

清华大学

综合论文训练

题目：神经网络液闪探测器重建研究

系 别：工程物理系

专 业：工程物理

姓 名：董正

指导教师：续本达 教授

二〇二〇年六月

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名： _____ 导师签名： _____ 日 期： _____

中文摘要

本文回顾了中微子物理和中微子实验的发展历史，综述了机器学习和神经网络的发展历程。针对锦屏 1 吨探测器，将神经网络回归算法应用于粒子空间信息重建，设计了一维卷积神经网络、单一输入 LSTM 神经网络和多输入 LSTM 神经网络结构，并调研了球面卷积神经网络在液闪探测器信号重建上的应用。利用模拟数据训练了一维卷积神经网络模型，重建精度为 $61.73mm$ ，训练了单一输入 LSTM 网络模型，重建精度为 $146.88mm$ 。与此同时，发现了一维卷积神经网络的重建结果与电荷加权平均法的重建结果在空间分布上具有相似性。

关键词：中微子；神经网络；探测器重建

ABSTRACT

In this paper, the history of neutrino physics is reviewed. The development of machine learning and neuro networks is reviewed. Neuro networks are applied to the data reconstruction process of Jinping 1t liquid scintillator detector prototype. 1d-Conv as well as single and multiple input LSTM net structures are designed to complete the reconstruction process. The development and appliance of spherical convolutional neuro network is also examined. 1d-Conv and single input LSTM model are trained using simulation data. The reconstruction accuracy is examined for both models. The 1d-Conv model yields an average reconstruction accuracy of $61.73mm$ while the single input LSTM model has an accuracy of $146.88mm$. Meanwhile, a similarity in sapcial distribution between the reconstruction outcome of 1d-Conv model and charge weighed average method is dicovered.

Keywords: Neutrino; Neuro Networks; Reconstruction

目 录

第 1 章 引言	1
1.1 中微子物理	1
1.2 中微子实验综述	3
1.3 锦屏中微子实验	3
1.4 论文工作概述	4
第 2 章 顶点重建方法	5
2.1 传统重建方法综述	5
2.2 神经网络发展综述	6
2.2.1 神经网络、机器学习与人工智能	6
2.2.2 神经网络的发展及原理	7
2.3 神经网络重建方法	9
2.3.1 一维卷积神经网络重建	10
2.3.2 LSTM 神经网络重建	11
2.3.3 S-CNN 重建方法简述	11
2.4 论文方法总结	12
第 3 章 数据重建结果及讨论	13
3.1 数据来源及数据结构	13
3.2 重建结果性能指标	14
3.3 一维卷积神经网络重建结果	14
3.3.1 神经网络结构	14
3.3.2 重建结果	15
3.4 单一输入 LSTM 神经网络重建结果	16
3.4.1 神经网络结构	16
3.4.2 重建结果	17
3.5 多输入 LSTM 神经网络重建结果	17

3.5.1 神经网络结构	17
3.5.2 重建结果	18
第 4 章 结论	21
4.1 论文工作总结	21
4.2 当前局限与改进思路	21
插图索引	23
表格索引	24
公式索引	25
参考文献	26
致 谢	28
声 明	29

主要符号对照表

JUNO	江门中微子实验
CNN	卷积神经网络
1d-Conv	一维卷积神经网络
s-CNN	球面卷积神经网络
PMT	光电倍增管

第 1 章 引言

1.1 中微子物理

中微子是基本粒子标准模型中的轻子，其在标准模型中的位置如图所示。

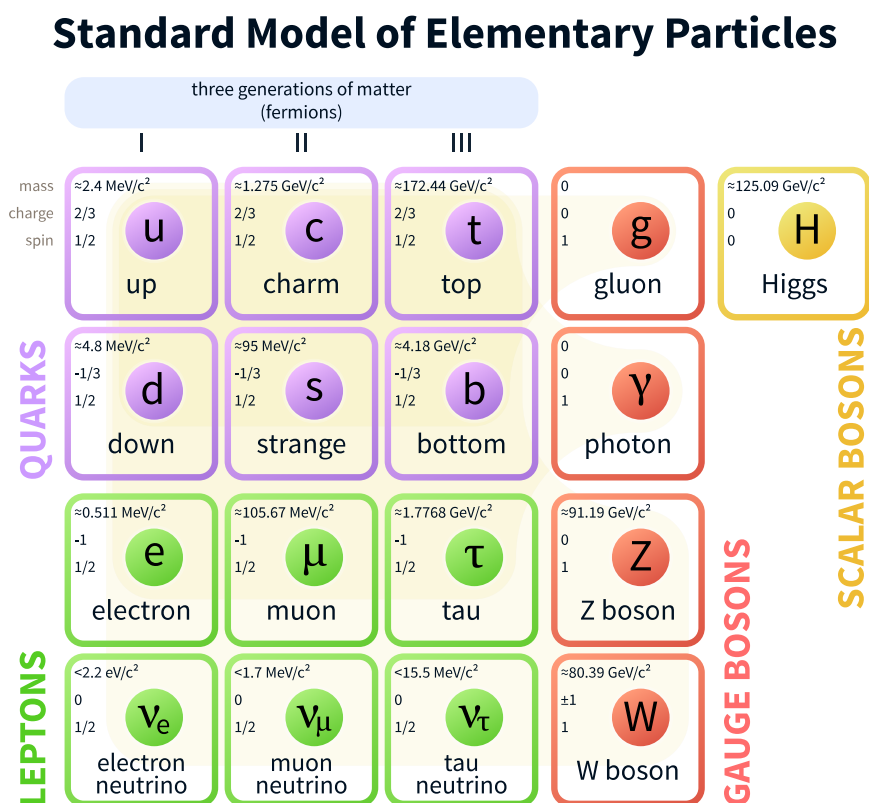


图 1.1 基本粒子标准模型

中微子的发现过程最早可以追溯至 1914 年。查德威克在测量 β 射线能谱时，发现 β 射线的能谱不同于 α 射线谱的分立分布，呈现连续分布，这一现象直观上违背了 β 衰变过程中的动量守恒和能量守恒^[1]。泡利在 1930 年尝试对上述现象做出解释，认为 β 衰变中部分能量被一种中性的、不易于探测的粒子带走了，并将这种粒子命名为 neutron^[2]。四年后，费米进一步发展了 β 衰变理论^[3]，指

出中子的 β 衰变过程按照如下过程发生：

$$n^0 \rightarrow p^+ + e^- + \bar{\nu}_e \quad (1-1)$$

其中 $\bar{\nu}_e$ 为泡利猜想中的中性粒子，费米将其命名为 **neutrino**，即中微子。

随着标准模型的发展，中微子被按照弱相互作用中的味道本征态划分为电子中微子、 μ 中微子和 τ 中微子。按照第三代粒子物理标准模型的预测，中微子没有静止质量，但是理论上并不排除中微子具有静止质量的可能性。随后的多项实验也分别证明中微子具有静止质量，这一发现基础物理的研究有深远的意义^[4]。

Pontecorvo 在 1957 年提出，根据量子力学原理，如果中微子静止质量不为零，则中微子的味道本征态和质量本征态不简并。中微子在参与弱相互作用时处于味道本征态，而在传播过程中处于质量本征态，中微子在传播过程中处于味道本征态的叠加态，不同类型的中微子可以相互转换，称为中微子震荡现象。用 $|\nu_\alpha\rangle$ 描述中微子的味道本征态，其中 α 可以是 e 、 μ 、 τ ，用 $|\nu_i\rangle, (i = 1, 2, 3)$ 描述中微子的质量本征态，则两个表象之间的变换可以用下列矩阵描述^[5]：

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} \\ 0 & e^{-i\delta} & 0 \\ -s_{13} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \quad (1-2)$$

