

## Requirement 1 Report - VPT Approach

This report is an analysis of Baker et al. [1] to see what we could learn of this popular machine behavior model for our own training purposes.

### ***How can VPT be implemented to achieve improvements in baseline AI behavior?***

The Visual Pre-Training model of training an AI agent via instilling behaviors within the agent through a semi-supervised imitation learning method. The purpose of this approach with Minecraft agents in particular is to train the agent to perform some degree of baseline behaviors out of the box for further training via reinforcement learning later on down the line. This method reduces the training time for reinforcement learning in tasks by instilling some basic actions into the model's preliminary behavior, cutting out a lot of time-wasting actions in training models by reinforcement learning alone such as moving around arbitrarily, not gathering resources to completion, and struggling to explore the environment.

### ***Can we utilize this approach to achieve results beyond the initial scope of the paper discussing it?***

Using this approach would likely not be feasible for our project. Baker et al. spent a great deal of time collecting and processing data for training, as well as using very complicated training techniques on complex hardware to generate their results. In their methods section, they discuss spending 9 days training their models using their VPT methodology on a complex series of 720 V100 GPUs, a feat which we will definitely not be able to acquire resources to emulate. However, all of the training material Baker et al. collected has been made available publicly so that others can utilize it. We can at least use this data for our own more feasible training purposes. Plus, they have made their VPT-trained model available to the public as well, allowing for people to build upon this VPT model with reinforcement learning to get a great deal further in the process of training their Minecraft agents. This is the entire basis for the MineRL 2022 BASALT competition, meaning that building upon this powerful model is not only openly accessible but likely a very feasible task as well.

### ***To implement this approach, would we need to make our own training videos, and how many would we need?***

Using the metrics described by Baker et al., it would likely not be feasible to construct our own datasets via the Inverse Dynamics Models (IDM) generation approach described in the paper. In order to construct their dataset, they scraped 270k hours of video, processed it to produce 70k hours of clean video, and then ran this data through their IDM algorithm to label it for use while training an AI agent. The process of cleaning the data involved removing segments with streamer overlays, and other such tools unavailable to a base Minecraft agent. Due to the strenuous and system intensive nature of collecting and processing data in this manner, it would likely not be worth our effort to try to imitate their approach in other domains, especially given

the already large amount of training data available for training in different tasks in Minecraft that already exists.

## **References**

- [1] B. Baker et al., "Video pretraining (VPT): Learning to act by watching unlabeled online videos," arXiv.org, <https://doi.org/10.48550/arXiv.2206.11795> (accessed Sep. 23, 2023)