 20 个,最终的分类准确率优于各 FSVM 子分类器。为了进一步表明本文所提出的决策融合方法的优越性,本文训练了一个包含所有输入特征的单 FSVM 分类器,并对其性能进行了测试。实验发现对于同样的测试样本集其错分样本数为 35(其中积雨云被错分为背景以及背景被错分为积雨云的样本数分别为 8 与 27),分类正确率为 94.2%,低于本文方法。从计算性能上看,将算法运行于 Window XP(Intel(R) Pentium CPU G2030 @3.0GHZ),2G 内存的电脑时,当采用光谱、亮温差、一阶纹理、GLCM、小波特征训练 5 个 FSVM 子分类器时,各子

分类器所需的训练时间分别为 75.20、60.88、80.26、73.11、87.48 s,而实现各子分类器的决策融合,额外耗时为 50.20 s。可以看出,虽然为了构造本文所提出的决策融合 FSVM,总的时间消耗达到了 427.13 s,但由于训练过程并不要求实时性,且利用训练好的 FSVM 进行预测时,算法能在 0.5 s 内完成,因此训练过程的时间消耗并不影响本文算法的实际应用价值。

为了测试本文方法在积雨云检测中的实际应用能力,选取 FY-2D 卫星 2012 年 8 月 4 日某时次的云图,截取云图中包含积雨云的区块进行试验。图 1(a)为原始 IR1 通道云图,图 1(b)为专家标注结果,图 1(c)为本文方法的检测结果,而图 1(d)为将全部 25 维特征训练为一个单一的 FSVM 分类器所得的检测结果。

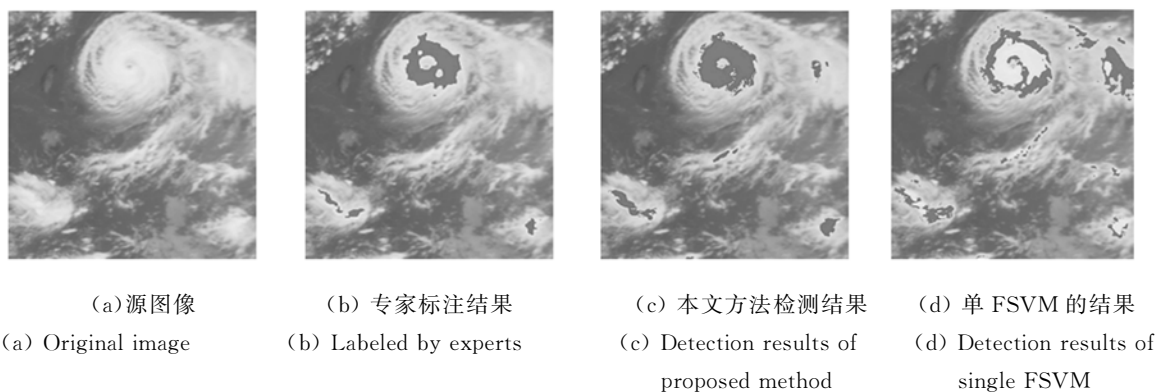


图 1 实测云图积雨云检测结果

Fig. 1 Cumulonimbus detection results of actual satellite cloud image

由上图可以看出,本文方法的积雨云检测结果与专家标注结果非常接近,而采用全部 25 维特征所训练的单一 FSVM 的检测结果却存在着严重的漏检及误检现象。这进一步印证了传统 FSVM 由于特征维数过高造成了“过拟合”现象,从而在实际云图处理时表现出较差的泛化能力。本文结合模糊不确定性理论与机器学习法,根据卫星云图不同特征对积雨云不同的鉴别能力,构造多个 FSVM 对云图进行分类识别,通过拟合输出模糊概率,采用多特征决策加权融合法得到最终的检测结果,这不仅解决了“过拟合”与准确率之间的矛盾,而且平衡了不同特征在分类识别中的作用,取得了较好的积雨云检测结果。

参考文献:

