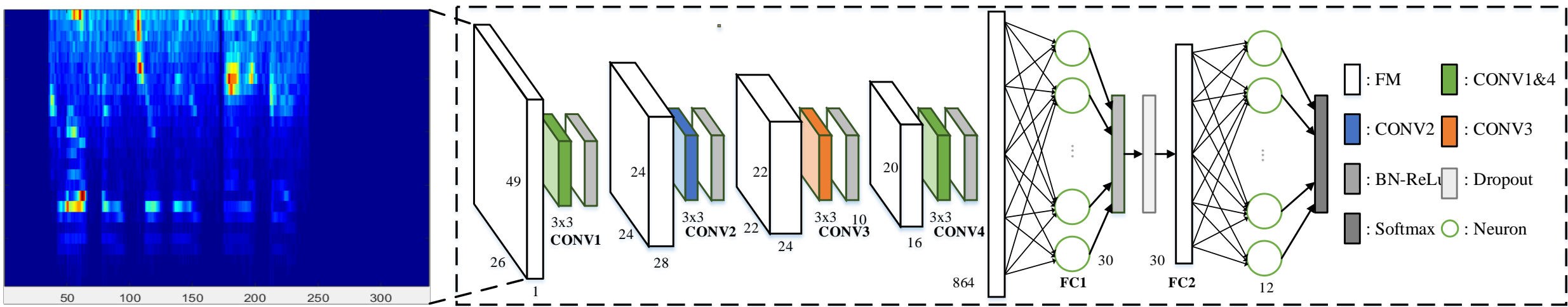


问题背景：当前人工智能，尤其是以人工神经网络为核心的深度学习已经成为多个领域学界、业界的研究热点。随着人工智能技术落地应用，智能语音识别作为人机交流接口的关键技术发展迅速。本题给出一个基于CNN（卷积神经网络）的语音关键词识别网络，网络结构如下：



Layer	Conv1	Conv2	Conv3	Conv4	FC1	FC2
Input	26×49×1	/	/	/	/	/
Output	24×24×28	22×22×24	10×20×16	8×9×12	30	12
Kernel size	3×3×28	3×3×24	3×3×16	3×3×12	30	12
Stride	(1, 2)	(1, 1)	(2, 1)	(1, 2)	/	/

Weight: 8bits  
Data: 16bits  
Bias: 16bits  
Activation Function: ReLu  
BN: Batch Normalization

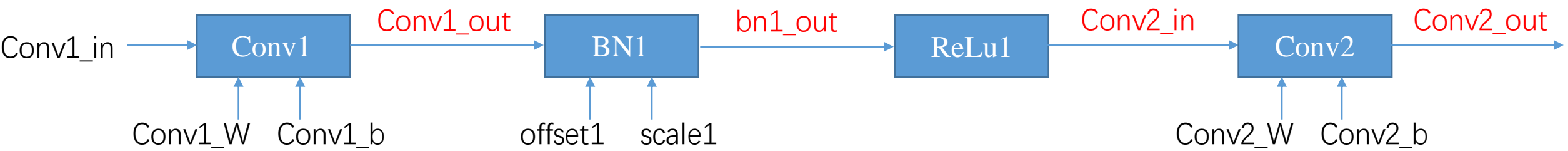
**问题1：** 根据提供的训练完成的神经网络，使用matlab复现第1层卷积正向过程。 (+30分)

**问题1\_plus：** 根据提供的训练完成的神经网络，使用matlab复现前两层卷积（以及之间的BN和Relu）正向过程。  
(附加分： +10分)

注意：数据位宽为 (1, 16, 14)，权重W的位宽为 (1, 8, 6)，bias的位宽为 (1, 16, 15)。其中括号内的内容表示 (符号位位宽，字长，小数位宽)。在复现正向的过程中，注意对每层计算后的数据进行定点化处理，以便于完成后续硬件代码。