式1-2中， $s_{ij} = \sin \theta_{ij}$ ， $c_{ij} = \cos \theta_{ij}$ ， $(i, j = 1, 2, 3)$ 。 θ_{ij} 称为混合角，代表中微子之间的相互转换，其中 θ_{23} 与大气中微子震荡相关， θ_{13} 和反应堆中微子及长基线加速器中微子震荡相关， θ_{12} 和太阳中微子相关。 δ 称为 CP 相角，同样与反应堆中微子及长基线加速器中微子震荡相关。

中微子震荡能够间接说明中微子具有静止质量，进而发现超出标准模型的物理现象，因此是当前的研究重点，世界上主要的中微子探测实验也都围绕对中微子震荡的研究展开^[4]。

1.2 中微子实验综述

如前文所述，中微子的发现起源于查德威克的实验，其后泡利、费米等物理学家均对中微子物理的理论建构做出了贡献。在此之后的许多实验均证明中微子与物质相互作用截面极低，因此如何对中微子进行探测成为了首要问题。为了提高探测率，可以提高中微子通量（即寻找高通量中微子源）或提高探测器效率。目前在实验中主要用到的中微子源有四种，分别是大气中微子、太阳中微子、反应堆中微子和加速器中微子。

1968 年，美国 Brookhaven 实验室的戴维斯在废旧金矿中通过四氯乙烯 (C_2Cl_4) 捕获太阳中微子。根据已有对太阳内部反应的研究可知太阳内部 $p-p$ 链聚变反应会产生大量的中微子，然而戴维斯观测到的中微子个数只有预期的三分之一^[6]。经过后续物理学家们的努力，这一反常现象通过太阳中微子震荡理论得到了解释。

1988 年，日本神冈实验测得大气中 μ 中微子的数目比预期要少，十年后，升级后的超级神冈探测器精确测量了大气中 μ 中微子数目随天顶角的变化关系，发现了大气中微子震荡。2002 年，日本 KamLAND 实验用反应堆中微子观测到了大气中微子的震荡模式^[7]，其后 2003 年日本的 K2K 实验^[8] 和 2006 年美国的 MINOS 实验^[9] 也分别用加速器中微子验证了大气中微子震荡模式。

至此，式 1-2 中描述太阳中微子震荡模式的 $\sin^2 2\theta_{12}$ 与描述大气中微子震荡模式的 $\sin^2 2\theta_{23}$ 均已被测定。随后大亚湾中微子实验于 2012 年测定了第三个混合角 $\sin^2 2\theta_{13}$ ^[10]，使得描述中微子震荡的参数只有 CP 相角 δ 还未被测定。除此之外，中微子质量顺序问题以及 θ_{23} 八重度问题等也均尚未得到解决^[9]。

1.3 锦屏中微子实验

中国锦屏实验室 (China Jing Ping Lab, CJPL) 依托四川省雅砻江水电站，建于最高高度达 2400m 的山体下，是目前世界上最深的地下实验室。山体的屏蔽使得锦屏实验室拥有低 μ 中微子通量和低反应堆中微子通量的独厚条件，有利于对低能中微子开展探测^[11]。CJPL 计划建设 2000 吨的液体闪烁体探测器用于中微子探测，实验初期阶段，建设了 1 吨球形液闪探测器原型机，原型机设计图如图 1.2 所示。

如图 1.2 所示，该原型机采用球形设计，球形液体闪烁体容器外有 30 个 PMT

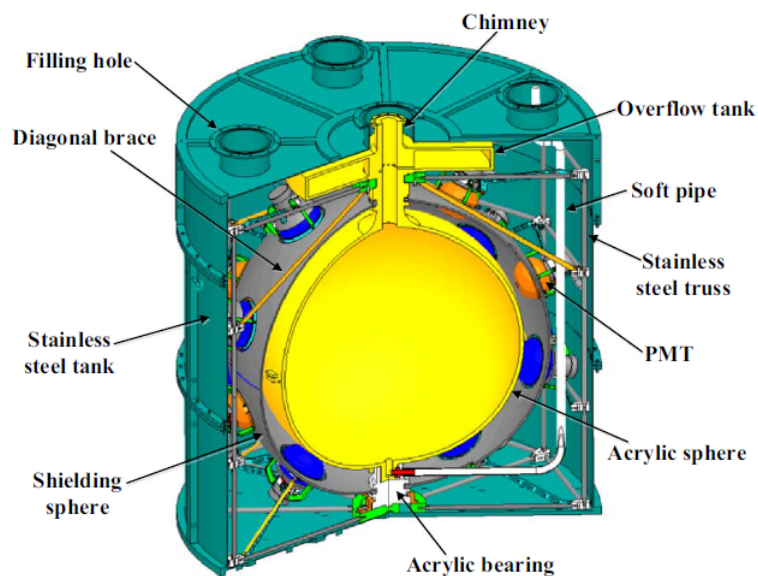


图 1.2 1 吨原型机示意图^[12]

(Photomultiplier Tube, 光电倍增管)^[12]。粒子在闪烁体中沉积能量，闪烁体发出的光子击中 PMT 后在 PMT 中倍增形成波形信号。

1.4 论文工作概述

本文的工作主要围绕锦屏中微子实验中的 1 吨探测器展开，基于 1 吨探测器上的模拟数据，调研了适用于探测器顶点重建的神经网络重建方法，并设计了相应的网络结构。利用 1 吨探测器的模拟数据作为训练数据，训练了 1d-Conv、单一输入 LSTM、多输入 LSTM 等 3 个神经网络模型，并以顶点重建精度为性能指标，评价了几个模型的重建效果。与此同时，还调研了球面卷积神经网络在球形探测器重建工作中的应用，为在体积更大的球形闪烁体探测器上应用神经网络重建方法奠定了基础。

第 2 章 顶点重建方法

对于液体闪烁体探测器，粒子在探测器中发生的一系列物理过程最终会形成光电倍增管（PMT）的波形信号，为了掌握探测器内部发生的物理事件，需要对波形信号进行分析，并基于波形信号中蕴含的电荷和时间信息对探测器内的物理事件进行重建。本章将讨论论文研究过程中应用的主要方法，包括回顾探测器顶点重建中常用的极大似然方法，并详细综述本文研究主要采用的神经网络重建方法。

2.1 传统重建方法综述

基于极大似然法的重建方法是探测器顶点重建中的常用方法，这一方法通过考察粒子在探测器灵敏体积内发生反应沉积能量、闪烁体受激发光、光子引起 PMT 响应等物理过程，计算得出在特定时空位置 $\{\vec{x}_0, t_0\}$ 处发生的具有特定能量的顶点事件对应的第 i 个 PMT（位于 \vec{x}_i 位置）在 t_i 时刻响应的似然函数：

$$\mathcal{L}(\vec{x}_0, t_0; \{(\vec{x}_i, t_i)\}) = \prod_{i=1}^N p(t_i - t_0 - \frac{s_i n}{c}) \frac{R - a \cos \theta_i}{s_i^3} \quad (2-1)$$

式 2-1 中， $p()$ 为闪烁体发光函数， n 为闪烁体折射率， c 为真空中光速， θ_i 、 s_i 等参数与事件位置和探测器几何信息有关，在此不详述^[13]。在重建过程中，对于给定的 PMT 响应信息，通过求取似然函数极大值反推顶点的时空位置和粒子能量进行重建。