- [1] LIU C C, SHYU T Y, CHAO C C, *et al.*. Analysis on typhoon longwang intensity changes over the ocean via satellite data [J]. *Journal of Marine Science and Technology*, 2009, 17: 23-28.
- [2] THOMAS F, REMY R. An algorithm for the detection and tracking of tropical mesoscale convective systems using infrared images from geostationary satellite [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(7): 4302-4315.
- [3] 刘延安, 魏鸣, 高炜, 等. FY-2 红外云图中强对流云团的短时自动预报算法[J]. *遥感学报*, 2012, 16(1): 86-92.
- LIU Y A, WEI M, GAO W, *et al.*. Short-term automatic algorithm of severe convective cloud identification using FY-2 IR images[J]. *Journal of Remote Sensing*, 2012, 16(1): 86-92. (in Chinese)
- [4] 谭熊, 余旭初, 张鹏强, 等. 基于多核支持向量机的高光谱影像非线性混合像元分解[J]. *光学精密工程*, 2014, 22(7): 1912-1920.
- TAN X, YU X CH, ZHANG P Q, *et al.*. Nonlinear mixed pixel decomposition of hyperspectral imagery based on multiple kernel[J]. *Opt. Precision Eng.*, 2014, 22(7): 1912-1920. (in Chinese)
- [5] 周涛, 陆惠玲, 陈志强, 等. 基于两阶段集成 SVM 的前列腺肿瘤识别[J]. *光学精密工程*, 2013, 21(8): 2137-2145.
- ZHOU T, LU H L, CHEN ZH Q, *et al.*. Prostate tumor recognition based on two-stage intergrating SVM[J]. *Opt. Precision Eng.*, 2013, 21(8): 2137-2145. (in Chinese)
- [6] HONG G, HEYGSTER G, KUNZI K. Intercomparison of deep convective cloud fractions from passive infrared and microwave radiance measurements [J]. *IEEE Trans. Geosci. Remote Sens.*, 2005, 2(1): 18-22.
- [7] 朱平, 李生辰, 王振会, 等. 青藏高原东部暴雨云团局地强降水响应特征[J]. *遥感学报*, 2014, 18(2): 405-431.
- ZHU P, LI SH CH, WANG ZH H, *et al.*. Characteristics of rainstorm cloud clusters to local heavy precipitation over the eastern Qinghai-Tibet plateau [J]. *Journal of Remote Sensing*, 2014, 18(2): 405-431. (in Chinese)
- [8] DING H, WEI Y, JIANG Q R, XIANBIN Z. Cloud type classification algorithm for CloudSat satellite based on support vector machine [J]. *Transactions of Atmospheric Sciences*, 2011, 34: 583-591.
- [9] MECIKALSKI J R, BEDKA K M. Forecasting convective initiation by monitoring the evolution of moving cumulus in daytime GOES imagery [J]. *American Meteorological Society*, 2006, 134(1): 49-78.
- [10] THOMAS F, REMY R. Composite life cycle of

5 结 论

气象卫星云图具有模糊性,对其准确分类识别难度较大。FSVM 算法通过输入模糊隶属度解决了输入样本的不确定性,但并未涉及输出类别的不确定性。本文针对云图的不同特征在积雨云检测中能力的不同,针对不同类的特征训练多个 FSVM 子分类器,通过引入输出模糊隶属度,并采用对多个子分类器进行加权决策融合的方法,实现了对传统 FSVM 的优化。实验表明,本文方法不仅解决了“过拟合”与预测准确率之间的矛盾,而且也能自适应地确定不同特征的权重,从而大大提升了积雨云检测的准确率,这不仅为强对流云团的识别提供了一种可靠的方法,而且为卫星云图的自动分类识别的提供了一种新思路。

- tropical mesoscale convective systems from geostationary and low earth orbit satellite observations: method and sampling considerations [J]. *Quarterly Journal of the Royal Meteorological Society*, 2013, 139(673):941-953.
- [11] 凌剑勇, 何昕, 李一芒, 等. 融合纹理特征与深度信息的足迹自动比对算法[J]. *光学精密工程*, 2014, 22(7):1946-1954.
- LING J Y, HE X, LI Y M, *et al.*. Footprint automatic alignment algorithm fusing texture features and depth information[J]. *Opt. Precision Eng.*, 2014, 22(7):1946-1954. (in Chinese)
- [12] 王勇, 杨公训, 路迈西, 等. 煤泥浮选泡沫图像灰度直方图及其统计纹理特征研究[J]. *选煤技术*, 2006, (1):13-16.
- WANG Y, YANG G X, LU M X, *et al.*. Research on grey scale histogram of coal slime flotation froth image and its statistic texture characteristic [J]. *Coal Preparation Technology*, 2006, (1):13-16. (in Chinese)
- [13] HARALICK R M, SHANMUGAM K, DINTON I. Textual features for image classification [J]. *IEEE Transactions on Systems Man and Cybernetics*, 1973, 3(6):610-621.
- [14] 张刚, 马宗民. 一种采用 Gabor 小波的纹理特征提取法[J]. *中国图像图形学报*, 2010, 15(2):247-254.
- ZHANG G, MA Z M. An approach of using Gabor wavelets for texture feature extraction[J]. *Journal of Image and Graphics*, 2010, 15(2):247-254. (in Chinese)
- [15] 刘三民, 王彩霞, 孙知信. 一种基于 SVM 后验概率的网络流量识别方法[J]. *计算机工程*, 2012, 38(17):171-173.
- LIU S M, WANG C X, SUN ZH X. A network flow identification method based on SVM posteriori probability [J]. *Computer Engineering*, 2012, 38(17):171-173. (in Chinese)
- [16] LOUISA L, SUEN C Y. Application of majority voting to pattern recognition: an analysis of its behavior and performance [J]. *IEEE Trans. on Systems, Man, and Cybernetics*, 1997, 27(5):553-568.

作者简介:



金 炜 (1969—), 男, 浙江兰溪人, 博士, 副教授, 硕士生导师, 2006 年于重庆大学获得博士学位, 主要从事遥感图像处理、多尺度几何分析、压缩感知、光电检测的研究。E-mail: xyjw1969@126.com



符冉迪 (1971—), 男, 浙江慈溪人, 硕士, 副教授, 2001 年于解放军信息工程大学获得硕士学位, 主要从事数字图像处理、摄影测量与遥感技术的研究。E-mail: furandi@nbu.edu.cn

(版权所有 未经许可 不得转载)