

Гиперсети (Hypernetworks) в поиске нейронных архитектур (NAS)

Презентация по теме NAS

На основе современных исследований в области Neural Architecture Search

9 ноября 2025 г.

Что такое гиперсеть (Hypernetwork)

Гиперсеть — это нейронная сеть, которая **генерирует веса** для **другой нейросети**. То есть гиперсеть сама не решает задачу напрямую, а создаёт параметры для модели, которая решает задачу.

$$H(\phi, a) \rightarrow \theta_a$$

- H — гиперсеть с параметрами ϕ .
- a — описание архитектуры (код структуры сети).
- θ_a — веса, сгенерированные гиперсетью для архитектуры a .

Далее создаётся модель $f_a(x; \theta_a)$ — архитектура с сгенерированными весами.

Как работает гиперсеть (пошагово)

- ❶ **Кодирование архитектуры:** архитектура преобразуется в вектор признаков a .
- ❷ **Генерация весов:** гиперсеть $H(a)$ возвращает веса θ_a .
- ❸ **Оценка модели:** веса θ_a используются для проверки точности f_a на валидационных данных.
- ❹ **Обучение гиперсети:** гиперсеть учится генерировать хорошие веса для разных архитектур.

Зачем нужны гиперсети в NAS

- Обучение каждой архитектуры с нуля слишком дорого.
- Гиперсеть позволяет **один раз обучить её и сразу генерировать веса для множества архитектур**.
- Можно быстро оценивать новые архитектуры без полного обучения.
- Превращает NAS в задачу генерации и проверки архитектур на лету.

Пример использования: SMASH (Brock et al., 2018)

- 1 Гиперсеть получает на вход код архитектуры.
- 2 Генерирует веса для соответствующей архитектуры.
- 3 Архитектура тестируется на валидационных данных для оценки качества.
- 4 Лучшие архитектуры сохраняются, слабые отбрасываются.

Итог

NAS с гиперсетью позволяет **ускорить поиск архитектур** и **снизить вычислительные затраты**.

- **Сильное ускорение NAS** — нет необходимости обучать каждую архитектуру с нуля.
- **Обобщение** — гиперсеть может генерировать хорошие веса для ранее не виденных архитектур.
- **Экономия ресурсов** — обучение одной гиперсети дешевле, чем тысячи отдельных моделей.
- **Гибкость** — легко оценивать большое пространство архитектур.





- Размер гиперсети может быть большим, особенно для сложных архитектур.
- Точность сгенерированных весов обычно ниже, чем после полного обучения.
- Требуется тщательно выбирать архитектуру гиперсети и кодирование архитектур.
- Иногда трудно масштабировать на очень большие пространства поиска.

Главная мысль

Гиперсети позволяют **генерировать веса** для любой архитектуры **NAS** без её полного обучения, что ускоряет поиск и снижает вычислительные затраты.

Плюсы: скорость, экономия ресурсов, обобщение. **Минусы:** возможная потеря точности, сложность масштабирования.

Список литературы I

-  Brock et al. (2018). *SMASH: One-Shot Model Architecture Search through HyperNetworks*.
-  Liu et al. (2018). *Progressive Neural Architecture Search*. ECCV.
-  Bender et al. (2018). *Understanding and Simplifying One-Shot Architecture Search*. ICML.
-  Zhang et al. (2019). *Graph HyperNetworks for Neural Architecture Search*. ICLR.