此外，还有一种在重建过程中被用于对顶点位置进行初步估计的方法，称为电荷加权平均方法。这一方法的思路是通过求 PMT 空间位置关于该 PMT 在单次事件中接收到的光子数的加权平均来大致估计顶点位置，假设第 i 个 PMT（位置为 \vec{x}_i ，共有 N 个）在单次事件中接收到的光子数为 q_i ，则顶点位置 v 可以由下式给出^[14]：

$$v = \frac{3}{2} \sum_{i=1}^N \frac{q_i \times \vec{x}_i}{q_i} \quad (2-2)$$

2.2 神经网络发展综述

2.2.1 神经网络、机器学习与人工智能

神经网络是人工智能研究的一个分支，是机器学习中使用到的核心方法之一。人工智能、机器学习、神经网络等概念术语的关系如图所示。

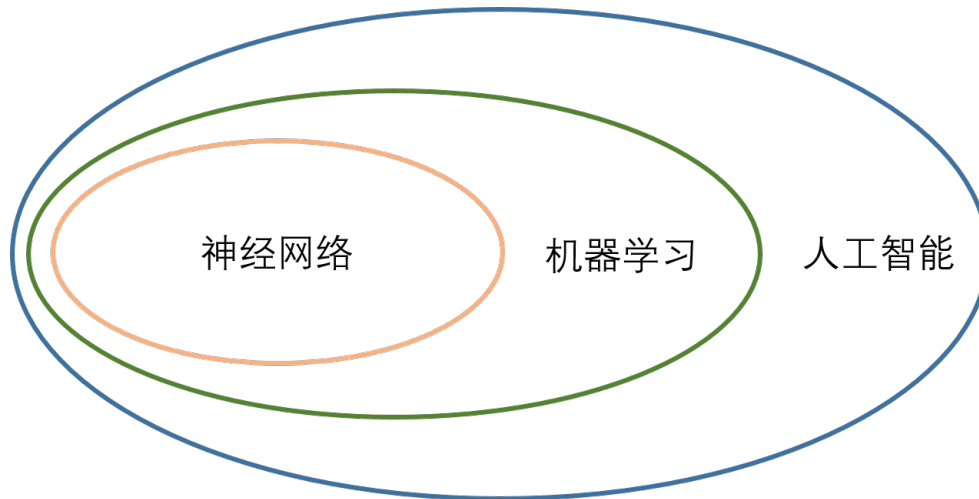


图 2.1 人工智能、机器学习和神经网络的概念关系

对于机器学习的研究最早可以追溯至 20 世纪 50 年代，那时对于人工智能的研究处于“推理期”，人们认为可以通过赋予机器逻辑推理能力进而赋予机器智能^[15]。在这一时段被开发出的很多程序都在证明数学定理方面取得了良好的成果，引起了科学界的兴趣。

然而随着研究的深入，人们逐渐发现单纯的推理能力不足以实现机器智能，要让机器拥有智能，就必须让机器拥有知识。自 20 世纪 70 年代中期，人工智能的研究进入了“知识期”^[16]，研究者开始尝试给智能系统以结构化的知识，这一阶段，大量“专家系统”问世，知识表示和实体建模也在这一时期得到了集中的发展。但是，对“专家系统”的进一步研究揭示出这样的模式会面临“知识工程瓶颈”，即人工给予机器的知识是结构化的、是有限的，将自然语言组成的知识“教给”机器复杂且低效，这样的知识无法使机器拥有智能。研究者们进而想到，如果能够让机器进行“自主学习”，那么机器就能够高效地学习进而利用知识。至此，现代的“机器学习”概念初步形成^[17]。

Michalski 等人将机器学习划分为“从样例中学习”、“在问题求解和规划中

学习”、“通过观察和发现学习”和“从指令中学习”等种类^[18]，自 20 世纪 80 年代以来，对于机器学习的研究主要集中与“从样例中学习”，近年来“在问题求解和规划中学习”也得到了一定的发展。目前主流的机器学习研究有三大分支，分别是监督学习、无监督学习和强化学习。监督学习和无监督学习对应 Michalski 提出的概念中的“从样例中学习”，区别在于监督学习中用到的样本数据带有标签，非监督学习中用到的样本数据没有标签；强化学习对应于前述概念中的“在问题求解和规划中学习”。

本文涉及的顶点重建问题，应当属于监督学习范畴，监督学习的过程包括从训练集数据和标签中学习规律，然后在测试集上基于数据对标签进行预测。依据标签的连续性，离散的标签对应分类问题，连续的标签对应回归问题。本文涉及的重建工作属于监督学习中的回归问题，粒子鉴别属于监督学习中的分类问题。

2.2.2 神经网络的发展及原理

神经网络是处理监督学习问题中的常用方法，对相关问题的研究也很几乎和对于机器学习的研究同时起源。1943 年，McCulloch 和 Pitts 基于已有的研究结果，将神经网络的最小单元抽象为“M-P 神经元模型”，模型模式图如下图所示。这一模型将单一神经元视为一个门电路，该电路的输入信号是来自 n 个其

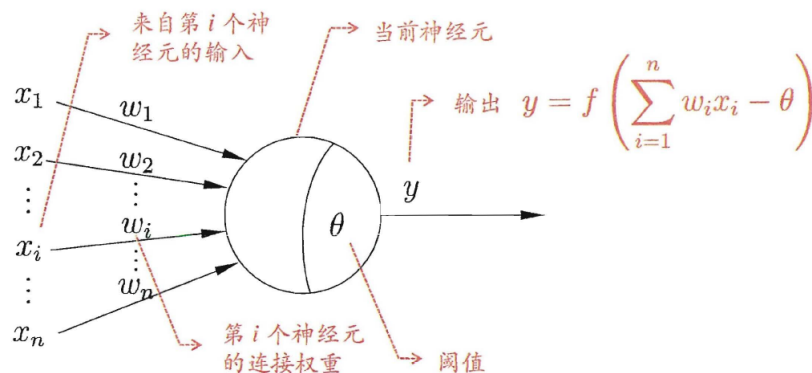


图 2.2 M-P 神经元模型^[17]

他神经元的信号，输出信号只有 0 或 1^[19]。神经元将输入信号赋予一定的权重并向下一级传递，如果神经元接受到的总输入值超过神经元的阈值，则向后输

出 1，反之则输出 0。

神经元通过相互连接、不断嵌套可以组成复杂的网络结构。两层神经元组成的网络结构称为感知机 (Perceptron)，是神经网络的雏形^[20]。单层感知机的功能十分简单，能够处理输入信号与阈值比较得到 0、1 输出，因此又被称为阈值逻辑单元。多层感知机相互连接即构成了神经网络，可以实现更为复杂的功能。

感知机中神经元对阈值的判断称为激活函数，典型的激活函数是阶跃函数，输出非 0 即 1，这样的激活函数不连续且不光滑，无法充分发挥多层神经网络的优势。因此，在实际应用中常选用更为平滑的激活函数实现输入与输出之间的映射，常用的激活函数有 ReLU、sigmoid 等，激活函数大致图像如下图所示。

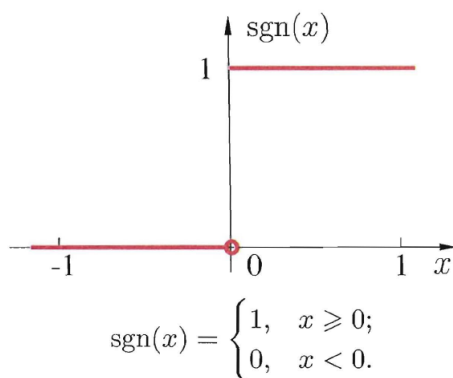


图 2.3 阶跃函数示意图^[17]

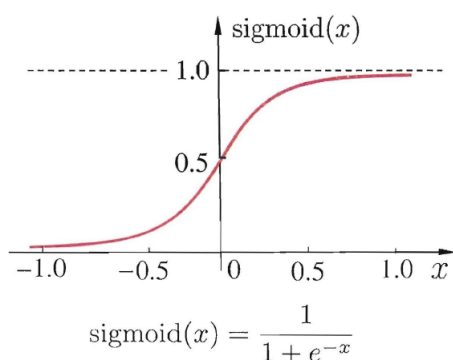


图 2.4 sigmoid 函数示意图^[17]

以上部分介绍了神经网络的基本结构，然而支撑神经网络实现机器学习功能的则是误差反向传播算法。从输入到输出的运算过程称为前向传播，对于监

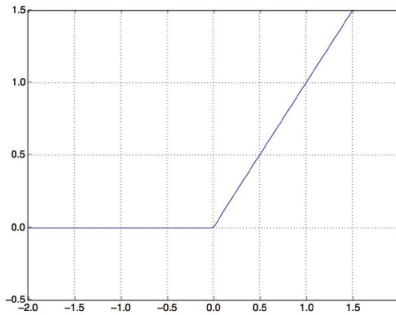


图 2.5 ReLU 函数示意图^[20]

