

基于深度学习的通信信号调制与识别研究

康素成

(盐城师范学院 物理与电子工程学院 江苏 盐城 224007)

摘要: 通信信号调制识别面临准确率低、特征提取计算复杂等问题,采用深度学习算法建立信号调制识别模型,可以利用卷积神经网络设计出准确率高、运算快速的信号识别分类器。通过建立样本集对模型展开训练和验证测试,可知在载噪比达到 6dB 以上时,采用该分类器对多数通信信号可以取得理想调制识别效果,能够为调制识别中频信号提供有效方法。

关键词: 深度学习;通信信号;调制与识别

引言: 在无线电、通信侦察等领域,通信信号调制识别为重要研究课题。采用该技术对未知信号进行自动识别,可以通过正确解调获得信息,对加强频谱监管具有重要意义。结合以往研究来看,主要通过特征提取模式进行信号调制模式分类,但采取人工提取方法导致计算过程复杂,导致方法分类性能不佳。而深度学习算法的出现,为提高信号调制识别准确率提供了新思路,因此应加强基于深度学习的通信信号调制识别研究,从而推动信号调制识别技术发展。

1 通信信号调制与识别理论

对通信信号进行调制识别,关键在于提取信号特征。而采取不同信号调制模式,包含的特征信息存在差异,因此提取信号特征应凸显调制信息,确保可以准确识别通信信号。现阶段,主要需要提取的信号特征包含谱特征和高阶累积特征。一方面,提取谱特征,主要是由于调制信号瞬时变化多样,包含相位、幅度、频率等,可用于对未处理信号参数进行估算。通过准确提取原始信号特征,则能提高信号调制识别准确率。对数据集信号展开分析,需要对瞬时幅度的谱密度最大值、绝对值标准差等特征值展开分析,并对瞬时相位非线性分量、瞬时频率绝对值标准偏差等数值进行提取,用于确认是否含有与幅度、相位等相关的信息^[1]。另一方面,对信号积累量特征进行提取,主要适用于复杂调制方式,需要对上述特征量进行重复高阶计算,排除瞬时信息受到的噪声等因素干扰。根据高阶计量特征完成信号分类,可知调制信号将受到高斯白噪声等信号干扰,但噪声在二阶以上统计量为零,因此能够通过高阶累积量分辨调制信号^[2]。在不同调制模式下,信号的高阶累积量特征存在差异,能够为准确识别调制方式提供依据。

实际在通信信号传输过程中,发射天线端多采用基带信号,具有频率低、波长长的特点,受传输距离限制,仅能发射和接收与自身尺寸接近的信号。将基带信号转换为适合通信的形式,一般采用高频载波调制技术,通过变化参数携带信息,有效提高信道利用率,增强信号传输抗衰落、抗干扰能力^[3]。目前,广泛应用数字调制方式简化调制流程,包含频移键控 MFSK、幅移键控 MASK、相移键控 MPSK 等多种模式。其中,采用 MFSK 模式将通过载波变化频率进行对应码元信息传递,在频率不同的情况下,载波拥有相同能量,确保信号幅值不变。利用频率间隔确定载频谱线分布情况,能够反映载波信号变化。采用 MASK 模式,利用载波幅度进行数字信息传递,基带信号和载波信号乘积将通过带通滤波器处理后传输。采取 MPSK 模式,可知信号载波拥有 M 种形式,能够根据基带码元对信号时域表达展开分析。从总体上来看,时域信号能量容易受噪声干扰,导致提取到的特征质量差,给信号识别带来困难。而频域信号无法根据时间确定变化趋势,在低信噪比下同样难以有效提取特征。

2 基于深度学习的通信信号调制与识别

2.1 深度学习算法

采用深度学习算法进行通信信号调制识别,应用较广的包含 BP (反向传播误差算法)网络、RNN (循环神经网络)和 CNN (卷积神经网络),具体还需结合调制信号分布形式合理建立网络模型结构和选择参数,确保可以准确识别信号调制模式。采用 BP 网络,包含一个输入层和输出层,剩余为隐含层,各层包含多个神经元,采用全连接方式。各神经元拥有相同计算方式,通过 BP 算法进行权重参数更新,可以减小学习误差,达到局部最优。利用梯度下降法进行网络参数迭代优化,减小函数损失,能够增强模型准确性。采用 RNN 网络,包含输入层、循环层、全连接层和输出层,利用循环层提取特征,通过全连接层进行拼接输出,达到分类识别目标^[4]。由于神经元输出依赖上一个输入,可以形成封闭传输形式。在内部设置隐藏节点保留上一个输出数据状态,参与各时刻计算,可以使网络模型获得较强处理序列样本能力。一旦上一刻信息存在误差,将引发网络误差积累问题。采用 CNN 网络,带有共享参数和局部感知优势,在信号特征提取上表现出较高敏感度。CNN 网络包含输入层、隐含层和输出层,其中隐含层由卷积层、池化层和全连接层构成,依靠卷积层提取数据特征,并通过池化层筛选冗余信息,降低数据维度,能够简化数据运算,最终通过全连接层将提取到的特征应设置分类空间。对载波信号进行调制,基带信息不受影响,仅对信号进行线性计算、移动等操作,信号分布不变,类似图像处理流程。相比较而言,采用 CNN 网络能够实现多函数共享参数,将各部分信息整合得到全局信息,降低运算参数数量的同时,自动、稳定提取调制信号特征,能够加快网络训练速度,获得较高识别性能。按照算法流程,在输入样本数据后向对低层特征进行提取,然后提取高层特征,完成从特征到目标的映射,实现结果输出。

