

Обзор:

1. Статья была написана весной 2021 года (самая ранняя версия на OpenReview от мая 2021), camera-ready версия вышла в октябре 2021. Работа опубликована на NeurIPS'21 в виде постера.

После публикации на конференции статью много освещали:

- первый автор обсуждал статью в [подкасте Янника Килчера](#)
- крупное независимое издание Quanta Magazine опубликовало [обзор](#), где статью очень разрекламировали

2. У статьи 4 автора

- a. *Boris Knyazev*. PhD студент в University of Guelph и Vector Institute (Онтарио, Канада), а также Research Scientist в Samsung AI Lab в Монреале. Занимается преимущественно графовыми нейросетями и компьютерным зрением, в частности генерацией сцен и распознаванием лиц и эмоций.
- b. *Michal Drozdal*. Research Scientist в FAIR Montreal. Довольно много времени занимался ML в медицине, но за последние пару лет сместился к общему компьютерному зрению, в частности к генерации сцен, реконструкции 3D моделей и генеративно-состязательным сетям.
- c. *Graham W. Taylor*. Руководитель ML Research группы в University of Guelph, научный руководитель Бориса. Занимается обучением с малым количеством размеченных данных, а также graph representation learning, generative modelling и немного ML в биологии. В общем, личность широких взглядов.
- d. *Adriana Romero-Soriano*. Research Scientist в FAIR, преподает в McGill University. В целом интересуется мультимодальными данными, но в последние 2-3 года чаще публикует работы про генерацию сцен и генеративно-состязательные сети. Есть несколько публикаций про графовые нейросети последним автором. Часто работает в соавторстве в Michal Drozdal.

Работа скорее похожа на случайную находку.

3. Наибольшее влияние оказала статья [Graph HyperNetworks for Neural Architecture Search \(Zhang et al.; 2019\)](#). Используется схожая архитектура, потому что она может принимать на вход нейросеть.

Другая базовая работа — [DARTS: Differentiable Architecture Search \(Liu et al.; 2018\)](#). На основе предложенных в этой статье базовых блоков авторы собирают свой набор данных.

4. Статью процитировали 4 раза. Две статьи также пробуют обучить нейросети для оптимизации:
 - [Teaching Networks to Solve Optimization Problems](#)
 - [Tutorial on amortized optimization for learning to optimize over continuous domains](#)

Еще в двух работах статьи касаются, хотя прямой связи нет:

- [From data to functa: Your data point is a function and you should treat it like one](#)
- [Federated Learning with Heterogeneous Architectures using Graph HyperNetworks](#)

5. Прямых конкурентов нет. Вероятно, это первая статья, решающая подобную задачу.
6. Есть две идеи:
 - a. Авторы посмотрели только на то, как обученная модель ведет себя в задаче transfer learning. Возможно, что будет интересно сравниться с методами для инициализации нейросетей
 - b. Пространство DARTS довольно мало и многие архитектуры в нем не представить. Авторы добавили несколько блоков, чтобы можно было вывести ResNet, VGG и ViT, но все архитектуры все равно сильно перекошены в сторону классификационных CV моделей. Интересно, что будет с задачами из других областей.
7. Метод позволяет очень быстро обучить модель, хоть качество и не слишком хорошее. В текущем состоянии можно попробовать использовать предсказания предложенной модели как первое приближение, а затем дообучаться классическими методами.

Еще среди экспериментов в статье была задача о предсказании финальной метрики, скорости на инференсе и скорости сходимости. Модель из статьи показывала на такой задаче хороший результат. Кажется, что такое тоже может быть полезно на практике.