

# 网络结构搜索 (NAS) 实验计划

Yi-Qi Hu

2019 年 6 月 3 日

## 1 前言

在网络结构搜索中可以被分为三个阶段，拓扑结构搜索、结构操作搜索和网络结构评估。先前的实验对这三个阶段进行了全过程探索，实验结果表明，网络结构搜索的效率不高。因此在此实验计划中，我们将对每个阶段分开进行探索，用以得到更加合理的修正计划。

为了方便的对实验进行描述，先给出一些简单的符号定义。用  $S \in \mathcal{S}$  表示一个网络拓扑结构，其中  $\mathcal{S}$  表示网络拓扑结构空间。用  $O_S \in \mathcal{O}_S$  表示在  $S$  上的一个操作集合。用  $D$  表示评估数据集 (因为验证测试集是固定的，所以  $D$  仅表示训练数据集)。用  $f(\cdot)$  表示评估准则，评判在数据集上网络结构的好坏，即  $f(O_S, D)$ 。

接下来的部分将分别就三个阶段进行叙述。

## 2 网络结构评估阶段

网络结构评估阶段的输入是一组网络结构配置，包括网络拓扑结构和依赖于网络拓扑结构的操作结合，用  $O_S$  来表示。网络结构评估的基础是数据集，在评估时，验证数据集固定，因此我们只需要考虑训练数据集。用  $D_{r,e}$  表示训练数据集，其中  $r$  表示 subsample ratio,  $e$  表示 epoch, 通过对  $r$  和  $e$  的设置能得到不同代价的评估数据集。

### 2.1 现有方法

现有实验中，训练数据量每轮递增，即  $r$  参数每轮按照固定的比例增长，训练轮数固定， $e = 30$ 。现有评估方式简单，实现较为容易，仅对实验中的数据进行少量操作即可完成，开销较小。缺点在于由于每轮中的评估数据不同，因此轮数之间的评估结果无法比较。

### 2.2 新实验：评估修正调查

#### 2.2.1 实验 1.1-评估修正调查

**目的：**探究评估结果能否在 subsample ratio 和 epoch 两个维度上，用很少的代价进行修正。

**方法：**修正问题是一个回归问题，输入是 subsample ratio 和 epoch，即  $r$  和  $e$ 。回归的目标是评估准则，如 accuracy 等，记做  $y$ 。回归中还需要考虑效率问题，即尽可能使用少量的观察样本，得到一个尽可能准确的回归结果。在本次实验中，采用先规定好回归函数形状，然后

进行微调的方式进行。令  $g(\cdot)$  为一个函数，则  $y = g(w_2r + w_1e + w_0)$ 。这种方式，仅需要 3 个观察样本便能得到回归参数  $(w_0, w_1, w_2)$  的闭式解。但是关键问题是如何得到函数  $g$ ， $g$  的选择对回归效果的影响很大，因此，此次实验主要的目的对回归函数形状进行探究。

**实验步骤：**

- 选定  $O_S$ ，在  $D_r$  上进行训练，对每一个 epoch 记录下  $y$ ，便得到了一组  $(r, e; y)$  对；
- 增加  $r$ ，重复上述步骤，得到另一组观察值对；
- 对同一个  $S$ ，调整  $O$ ，重复上述两步，得到观察结果；
- 对不同的  $S$ ，重复上述三步，得到观察结果；
- 将采样结果绘制成图，观察函数变化形状。

本次实验的结果用三维图表示，每一张图上均是  $y$  随着  $r$  和  $e$  变化的结果，对每个  $O_S$  均有一张图。

**预计时间：**2 周

### 3 结构操作搜索阶段

操作搜索的目的是为每一个拓扑结构配置一组操作序列，构成一个完整的网络结构，使其具有最好的表现。可归结为如下优化问题：

$$O_S^* = \operatorname{argmax}_{O_S \in \mathcal{O}_S} f(O_S, D). \quad (3.1)$$

#### 3.1 现有方法

对拓扑结构上的每个节点做一次独立的赋值，赋值的方式是在一个树形结构上，用一次随机过程决定此节点的操作及其具体的参数。树形结构的分支上携带一概率，表示选中此操作/参数的概率，这样优化目标转移为概率的设置空间，每一轮拓扑结构筛选更新一次概率配置。现有方法的优势在于优化空间归一化为一个固定的概率搜索空间，搜索空间的维度不随着拓扑空间的变化而变化。缺点在于配置选择依赖于随机过程，没有考虑连续两个操作之间的关系，可靠性效率均不足；每轮拓扑结构筛选更新一次概率配置，优化目标不正确。

