

文章编号: 2096-1618(2017)06-0590-07

# 双偏振天气雷达水凝物识别研究

冉元波, 孙 敏, 高梦清, 王海江  
(成都信息工程大学电子工程学院, 四川 成都 610225)

**摘要:**对双偏振天气雷达回波的水凝物分类,是利用降水粒子对极化电磁波的散射特性对水凝物相态进行识别的过程。不同相态降水粒子由于在形状、大小和空间取向等方面存在差异,对一定极化状态的电磁波会产生不同的散射特性,导致与这些特性密切相关的雷达偏振参量也各不相同,综合利用这些偏振参量,可以有效地识别出各种水凝物的相态。提出一种利用深度学习和模糊逻辑算法进行联合判别的水凝物相态识别方法:首先,采用深度学习算法对降水粒子所对应的雷达回波产品数据进行特征提取,并利用 Softmax 分类器对提取到的特征进行分类,实现降水粒子所对应降水模式的识别;其次,在已知降水模式的情况下,利用模糊逻辑算法,实现对降水粒子类型的最优判别;最后,结合其他雷达产品对分类结果进行分析与比较,发现两者能够达到极好的吻合。这种采用深度学习对降水粒子进行初次聚类,再利用模糊逻辑算法实现精确聚类的方法,大大提高了水凝物识别的准确性。

**关键词:**双偏振天气雷达;水凝物分类;偏振参量;深度学习;模糊逻辑;聚类

**中图分类号:**TN959.4

**文献标志码:**A

**doi:**10.16836/j.cnki.jcuit.2017.06.003

## 0 引言

双线偏振雷达通过交替发射或同时发射水平和垂直偏振波,并对两个偏振方向的回波信号采用不同方式的信号处理,从而得到差分反射率因子、差分相移、退偏振因子、相关系数等多个偏振参量<sup>[1-2]</sup>。这些偏振参量,反映了大气空间中降水粒子的大小、形状、相态以及取向等特征,促进了水文气象分类等研究领域的发展。与常规多普勒天气雷达系统相比,双偏振多普勒天气雷达探测系统提高了降水估计与水凝物相态识别的能力,也是人工影响天气、航空预警以及灾害监测等领域的重要工具<sup>[3-6]</sup>。

以升级改造后的 WSR-98D/XD 型双偏振多普勒天气雷达的探测数据为资料,设计了一个基于深度学习与模糊逻辑的水凝物识别系统。在文中,对每一个距离库,均采用其附近 1 平方公里内所有距离库的偏振参量所构成的偏振参量矩阵来反应该距离库的降水粒子信息。首先,系统通过 CNN 算法对偏振参量矩阵进行特征提取;其次,采用 Softmax 分类器对这些特征参数进行降水模式分类;最后,采用模糊逻辑方法对每一类降水模式中的降水粒子类型进行判断,得出最优的识别结果。这种采用深度学习对降水粒子进行初次聚类,再利用模糊逻辑算法实现精确聚类的方法,大大提高了水凝物分类的准确性。

## 1 深度学习

深度学习旨在通过模拟大脑的学习过程,构建深层次的神经网络模型,结合大量的数据训练,找到数据的内部结构与变量之间的关系,常见的网络模型有自动编码器神经网络、深度置信神经网络以及卷积神经网络<sup>[7]</sup>。文章所采用的模型为卷积神经网络模型。

### 1.1 卷积神经网络

卷积神经网络(CNN)是一种受生物视觉神经机制的启发而设计的多层感知器。CNN 由多个卷积层和降采样层交替排列组成,具备自动提取样本特征的能力。在卷积神经网络中,相邻的卷积层与降采样层可以被看作为一个特征提取器。在卷积层中,每个神经元与部分邻层的神经元相连接,通过局部感知器提取初级的原始信号特征,同时维持原始信号的空间结构关系,从而避免了传统识别算法中复杂的特征提取与数据重建过程。在降采样层中,通过采样对原始数据进行压缩,减少了计算复杂度和建立空间结构上的不变性。与生物的神经网络类似,CNN 采用了权值共享的方式,减少了权值参数的数量,大大降低了网络模型的复杂度,从而在图像分类以及语音识别等领域得到了广泛的运用。

