

# GradInit: Learning to Initialize Neural Networks for Stable and Efficient Training

Никита Конодюк, практик-исследователь

26 января 2022 г.

Статья опубликована на NIPS 2021 как постер, первая версия выпущена в феврале 2021 года.

Авторы:

- [Chen Zhu](#) – PhD student в Maryland и research intern в Google Research. Научные интересы широки, но наиболее цитируемые работы – про data poisoning и adversarial training. На последний NIPS принято три его постера.
- Renkun Ni – PhD student в Maryland.
- Tom Goldstein – профессор в Maryland, научный руководитель Chen Zhu и Renkun Ni.
- Zheng Xu, W. Ronny Huang – Google Research.

Авторы не упоминают ни одной статьи, оказавшей непосредственное влияние. Ближайший по идее метод – MetaInit, однако он достаточно сильно отличается от предложенного и не может считаться его предшественником.

В силу относительной свежести статья мало цитируется, однако уже есть интересный результат: статья [A Loss Curvature Perspective on Training Instability in Deep Learning](#) приходит к выводу, что правильный подбор гиперпараметров и эвристик (learning rate, warmup, gradient clipping) позволяет добиться тех же результатов, что и GradInit и другие продвинутые методы инициализации.

Метод заявляется как architecture agnostic и потенциально широко применим в индустрии. Тем не менее, следует учитывать, что проведенные эксперименты недостаточно полны (см. рецензию), что может ограничить практическую применимость метода.