



Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos

Подготовила Жумлякова Светлана, БПМИ203



VPT: Основная идея

Цель: обучить foundation model для игры в Minecraft на размеченных видео игры

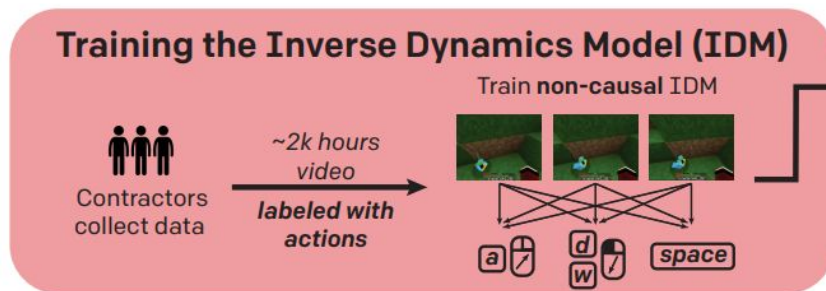
- 1) Собрать немного размеченных видео
- 2) Собрать много неразмеченных видео
- 3) Обучить модель разметки и разметить данные
- 4) На полученных данных обучить **VPT foundation model**



Почему Minecraft?

- Популярная игра -> много видео
- Много возможных стратегий игры
- Сложность игры для RL алгоритмов (без предобучения плохие результаты на многоступенчатых задачах)

Пайплайн VPT: IDM



- IDM: $t \rightarrow p_{\text{IDM}}(a_t | o_{1...T})$
(a - action, o - observation)
- Non-causal: $T > t$
- 3d conv + ResNet + Transformer
- Loss: сумма negative log-likelihood по предсказаниям

Пайплайн VPT: Сбор данных для ВС

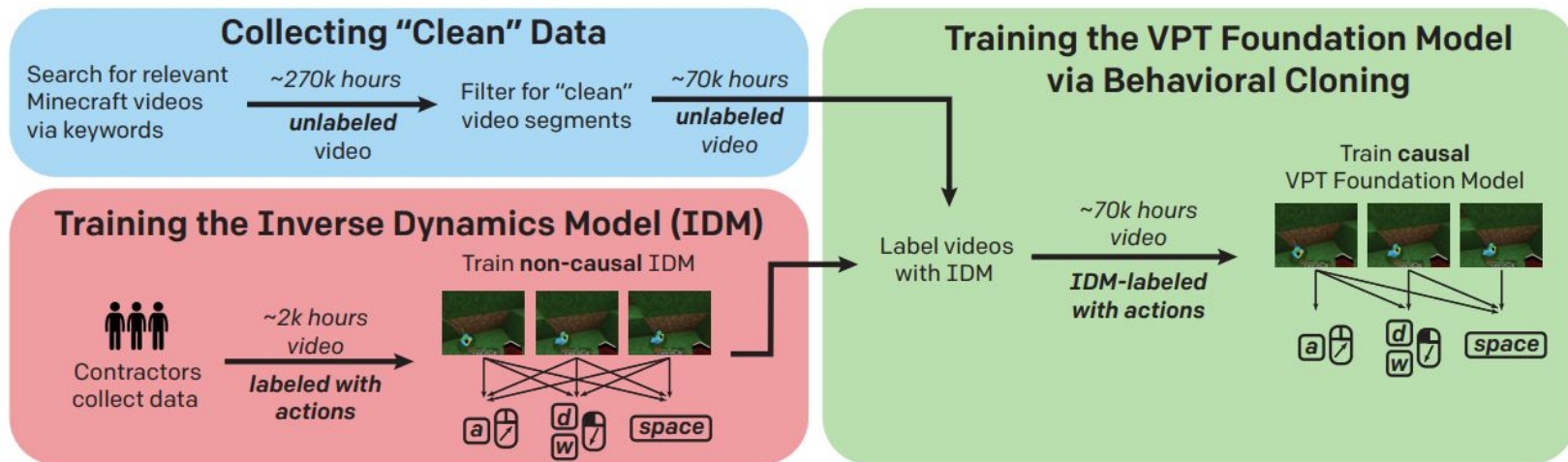
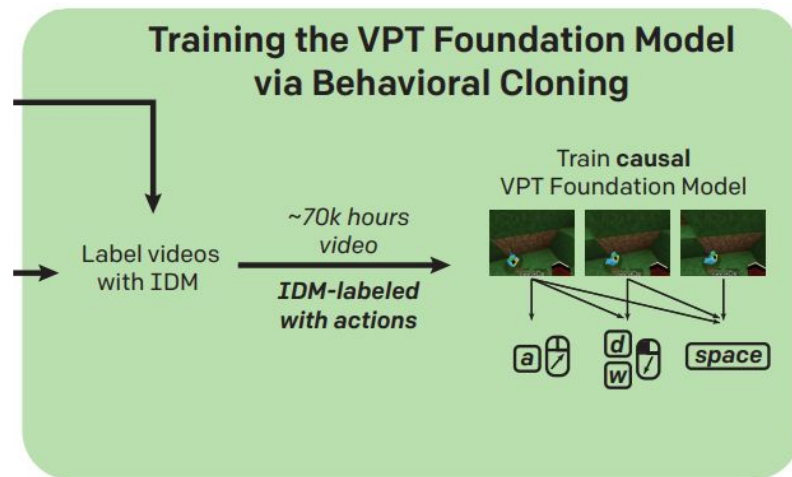