2.2 训练样本集建立

运用深度学习算法,首先需要做好样本收集和制作,确保建立的样本集数量足够,可以全面覆盖通信技术应用场景。正确标注样本,使训练处的模型获得良好鲁棒性,能够提高网络模型输出数据准确度。但在现实生活中,通信信号参数范围较广,接收到的信号数量有限,其中包含许多不完整的信息,难以达到 CNN 网络训练要求。为解决问题需对真实信号进行模拟,采用 Python 软件生成模拟信号和实现数据处理,随机设置载波频率、信噪比等参数,保证样本集完整、有效。从通用性角度确认信号频率参数范围,对信号频率进行归一化处理,将载波和采样频率比值设定为 1/4 定值,可以得到采样频率为 93.3kHz,得到中频信号频谱。考虑载波频率测量将受到信道、接收机等因素影响,设置误差为 $\pm 8\%$ 的符号速率。采样率和符号速率比值在 4~20 倍范围,符号速率在 4~24kHz 之间,载噪比在 2~20dB 范围。信号强弱存在差异,无法取固定振幅值,采用零均值标准化线性函数进行归一化处理,能够得到均值为 0、方差为 1 的数据集。接收到信号频带宽度可能较大,存在多个信号和高斯白噪声。采用 FIR 带通滤波器进行滤波处理,能够通过非递归的滤波处理保留信号的幅频和相频特性。根据频带宽度设置频率

上下限,可以利用 MATLAB 中的滤波器函数实现滤波操作,最终得到包含 280000 个样本的数据集。

2.3 分类器设计与训练

采用 CNN 网络进行信号特征提取,将相应的分类器模型,可知卫星通信信号样本仅为二维时间序列,可以建立二维网络模型,符号数量需要达到 1 万以上,因此需要设置 5 万个数据采样点进行分类器训练,确保可以提取全面的信号调制特征。网络输入长在 5000~50000 之间,对不同卷积层数网络性能进行分析,最终确认卷积核长在 10~20 间的网络性能最佳。使用 ReLU 激活函数对各层卷积层进行激活,利用函数常数抑制训练中的梯度消失问题,建立的网络模型自动完成时域信号隐含特征提取。经过平均池化层处理,将卷积核运算后得到全局平均值,即一维的特征向量,降低参数数量。全连接层采用 softmax 激活函数进行特征整合,可以完成信号调制特征分类。在分类器设计上,采用 4 层 CNN 网络结构,网络输入长为 5000,首层卷积层包含 64 个长为 20 的卷积核,2~4 层的卷积核结构相同,均包含 128 个长 10 的卷积核,并加入 dropout 层每次将 30%参数置零。采用 Keras 深度学习框架,将 Tensorflow 当成是后端进行计算,利用 Python 接口进行编译操作。采用的框架属于高层神经网络 API,能够调用 GPU 加速计算,快速完成模型设计和训练,为 CNN 和 RNN 等模型结合提供支持。采用开源深度学习框架,利用交叉熵损失函数进行分类,如式(1)所示,通过降低损失函数值对隐藏层参数进行更新,达到优化分类结果的目标。

$$L_i = -\sum_j t_{i,j} \log(p_{i,j}) \quad (1)$$

式中, L_i 为损失函数, t 为真实结果, i 为输入数据, j 为类别, p 为预测结果。应用 Adam 优化器降低损失值,需要将学习率设定为 0.001,dropout 设定 0.5 随机失活率,避免因全连接层参数过多引发拟合问题。

对分类器进行训练,可知样本集包含 BPSK、8PSK、16QAM 等 9 类卫星信号,每类取 2 万个样本数据进行训练,网络输入维度为 (None,length,1),首个维度 None 用于表述样本数量,length 表示网络输入长,1 指的是通道数。在输入长为 5000 的情况下,将样本集划分为训练集和验证集两部分,前者占总样本数 80%,后者占 20%。利用训练集对分类器进行训练,可知将对误差的反向传播参数进行更新。按照流程向对信号幅值进行归一化处理后,使用滤波器完成载波频率和符号速率滤波,可以得到深度学习框架。将处理后的数据送入网络,使用 Adam 方法进行训练,将 batch-size 设定为 1,指的是一次送入 16 个样本,训练周期数为 70 次。将四层神经网络当成是基础开展训练,实际需要对卷积核尺寸进行调整,使卷积核尺寸分别达到 5、10、50、100,分别对不同网络结构展开训练,得到对应分类器模型。此外,需要对网络输入长度进行调整,使输入长度分别达到 5000、10000、20000、50000,训练得到相应模型。

