

# Дифференцируемый алгоритм поиска архитектуры модели с контролем её сложности

*К. Д. Яковлев, О. С. Гребенькова, О. Ю. Бактеев*

iakovlev.kd@phystech.edu; grebenkova.os@phystech.edu; bakhteev@phystech.edu

В работе исследуется задача построения модели глубокого обучения. Предлагается метод поиска архитектуры модели, позволяющий контролировать её сложность с небольшими вычислительными затратами. Под сложностью модели понимается минимальная длина описания, минимальное количество информации, которое требуется для передачи информации о модели и о выборке. В основе метода лежит дифференцируемый алгоритм поиска архитектуры модели (DARTS). Контроль сложности параметров производится гиперсетью. Под гиперсетью понимается модель, генерирующая параметры другой модели. Предлагается использовать гиперсеть в качестве функции релаксации. Предложенный метод позволяет контролировать сложность модели в процессе поиска архитектуры. Для оценки качества предлагаемого алгоритма проводятся эксперименты на выборках MNIST и CIFAR-10.

**Ключевые слова:**

## 1 Введение

В последнее время растет интерес к разработке алгоритмических решений для автоматизации процесса проектирования архитектуры. Лучшие существующие алгоритмы поиска архитектуры требуют больших вычислительных затрат, несмотря на их высокую производительность.

В данной работе рассматривается задача поиска архитектуры модели глубокого обучения с контролем её сложности. В качестве базового алгоритма используется дифференцируемый алгоритм поиска архитектуры (DARTS) [1]. Данный метод решает задачу поиска архитектуры модели путем перевода пространства поиска из дискретного в непрерывное представление. Градиентные методы оптимизации позволяют использовать меньше вычислительных ресурсов. Данный алгоритм универсален для работы как со сверточными, так и с рекуррентными нейронными сетями.

В работе [2] стабильность алгоритма DARTS была поставлена под сомнение. Одним из источников нестабильности является этап получения фактической дискретной архитектуры из архитектуры непрерывной смеси. На этом этапе часто наблюдается снижение качества модели. В данной работе веса модели формируются как минимизатор случайно сглаженной функции, определяемой как ожидаемая потеря в окрестности текущей архитектуры. В работе [3] предлагается использовать 0-1 функцию потерь для уменьшения расхождения между дискретной архитектурой и архитектурой непрерывной смеси.

Предлагаются альтернативные подходы к решению задачи поиска архитектуры модели. В работе [4] формулируется задача обучения распределению с ограничениями. Предложенный метод может быть эффективно оптимизирован и обладает теоретическими преимуществами для повышения способности к обобщению.

В работе [5] строится алгоритм поиска нейронной архитектуры с ограниченным ресурсом (RC-DARTS). К базовому алгоритму DARTS добавляются ресурсные ограничения, такие как размер модели и вычислительная сложность. Для решения задачи условной оптимизации вводится алгоритм итерационной проекции.

Работе [6] исследует гиперсети. Подход заключается в использовании небольшой сети для генерации весов более крупной сети. Рассматривались два варианта использования гиперсетей: статические гиперсети для генерации весов для сверточной сети и динамические гиперсети для генерации весов рекуррентной сети.

Вычислительный эксперимент будет проводиться на выборках MNIST [7] и CIFAR-10 [8].

## 2 Постановка задачи

### 2.1 Дифференцируемый алгоритм поиска архитектуры

Пусть  $\mathcal{O}$  – множество операций. Введем смешанную операцию для каждого ребра  $(i, j)$ :

$$\hat{o}^{(i,j)}(x) = \sum_{o \in \mathcal{O}} \frac{\exp(\alpha_o^{(i,j)})}{\sum_{o' \in \mathcal{O}} \exp(\alpha_{o'}^{(i,j)})} o(x), \quad (1)$$

где  $\alpha_o^{(i,j)}$  обозначает соответствующий вес операции  $o$  на ребре  $(i, j)$ . Таким образом, каждому ребру  $(i, j)$  ставится в соответствие вектор  $\alpha^{(i,j)}$  размерности  $|\mathcal{O}|$ . Пусть  $\alpha = [\alpha^{(i,j)}]$ . Сформулируем двухуровневую задачу оптимизации:

$$\begin{aligned} & \min_{\alpha} \mathcal{L}_{val}(\mathbf{w}^*(\alpha), \alpha), \\ \text{s.t. } & \mathbf{w}^* = \arg \min_{\mathbf{w}} \mathcal{L}_{train}(\mathbf{w}, \alpha) \end{aligned} \quad (2)$$

Здесь  $\mathcal{L}_{val}$  и  $\mathcal{L}_{train}$  функции потерь модели на валидации и на обучении соответственно.

