

1. Да, вполне себе можно. Вот [пример научной статьи](#), коллега решает задачу object recognition на изображениях рекуррентной сверточной сетью.
2. Чаще всего это эффективно в случаях, когда датасет небольшой. Тогда есть смысл сделать такую проверку, которая даст довольно точное среднее для прототипируемой модели
3. С вероятностью 1-rate, т.е в моём случае 0.7
4. Обычно для Адама используют такой learn rate, он наиболее сбалансированный. Однако да, никто не запрещает покрутить его и поэкспериментировать