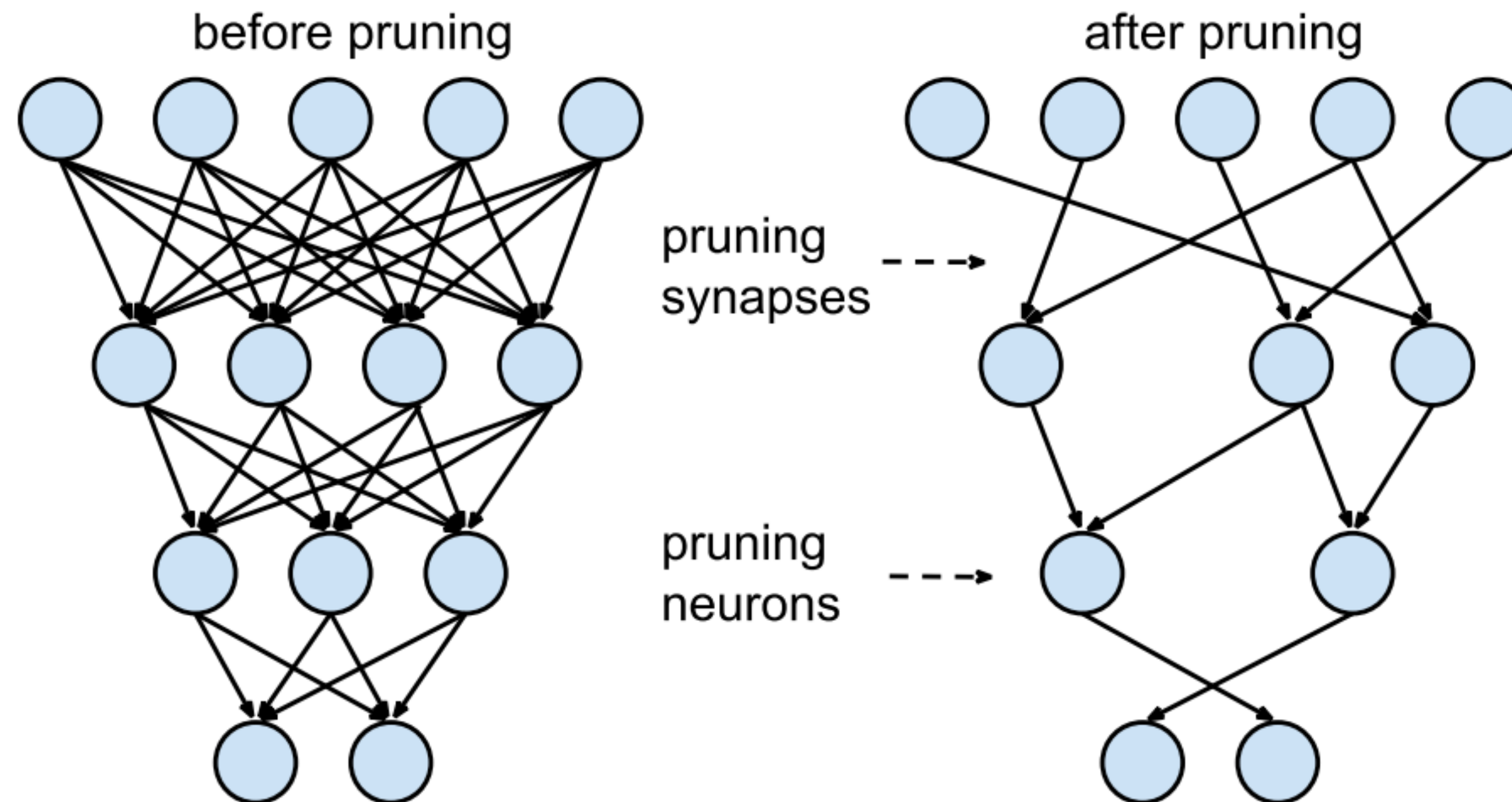


The Lottery Ticket Hypothesis

Морозов Дмитрий, БПМИ172

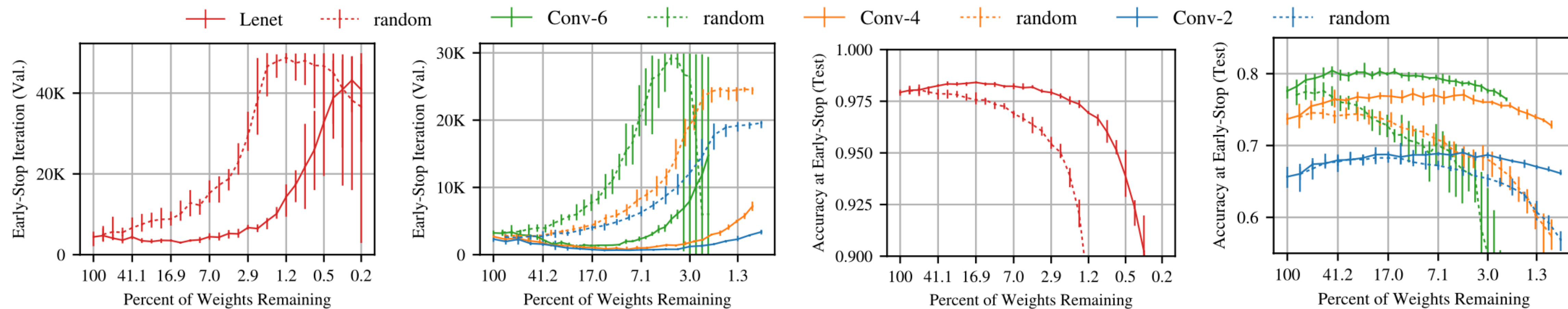
Pruning — обрезка, сокращение



Гипотеза лотерейного билета

Плотные, случайно инициализированные сети прямой связи содержат подсети (*выигрышные билеты*), которые, при обучении в изоляции, достигают точности теста, сравнимой с исходной сетью за аналогичное количество итераций.

