

文章编号: 2096-1618(2021)05-0512-06

基于人工智能的冰雹天气识别方法研究

陈夏¹, 王海江¹, 周淑玥¹, 洁拉曲铁¹, 李静², 徐自励²

(1. 成都信息工程大学电子工程学院, 四川 成都 610225; 2. 中国民航局第二研究所, 四川 成都 610041)

摘要:冰雹天气是强对流天气系统产生的一种强天气现象,具有历时短、局地性强、破坏力大等特点,是主要的灾害性天气之一,因此对冰雹天气的识别一直是气象领域一个重要的研究课题。提出一种基于人工智能的冰雹天气识别方法,通过使用 Faster R-CNN 深度学习对冰雹天气区域进行识别,在测试集上的识别率和准确率分别达到 87% 和 97%。通过理论分析,结合实际降雹情况及与传统冰雹天气识别方法对比验证的方式,验证了识别结果的准确性及可靠性。实验结果对冰雹天气这种灾害性天气的识别与预警具有重要的参考意义。

关键词:冰雹天气;人工智能;识别

中图分类号:TP183

文献标志码:A

doi:10.16836/j.cnki.jcuit.2021.05.007

0 引言

冰雹天气是发生在大气中的一种剧烈的自然现象,降雹时十分激烈,且具有历时短、局地性强等特征^[1],每年都给人民的生命及财产造成重大损失。

多普勒天气雷达作为一种遥感探测设备,利用发射电磁波的方式探测一定范围内的气象状况,在一定程度上减少了冰雹等灾害性天气带来的危害。然而,天气雷达无法直接告知气象预报员强天气信息,需要依靠人为识别与监测。在人工智能技术不断发展的背景下,人工智能在气象领域的研究与应用成为气象研究者的一种趋势^[2]。

深度学习具有强大的目标提取能力和自学能力,其本质上是具有深层次的神经网络。Hinton 是深度学习的积极推动者,2006 年在科学杂志 Science 上发表了一篇训练 7 层或更多层的深度神经网络的算法^[3],极大地推动了深度学习的发展。特别是深度学习在目标检测与识别领域上,表现出与人类不分伯仲的能力。Sermanet 等^[4]为准确实现目标的检测,提出基于 DCNN 的识别、定位和检测的集成框架,其核心思想是采用多尺度滑动窗口方法对目标进行检测,但该方法检测效率低下。为进一步提升目标检测效率,Girshick 等^[5]提出一种基于区域的卷积神经网络(R-CNN)算法,该算法采取选择性搜索(selective search)方法生成可能存在目标的候选区域,一定程度上解决了目标检测效率低下的问题。虽然 R-CNN 一定程度上提升了目标检测的效率,但是 R-CNN 生成的候选框具有大量重复的问题,存在着重复计算的问题,降低了算法的检

测效率。针对 R-CNN 的一些不足,Girshick^[6]对这两者进行进一步改进,再次提出了一种名为 Fast R-CNN 的快速目标检测方法,大大提升了模型训练和测试速度,同时结果也表明检测性能优于 R-CNN。但该方法也存在着生成目标候选框时耗时较长的问题。在此基础上,Ren S 等^[7]提出了检测速度更快且精度更高的 Faster R-CNN 算法,通过引入 RPN(region proposal network)来替换选择性搜索方法,RPN 能生成高质量的目标候选区域,同时采用共享卷积网络的方式对目标进行特征的提取,大大提高了目标检测效率与准确度,能满足实时性的要求,从而使该算法在目标检测与识别上比其他算法极具优势。2016 年,代季峰等^[8]提出一种基于区域的全卷积网络(R-FCN)方法,实现了大部分参数的共享,在目标检测上与 Ren 等提出的方法相比,检测速度上得到了明显的提升。

多普勒天气雷达数据资料对于冰雹天气的识别具有非常重要的作用。本文提出一种全新的人工智能冰雹识别方法,在冰雹天气在雷达回波上具有相似特征的基础上,使用大量的冰雹天气实测资料分析回波特征,如组合反射率(composite reflectivity, CR)、回波顶高(echo tops, ET)和垂直积分液态水含量(vertical integrated liquid, VIL),以及双偏振雷达的偏振参量,各偏振参量的取值情况。冰雹天气识别原理如图 1 所示。

