

Морфизмы сети (Network Morphisms) в поиске нейронных архитектур

Презентация по теме NAS

На основе современных исследований в области Neural Architecture Search

9 ноября 2025 г.

Что такое морфизмы сети

Network Morphism — это метод изменения структуры нейронной сети без разрушения её текущей функции. Если сеть $f(x; \theta)$ реализует функцию, то после модификации $f'(x; \theta')$:

$$f'(x; \theta') = f(x; \theta)$$

при корректной инициализации параметров θ' .

Идея

Менять архитектуру сети (добавлять слои, фильтры, нейроны), так чтобы поведение модели оставалось прежним — и затем продолжать обучение уже обновлённой сети.

Примеры морфизмов

- **Добавление слоя** Новый слой инициализируется как тождественное преобразование ($W' = I$), чтобы не изменить выход сети.
- **Расширение слоя** Увеличение числа нейронов или фильтров. Новые веса устанавливаются в нули, чтобы не влиять на выход.
- **Дублирование фильтров** Копирование существующих фильтров с небольшими вариациями — позволяет увеличить ёмкость модели.

Результат

После модификации сеть остаётся функционально эквивалентной исходной, но становится способной учиться дальше, используя новую структуру.

Зачем нужны морфизмы сети

Проблема: обучение каждой новой архитектуры с нуля требует больших вычислительных ресурсов.

Решение: использовать веса ранее обученных моделей, изменяя их структуру без потери выученной функции.

- Сеть может **расти постепенно**, сохраняя уже выученные знания.
- Не требуется начинать обучение заново — достаточно дообучить после модификации.
- Позволяет NAS **эволюционировать** архитектуру, а не строить новую с нуля.

Преимущества Network Morphisms

- **Сохранение знаний:** новая сеть наследует функциональность старой.
- **Ускорение поиска:** меньше времени на обучение новых архитектур.
- **Гладкая эволюция:** постепенное усложнение модели без потери качества.

Применение

Метод активно используется в эволюционных стратегиях NAS и при переносе знаний (*Net2Net*, *NetMorph*).

- **Рост сложности:** строгие морфизмы только добавляют элементы, но не позволяют их удалять.
- **Риск переусложнения:** архитектура может стать слишком громоздкой.
- **Ограниченность операций:** не все архитектурные изменения можно реализовать без потери функции.

Решение: использование *приближённых морфизмов*, допускающих минимальное изменение функции для возможности уменьшения архитектуры.

Пример: добавление слоя через морфизм

example_network_morphism.png

Главная идея






Морфизмы сети позволяют **изменять архитектуру нейронной сети без потери уже выученных знаний**.

Ключевые преимущества:

- Сокращение затрат на обучение.
- Возможность непрерывного роста сети.
- Повышение эффективности NAS за счёт переиспользования весов.

Ключевая трудность: Контроль сложности модели и разработка приближённых морфизмов, допускающих как расширение, так и сокращение сети.

Список литературы I

-  Chen et al. (2016). *Net2Net: Accelerating Learning via Knowledge Transfer.*
-  Wei et al. (2016). *Network Morphism.*
-  Elsken et al. (2019). *Efficient Multi-objective Neural Architecture Search via Lamarckian Evolution.*
-  Elsken et al. (2018). *Simple and Efficient Architecture Search for Convolutional Neural Networks.*
-  Chen et al. (2016). *Net2Net: Accelerating Learning via Knowledge Transfer.*