CNN 作为一种典型的有监督学习算法,依赖于大量有标记的样本对模型进行训练。在训练过程中,主

收稿日期:2017-09-11

基金项目:四川省教育厅基金资助项目(16ZA0209);四川省科技厅基金资助项目(2016JY0106)

要采用反向传播算法 (BP)。样本数据通过多次的卷积与采样运算得到相应的输出结果,如果输出结果与样本标签不相符,则系统将输出误差以某种形式通过隐藏层向输入层逐层反向传播,同时将误差分配到各层的所有单元,从而得到各个单元的误差信号,并以此误差信号来修正各个单元的权值。因此,CNN 的训练过程可以归纳为原始信号的正向传播与误差信号的反向传播并调整各神经元权值的过程<sup>[8-10]</sup>。

1.2 样本数据的选取

此次试验的数据来源于 WSR-98D/XD 型双偏振天气雷达,其主要性能参数如表 1 所示。

表 1 WSR-98D/XD 雷达主要参数指标		
序号	项目	详细说明
1	天线系统	抛物面天线,波瓣宽度 1°,增益43 dB,水平/垂直偏振
2	发射系统	X 波段,频率9420 MHz,峰值功率75 kW,脉冲宽度0.5 μs
3	接收系统	通道数 2,带宽1.1 MHz,噪声系数<3.5 dB,动态范围>70 dB
4	伺服系统	支持 PPI、RHI、VOI 及扇形扫描
5	雷达产品	水平反射率、多普勒速度、谱宽、差分反射率、差分传播相移、零相关系数等

在此次深度学习的试验中,将每个距离库上的多个偏振参量所构成的偏振特征向量作为一个基本的数据单元,其中每个基本的数据单元包含水平反射率  $Z_H$ 、差分反射率  $Z_{DR}$ 、差分传播相移  $K_{DP}$  以及相关系数  $\rho_{HV}$  等 4 个偏振参量。考虑到相邻距离库之间数据的高度相关性,故在对某一距离库的数据进行选择时,以该距离库为中心,用附近一平方公里内的数据表示该距离库,从而大大提高了结果的准确性。其具体操作如下:

(1)构建一个样本数据提取窗,其由 21×21 个单元格构成,每个单元格的长度为50 m(与每个距离库的长度近似相等)。

(2)将上述数据提取窗放入 PPI 数据产品中,其中心与待选择距离库的中心重合,对于数据提取窗中的每个方格,若单元格落入哪一个距离库,则用落入的距离库的数据作为该单元格数据。对于距离雷达站中心处较近的距离库,由于每个距离库对应的网格较小,故导致多个距离库可能落入同一单元格,故将多个距离库数据产品的均值作为该单元格的数据。

(3)通过以上步骤,可以得到一个大小为 21×21×4 的样本数据。对于样本标签,采用投票规则,利用 21×21 个单元格所对应的降水模式类型进行投票,将得票最高的降水模式类型作为该样本的降水模式类型。

综上,对于每一个距离库,可以得到一个 21×21×4 的样本数据(偏振参量矩阵)和一个样本标签数据。

2 模糊分类器

1965 年,Zadeh 等<sup>[11]</sup>首次提出模糊逻辑的概念,其主要包括模糊化、规则推断、集成以及退模糊 4 个过程。模糊逻辑分类器使用  $Z_H$ 、 $Z_{DR}$ 、 $K_{DP}$  和  $\rho_{HV}$  等 4 个偏振参量作为输入,并依次通过模糊化、规则推断、集成以及退模糊化的处理,最后将得到的结果转化成单一的降水粒子类型。利用深度学习算法对降水粒子所属的降水模式进行识别,在已知降水模式的基础上,利用模糊逻辑方法对降水粒子的类别进行判别,由于在不同降水模式下所对应的降水粒子各异,故各个降水模式所设计的模糊逻辑方法也各不相同,其对应关系如表 2 所示。下面将对降雨模式下所对应模糊逻辑方法的设计过程进行详细介绍,其最终输出结果为 2(小雨)、3(中雨)和 4(大雨)。