Гипотеза лотерейного билета



Определение выигрышных билетов

Алгоритм:

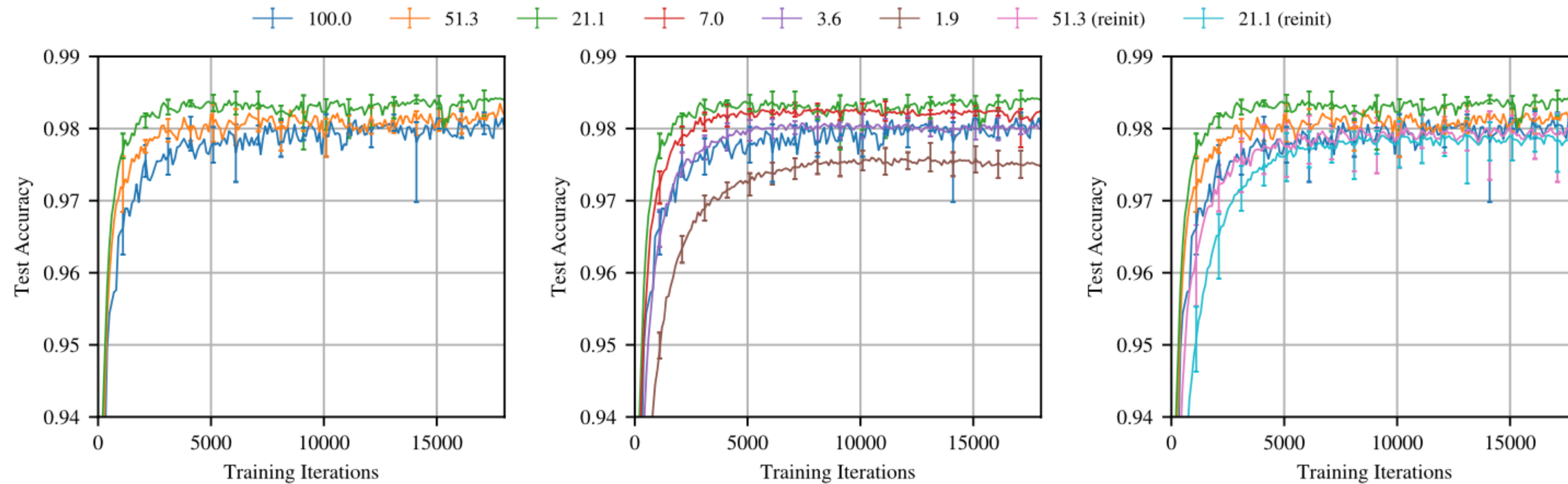
1. Произвольно инициализируйте нейронную сеть $f(x; \theta_0)$.
2. Обучите сеть для j итераций, приводя к параметрам θ_j .
3. Сократите $p\%$ параметров в θ_j , создав маску m .
4. Сбросьте остальные параметры до их значений в θ_0 , создав выигрышный билет $f(x; m \odot \theta_0)$.