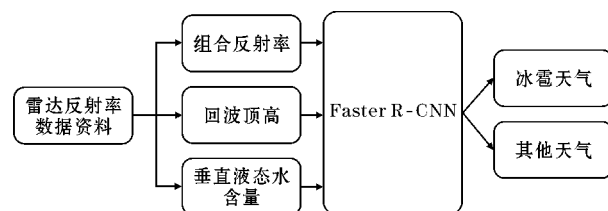


图 1 基于人工智能的冰雹天气识别方法

收稿日期:2021-09-08

基金项目:国家自然科学基金资助项目(U1733103)

1 数据与方法

1.1 数据

本文构建了较丰富的样本数据集,收集了中国 2002–2010 年一些地区的 CINRAD_SA 单极化冰雹天气历史数据;以及 2016 年 3 月 17 日广东省和 2021 年 3 月 30 日重庆的双极化天气雷达冰雹天气历史数据。从大量的冰雹天气历史数据集中,筛选了约 2000 个冰雹天气过程,用于深度学习模型算法的训练、验证和测试。

以反射率因子数据资料为基础,根据水文气象算法,反演生成组合反射率、回波顶高和垂直积分液态水含量产品,作为深度学习的输入数据,实现冰雹天气的准确识别。对 CR、ET 和 VIL 数据进行预处理后,将 3 个二维的 400×400 数据合成 $3\times 400\times 400$ 的数据,类似灰度图的 R、G、B 三通道,作为深度学习冰雹天气识别模型的输入数据,合成操作过程如图 2 所示。

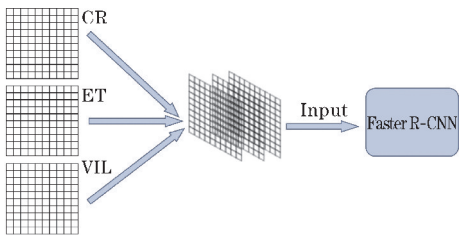


图 2 数据合成过程

深度学习输入数据多为人为标记目标后的数据,本次根据冰雹天气在 CR、ET 和 VIL 的特征标志,结合冰雹天气发生的实际情况,利用标记软件对冰雹天气发生区域进行人工标记;由于冰雹天气本身的特点,局地性强且历时短,所以在 2000 个冰雹天气过程中,满足冰雹发生特征的雷达数据只有 1~3 个体扫。最终形成 4000 个训练数据集,500 个测试集,样本数据集如表 1 所示。

表 1 样本数据集

数据名	数据大小	数据量/个
训练集	$400\times 400\times 3$	4000
测试集	$400\times 400\times 3$	500

1.2 方法

1.2.1 天气雷达产品

根据降水粒子对回波散射的原理,粒子大小对回

波强度的贡献更大,粒子多不如粒子大。在含有大粒子的强对流降水天气,如冰雹天气含有冰雹粒子,此时后向散射强度很大,导致冰雹天气的反射率因子异常大。冰雹云区域的雷达反射率因子,在 CR 上的反映一般能高达 50 dBZ 以上,但仅用 CR 不能很好地区分冰雹天气。ET 是衡量强对流天气强度的一个重要产品,能很好地反映降水云体内上升气流的强弱情况。ET 值很大的区域,说明降水云体内的上升气流很强,而上升气流能决定冰雹直径,强上升气流能托住大雹粒。因此,ET 也是冰雹天气的一个重要特征指标,冰雹天气的 ET 至少能达到 12 km,最高可达 20 km 以上,可以作为是否发生冰雹天气的一个重要判断依据。除 CR 和 ET 外,VIL 也是判别冰雹天气的有效工具,是冰雹识别业务中使用率最高的产品之一。VIL 值代表在某确定的底面积上垂直柱体内的液态水总含量,由于冰雹单体并非由液态水滴构成,使冰雹天气区域的 VIL 值异常大。冰雹粒子的散射能力远比雨滴粒子大得多,在冰雹天气发生的区域,VIL 可高达 40 kg/m^2 以上。因此,VIL 有助于识别较大的冰雹单体,是识别冰雹云的一个重要指标,也是识别冰雹最直观的产品。

1.2.2 Faster R-CNN

Faster R-CNN 算法是目标检测与识别领域较优异的算法,由两大模块构成,RPN 候选区域提取模块和 Fast R-CNN 检测网络模块。在结构上,Faster R-CNN 可以看成由特征提取(feature extraction),区域建议生成网络 RPN,区域池化(Roi pooling)及分类(classification)构成。将这些模块整合到一个网络中,使 Faster R-CNN 在检测速度和精度上极具明显的优势。