表 2 降水模式与降水粒子类型之间的对应关系		
降水模式	降水粒子类型	最终输出
雨	毛毛雨(LR)	2
	雨(RN)	3
	大雨(HR)	4
	湿雪(WS)	5
雪	干雪(DS)	6
	晶体(CR)	1
	冰晶(IC)	7
	霰(GR)	8
雹	冰雹(IH)	9
	雨加雹(RH)	10
	晶体(CR)	1
	冰晶(IC)	7

2.1 模糊化及隶属函数的选取

模糊逻辑算法用于水凝物粒子的分类,首先需要将 4 个输入的特征变量进行模糊化处理,即把原始的输入数据以隶属度形式转化为模糊基。每一个特征参数对应着多个不同的模糊基,从而得到多个不同的隶属度,正是这些不同的隶属度决定了最终输出结果,故隶属函数的选取成了模糊化过程中最重要的环节。通过大量实验对比,发现 Beta 型隶属度函数对水凝物粒子的分类效果最佳,其表达式为

$$\beta(x,a,b,m) = \frac{1}{1 + \left| \frac{x - m}{a} \right|^{2b}} \tag{1}$$

其中,  $x$  为输入变量,  $a$  为函数宽度,  $b$  为斜率, 函数关于  $x = m$  对称且取值范围为  $0 \sim 1$ 。

2.2 规则推断

经过前人的分析, 基于模糊逻辑的分类器核心主要在于隶属函数与规则的构建, 其规则可以用逻辑语言描述如下:

IF ( $X_1$  IS  $MBF_{1j}$  AND  $X_2$  IS  $MBF_{2j}$  AND  $X_3$  IS  $MBF_{3j}$ )  
THEN Hydrometeor IS  $j$

其中,  $j = 1, 2, 3$ , 对应降水模式下的毛毛雨、雨和大雨等 3 种降水粒子类型, 因此  $j$  类降水粒子的强度  $R_j$  可以表示为

$$R_j = \sum_{i=1}^4 [W_{ij} \cdot MBf_{ij}(X_i)] \quad (2)$$

其中,  $W_{ij}$  表示第  $i$  个参数对  $j$  类降水粒子的贡献强度,  $MBf_{ij}(X_i)$  表示特征参数  $X_i$  对应第  $j$  类粒子的隶属度值。

2.3 集成

从各个独立的规则推断中得到的结论采用最大集成法进行集成, 最大集成法仅把最大真值的结果作为最终结果, 即  $C = \max_j R_j$ 。

2.4 退模糊化

退模糊过程, 即为寻找上述最大规则强度对应的索引值, 并将其作为结果输出。

3 方法实现

构建了一个基于深度学习与模糊逻辑的水凝物类型识别系统, 其系统原理框图如图 1 所示。在该系统中, 主要利用 CNN 算法对所得到的偏振参量矩阵进行特征提取, 并将提取到的特征送入 Softmax 分类器中进行降水模式的识别。系统在已知降水模式的条件下, 采用模糊逻辑的方法对降水粒子的类型做出最优的判断。

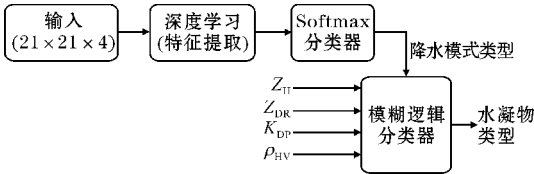


图 1 水凝物类型识别系统原理框图

在该降水粒子类别的识别系统中, 主要由基于深度学习的降水模式分类系统与基于模糊逻辑的降水粒子类型判别系统两个模块构成, 下面将对两个子系统的实现过程进行详细介绍。

3.1 降水模式分类系统

该系统以 4 个偏振参量所构成的特征矩阵作为输入, 通过卷积神经网络对其进行特征提取, 并采用 Softmax 分类器对提取到的特征进行分类, 最后得到降水粒子所对应的降水模式(雨、雪、雹), 系统结构框图如图 2 所示。