Итог:

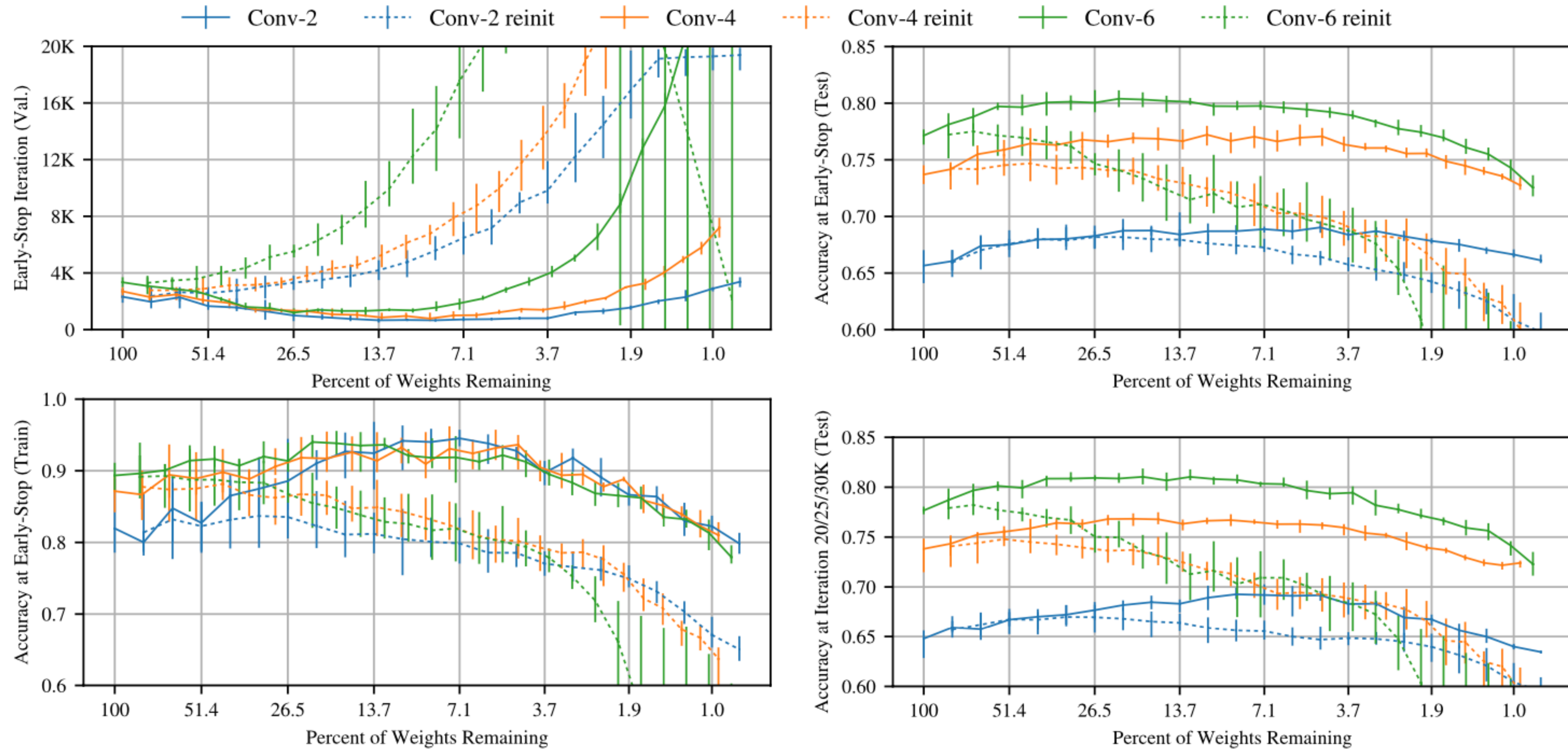
- Обрезка обнаруживает обучаемые подсети, которые достигают точности тестирования, сопоставимой с исходными сетями, из которых они получены при сопоставимом количестве итераций.
- Обрезка находит выигрышные билеты, которые учатся быстрее, чем оригинальная сеть, в то же время достигая более высокой точности на тесте.
- Гипотезу лотерейного билета можно рассматривать как новый взгляд на процесс составления нейронных сетей.

