



基于毫米波特征图的行为感知的研究与实现

1023041002 王施涛

南京邮电大学 计算机学院、软件学院、网络空间安全学院 软件工程

目录

CONTENTS

01

研究背景

02

数据处理与系统设计

03

实验结果与分析

04

总结与展望

厚德 弘毅 求是 笃行



PART 01

研究背景

研究背景

随着社会的发展和进步，人机交互技术变得越来越普及。人类的行为识别技术是人工智能领域的重要研究方向，有着举足轻重的地位，在智能家居、医疗保健和健身跟踪等诸多领域中发挥着重要作用，具有广泛的发展前景。传统的基于视觉的行为识别技术虽然已经相对成熟，但摄像头等设备采集到的图像或视频数据在光照较强或光线昏暗的条件下，图像的识别精准度会下降，同时视频数据往往会造成隐私问题。而无线方式可以有效避免这些问题。使用毫米波雷达是通过雷达的回波信号获得人体行为动作的运动特征，实现行为的感知与识别。该识别方式具有不接触物体进行识别的优点，且具有较高的采集精度，因此它可以将信息的交互变得便捷高效，在物体检测，距离、角度以及速度等信息的测量上具有重要意义。



PART 02

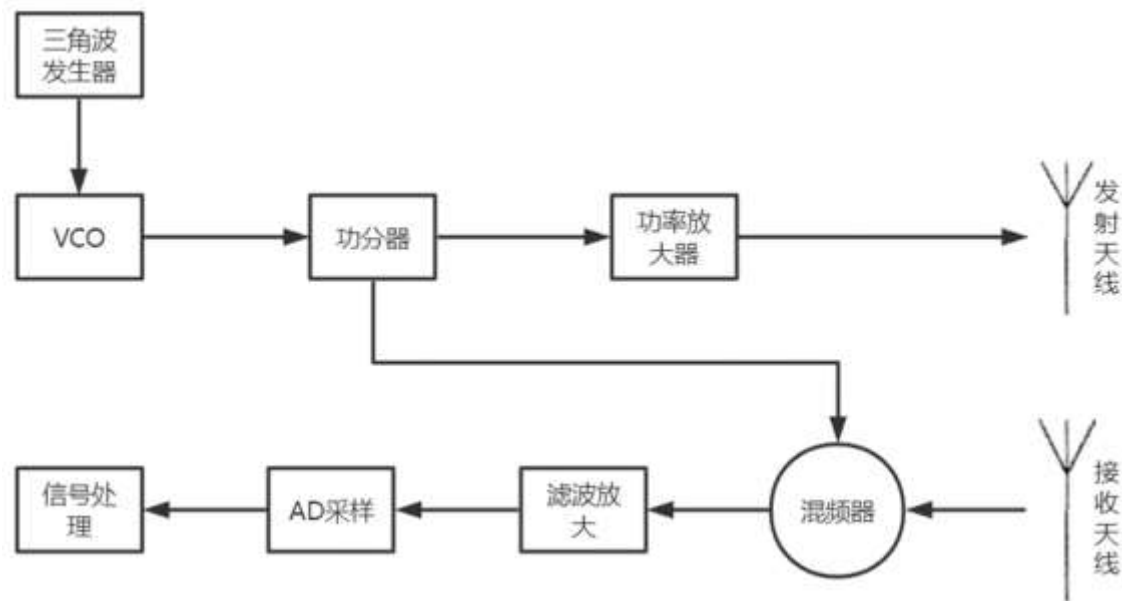
数据处理与系统设计

实验设备

本文实验中所采用的设备是DCA1000采集卡，其集成了FMCW雷达的特点，采集范围为水平方向 120° ，垂直方向 120° 。使用mmWave Studio软件进行采集参数的配置，包括静态设置和波形设置等，采用三发四收的机制，进行数据的采集。该设备具有较高的距离分辨率、不存在距离盲区且不易被截获。