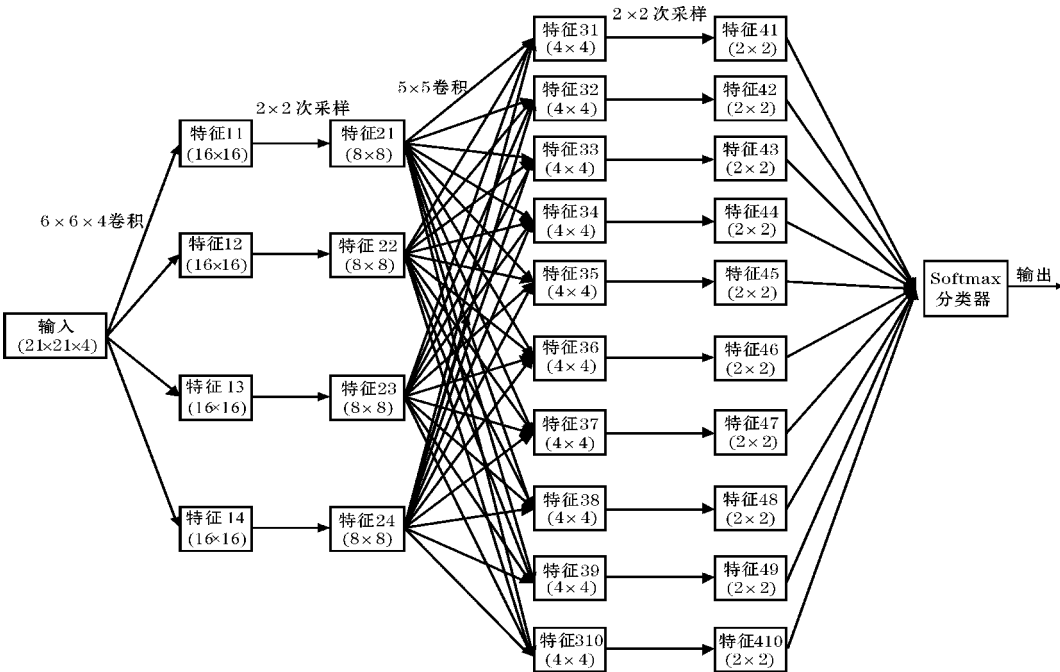


图 2 基于深度学习的降水模式识别系统结构框图



在系统中,主要由输入输出层、两个卷积层和降采样层以及 Softmax 分类器 4 部分构成。其工作流程主要分为训练过程与测试过程。

3.1.1 训练过程

训练过程主要是通过大量带标签的样本数据,并结合 BP 算法对系统中的参数进行自适应调整,使系统中的各个参数达到能够准确分类的要求。其具体步骤如下:

(1)将带标签的样本数据输入到该识别系统中,其中每个样本数据都是  $21 \times 21 \times 4$  的偏振参量矩阵,将其与卷积层进行卷积运算,并利用次采样层对卷积结果进行采样,重复以上过程,最终得到 40 个特征数据。

(2)将得到的 40 个特征数据转化为一个  $1 \times 40$  的特征向量,并利用 Softmax 分类器对其进行分类处理,得到输入数据所对应的降水模式类型。

(3)将所得的降水模式识别结果与样本标签进行比较分析,得到相应的输出误差,将输出误差通过隐藏层向输入层逐层反向传播,并将误差分配到各层的所有单元,从而得到所有单元的误差信号,并以此来修正各个单元的权值。

3.1.2 测试过程

测试过程主要是将少量的样本数据利用该系统进

行水凝物识别,并与样本标签进行比较,从而对系统的准确度进行一个整体评价。

对于训练完成的卷积神经网络系统,其各个单元的参数已得到确定,只需将测试样本数据输入到系统中,逐层传播直到输出层输出分类结果。在这里,利用能代表雨、雪以及雹等特殊天气气候的样本数据进行训练,测试过程所得到的准确率可达到 80.32 %。

3.2 降水粒子类型判别系统

通过基于深度学习的降水模式识别系统,得到粒子对应的降水模式类型。在不同的降水模式下,粒子所属的类型也各有差异,故对于不同的降水模式,将采用不同的基于模糊逻辑的降水粒子判别系统,实现降水粒子类型的最优判别。这里将只对降雨模式下的粒子判别系统进行详细说明,其原理结构如图 3 所示。