督学习问题，神经网络可以处理输入的数据，得到输出结果作为标签，为了实现机器的自主学习，就需要将神经网络的预测值与训练集数据标签的实际值进行比较，然后更新权重使得预测值与实际值更加接近。在训练集中，对于每一个输入 $X(x_1, x_2, \dots, x_n)$ ，都有一个对应的标签 $Y(y_1, y_2, \dots, y_m)$ ，将 X 输入神经网络后，会得到一个由神经网络计算出的预测值 \tilde{Y} 。实际值与预测值之间的差异 $bias = \tilde{Y} - Y$ 称为误差。

训练神经网络的目标是尽可能减小误差，这一过程通过误差反向传播算法实现。对于神经网络中的每一个神经元，其赋予若干个输入信号的权重构成一个权重矩阵 W ，误差反向传播算法通过计算误差 $bias$ 关于权重的梯度 $\frac{\partial bias}{\partial W}$ ，对新一轮训练的权重矩阵进行更新，从而使得新一轮训练的误差较上一轮减小。比较典型的权重更新方法有随机梯度下降、动量法、Adam 法等，其主要思路均为将权重矩阵向结果误差更小的方向更新，具体实现在本文中不详述^[20]。

2.3 神经网络重建方法

由前文叙述可知，神经网络可以被用于探测器的顶点重建和粒子鉴别。对于顶点重建工作，可以将 PMT 的响应信号（包含电荷和时间信息）作为输入 X ，将顶点事件的空间位置作为标签 Y ，然后通过神经网络训练回归模型，实现重建功能。类似的，对于粒子鉴别工作，可以同样将 PMT 响应作为输入，将粒子类别作为输出，通过神经网络训练分类模型。

已经有研究将神经网络用于粒子探测相关的工作，比较有代表性的工作有 NoVA 探测器上的中微子事件分类器^[21] 和 KamLAND 实验中用于消除宇宙线本底的神经网络算法^[22]。尽笔者所知，将神经网络用于粒子重建的工作较少，NoVA

实验中使用神经网络进行中微子能量重建^[22]，而几乎没有见到关于将神经网络用于粒子空间重建的报道。因此本文将尝试利用神经网络算法对探测器粒子空间信息进行重建。

由于神经网络的模型较为多样，针对不同的数据结构也分别有针对性的模型结构，下文将分别对本文工作中用到的网络结构进行说明。

2.3.1 一维卷积神经网络重建

卷积神经网络在处理计算机视觉相关问题上有着出色的表现。对于卷积神经网络，输入与权重之间的运算不再是简单地向量点积，而是更为复杂的卷积（也有文献或教材中成为互相关）运算。对于输入信号是二维的情形，卷积运算通过一个 $m \times m$ 大小的矩阵（称为卷积窗或卷积核， m 称为窗宽）在 $n \times n$ 的输入信号上滑动，在每一个位置上通过矩阵乘法计算得到输出值，由输出值组成的二维图称为特征图。二维卷积运算的大致过程如图所示。

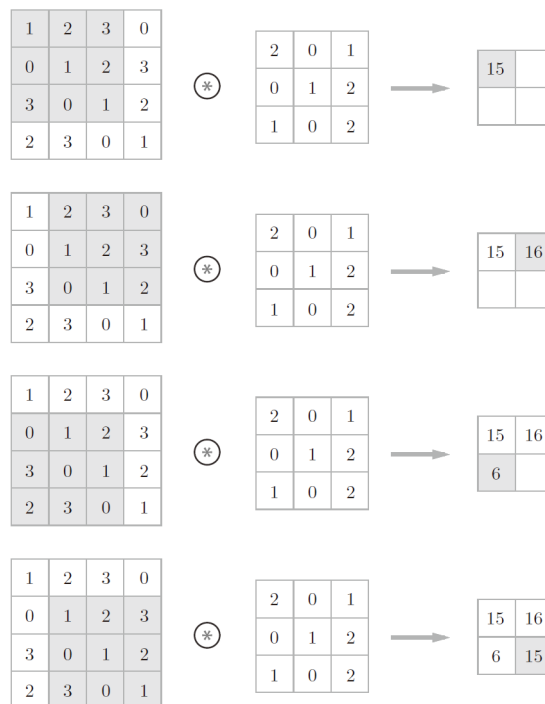


图 2.6 二维卷积及运算示意图^[23]

对于输入数据是一维（序列）的情形，二维卷积窗退化为 $1 \times m$ 的一维窗口，在输入序列上滑动进行卷积运算，得到的输出序列也是一维的。一维的卷积核

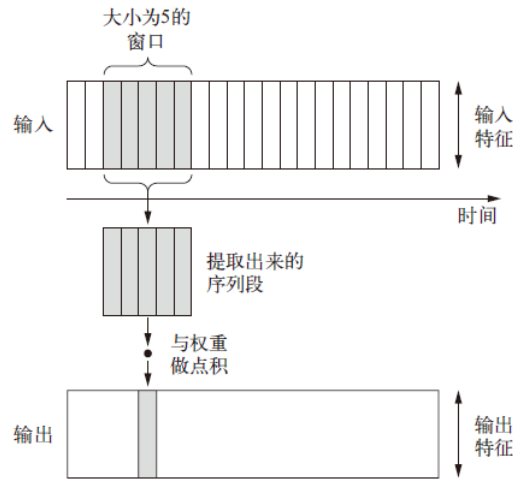


图 2.7 一维卷积示意图^[20]

可以识别序列中认定的局部模式，并对整个序列进行同样的操作。

2.3.2 LSTM 神经网络重建

对于处理序列信息，一维卷积神经网络善于发现序列中的局部模式，这一点也让一维卷积神经网络处理序列信息是具有优势。但一维卷积神经网络对于序列整体的信息的利用不够充分，同时，在训练网络时，单个序列被当做一个整体，作为一个向量输入网络，这容易使网络忽略序列中包含的整体信息。为了解决这一问题，科学家提出了循环神经网络的概念，LSTM（Long Short-Term Memory，长短期记忆）是循环网络中比较有代表性的一种结构。这一网络结构由 Hochreiter 和 Schmidhuber 在 1997 年提出^[24]，其主要思路是在平行于处理序列的方向上增加一条“传送带”，用于携带过去时刻的状态信息，方便在未来时刻随时调用。对于 SLTM 网络，参与训练的参数主要是不同时刻的状态信息在被调用时被赋予的权重。

2.3.3 S-CNN 重建方法简述

目前世界上主要的中微子探测器如 JUNO、超级神冈等均采用球形结构，其产生的信号一定程度上可以认为是球面信号。上述两种神经网络重建方法使用到的网络结构均未考虑球面信号的特殊性，一定程度上显性地损失了探测器的几何信息（神经网络的性质使得模型在训练过程中获得了探测器的几何信息，有关内容将在第三章报告）。为了充分利用球面信号的几何信息，可以考虑将探测

器产生的信号直接当做球面信号输入神经网络进行处理。

当前采用神经网络处理球面信号主要有两种思路：对球面信号进行变化和对卷积操作进行变换。前者的思路是通过共形映射将球面信号映射为平面信号，再将平面信号输入到传统的平面卷积神经网络中进行处理。这一思路的代表性网络是 SphereNet，这样操作的好处是当前已经有许多成熟的平面神经网络结构，可以直接应用于映射到平面的球面信号；该方案的局限性在于共形映射变换一定程度上牺牲了球面信号的旋转不变性。

处理球面信号的另一种思路是对卷积运算进行变换，这一思路的代表网络是 Cohen 于 2018 年提出的 Spherical-CNN，这一网络结构从群论入手，首先分析了平面卷积运算的本质是保障了平移等变性（即输入信号上的平移变化在输出信号上同样对应平移变化）。基于这一特征，网络设计者推广了球面卷积的定义，通过推广的快速傅里叶变换使得球面卷积运算具有旋转等变性，直接作用于球面信号，输出球面特征图^[25]。这一思路的优势的充分理解和运用了卷积运算善于发现局部特征的优势，局限性在于该网络结构内部高度自洽，但是不便于使用者针对特定的情形进行修改。