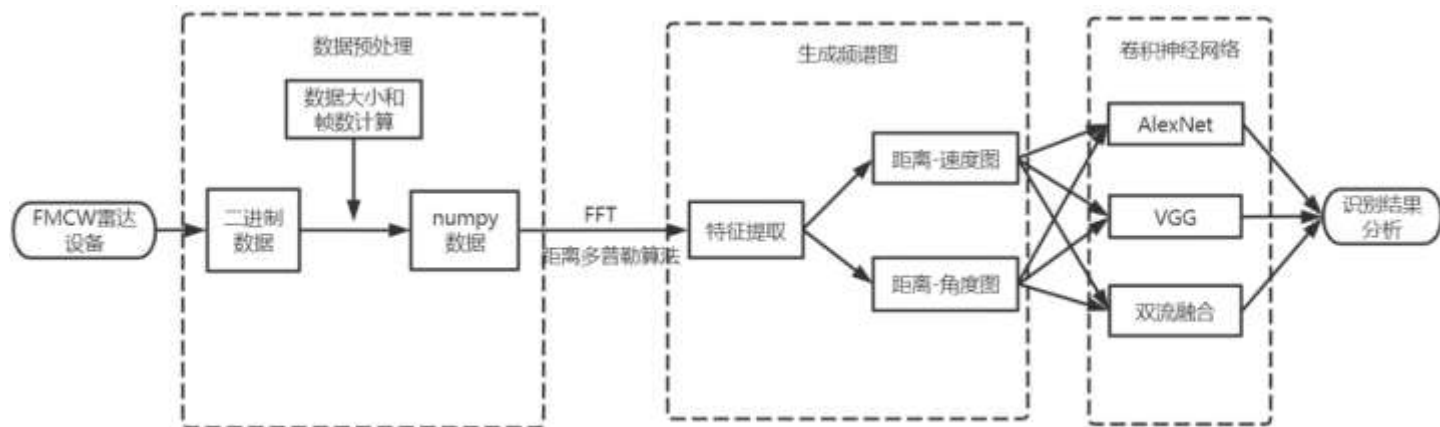
DCA1000采集卡



FMCW雷达系统框架



系统框架设计



系统设计框架

第一步：数据采集。调整好设备参数，利用FMCW雷达采集数据，志愿者在设备前两米左右做动作。

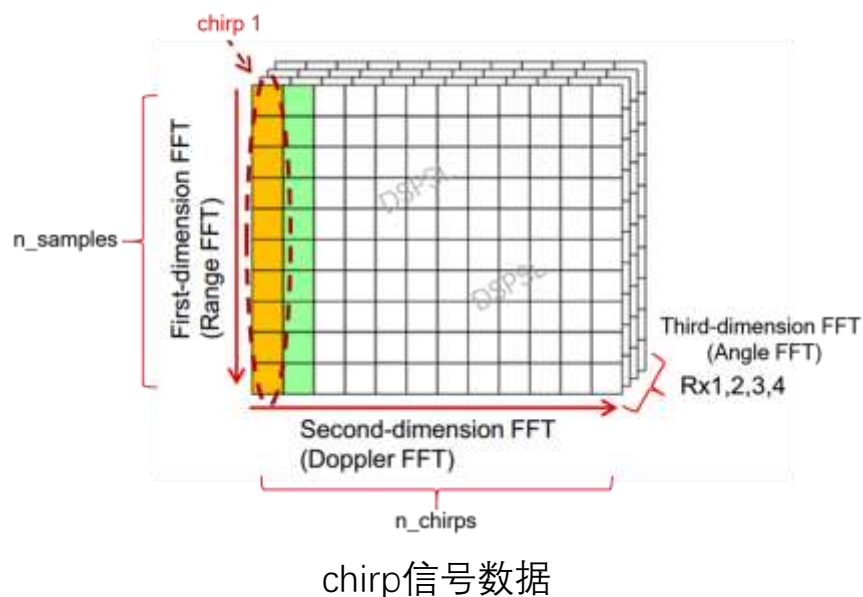
第二步：数据预处理。根据设备的参数，计算出数据大小和帧数，将采集数据由二进制类型转变成numpy类型。