---

#### Алгоритм 1 DARTS – Differentiable Architecture Search

---

- 1: Для каждого узла создадим смешанную операцию  $\hat{o}^{(i,j)}$ , параметризованную  $\alpha^{(i,j)}$
  - 2: **пока** алгоритм не сошелся
  - 3:   обновим  $\alpha$ , сделав градиентный шаг вдоль  $\nabla_{\alpha} \mathcal{L}_{val}(\mathbf{w} - \xi \nabla_{\mathbf{w}} \mathcal{L}_{train}(\mathbf{w}, \alpha), \alpha)$
  - 4:   обновим веса  $\mathbf{w}$ , делая градиентный шаг вдоль  $\nabla_{\mathbf{w}} \mathcal{L}_{train}(\mathbf{w}, \alpha)$
  - 5: получить окончательную архитектуру из полученного в результате алгоритма  $\alpha$
- 

### 2.2 Линейная гиперсеть

Пусть  $\Lambda$  – множество параметров, контролирующие сложность модели. Под гиперсетью мы будем понимать следующее отображение:

$$\mathbf{G} : \Lambda \times \mathbb{U} \rightarrow \mathbb{W}, \quad (3)$$

где  $\mathbb{W}$  – множество параметров модели, а  $\mathbb{U}$  – множество параметров гиперсети.

В данной работе рассматривается линейная гиперсеть:

$$\mathbf{G}_{linear}(\lambda) = \lambda \mathbf{b}_1 + \mathbf{b}_2, \quad (4)$$

где  $\mathbf{b}_1, \mathbf{b}_2$  – не зависящие от  $\lambda$  константы.

## Литература

- [1] *Liu Hanxiao, Simonyan Karen, Yang Yiming*. Darts: Differentiable architecture search // CoRR, 2018. Vol. abs/1806.09055. URL: <http://arxiv.org/abs/1806.09055>.
- [2] *Chen Xiangning, Hsieh Cho-Jui*. Stabilizing differentiable architecture search via perturbation-based regularization // CoRR, 2020. Vol. abs/2002.05283.
- [3] *Chu Xiangxiang, Zhou Tianbao, 0046 Bo Zhang, Li Jixiang*. Fair darts: Eliminating unfair advantages in differentiable architecture search // CoRR, 2019. Vol. abs/1911.12126. URL: <http://arxiv.org/abs/1911.12126>.
- [4] *Chen Xiangning, Wang Ruochen, Cheng Minhao, Tang Xiaocheng, Hsieh Cho-Jui*. Drnas: Dirichlet neural architecture search // CoRR, 2020. Vol. abs/2006.10355.
- [5] *Jin Xiaojie, Wang Jiang, Slocum Joshua, 0001 Ming-Hsuan Yang, Dai Shengyang et al*. Rc-darts: Resource constrained differentiable architecture search // CoRR, 2019. Vol. abs/1912.12814. URL: <http://arxiv.org/abs/1912.12814>.
- [6] *Ha David, Dai Andrew M., Le Quoc V*. Hypernetworks // CoRR, 2016. Vol. abs/1609.09106. URL: <http://arxiv.org/abs/1609.09106>.
- [7] *LeCun Yann, Cortes Corinna*. MNIST handwritten digit database, 2010. URL: <http://yann.lecun.com/exdb/mnist/>.
- [8] *Krizhevsky Alex, Nair Vinod, Hinton Geoffrey*. Cifar-10 (canadian institute for advanced research). URL: <http://www.cs.toronto.edu/~kriz/cifar.html>.

*Received February 25, 2021*