#### 3.2 新实验：结构操作搜索优化

##### 3.2.1 实验 2.1-操作配置优化

**目的：**解决结构操作搜索现有方法中，优化目标不正确的问题。

**方法：**在此实验中尝试两种方法，第一种仍然采用现有方法中的树形结构对操作/参数进行配置，但是对于每个拓扑结构均启用一个优化过程，优化概率空间设置；第二种摒弃树形结构，直接对操作/参数配置空间进行优化，每一个节点对应一组配置参数表，所以总体的优化维度和节点个数相关。

**实验步骤：**

- 选取数个拓扑结构,  $\{S_1, S_2, \dots\}$ , 在每一个拓扑结构上启动优化过程, 评估过程中固定  $D_{r,e}$ , 优化方法选用 RACOS;
- 采用方法描述中第一种实验方式, 设置优化空间 (树形结构的概率空间), 进行优化;
- 采用方法描述中第一种实验方式, 设置优化空间 (操作/参数空间), 进行优化;
- 重复实验, 观察两种方式的实验结果。

此实验的结果可用表表示, 记录在不同的拓扑结构下, 两种方法的实验结果和效率。

预计时间: 3 周

### 3.2.2 实验 2.2-利用先验进行操作配置

**目的:** 探究是否能够更加有效的利用先验知识对操作进行配置。

**方法:** 暂时能够想到两种方法, 第一种是学习的方法, 用现有的网络结构作为训练数据, 采用时序模型, 如 HMM or RNN 等对网络操作序列进行建模, 用建模后的模型, 对拓扑结构上的节点进行预测, 并将预测结果作为初始化; 第二种是通过扩展现有的网络操作配置得到较好的初始化操作配置。

**实验步骤:** 具体的实验步骤需要根据任务归约确定, 还需要进行进一步的实验设计。

预计时间: 4 周

## 4 拓扑结构搜索阶段

拓扑结构搜索是为了得到合适的网络拓扑结构, 即一个合理的数据流程图。

### 4.1 现有方法

现有的方法是在一定约束内 (如深度、宽度等) 穷举所有的网络拓扑, 然后通过筛选得到最后的拓扑结构。此方法的优点在于能够对所有潜在的网络结构进行探索; 缺点是穷举的结构数量太多, 产生异常繁重的评估开销。

### 4.2 新实验: 拓扑结构调查

#### 4.2.1 实验 3.1-拓扑结构搜索收益调查

**目的:** 拓扑结构搜索和结构操作搜索均需要消耗评估资源, 此实验为了探索拓扑结构搜索和定拓扑结构对操作进行搜索, 两者之间哪个更加重要, 即哪个阶段所能获得的收益更大。

**方法:** 对于此实验, 我们要探究的是对于一个拓扑结构, 其操作设置决定的评估值是服从一个什么样的分布。以及, 对于不同拓扑结构, 其评估值分布的差异性是否存在。如果存在, 进一步探索差异性是由什么因素决定的 (深度, 宽度, 链接方式 or 其他)。

**实验步骤:**

- 对于一个确定的网络拓扑结构  $S$ , 尽可能穷举出所有的参数配置, 并在数据集上进行评估, 记录下对应的评估值;

表 1: 实验设计总结。

实验	内容	使用技术	时间	负责人
实验 1.1	评估修正调查	CNN	2 周	XXX
实验 2.1	操作配置优化	CNN, RACOS	3 周	XXX
实验 2.2	利用先验进行操作配置	CNN, HMM, RNN 等	4 周	XXX
实验 3.1	拓扑结构搜索收益调查	CNN	3 周	XXX

- 选取不同的网络拓扑结构，注意选取的拓扑结构需要具有特殊的特性，比如说深度，宽度，链接方式上需要具有不同的特点，重复上一步骤；
- 将结果绘制成图表，方便观察和比较。

此实验结果可用图表示，可以将穷举/随机生成的结果和人为设计的操作设置进行比较。

预计时间：3 周

## 5 总结

本次实验设计总结如表 1 所示，本文仅给出了本次实验的目的，大致方法和结果要求，每个实验均需要根据情况具体设计并实现。在实验过程中需要记录详细的设计和实验结果，方法实现的每个模块需要有详细的测试计划和报告。