2.4 仿真测试分析

利用深度学习算法建立 CNN 分类器,为验证算法的有效性需要将验证集样本数据输入到卷积核尺寸和网络输入长度不同的网络模型中,对信号调制识别性能进行比较。从分析结果来看,随着卷积核尺寸从 5 增长为 100,信号调制识别准确率经历了先增加后减小的变化趋势,在卷积核尺寸为 10 时准确率最大,能够达到 82%,在卷积核尺寸达到 100 时准确率最小,仅为 66%。由此可见,为优化分类器性能,需要将网络模型卷积核尺寸设定为 10。从网络输入长度变化情况来看,随着输入长度由 5000 逐渐增加至 50000,信号调制识别准确率逐步提升,从 82%提升至 91%。由此可见,输入信号长度越大,识别准确率越大。分析原因可知,信号长度更大意味着将包含更多特征信息,使网络模型提取更多隐含特征,能够为准确识别信号调制模式提供支持。而随着训练数据的增加,需要适当

延长训练周期。从算法收敛情况来看,经过 100 次循环迭代后模型能够达到收敛状态。

通过反复测试验证,可以得到最优的网络模型分类器,为四层 CNN 网络,卷积核尺寸为 10,输入长度为 50000。对网络性能展开测试,重新输入 1300 个样本数据,包含 9 类卫星调制信号,且每类信号在 2~20dB 载噪比下产生 150 个样本,通过测试确认不同载噪比下信号调制识别情况。测试结果如表 1 所示,根据表中数据可知,BPSK 和 OQPSK 信号在载噪比达到 2dB 时即可达到 100%的准确率,而 P14QPSK 和 QPSK 在载噪比达到 6dB 时能够达到 100%准确率,其余信号也能在载噪比达到 6dB 或 10dB 时达到 90%以上准确率,因此设计的分类器拥有良好信号调制识别性能,在载噪比较高时可以获得更高准确率。从总体上来看,BPSK 和 OQPSK 等信号属于相位调制信号,准确率稍高,而正交振幅调制信号准确率稍差,主要是由于星座图接近、容易出现相互混淆问题,进而导致识别准确率下降。采用建立的网络模型对超短波信号进行调制识别,可知该类信号符号数有限,多数不超 1 千个,采样点数可以设定为 5000。具体来讲,就是将网络输入长设定为 5000 后重新进行训练和验证,利用得到的分类器进行信号调制识别验证。测试信号除了包含上述 9 类,还包含 2ASK、2FSK、4FSK、MASK、GMSK,在 2~20dB 载噪比下产生 150 个样本,然后单独进行测试。从测试结果来看,各类信号调制识别准确率有所下降,但对 2ASK、2FSK、4FSK 信号进行识别,在 2dB 载噪比下可以达到 100%准确率,MASK 信号在 6dB 信噪比下达到 90%以上准确率,对 GMSK 信号识别效果不理想。分析原因可知,与采样点数减少和识别类型增加有关。从总体来看,在载噪比达到 6dB 以上条件下,用于通信信号调制识别效果较好。

表 1 卫星信号调制识别准确率

载噪比/dB	2	6	10	14	18	20
BPSK	100	100	100	100	100	100
QPSK	38	100	100	100	100	100
8PSK	0	90	100	100	100	100
OQPSK	100	100	100	100	100	100
P14QPSK	77	100	100	100	100	100
16QAM	2	40	88	100	100	100
32QAM	1	37	98	100	100	100
16APSK	0	37	98	97	100	100
32APSK	0	51	95	100	100	100

结论:在通信信号调制识别方面,为准确提取信号特征,采用深度学习算法自动提取主要特征信号。实际采用算法建立 CNN 网络模型和样本集,通过做好数据归一化处理加快运算速度,并在分类器设计和训练中进行网络性能测试,筛选出最优网络参数。根据仿真结果可知,算法拥有理想分类识别能力,对完善相关理论和发展信号调制识别技术具有重要意义。

参考文献:

- [1]张海燕,闫文君,张立民,等.通信信号调制识别综述[J].海军航空大学学报,2022,37(01):126-132.
- [2]张磊,吴颖.基于深度学习算法的 HPLC 通信信号自动调制识别研究[J].通信电源技术,2020,37(10):46-48.
- [3]侯涛,郑郁正.基于深度学习的通信信号调制方式识别[J].无线电工程,2019,49(09):796-800.
- [4]李唱白,杨杰,黄知涛,等.基于深度学习的通信信号调制识别算法[J].空间电子技术,2019,16(01):49-54+74.