特征提取层采用了由牛津大学计算机视觉团队(visual geometry group,VGG)提出的 VGG16 网络^[9],对本文的输入数据进行特征提取。该网络作为 CNN 的一种,在迁移学习任务中表现优异。VGG16 具有 13 个卷积层,13 个激励层,4 个池化层。

RPN 可以快速且准确地生成高质量的检测框,代替传统检测框的生成方式,极大地提高了目标检测效率。特征图经过 RPN 后,先利用大小为的卷积核进行卷积处理,相当于每个点又融合了周围的空间信息,同时输出深度 512 不变,即生成 512 张特征图,相当于特征图上的每个点对应的数据大小为 $1\times 1\times 512$ 。RPN 网络对每个特征点配备 k 个矩形窗口,称为 anchor。每个 anchor 对应 4 个坐标值(x_1, y_1, x_2, y_2),表示矩形左上和右下角点坐标,从而确定了每个矩形框的位置。RPN 分成两条线,一条通过卷积后,把特征图的

数据转换为 $1 \times 2 \text{ k}$ 大小的图像,这对应特征图上每一个点都有 k 个 anchors,而且每个 anchor 可能是含有目标的正候选矩形框(前景)和无目标的负候选矩形框(背景),形成 $1 \times 2 \text{ k}$ 的数据送入 Softmax 进行分类计算属于前景和背景的分数值。RPN 网络的第二条线,经过卷积后,把特征图的数据转换为 $1 \times 4 \text{ k}$ 大小的图像,对应特征图上每个点都有 k 个 anchors,每个 anchors 有 4 个用于回归的变换量(d_x, d_y, d_w, d_h)。最后综合所有的位置回归变换量和前景检测框,同时剔除尺寸非常小的前景检测框,对最终候选区域做出建议。

针对不同大小的候选区域, Roi pooling 层收集 VGG16 最后一层得到的特征图和 RPN 生成的候选区域作为输入数据,将后者映射到特征图上,得到特征图上候选区域对应的特征区域;同时将特征图上的这些区域进行水平分和垂直平分固定大小的网格;然后利用最大值池化对每一个网格进行采样,实现固定长度的输出,最终输出固定大小为 $7 \times 7 \times 512$ 的特征图,送入到后面的全连接层来进行目标识别和定位。采用两个全连接层对已固定维度的特征图进行特征的综合整理,用于类别的分类和位置的回归,再通过使用 Softmax 计算每个候选区域所属类别的分数值,并输出概率向量;同时,再次对候选框进行位置调整,最终获得最精确的目标识别框。

2 实验设计

在本次深度学习实验中,采用的开发平台为 Python+Tensorflow,运行环境为:NVIDIA Tesla P40, CUDA 9.1, Ubuntu 16.04, memory 24 GB。采取了交替优化方法(alternating optimization)训练 Faster R-CNN,即把整个网络模型分成两个部分进行训练,一部分是检测网络,另一部分是 RPN,详细的训练步骤如下:

步骤 1,获取预训练好的 VGG16,迭代训练 RPN 网络;步骤 2,使用步骤 1 训练好的 RPN 网络,对检测网络进行训练;步骤 3,第二次训练 RPN 网络,微调权值参数;步骤 4,利用步骤 3 中训练好的 RPN 网络,再次训练检测网络,微调检测网络中的权值参数。

对于 RPN 网络的训练,并不是用全部 anchors 拿去训练,而是利用 IoU(intersection-over-union)^[10] 为衡量标准,选取一部分正候选框和负候选框进行训练。

测试过程使用一些带标签的样本数据,通过将冰雹识别系统的输出与标签进行比较来评估系统的准确性。在本研究中,使用大量的冰雹天气有用样本数据

对该系统进行了测试,发现该系统在测试集上的识别率达到 87%,准确率达到 97%。

3 实验结果与分析

以 2008 年 6 月 25 日天津一次冰雹天气过程为例,采取理论分析并结合实际降雹情况,对基于人工智能的冰雹天气识别方法的识别结果进行验证分析。为进一步验证识别冰雹天气的准确性及合理性,还收集了双偏振雷达冰雹天气历史数据,2016 年 3 月 17 日广东省韶关市的冰雹天气过程为例,结合曹俊武等^[11]提出的双偏振多普勒天气雷达识别冰雹区方法,进一步验证冰雹天气识别方法的识别结果。