Выигрышные билеты в полносвязных нейронных сетях

Замечание: $P_m = \frac{\|m\|_0}{|\theta|}$ является разреженностью маски m , например, $P_m = 25\%$ при обрезке 75% веса.

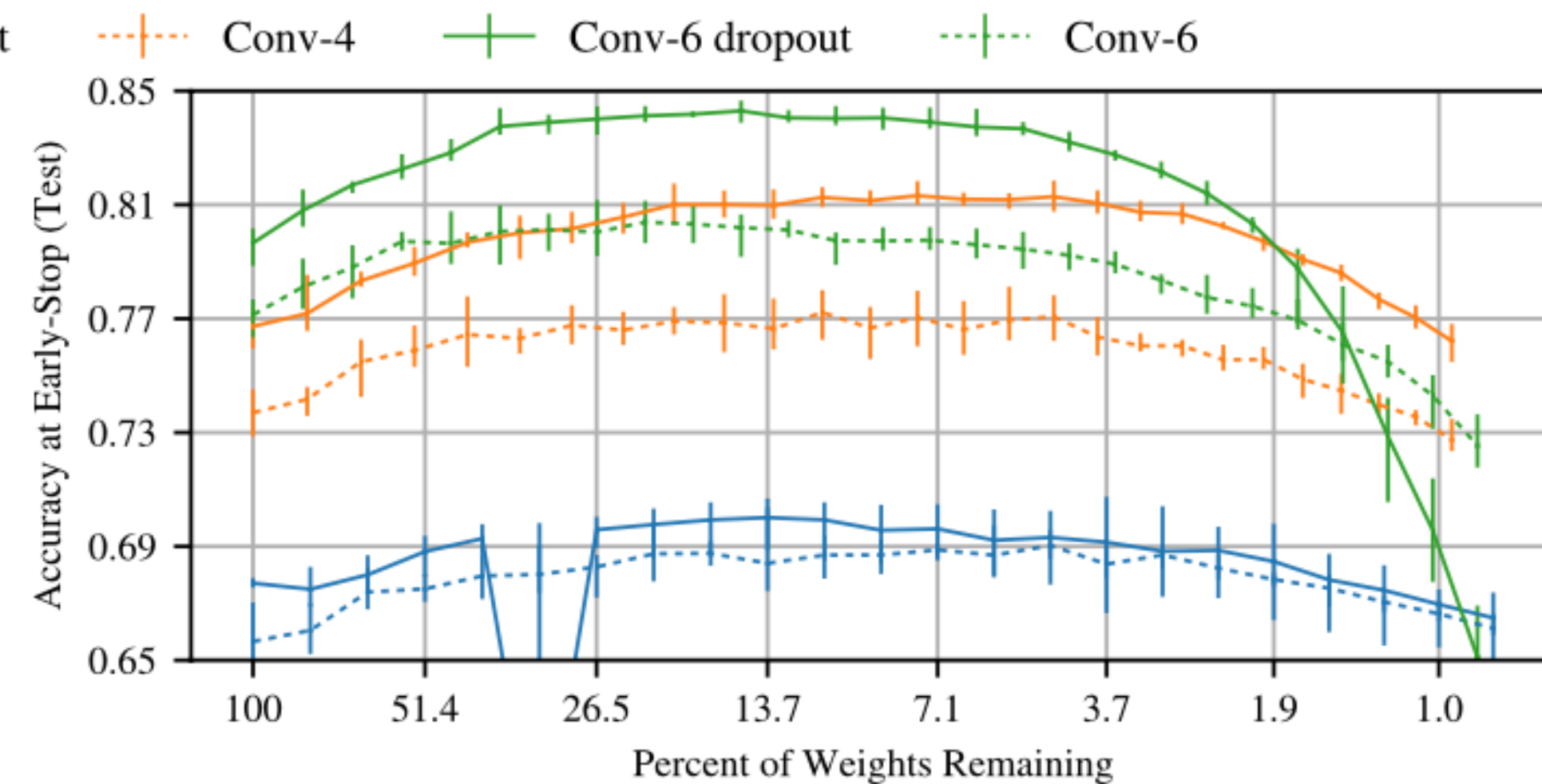
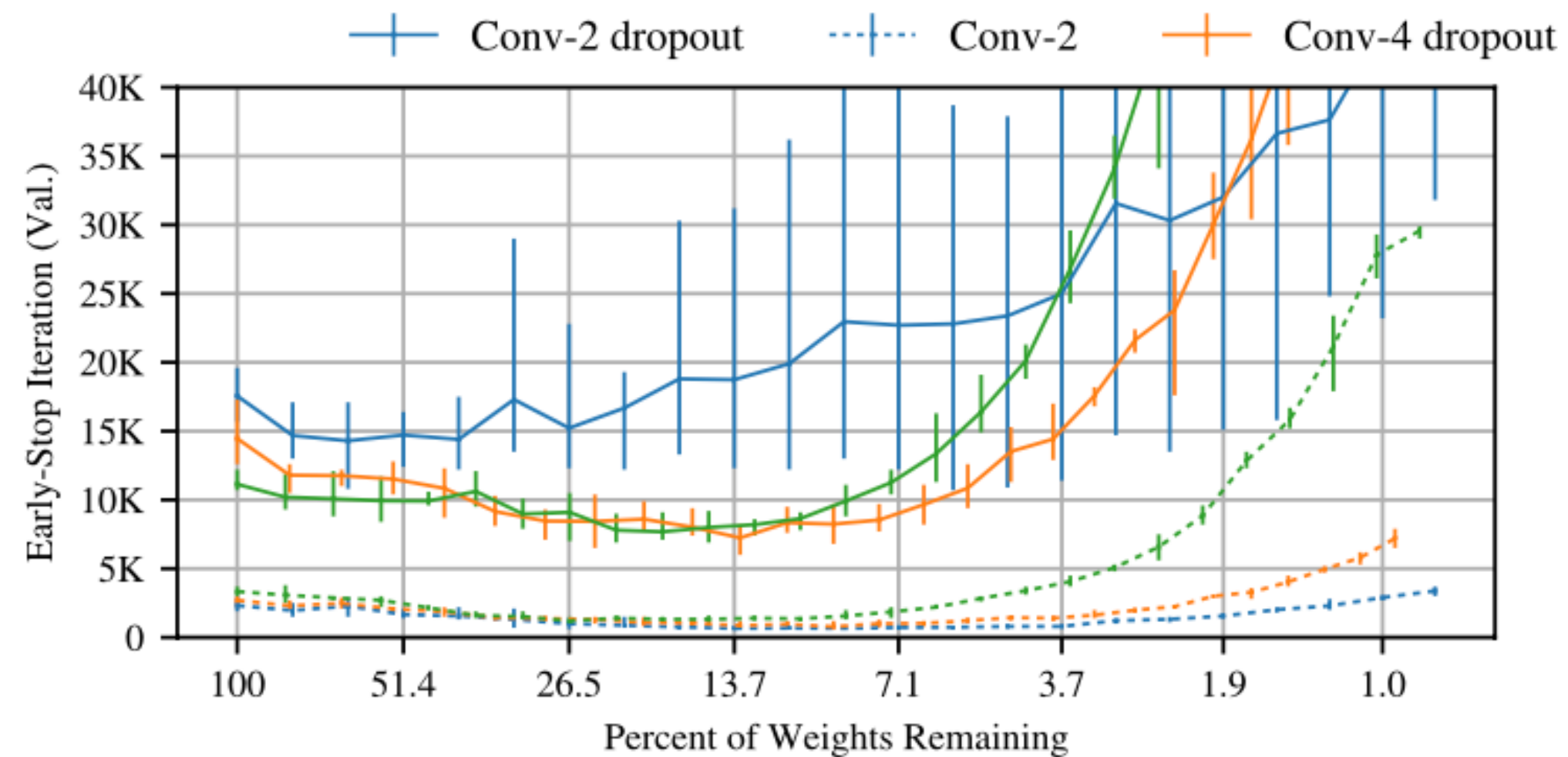