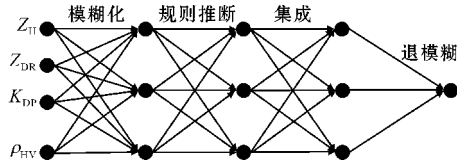


图 3 基于降雨模式的降水粒子类型的模糊判别系统

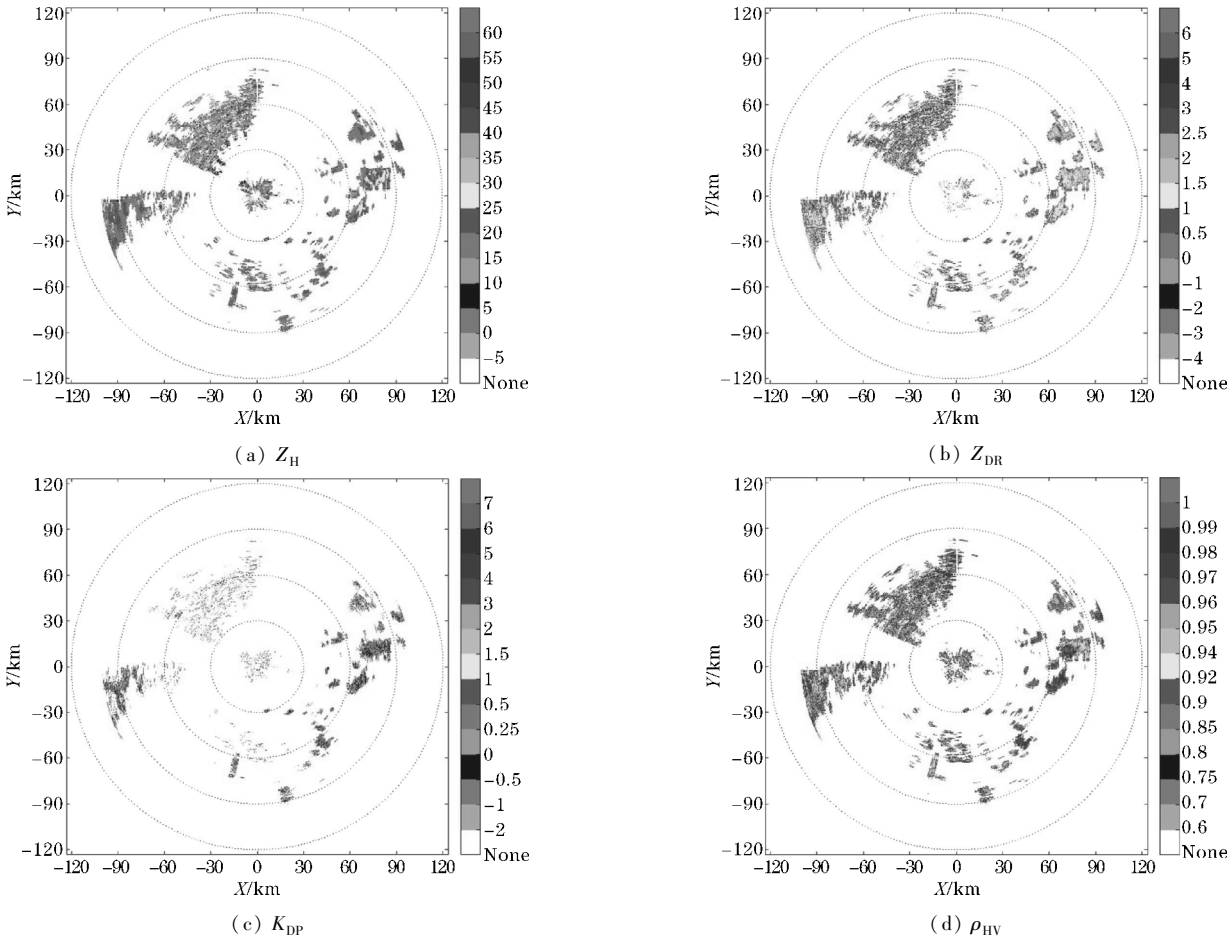


图 4 层状云降水过程的偏振参量产品分布图

该系统以水平反射率  $Z_H$ 、差分反射率因子  $Z_{DR}$ 、差分传播相移  $K_{DP}$  以及相关系数  $\rho_{HV}$  4 个参量为输入,依次通过模糊化、规则推断、集成以及退模糊 4 个过程,得到粒子的最优判别类型,其输出结果与粒子类型的对应关系如表 3 所示。