天津冰雹天气的数据来源于 2008 年 6 月 25 日 0754UTC 至 2008 年 6 月 25 日 0854UTC 的塘沽 SA 雷达连续体扫数据。据滨海新区气象中心当日最新数据资料描述,在天津塘沽区、开发区等局部地区在下午不同时间段的不同地点发生了强烈的冰雹天气,并伴有短时大风天气,在不同地区都出现了一定的灾情。0754UTC ~ 0854UTC 的天气过程是一次典型的冰雹天气过程,此次冰雹天气强度大、密度高、持续时间相对较长,其最大冰雹直径达 40 mm。

选取 0754UTC ~ 0800UTC 的雷达体扫反射率因子资料,根据气象算法理论,反演出 CR、ET 和 VIL 产品。体扫时间内的 3 种产品分布结果如图 3 所示,图 3(a)、(b)、(c)分别对应 CR、ET、VIL 的图像,根据 CR、ET、VIL 资料图,在 0754UTC ~ 0800UTC,塘沽 SA 雷达所在位置的右上方区域发生着冰雹天气。这段时间内发生的冰雹天气,其最强回波区域的最大 CR 达 60 dBZ 以上,ET 达到 12 km 以上,其中 ET 设置的阈值为 30 dBZ,同时最强回波区域的最大 VIL 也达 70 kg/m^2 ,反映了此次冰雹天气的特征。通过此次天气过程在 3 种产品上的特征,结合当地的气象资料描述情况,该区域对应的天气过程可以确定为冰雹天气过程。

利用该时间段内的雷达体扫数据资料为基础,以 CR、ET、VIL 产品作为输入数据,利用深度学习算法 Faster R-CNN 对以上不同时间段内的冰雹天气过程进行识别,其识别结果如图 4 所示。Faster R-CNN 通过利用提取的冰雹天气特征准确识别冰雹天气,其最终输出最为精确的冰雹天气识别矩形区域框,并给出冰雹天气发生的置信概率,把该识别框映射到以 CR、ET 和 VIL 为基础的图像上,如图中红色边框区域。