Выигрышные билеты в свёрточных нейронных сетях

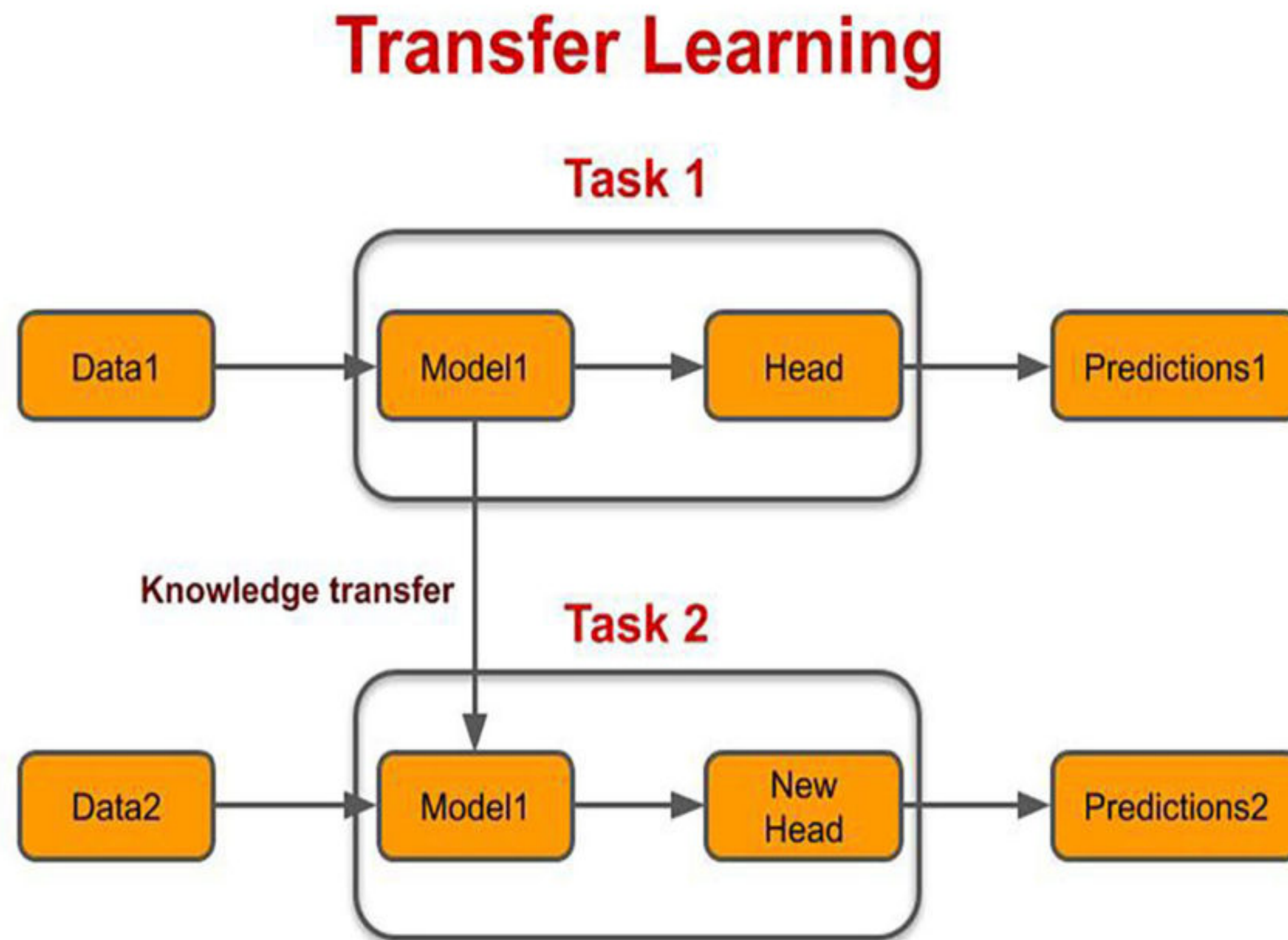


Выигрышные билеты в свёрточных нейронных сетях

Взаимодействуют ли Dropout и наша стратегия поиска выигрышных билетов?