2.4 论文方法总结

本文将主要采用一维卷积神经网络和 LSTM 神经网络结构，基于锦屏 1 吨探测器上模拟数据训练相应模型，进行顶点重建工作。由于锦屏 1 吨探测器几何尺寸较小，同时仅有 30 个 PMT，产生的球面信号较为稀疏，因此本文仅针对球面卷积神经网络进行初步调研，不进行实验。

第 3 章 数据重建结果及讨论

3.1 数据来源及数据结构

本文工作中用到的训练集和测试集数据均来自 JSAP 生成的模拟数据。JSAP (Jinping Simulation and Analysis Package) 是基于 geant4 编写的用于模拟 CJPL 中微子探测装置的软件包。建构探测器的几何信息后, 可以对探测器中发生的物理事件进行模拟, 得到与事件对应的 PMT 模拟输出信号。

进行顶点重建时使用的锦屏 1 吨探测器模拟数据为 HDF5 格式 (.H5), 数据中包含 GroundTruth 和 Truth 两个数据表, 其中 GroundTruth 存储粒子的能量、顶点事件的空间坐标, Truth 用于存储光子击中 PMT 的时间。数据内容及数据类型如下表所示。

表 3.1 GroundTruth 表数据结构

变量名	变量类型	变量含义
EventID	int32	事件编号
x	int64	顶点 x 坐标
y	int64	顶点 y 坐标
z	int64	顶点 z 坐标

表 3.2 Truth 表数据结构

变量名	变量类型	变量含义
EventID	int32	事件编号
PEChannel	int64	PMT 通道编号
PETime	int64	光子击中 PMT 的时间

其中表 3.2 中的信息是基于探测器输出信号得到的, 将被作为模型训练中的输入信号 X , 表 3.1 中的信息是粒子在探测器中产生顶点事件的信息, 将被作为训练模型中的数据标签 Y 。

对于单个事件（即一个 EvnetID），会有多个光子在不同时刻击中不同的 PMT，即有多个 PEChannel 和对应的 PETIME。如果要将这些信息输入神经网络中，需要对数据进行相应的编码和整合，针对不同的网络结构的数据编码方式将在下文分别叙述。

3.2 重建结果性能指标

对于顶点空间位置重建，采用重建精度来衡量重建结果。对于单个事件，GroundTruth 表中存储的空间信息为 $\vec{r}(x, y, z)$ ，通过神经网络训练得到的重建结果为 $\tilde{\vec{r}}(\tilde{x}, \tilde{y}, \tilde{z})$ ，则重建精度定义为二者之间的绝对误差 $bias = \|\tilde{\vec{r}} - \vec{r}\|$ ，对于多个事件，重建精度为单次事件重建精度的绝对值。

与此同时，还将绘制重建结果 z 坐标关于 $x^2 + y^2$ 的关系图像，用于衡量重建结果在探测器中的空间分布。

3.3 一维卷积神经网络重建结果

论文工作采用基于 TensorFlow 的 Keras 深度学习框架进行，本节将汇报重建结果。

3.3.1 神经网络结构

一维卷积神经网络结构如下图所示。

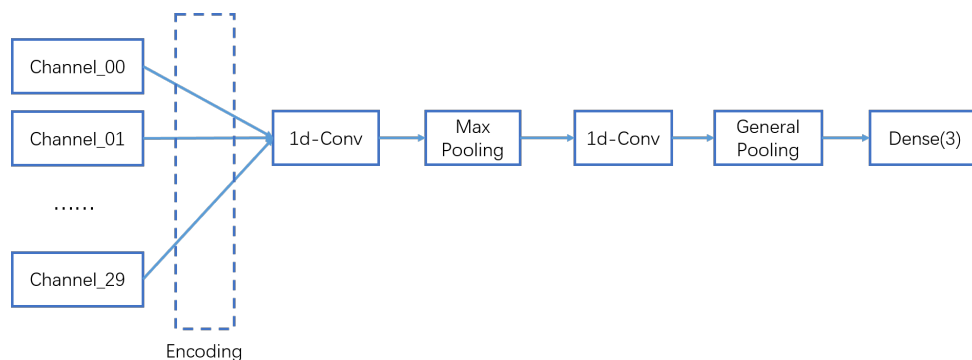


图 3.1 一维卷积神经网络结构示意图

如图，一维卷积神经网络采用单一输入，数据经过编码后分别输入两个卷

积层和池化层（池化层，Pooling，用于减少信号在网络中传播的数据量，提高训练效率），然后通过密集连接层（Dense，即多层感知机）与输出相连。

对输入数据的编码规则采用类似 One Hot 编码的方式，对每一个 EventID，将该事件对应的 PEChannel（编号为 1 ~ 30）按照 PETime 排序，然后将 PETime 对应的 PEChannel 嵌入到完整的时间序列中，对于没有 PMT 响应的时刻，该点取值为 0。单个输入信息如下图所示。

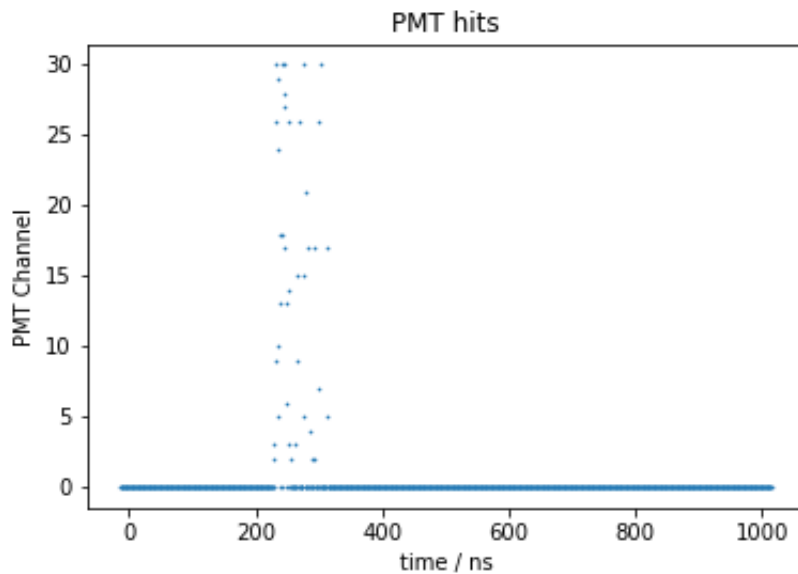


图 3.2 单次事件输入信号示例

由于输入序列中有大部分 0 值，为了提高训练效率，做全部 PETime 的直方图如下图。

由图可知，绝大部分 PMT 相应集中在 150 ~ 450ns 范围内，因此在训练时只截取该时间范围内的序列作为输入。

3.3.2 重建结果

将一维卷积神经网络模型训练至性能最优后，在验证集上进行测试，得到卷积神经网络的重建精度为 61.73mm，即对于每一个事件，其重建位置和实际位置之间差别的绝对值平均为 61.73mm。