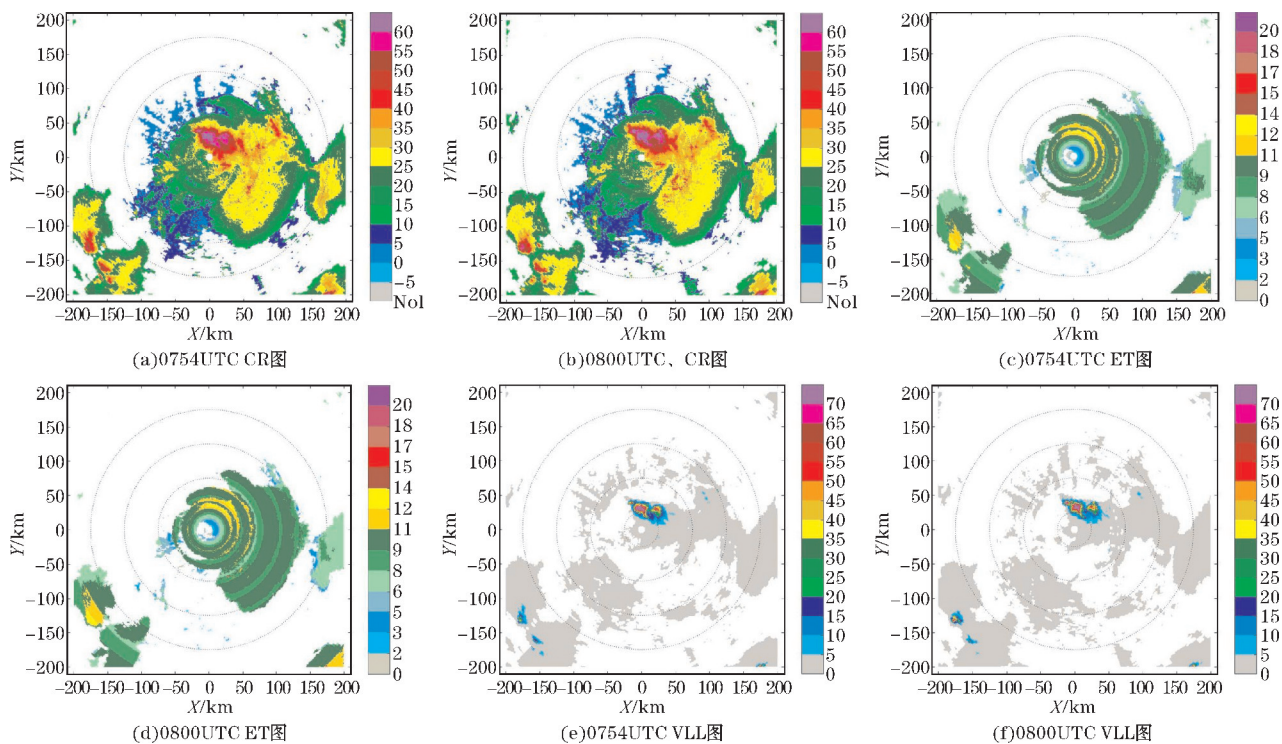


图3 0754UTC、0800UTC 的雷达数据资料



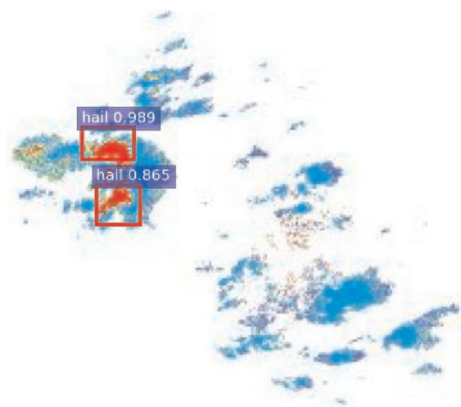
图4 0754UTC ~0800UTC 冰雹天气识别结果

图4中的红色区域为CR、ET和VIL上符合冰雹天气特征的区域,该区域CR至少满足50 dBZ以上,ET至少满足12 km以上及VIL大于 45 kg/m^2 ;其他区域则认为是其他天气过程,图中用绿色表示这些区域。从图4中可以看出,符合冰雹天气特征的区域,Faster R-CNN能快速、精确地识别发生冰雹天气的区域位置,识别结果同理论分析的结果基本一致,验证了该识别方法的可靠性。在0754UTC~0800UTC识别出两个冰雹天气过程区域,该区域也与图3确定的冰雹天气区域相吻合,且两个区域相距较近且范围较小,可见Faster R-CNN冰雹天气识别精度较高。此外,也准确识

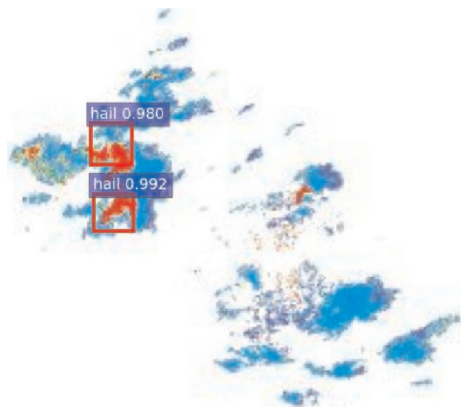
别了新出现的冰雹天气区域。

为进一步验证识别结果的正确性及合理性,还收集了双偏振雷达冰雹天气历史数据,数据来源于2016年3月17日广东省韶关市的冰雹天气过程为个例,结合传统的冰雹天气识别方法,对方法进行验证分析。这次天气过程具有明显的冰雹天气特征,是典型的冰雹天气过程。传统的冰雹天气识别方法以双偏振天气雷达某一层的测量参量作为输入数据来识别冰雹天气区域,其输入变量分别为水平反射率因子 Z_H ,差分反射率因子(ZDR),零滞后相关系数(CC),差分传播相移率 K_{DP} ,然后通过模糊逻辑算法识别冰雹天气