第三步：特征提取，生成频谱图。将第二步中转换后的数据进行归一化、背景减除、去除零帧等操作，同时进行快速傅里叶变换，通过距离多普勒算法，生成距离-角度和距离-速度的多普勒频谱图。

第四步：设计神经网络模型并训练。参考AlexNet和VGG进行网络模型的设计，将第三步中得到的特征图数据按照8:2的比例划分为训练集和测试集，然后将训练集输入到网络模型中进行训练，此外还采用双流融合的方式，输入两种频谱图，融合特征进行训练，训练后观察并分析识别正确率。

频谱图生成

处理后的数据是一个三维数组，[距离， dop， 天线个数]， 分别对应着采样点、 chirps和天线的维度， 分别对这三个维度做FFT(快速傅里叶变换) 就可以得到距离， 速度， 角度的参数信息。



chirp信号数据

(1) 距离FFT

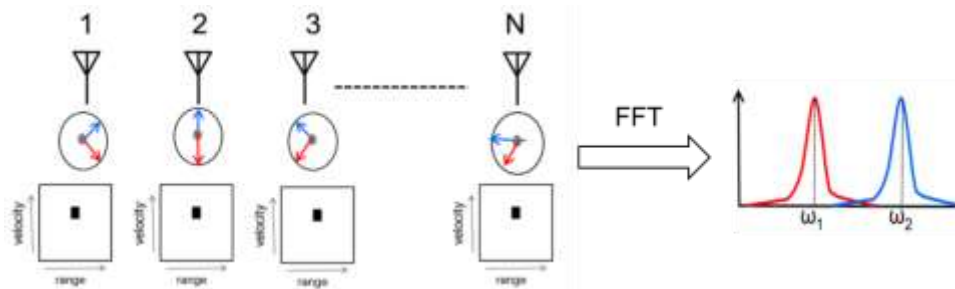
首先在采样点维度进行FFT，由采样点表示的时域信号就转变成由距离表示的频域信号，就可以得到物体的距离信息。

(2) 速度FFT

对距离FFT的结果在chirp维度作1DFFT，即2DFFT，提取2DFFT的峰值即可得到目标的差拍频率和多普勒频率，根据频率和速度的转换公式，频率就可以被转换为速度，得到距离-速度图。

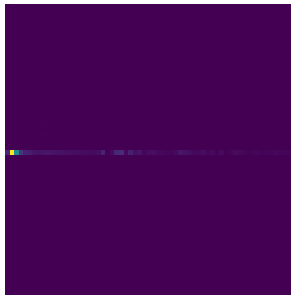
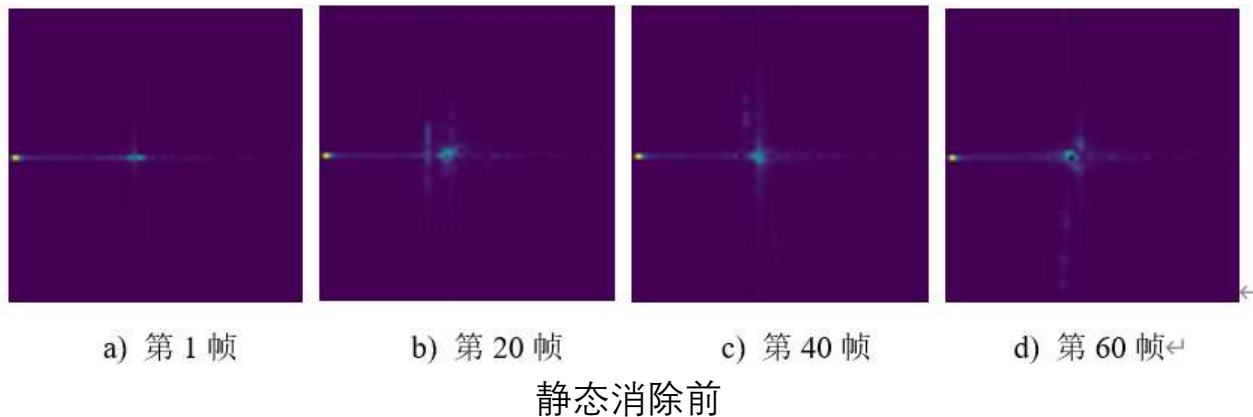
(3) 角度FFT