Figure 2: Video Pretraining (VPT) Method Overview.

Пайплайн VPT: Behavioral cloning (BC)

- Минимизируем negative log-likelihood действий на размеченных данных
- Causal - смотрим только в прошлое
- Архитектура IDM + маскирование



Пайплайн VPT: Вопросы?

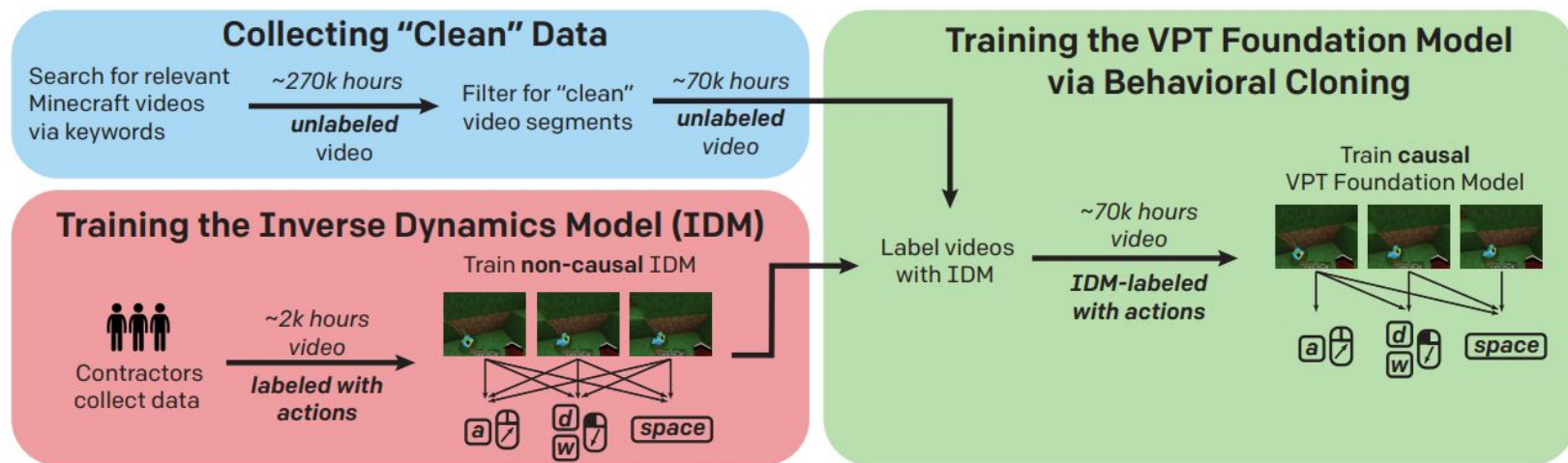
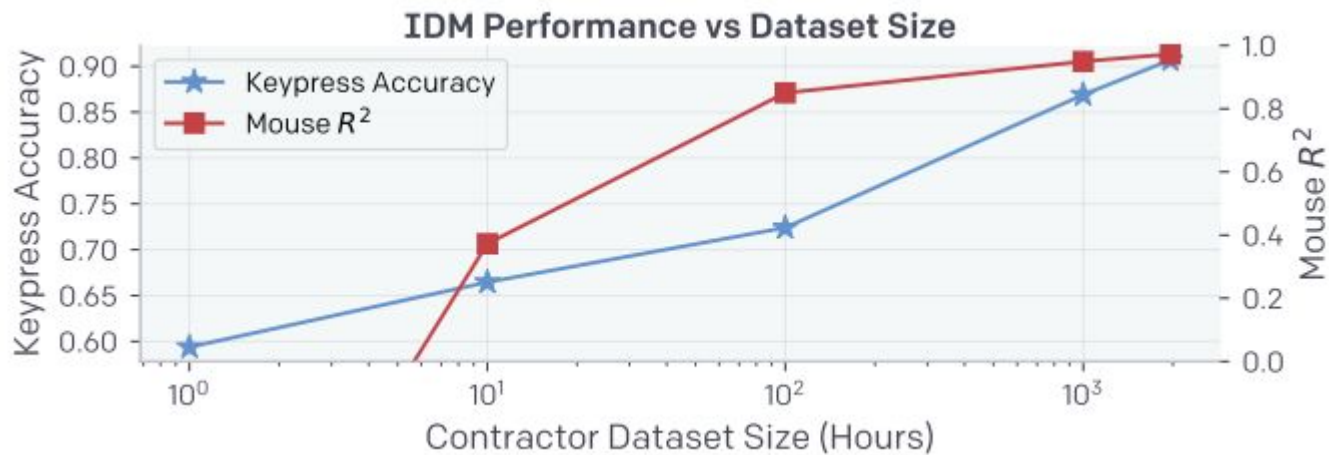
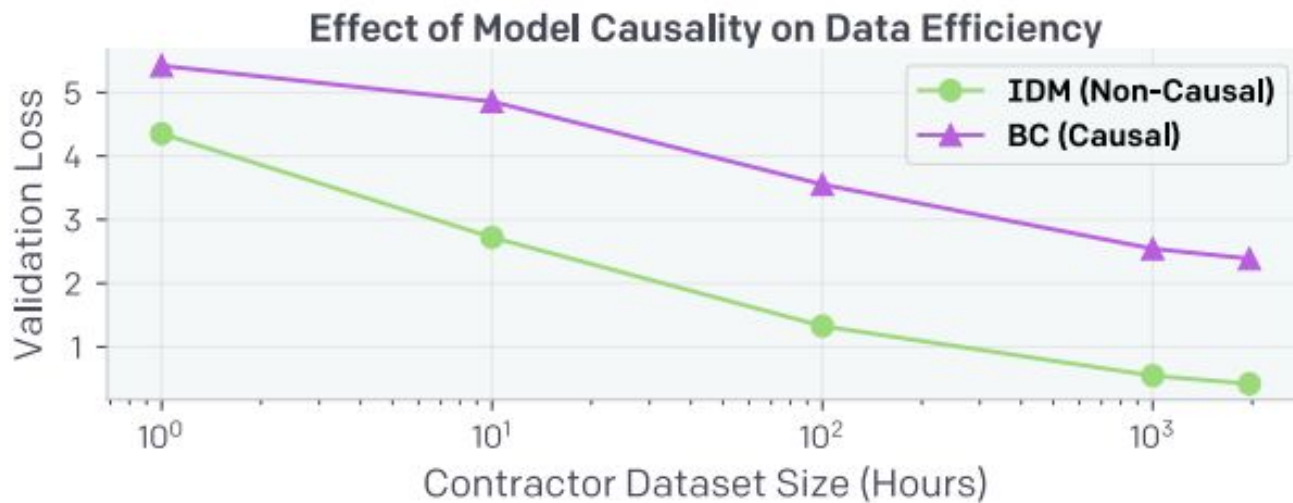


Figure 2: Video Pretraining (VPT) Method Overview.