同时，绘制重建结果 z 坐标关于 $x^2 + y^2$ 的关系图像如下。

绘制真实事件位置的 z 坐标关于 $x^2 + y^2$ 的关系图像如下。

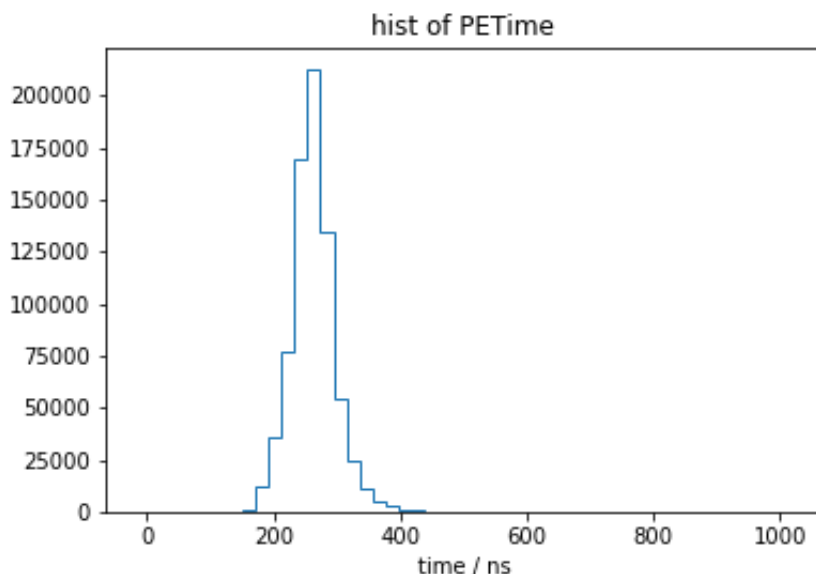


图 3.3 PETime 的直方图分布

将图 3.4与图 3.5对比可见，一维神经网络的重建结果保留了事件的空间对称性，但是对于处于探测器边缘距离 PMT 较近的事件，其重建效果并不理想。

与此同时，注意到图 3.4的重接结果与电荷加权平均法的重建类似。使用同样的数据，进行电荷加权平均重建结果如下图。

结合式 2-2，初步分析认为，一维卷积神经网络善于识别序列中的局部模式，同时空间位置上临近的 PMT 在接收光子的时间上也比较相近，因此这样的模式会被神经网络识别，相当于神经网络发现了隐藏在输入序列当中的空间信息，因此在重建结果上会与电荷加权平均法有一定相似性。

3.4 单一输入 LSTM 神经网络重建结果

3.4.1 神经网络结构

单一输入 LSTM 网络的结构如图所示，对于该网络，输入数据的处理与一维卷积神经网络相同。输入信号经过编码后，输入一层 32 维的 LSTM 神经网络，然后输入密集连接层与输出相连。

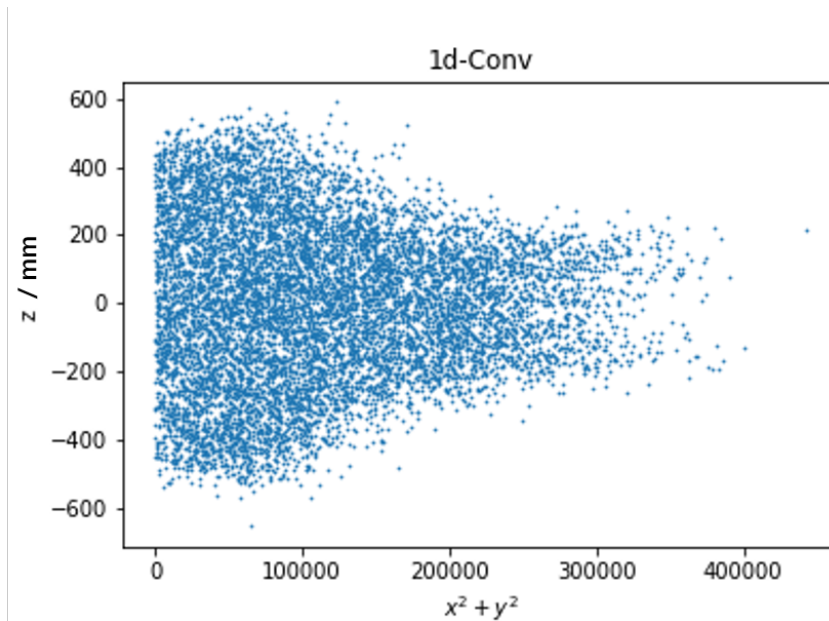


图 3.4 一维卷积神经网络重建结果空间分布

3.4.2 重建结果

将 LSTM 网络模型训练至性能最优后，在验证集上进行测试，得到卷积神经网络的重建精度为 146.88mm ，与一维卷积神经网络重建结果相比误差更大。

同时，绘制重建结果 z 坐标关于 $x^2 + y^2$ 的关系图像如下。

由图可见，LSTM 神经网络的重建结果与图 3.4 相比，损失了较多的对称性，重建结果在空间中的分布更加不均匀。

3.5 多输入 LSTM 神经网络重建结果

3.5.1 神经网络结构

上文报告的单一输入神经网络模型将 30 个 PMT 的信号整合成为了一个序列。从探测器产生数据的视角出发，可以尝试采用多输入 LSTM 神经网络，即将 30 个 PMT 的输出分别进行 One Hot 编码，分别输入 30 个独立的 LSTM 神经网络当中，最后使用一个密集连接层汇总输出。网络结构如下图。

这样的设计可以分别处理每个 PMT 上的信号，然后在密集连接层中寻找共同的模式或者 PMT 之间的几何关联，因此也有一定的应有前景。

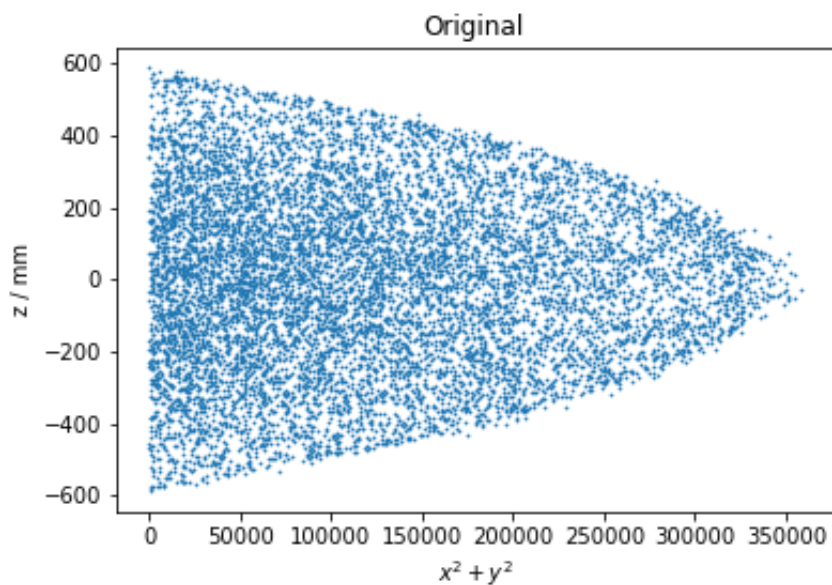


图 3.5 原始数据空间分布

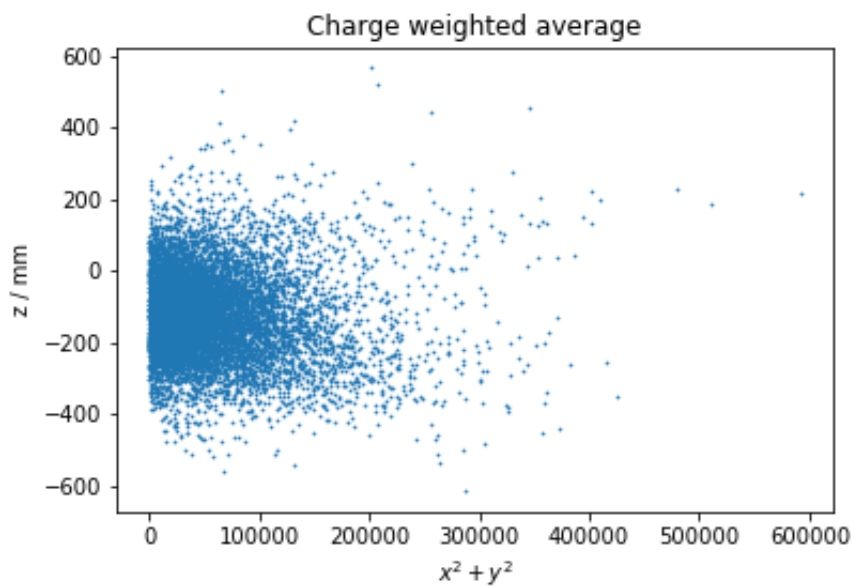


图 3.6 电荷加权平均重建结果空间分布

3.5.2 重建结果

根据上述结构设计网络并使用训练集数据进行训练，发现网络不收敛，训练误差持续维持在较高水平。对单一路输入信号绘图，得到结果如下。由图 3.10 初步分析，对于锦屏 1t 探测器，单个事件中单个 PMT 上产生的数据过于稀疏（平