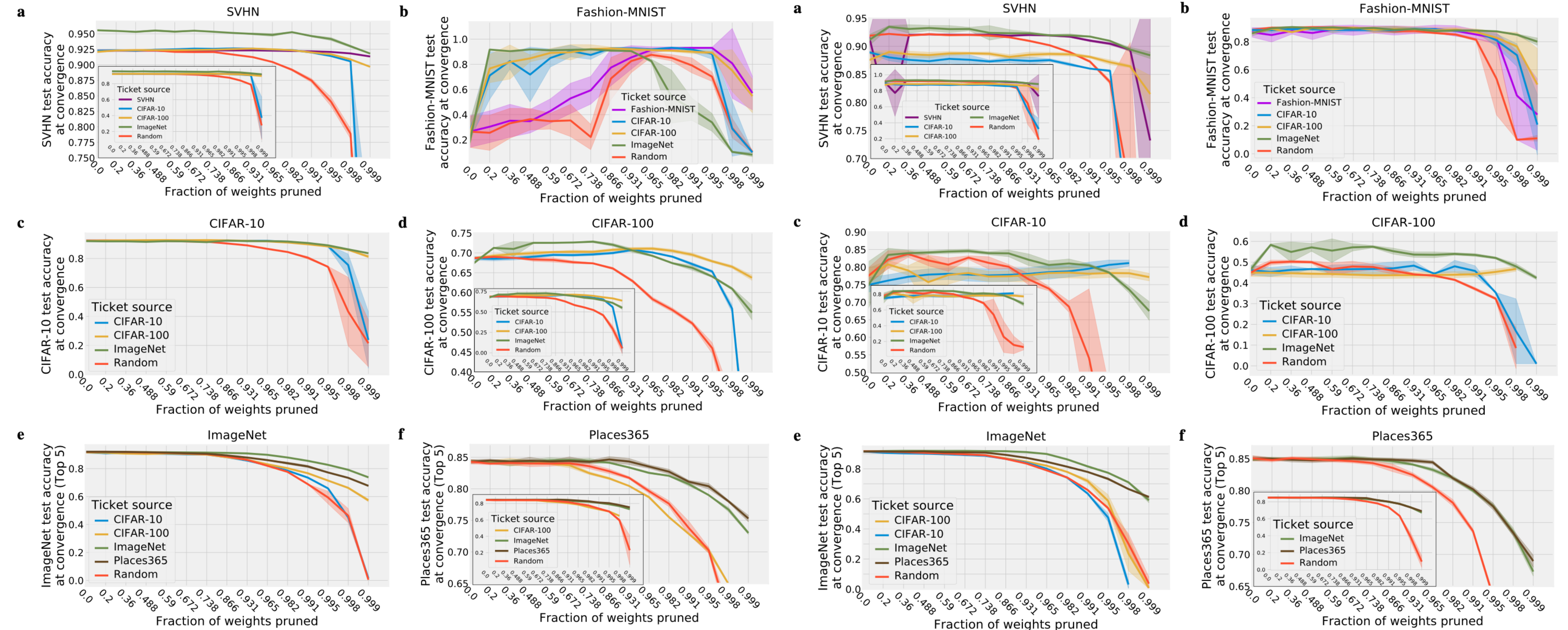
Использование выигрышных билетов в различных наборах данных



Модификации алгоритма нахождения выигрышного билета

- **Итеративная обрезка** — выполнение множества итераций попеременного обучения и обрезки, причем каждая итерация сокращает только небольшую часть весов. Такой подход помогает ослабить шум процесса сокращения и производить существенно лучшие модели сокращения и выигрышные билеты.
- **Поздний сброс** — сброс весов выигрышных билетов к их значениям на итерации намного меньшей, чем число общих итераций обучения, последовательно приводящий к получению лучших выигрышных билетов.
- **Глобальная обрезка** — объединение всех слоев вместе перед сокращением, что позволяет доли обрезанных параметров изменяться по слоям. Интересно, что глобальная обрезка предпочтительно сокращает веса в более глубоких слоях, оставляя первый слой относительно необрезанным.

Перенос выигрышного билета по наборам данных



Перенос выигрышных билетов VGG19 по наборам данных

Перенос выигрышных билетов ResNet50 по наборам данных

Список источников

- <https://arxiv.org/abs/1803.03635>
- <https://arxiv.org/abs/1906.02773>

Список вопросов

1. Опишите алгоритм идентификации выигрышных билетов.
2. Что такое Transfer learning и его применение к гипотезе лотерейного билета?
3. Какими могут быть модификации алгоритма нахождения выигрышного билета?
4. Каковы результаты экспериментов по переносу билетов по наборам данных для VGG19 и ResNet50?