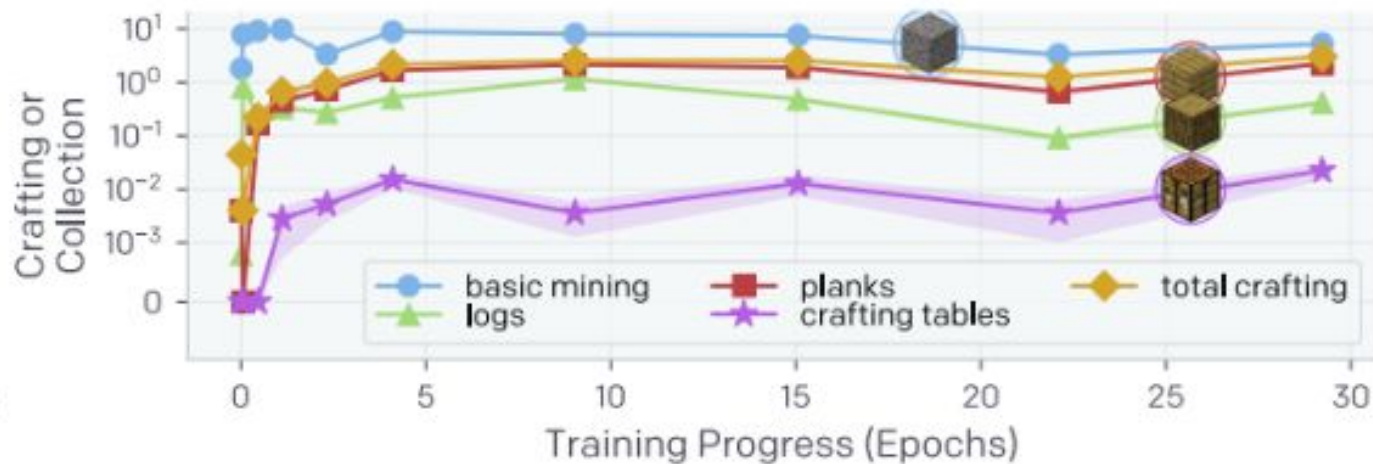
Обучение IDM



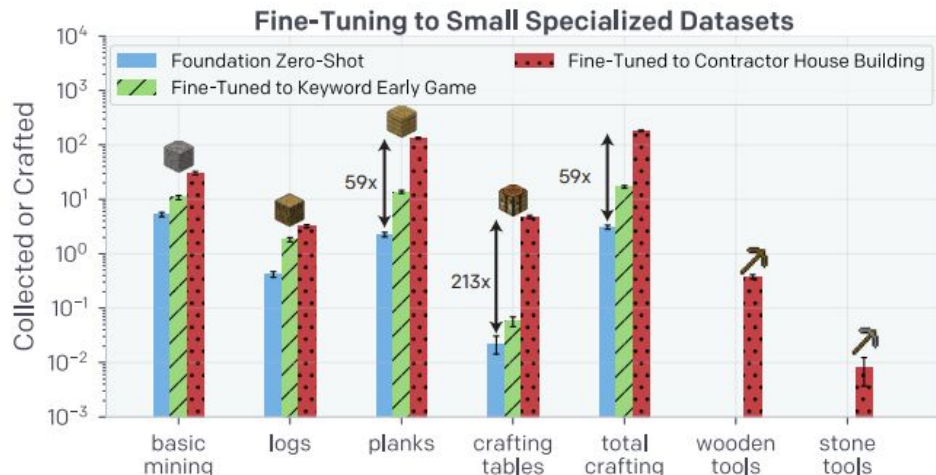
BC vs IDM на данных контрагентов



Обучение ВС: Первые результаты



Fine-Tuning



Дообучаем на данных:

- 1) Early Game датасет - видео первых минут игры
- 2) Contractor House Building - видео постройки дома за 10 мин с нуля

Итог:

- Больше созданных предметов
- Создание новых предметов

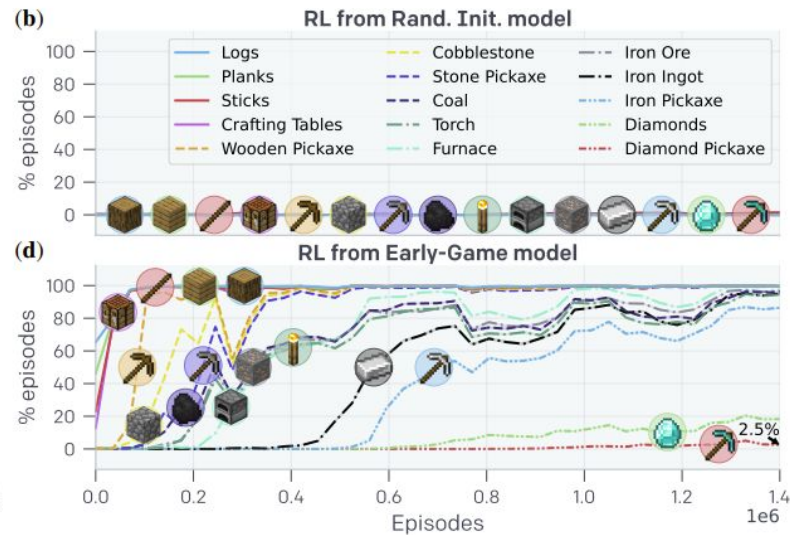
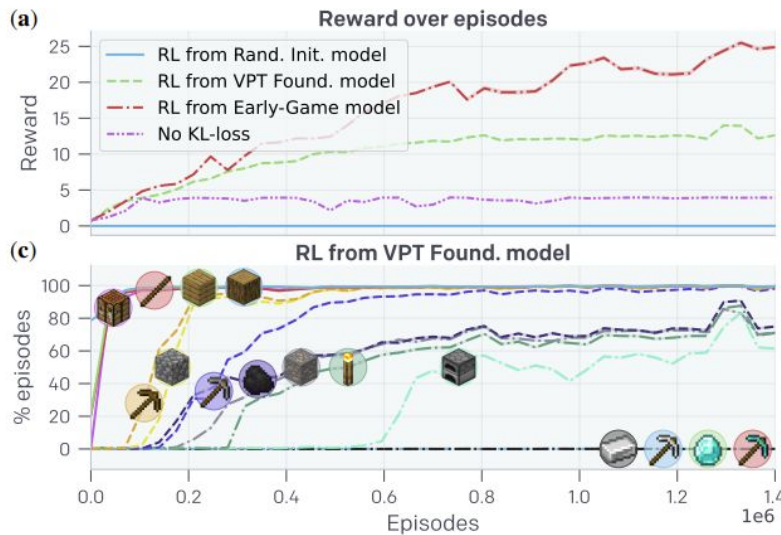
Fine-tuning with RL



Цель: создать алмазную кирку за 10 мин

- Чем правее предмет в цепочке, тем больше награда за него
- Забывчивость -> KL-дивергенция RL модели и frozen BC

Fine-tuning with RL: Результаты





Выводы

- VPT улучшает результат RL
- IDM - эффективный способ увеличения числа данных для обучения ВС



Сильные и слабые стороны статьи

- + Новизна в области RL
 - + Реализация от авторов в открытом доступе
 - + Достижение поставленных целей
-
- + - Требуется много вычислительных мощностей



Источники

- <https://arxiv.org/pdf/2206.11795.pdf>
- <https://openai.com/research/vpt>