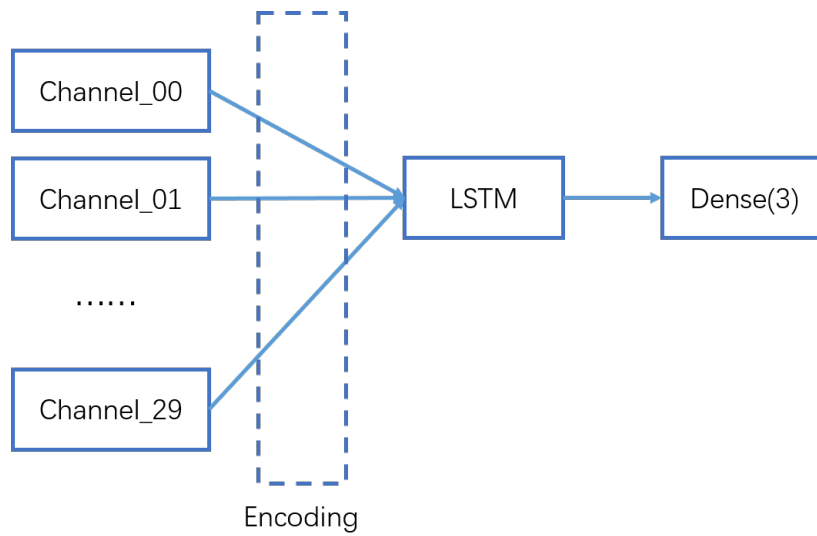


图 3.7 单一输入 LSTM 网络结构示意图

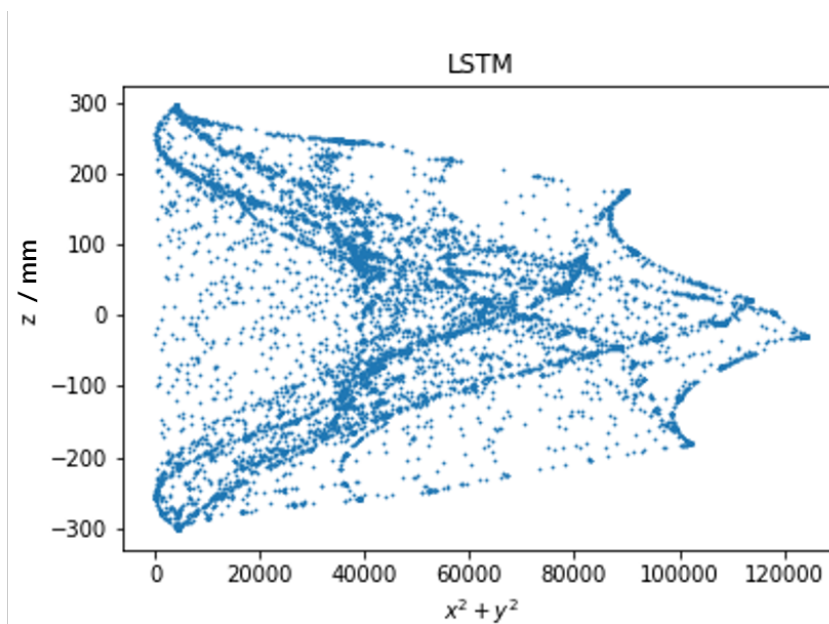


图 3.8 LSTM 神经网络重建结果空间分布

均 5 ~ 6 次响应), 输入 LSTM 网络中无法提供足够的信息用于网络参数更新。

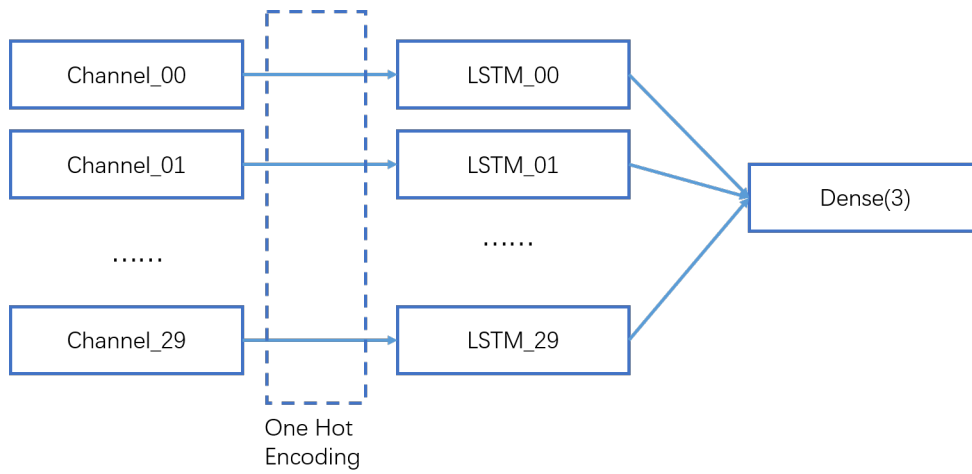


图 3.9 多输入 LSTM 网络结构示意图

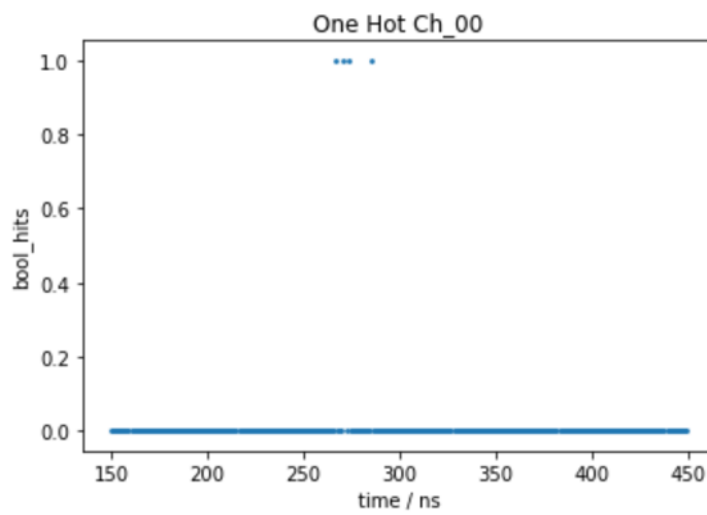


图 3.10 单通道输入信号

第 4 章 结论

4.1 论文工作总结

本文利用锦屏 1 吨探测器上的模拟数据，训练了一维卷积神经网络和 LSTM 神经网络。以重建精度为评价指标，衡量了两个网络的重建效果。其中一维卷积神经网络的重建精度在 61.73mm ，LSTM 神经网络的重建精度在 146.88mm ，两个网络结构的重建精度均有一定的提升空间。与此同时，本文还提出了一种多输入 LSTM 网络结构，但是训练过程暂不收敛，尚未得到可在文中报告的结果。此外，本文还对球面卷积神经网络的结构和实现方法进行了初步的调研，为之后的工作奠定了基础。