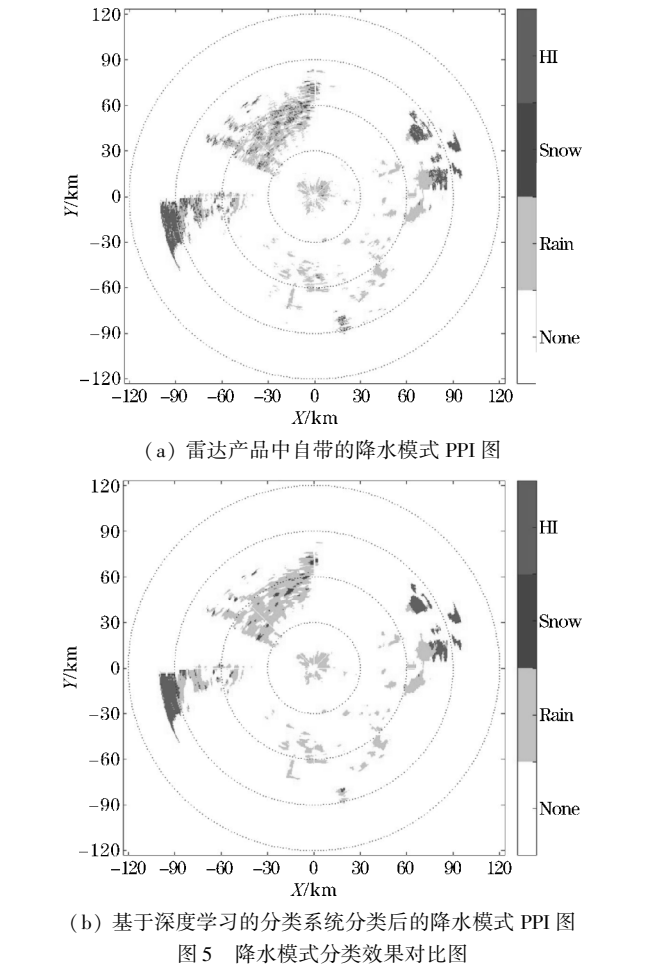
表 3 降雨模式下各降水粒子类型对应的输出结果

降水粒子类型	输出结果
小雨 (LR)	1
中雨 (RN)	2
大雨 (HR)	3

4 仿真与实现

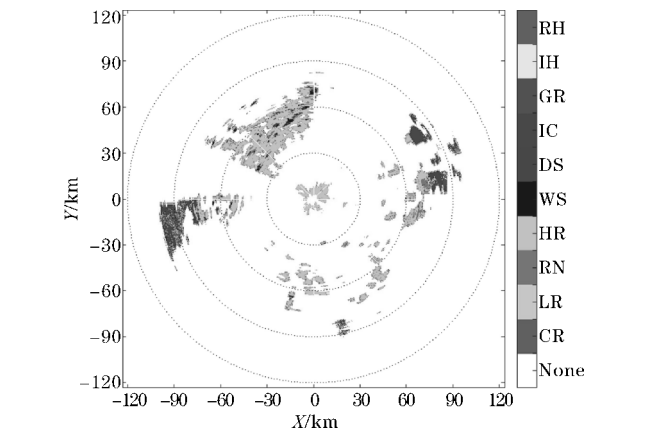
对基于深度学习的降水粒子识别系统的性能进行评估,评估数据来源于 2015 年 7 月 27 日中国四川省成都市双流机场附近区域所探测到的一次层状云降水过程。图 4 为对应探测仰角为  $1.4^\circ$  时,水平射率因子  $Z_H$ 、差分反射率因子  $Z_{DR}$ 、差分传播相移  $K_{DP}$  以及零滞后相关系数  $\rho_{HV}$  的 PPI 图,只针对 100 km 范围内的数据进行处理,这里每个距离圈的间隔为 30 km。