题1和附加题1\_plus要求复现第1层卷积网络/前两层卷积网络，对应数据在data.mat文件中给出，共11个文件，数据流如下图所示。

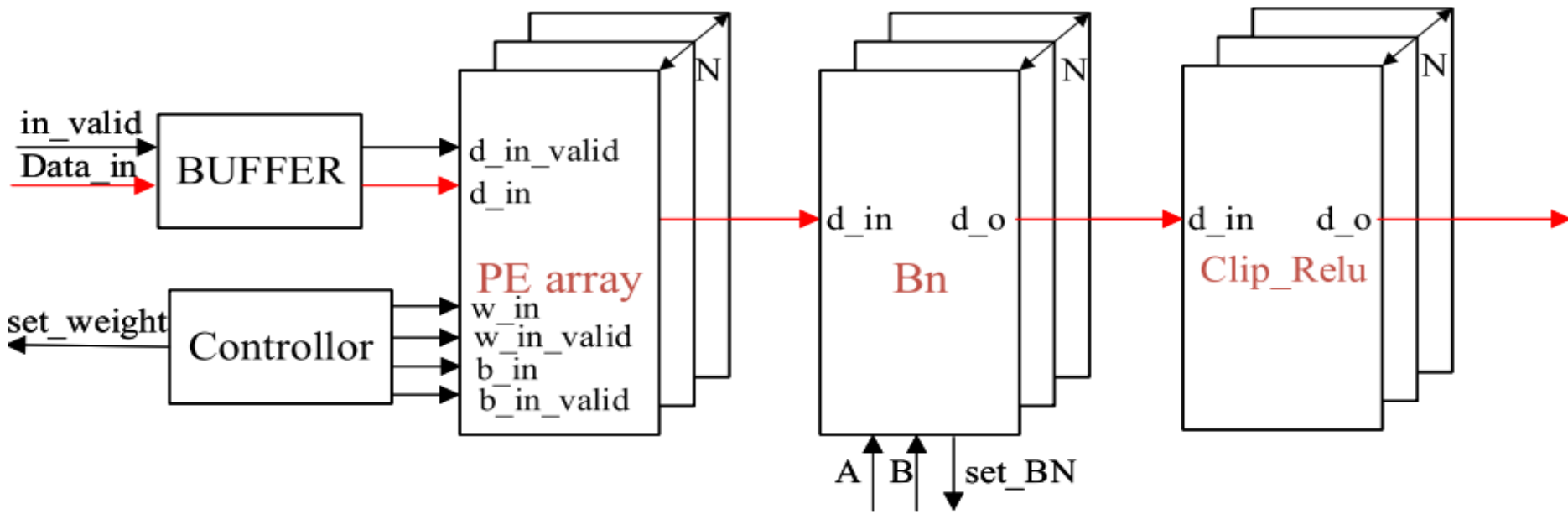
Conv1\_in为输入数据，  
Conv1\_W,Conv1\_b,offset1,scale1,Conv2\_W,Conv2\_b为网络权重（参数）  
Conv1\_out,bn1\_out,Conv2\_in,Conv2\_out为中间数据，是需要比对的数据



**问题2：**根据提供的神经网络，使用Verilog复现第1层卷积正向过程，完成RTL级代码，并进行仿真。（+70分）

**问题2\_plus：**根据提供的神经网络，使用Verilog复现前两层卷积（以及之间的BN和Relu）正向过程，完成RTL级代码，并进行功能仿真。（附加分：+10分）

注意：在运算过程中，每层计算完成的数据应当如何存储，如何调度。特别是对于附加题2\_plus，计算单元应该如何调度计算不同的网络层。（注意：两层卷积需要复用同一块硬件计算单元PE\_array）



**附加题3**：对问题2中用于对前两层卷积进行计算的硬件单元进行优化。（附加分：+10分）

优化方向：对PE\_array中的计算单元（乘法器、加法器，等）进行结构定制，考虑采用近似乘法器、加法器等电路结构，进一步减少所需要的硬件资源（晶体管数），并分析在不同的近似电路架构和参数配置下，计算误差的变化情况（例如，mean relative error distance (*MRED*), normalized mean error distance (*NMED*), probability of relative error distance smaller than 2% (*PRED*), 等）。

近似乘法器和加法器的结构可以参考如下文献：

- H. Jiang et al. "Approximate Arithmetic Circuits: A Survey, Characterization, and Recent Applications," Proceedings of the IEEE, 2020, doi:10.1109/JPROC.2020.3006451

- ❑ 题1和2共计100分，附加题30分，题目分共130分，最后得分总分不超过100分（超过100分，以100分计）
- ❑ 作业提交时间：2021.06.21，发送到邮箱：liubo\_seu@foxmail.com
- ❑ 作业提交方式：1个压缩包（学号+姓名+数字EDA设计题），压缩包中至少包含：Matlab代码，RTL代码（Verilog，+testbench），设计报告(word或pdf)。