对多个RX的2DFFT结果在天线维作1DFFT，得到距离-角度图。

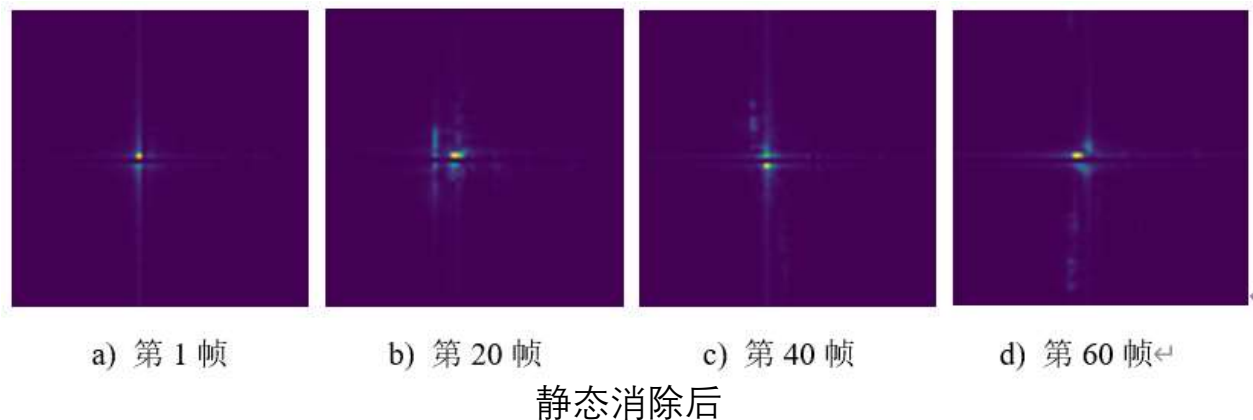


静态消除

采集环境中如墙壁、天花板等静态物体的反射在后续的定位或是行为识别过程中，会对人体的探测造成一定的干扰，因此在生成特征频谱图时，需要对这些背景的信号数据进行减除。



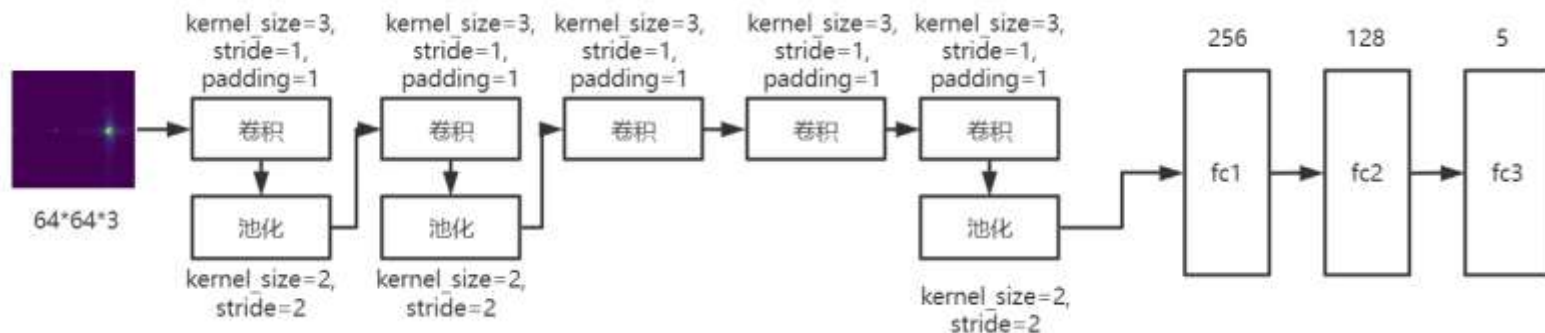
静态物体干扰图



以handwave动作的距离-速度图为例，图中间较高的能量线即为环境中的静态物体反射。实验中通过了设置零帧的方式进行了直接减除。将60帧的数据进行叠加，所得的热图作为后续训练的频谱图数据。