如图 4 所示,深度学习系统以  $Z_H$ 、 $Z_{DR}$ 、 $K_{DP}$  以及  $\rho_{HV}$  的 PPI 数据所构建的特征矩阵为输入,采用卷积神经网络提取相应特征,并结合 Softmax 分类器对粒子的类别进行了最优的判别,其判别结果如图 5 所示。



通过对比分析可见,图 5 (a)、(b) 分布大体一致,其差异主要体现在图 5 (b) 比图 5 (a) 更加平滑,去除了一些杂波的干扰,使分类效果更佳。其主要是由于在深度学习中,所选择的样本数据均来自于该距离库附近 1 平方公里的所有数据,通过对该距离库与附近所有数据的分析,再做出粒子类型的最优判断。虽然这一方式将使系统对数据处理的复杂度加大,但能够去除大量杂波和杂点的干扰,从而大大提高了判别的准确度。

如前所述,通过深度学习系统,得到了降水粒子所对应的降水模式,并选取该降水模式所对应的模糊逻辑判别系统对粒子的类别进行最优判别,最终得到晶体 (CR)、毛毛雨 (LR)、雨 (RN)、大雨 (HR)、湿雪 (WS)、干雪 (DS)、冰晶 (IC)、霰 (GR)、冰雹 (IH) 和雨加雹 (RH) 等降水粒子的 PPI 图,如图 6 所示。



系统主要对 100 km 范围内的水凝物进行分类,其中相邻距离圈之间的间隔为 30 km。如图 6 所示,在距离雷达站 70 km 范围内,其降水粒子多为雨滴。由图 4 可知,该区域内  $Z_H$  处于 10 ~ 25 dBZ,  $Z_{DR}$  处于 -1.5 ~ 1.5 dB,  $K_{DP}$  处于  $-1^\circ \sim 1^\circ/\text{km}$ , 以及  $\rho_{HV}$  处于 0.95 ~ 1, 由这些偏振参量的分布可以推测该区域的降水粒子应为雨滴,与系统最终的分类结果吻合。

如图 6 所示,在距离雷达站 70 km, 高度约为 1.7 km 处开始出现冰晶、冰雹以及霰等降水粒子,如图中的第 3 个距离圈附近区域所示,该区域的下方的降水粒子类型为雨滴,而上方区域为冰雹与冰晶等粒子,由此可判断该区域为融化层。由图 4 可知,该区域的  $Z_H$  主要分布于 25 ~ 40 dBZ、 $Z_{DR}$  主要分布于 1 ~ 2.5 dB、以及  $\rho_{HV}$  主要分布于 0.92 ~ 0.95, 通过这些偏振参量的分布,可以推测该区域为融化层,与系统的分类结果吻合。

## 5 结束语

构建了一个基于深度学习的水凝物识别系统,系统以双偏振雷达产品中的水平反射率  $Z_H$ 、差分反射率  $Z_{DR}$ 、差分传播相移  $K_{DP}$  以及相关系数  $\rho_{HV}$  等多个极化参数作为输入,结合深度学习与模糊逻辑两种方法对输入数进行处理、分析和判别,实现了晶体(CR)、毛毛雨(LR)、雨(RN)、大雨(HR)、湿雪(WS)、干雪(DS)、冰晶(IC)、霰(GR)、冰雹(IH)和雨加雹(RH)等降水粒子的识别。该水凝物识别系统主要由基于深度学习的降水模式分类系统(模块一)和基于模糊逻辑的降水粒子判别系统(模块二)两个模块构成。在模块一中,采用深度学习算法对特征参量矩阵进行特征提取,并利用 Softmax 分类器对所得到的特征参数进行分类判别,实现了雨、雪以及雹等降水模式的分类。在模块二中,考虑到对于不同的降水模式,其包含的降水粒子类型不同,而且同一类型的降水粒子在不同的降水模式下,其表现的特性也存在差异,故对不同的降水模式设计了不同的模糊逻辑判别系统,从而实现了水凝物类型的精准判别。