发生区域。实验表明,该传统方法能有效地识别冰雹的具体位置,具有较高的可信度。本次选取了偏振雷达第一个仰角层的测量参量作为传统方法的输入数据,将本文的识别方法与文献[11]的识别方法进行对比验证,冰雹天气过程识别结果对比如图5所示。图中红色框为本文方法的识别结果,同时给出了置信概率,红色区域为文献[11]方法识别结果,其他区域则认为是其他天气过程。可以看出,传统冰雹天气识别方法与本文方法的识别结果基本一致,进一步验证了基于人工智能的冰雹天气识别方法的准确性和可靠性。但准确率和识别效率上,本文方法具有较大的优势,且避免了传统方法需人为选取特征的问题,具有强大的特征提取能力和较强的适用性。



(a) 1030UTC 识别结果



(b) 1036UTC 识别结果

图5 识别结果对比

4 结束语

根据冰雹天气在组合反射率、回波顶高及垂直积分液态水含量产品资料上的主要特征,收集中国大量的冰雹天气历史数据,结合现有的深度学习目标检测

与识别算法,建立基于 Faster R-CNN 的冰雹天气识别方法。该方法以组合反射率、回波顶高及垂直液态水含量作为输入数据,采用 VGG16 对输入数据进行特征提取,RPN 网络生成候选区域,Roi Pooling 实现输入图像大小的固定,最后利用全连接层和 Softmax 分类器对冰雹天气进行判断,最终实现冰雹天气的准确识别。该方法在测试集上的识别率达 87%,准确率为 97%。为验证本文识别方法的准确性及可靠性,采取理论分析、与实际降雹情况以及与传统冰雹识别方法对比验证的方式,验证了本方法的准确性及可靠性,识别的结果与实际降雹情况相符,也与传统的冰雹天气识别方法结果相一致。同时,提出的基于 Faster R-CNN 的冰雹天气识别方法,可采取 GPU 等外部硬件进行加速,能有效减少训练的开销,可实时做到冰雹天气的临近预报要求;此外,该方法具有强大的特征提取能力和较强的自学习能力,具有较强的适用性,可推广使用到冰雹天气的自动识别与预警中。

致谢:感谢民航专业项目(ASSA2021-61)对本文的资助

参考文献:

- [1] 慕建利. 陕西关中强暴雨中尺度对流系统研究[D]. 江苏:南京信息工程大学, 2009.
- [2] 赵泽皓, 孔誉旋. 人工智能应用于气象业务的现状及发展趋势[J]. 河南科技, 2019, 690(28):152-154.
- [3] Hinton G E. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 37(6):104-112.
- [4] P Sermanet, D Eigen, X Zhang et al. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks[J]. arXiv preprint arXiv, 2013, 1312:6229.
- [5] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2014.
- [6] Girshick R. Fast R-CNN[J]. Computer Science, 2015, 34(6):904-917.
- [7] Ren Shaoqing, He Kaiming, Girshick Ross, et al. Faster R-CNN: Towards Real-Time Object Detec-

- tion with Region Proposal Networks [J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6).
- [8] Dai J, Li Y, He K, et al. R-FCN: Object Detection via Region-based Fully Convolutional Networks[J]. Advances in Neural Information Processing Systems, 2016, 30(2):190–200.
- [9] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science, 2015, 34(3):186–196.
- [10] Jiang B, Luo R, Mao J, et al. Acquisition of Localization Confidence for Accurate Object Detection[C]. European Conference on Computer Vision (ECCV), 2018.
- [11] 曹俊武,刘黎平. 双线偏振多普勒天气雷达识别冰雹区方法研究[J]. 气象, 2006(6):13–19.

A Hail Weather Recognition Method based on Artificial Intelligence

CHEN Xia¹, WANG Haijiang¹, ZHOU Shuyue¹, JIE Laqutie¹, LI Jing², XU Zili²

(1. College of Electronic Engineering Chengdu University of Information Technology, Chengdu 610225, China; 2. The Second Research Institute of CAAC, Chengdu 610041, China)

Abstract: Hail weather is a strong weather phenomenon produced by the system of severe convective weather, which has the characteristics of short duration, strong locality and great destructive power. Therefore, the detection of hail weather has always been a significant research topic in the meteorological field. This paper presents a hail weather recognition method based on artificial intelligence. In this method, the Faster R-CNN deep learning is applied to detect hail, and accuracy and precision on the test set are 87% and 97% respectively. The reliability and accuracy of the method are verified by adopting theoretical analysis and comparing with the actual situation and the traditional hail weather identification algorithm. The experimental results have significant reference for the identification and early warning of hail weather.

Keywords: hail weather; artificial intelligence; tracking