网络模型设计——AlexNet

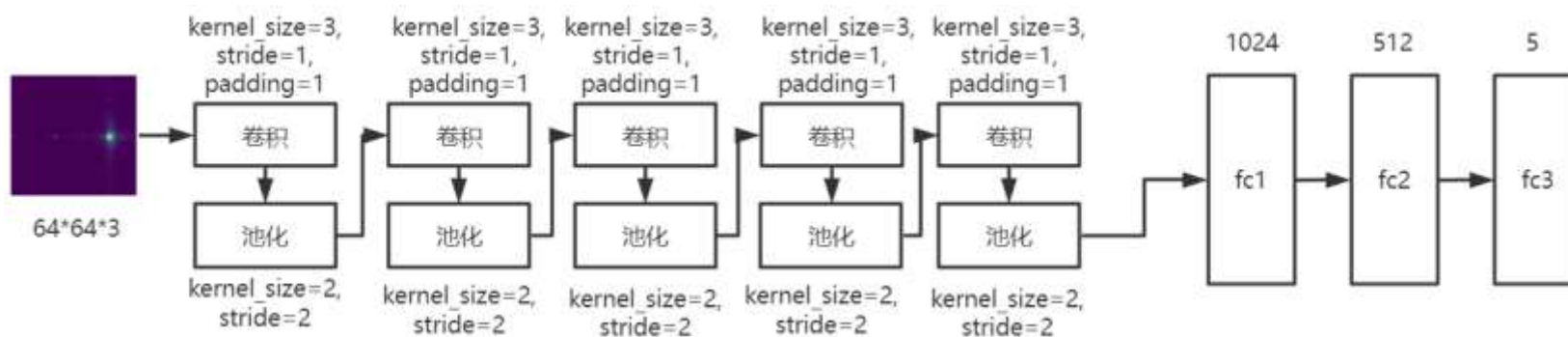


第一个网络模型是基于AlexNet的。在本实验中，生成的距离频谱图都是64*64的大小，图片尺寸不及原AlexNet网络输入的图片大小，因此在AlexNet网络结构的基础上，需要将卷积层的参数进行修改，将卷积核的大小修改为3*3，步长为1且padding为1，避免因过多层数的卷积导致输入图像过小。其次，池化层并没有使用3*3的池化核，而是选择了较为常见的2*2的池化核，也在一定程度上使得图片的缩小得到减缓。全连接层的输入没有选择较多的4096个神经元，而是将其缩小为256，并进一步缩小为128，以适应5分类的行为动作。这样的修改在一定程度上优化了最终的收敛速度和识别效果。

该网络采用cuda和ReLu激活函数来提高神经网络的训练速度，在数据处理时使用了Dropout和数据增强(data augmentation)的方式来防止过拟合。数据增强的方法包括旋转和水平翻转。