采用深度学习的算法对水凝物进行初次聚类,随后利用模糊逻辑算法对其进行精确聚类。这种利用深度学习与模糊逻辑两种算法进行联合分类的方式大大提高了水凝物类型识别的准确度,但同时也加大了运算的复杂度。

## 参考文献:

- [1] 刘黎平,王致君,徐宝祥,等.我国双线偏振雷达探测理论及应用研究[J].高原气象,1997,16(1):99-104.
- [2] Bringi V N, Thurai M, Hanesen R. Dual-polarization weather radar handbook[J]. AMS-Gematronik GmbH, 2007(2).
- [3] Schuur T, Ryzhkov A, Heinselman P, et al. Observations and Classification of Echoes with the Polarimetric WSR-88D radar[J]. 2003(1).
- [4] Zhang Z, Zheng G, Wei M, et al. Research on the system of advanced X-band Doppler weather radar with dual-linear polarization capability[C]. Advanced Computational Intelligence (ICACI), 2012 IEEE Fifth International Conference on. IEEE, 2012:739-741.
- [5] 汪月霞,林伙海,何建新,等.双偏振天气雷达降水粒子相态识别研究[C].中国气象学会年会, 2013.
- [6] Galati G, Pavan G. Estimation techniques for rainfall rate using differential phase shift in X-band weather radar[C]. Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. 1998 IEEE International. IEEE, 1998, 1:138-140.
- [7] Erhan D, Courville A, Bengio Y. Understanding representations learned in deep architectures[J]. Department d'Informatique et Recherche Operationnelle, University of Montreal, QC, Canada, Tech. Rep, 2010, 1355.
- [8] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436-444.
- [9] Schmidhuber J. Deep learning in neural networks: an overview[J]. Neural Networks the Official Journal of the International Neural Network Society, 2014, 61:85-117.
- [10] Sutskever I, Martens J, Dahl G, et al. On the importance of initialization and momentum in deep learning[C]. International Conference on International Conference on Machine Learning. JMLR. org, 2013:III-1139.
- [11] Liu H, Chandrasekar V. Classification of hydrometeors based on polarimetric radar measurements: Development of fuzzy logic and neuro-fuzzy systems, and in situ verification[J]. Journal of Atmospheric and Oceanic Technology, 2000, 17(2):140-164.
- [12] Chandrasekar V, Keränen R, Lim S, et al. Recent advances in classification of observations from dual polarization weather radars[J]. Atmospheric Research, 2013, 119:97-111.
- [13] 刘亚男,肖辉,姚振东,等.X波段双极化雷达对云中水凝物粒子的相态识别[J].气候与环境研究,2012,17(6):925-936.

# Study on Hydrometeor Identification based on Deep Learning

RAN Yuan-bo, SUN Min, GAO Meng-qing, WANG Hai-jiang

( College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China)

**Abstract:** Hydrometeor classification for Dual Polarization Doppler Weather Radar echo is a procedure that identifies hydrometeor types basing on the scattering properties of precipitation particles to polarized electromagnetic waves. The difference in shape, size or spatial orientation between different types of hydrometeor will produce different scattering characteristics for the electromagnetic waves in a certain polarization state, Moreover, the polarized parameters, which are calculated from the radar data and closely associated with these characteristics, are also different. The comprehensive use of these polarized parameters can effectively improve the identification accuracy of the phase of various hydrometeors. In this paper, a new identification method of the hydrometeor type basing on deep learning and fuzzy logic methods is proposed; Firstly, the deep learning (DL) approaches is used for extracting the features from the polarized products of the hydrometeors, and the Softmax classifier is applied to classify the pattern of precipitations about rain, snow and hail based on the features extracted by deep learning algorithm. Secondly, the adaptive fuzzy logic algorithm is adoptive for identify the precipitation particles in various precipitation patterns. Finally, the hydrometeor classifier has been applied to astratiform cloud precipitation process, and it is found that the identification result agrees with the other polarized products. The deep learning method is used for initial clustering, then using the fuzzy logic method for accurate clustering results, that greatly improves the accuracy of hydrometeor classification.

**Keywords:** dual polarization doppler weather radar; hydrometeor classification; polarized parameter; deep learning; fuzzy logic; clustering