笔者在完成毕业设计的过程中，在探测器信号的初步处理、神经网络的设计以及网络参数的调节都积累了初步的经验，相关的经验和技能也将被应用于后续的学习和科研工作中。

4.2 当前局限与改进思路

由于笔者水平有限，没能在进行毕业设计的时间范围内围绕论文工作取得足够优秀的成果。笔者认为论文工作还可以围绕如下几方面进行改进：

首先，本文工作中使用的 Keras 框架是基于 TensorFlow 的高级语言，省略了很多底层的实现细节。该框架的优势是对于初学者较为友好，可以迅速上手开展工作；但是该框架的局限是在微调参数时比较受限，同时由于省略了过多的细节，使用者很能囫圇吞枣而不加深究，难以对深度学习的算法形成更深入的理解。与此同时，Keras 依赖的 TensorFlow 框架也存在兼容性方面的问题，在使用过程中可能会引入不必要的麻烦。相比之下，PyTorch 框架更为底层，当前应用也相对广泛。如果能够将工作迁移至 PyTorch 框架下，则可以对网络参数进行更细致的调节和操作，以获得更良好的结果。

其次，本文的主要工作均围绕锦屏 1 吨探测器展开，该探测器几何尺寸和 PMT 数量均有限，因此能够使用的神经网络结构也相对局限。如果能将工作迁移至更大的球形液闪探测器（如 JUNO 或 CJPL2000 吨），则会获得规模更大的

训练数据，能够采用的网络结构也可能更为多样，前文中报告的球面卷积神经网络与多输入 LSTM 网络会有更大的应用空间。

插图索引

图 1.1	基本粒子标准模型	1
图 1.2	1 吨原型机示意图 ^[12]	4
图 2.1	人工智能、机器学习和神经网络的概念关系	6
图 2.2	M-P 神经元模型 ^[17]	7
图 2.3	阶跃函数示意图 ^[17]	8
图 2.4	sigmoid 函数示意图 ^[17]	8
图 2.5	ReLU 函数示意图 ^[20]	9
图 2.6	二维卷积运算示意图 ^[23]	10
图 2.7	一维卷积示意图 ^[20]	11
图 3.1	一维卷积神经网络结构示意图	14
图 3.2	单次事件输入信号示例	15
图 3.3	PETime 的直方图分布	16
图 3.4	一维卷积神经网络重建结果空间分布	17
图 3.5	原始数据空间分布	18
图 3.6	电荷加权平均重建结果空间分布	18
图 3.7	单一输入 LSTM 网络结构示意图	19
图 3.8	LSTM 神经网络重建结果空间分布	19
图 3.9	多输入 LSTM 网络结构示意图	20
图 3.10	单通道输入信号	20

表格索引

表 3.1	GroundTruth 表数据结构	13
表 3.2	Truth 表数据结构	13

公式索引

公式 1-1	2
公式 1-2	2
公式 2-1	5
公式 2-2	5

参考文献

- [1] Chadwick J. Intensit atsverteilung im magnetischen spectrum der beta-strahlen von radium b+ c. Verhandl. Dtsc. Phys. Ges., 1914, 16:383
- [2] Enz C P. No time to be brief: A scientific biography of Wolfgang Pauli. Oxford University Press on Demand, 2010
- [3] Fermi E. An attempt of a theory of beta radiation. 1. Z. Phys., 1934, 88(UCRL-TRANS-726):161–177
- [4] 王贻芳. 大亚湾反应堆中微子实验. 物理, 2007, 36(03):0–0
- [5] Bilenky S M, Pontecorvo B. Lepton mixing and neutrino oscillations. Physics Reports, 1978, 41(4):225–261
- [6] Davis R. A review of the homestake solar neutrino experiment. Progress in Particle and Nuclear Physics, 1994, 32:13
- [7] Fogli G, Lisi E, Marrone A, et al. Solar neutrino oscillation parameters after first kamland results. Physical Review D, 2003, 67(7):073002
- [8] Fogli G, Lisi E, Marrone A, et al. Status of atmospheric neutrino $\nu \mu \nu \tau$ oscillations and decoherence after the first k2k spectral data. Physical Review D, 2003, 67(9):093006
- [9] 曹俊. 大亚湾与江门中微子实验. 中国科学: 物理学力学天文学, 2014, 44(10):1025–1040
- [10] An F, Balantekin A, Band H, et al. Spectral measurement of electron antineutrino oscillation amplitude and frequency at daya bay. Physical review letters, 2014, 112(6):061801
- [11] Beacom J F, Chen S, Cheng J, et al. Physics prospects of the jinping neutrino experiment. Chinese physics C, 2017, 41(2):023002
- [12] Wang Z, Wang Y, Wang Z, et al. Design and analysis of a 1-ton prototype of the jinping neutrino experiment. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2017, 855:81–87
- [13] Galbiati C, McCarty K. Time and space reconstruction in optical, non-imaging, scintillator-based particle detectors. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2006, 568(2):700–709
- [14] Liu Q, He M, Ding X, et al. A vertex reconstruction algorithm in the central detector of jun0. Journal of Instrumentation, 2018, 13(09):T09005
- [15] Simon H A, Newell A. Human problem solving: The state of the theory in 1970. American Psychologist, 1971, 26(2):145
- [16] Barr A, Feigenbaum E A, Cohen P R, et al. The handbook of artificial intelligence, volume 1. Addison-wesley Reading, MA, 1981

- [17] 周志华. 机器学习. 清华大学出版社, 2016
- [18] Michalski R S. A theory and methodology of inductive learning. *Machine learning*. Springer, 1983: 83–134
- [19] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 1943, 5(4):115–133
- [20] 弗朗索瓦·肖莱. Python 深度学习. 人民邮电出版社, 2018
- [21] Aurisano A, Radovic A, Rocco D, et al. A convolutional neural network neutrino event classifier. *Journal of Instrumentation*, 2016, 11(09):P09001
- [22] Hertel L, Li L, Baldi P, et al. Convolutional neural networks for electron neutrino and electron shower energy reconstruction in the nova detectors. *Cell*, 2017, 60(80):100
- [23] 斋藤康毅. 深度学习：基于 Python 的理论与实现. 人民邮电出版社, 2018
- [24] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural computation*, 1997, 9(8):1735–1780
- [25] Cohen T S, Geiger M, Köhler J, et al. Spherical cnns. arXiv preprint arXiv:1801.10130, 2018.

致 谢

首先感谢续老师同意指导我的完成毕业设计，并在神经网络入门、**LINUX** 服务器基本操作等方面给予了我细致的指导。学生愚钝，工作成果也有限，实不足报答老师的教诲，不禁深感愧疚。

其次感谢我父母在我居家期间给予我的关怀、包容与支持，线上教学期间的经历都十分难忘，能够长时间和父母一起在家中生活也十分幸福。

然后感谢课题组窦威师兄帮助我利用 **JSAP** 生成模拟数据，并在工作全过程中解答我或大或小的疑问；感谢课题组的各位老师和同学们在组会上对我的工作提出宝贵的意见。

与此同时，感谢高一川同学向续老师推荐我；感谢邹铨、王子涵同学开办 **L^AT_EX** 讲座；感谢吕加飞、齐博涵同学在我使用 **L^AT_EX** 排版过程中给予我极大的帮助。感谢清华大学 **TUNA** 协会维护的 **ThuThesis** 模板，让我在使用 **L^AT_EX** 排版时的效率显著提升。

笔者在完成毕业设计过程中受到的帮助远不止于此，对于所有给予我帮助的师友在此一并表示感谢，并祝愿大家身体健康，诸事顺遂。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：_____ 日 期：_____

综合论文训练记录表

学生姓名		学号		班级	
论文题目					
主要内容以及进度安排	<div style="text-align: right; margin-top: 20px;"> 指导教师签字：_____ </div> <div style="text-align: right; margin-top: 10px;"> 考核组组长签字：_____ </div> <div style="text-align: right; margin-top: 10px;"> 年 月 日 </div>				
中期考核意见	<div style="text-align: right; margin-top: 20px;"> 考核组组长签字：_____ </div> <div style="text-align: right; margin-top: 10px;"> 年 月 日 </div>				

指导教师评语	指导教师签字： _____ 年 月 日
评阅教师评语	评阅教师签字： _____ 年 月 日
答辩小组评语	答辩小组组长签字： _____ 年 月 日

总成绩： _____

教学负责人签字： _____

年 月 日