网络模型设计——VGG

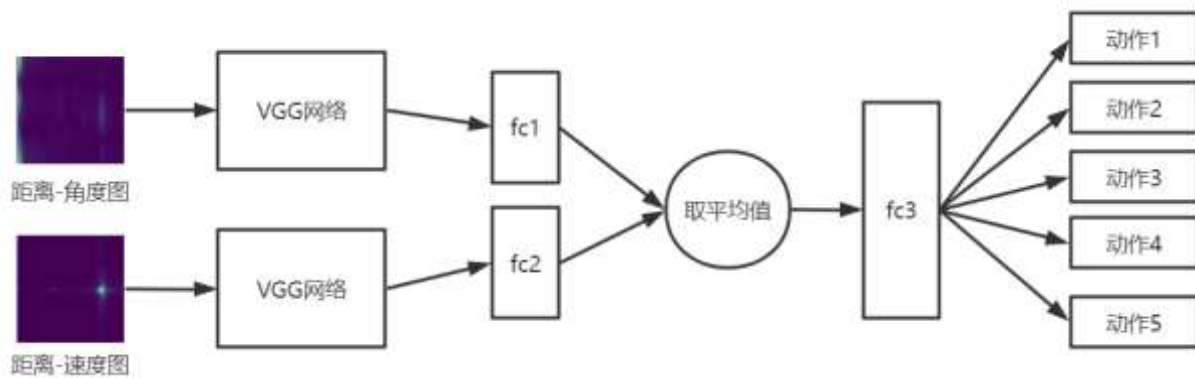


第二个网络模型是基于VGG16的。在本文实验中，由于部分行为动作运动范围较小，导致该动作在距离-角度图中运动幅度不大，图像数据存在着大量的空白信息，所以在设计网络结构时，将VGG的两层卷积简化为一次。在全连接层输入的神经元个数也变成了1024和512。

小卷积核的特点使得模型的非线性拟合能力得到加强，同时也减少了参数的数量，使得计算量得到降低。其次该网络结构通道数更多，第一层为64，之后逐层翻倍，最多达到512个通道，通过增加通道数，可以从输入数据中提取到更多的信息。



网络模型设计——双流融合



传统的基于视觉的和基于非视觉的行为识别技术，在特征提取上，总存在描述算法实时性不高，数据的特征信息提取不全等问题，从而降低了人体行为识别的准确率。因此本文实验尝试了融合两类频谱图特征的双流融合神经网络。

实验中进行训练的网络都选取了VGG，将两类图分别输入到网络中进行训练，并在全连接层进行融合。在全连接层进行连接时，选取行为识别率的平均值作为最终全连接层的输入，然后将最终全连接层的输出作为行为动作的识别率。

实验设计了两类DataSet实现分步读取。其中第一类DataSet正常读取数据，第二类DataSet（如右图）通过同一个索引，实现图片都一一对应于同一个动作。

```
# 实现两类图像数据读取时的一一对应
class TwoDataset(Dataset):
    def __init__(self, dataset1, dataset2):
        self.dataset1 = dataset1
        self.dataset2 = dataset2

    def __len__(self):
        return len(self.dataset1)

    def __getitem__(self, index):
        x1 = self.dataset1[index]
        x2 = self.dataset2[index]
        return x1, x2
```

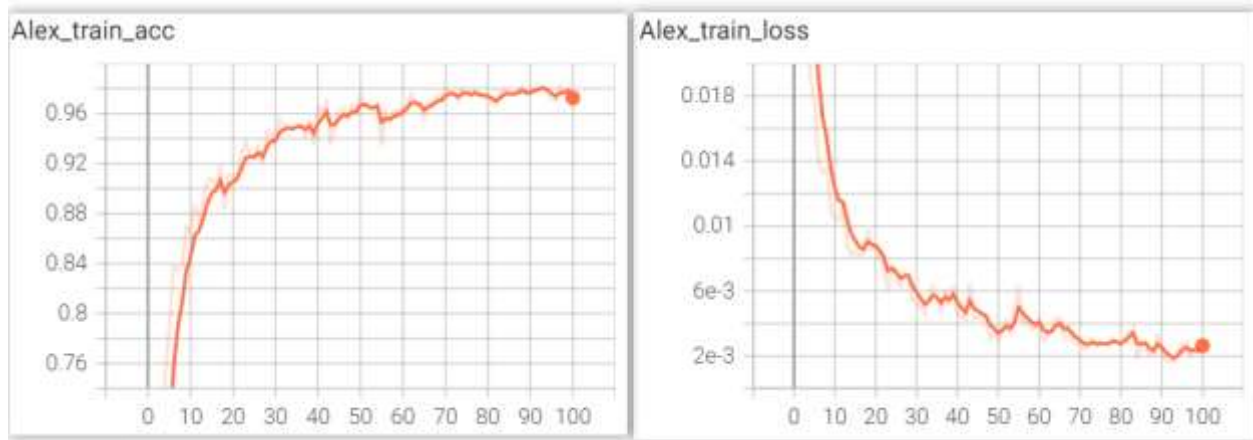


PART 03

实验结果与分析



距离-角度特征分析



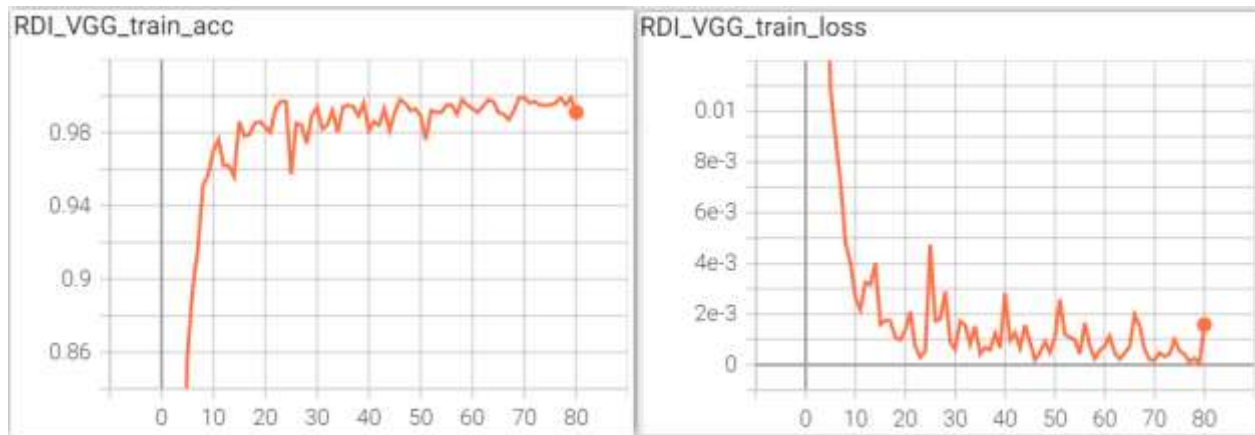
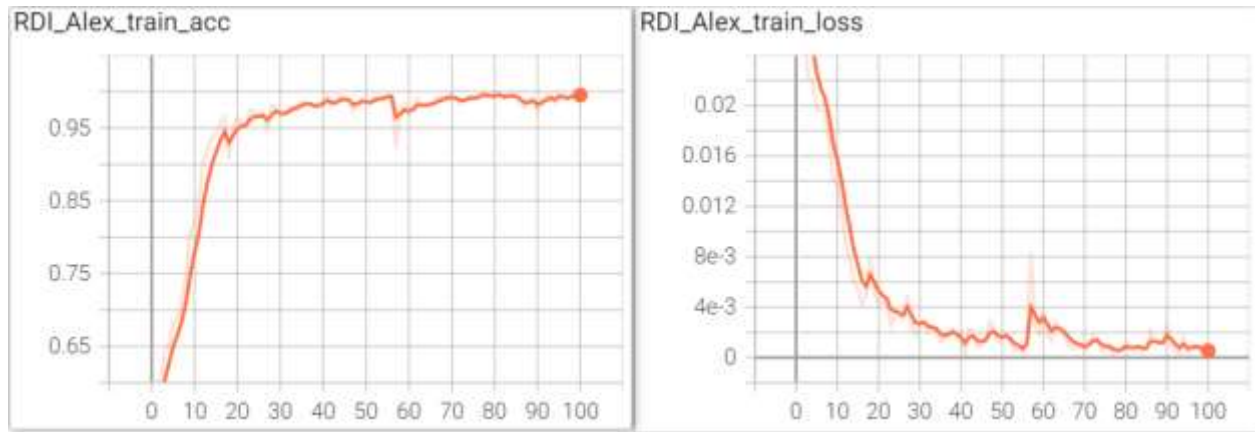
AlexNet训练集上的准确率为96.4%，测试集上的准确率为96.02%；
VGG训练集上的准确率为98.9%，测试集上的准确率为96.41%。



对于距离-角度图，AlexNet网络和VGG网络的行为识别效果差不多，测试集上的准确率均能达到96%左右。从训练过程的收敛速度上看，VGG网络要比AlexNet快一些。



距离-速度特征分析



AlexNet训练集上的准确率为99.5%，测试集上的准确率为96.81%；
VGG训练集上的准确率为99.1%，测试集上的准确率为89.24%。

对于距离-速度图，AlexNet网络与VGG网络在测试集上的识别准确率稍有差距，AlexNet网络仍然能达到96%左右的准确率，而VGG网络却下降到了89%，且从图像的收敛速度上来讲，VGG网路的收敛速度仍然更快，AlexNet网络的行为识别率起伏较大。



单流网络总结对比

AlexNet识别效果

特征图	train_accuracy	test_accuracy
距离-角度图	96.4%	96.02%
距离-速度图	99.5%	96.81%

VGG识别效果

特征图	train_accuracy	test_accuracy
距离-角度图	98.9%	96.41%
距离-速度图	99.1%	89.24%

从表中可以看出，VGG网络在识别距离-角度图的数据时，识别效果要比识别距离-速度图的效果好，说明该网络对角度特征的识别率更高。而AlexNet网络对于两种特征图的识别准确率相当且均较高，说明本文的实验当中，基于AlexNet网络模型架构修改的网路模型对于识别实验中的这些动作时，效果更好。所以卷积神经网络的结构、特征提取的类型等都会对人体行为动作的识别产生一定的影响。

双流融合分析

双流融合后得到训练集上的准确率为98.9%，测试集上的准确率大致为99.6%。



将上图中的数据和前面单独特征图训练的测试准确率对比，可以看出，将两种距离频谱图的训练数据在全连接层拼接效果比较理想。本文实验采取的融合方式是取平均值。从识别准确率的角度上分析，双流融合网络将两类图的特征融合后，识别准确率达到99.6%，比前面单独特征图的训练网络识别效果都要好；从收敛速度上分析，测试时在十几轮的时候就有了很好的收敛，和前面两类网路对比，收敛速度相对较快。



总结

综上所述，对于本课题基于毫米波雷达的人体行为感知识别的实验，针对设计的5个动作，基于AlexNet网络修改的卷积神经网络模型比较适合这些动作的识别，无论是对于距离-角度图还是距离-速度图，都有不错的识别效果。而双流融合网络在此基础上将两类输入图的特征进行融合，通过取平均值的方式输出最后结果，同样也达到了不错的效果，且比两个单流网络的识别准确率要更高。



PART 04

总结与展望



总结与展望

总结

本文实现了VGG和AlexNet两种神经网络，同时也生成了距离-角度和距离-速度两种频谱图。将频谱图数据输入到神经网络中进行学习和训练，通过调整参数，最终得到较为理想的识别概率，并和双流融合网络进行对比实验。实验结果表明，提取的特征不同、网络模型不同都会对人体行为动作的识别产生影响，双流融合输入两类特征的方式可以提高识别的准确率。

展望

- (1) 在采集数据时，采集时间较短。期望在下一步的实验中，可以加长采集时间，增加数据长度，观察行为识别效果如何。
- (2) 融合时机、融合方式、网络设计等的不同，会对基于双流融合的网络模型对行为动作的识别效果产生不同的影响，期望下一步可以选取其他的融合方式和时机，同时尝试其他更适合双流融合的网络，再一次对比分析双流融合网络的识别准确率。



感谢聆听

1023041002 王施涛

南京邮电大学 计算机学院、软件学院、网络空间安全学院 软件工程