

teachingMarioNew

December 3, 2023

1 Setup Mario

```
[1]: import gym_super_mario_bros
      from nes_py.wrappers import JoypadSpace
      from gym_super_mario_bros.actions import SIMPLE_MOVEMENT
```

```
[2]: SIMPLE_MOVEMENT # all actions that our ai will take
```

```
[2]: [['NOOP'],
      ['right'],
      ['right', 'A'],
      ['right', 'B'],
      ['right', 'A', 'B'],
      ['A'],
      ['left']]
```

```
[3]: # Setup Game
      env = gym_super_mario_bros.make('SuperMarioBros-v0')
      env = JoypadSpace(env, SIMPLE_MOVEMENT)
```

/home/joy/.local/lib/python3.10/site-packages/gym/envs/registration.py:593:

UserWarning: WARN: The environment SuperMarioBros-v0 is out of date. You

should consider upgrading to version `v3`.

logger.warn(

/home/joy/.local/lib/python3.10/site-packages/gym/core.py:329:

DeprecationWarning: WARN: Initializing wrapper in old step API which

returns one bool instead of two. It is recommended to set `new_step_api=True` to

use new step API. This will be the default behaviour in future.

deprecation(

/home/joy/.local/lib/python3.10/site-

packages/gym/wrappers/step_api_compatibility.py:39: DeprecationWarning:

WARN: Initializing environment in old step API which returns one bool

instead of two. It is recommended to set `new_step_api=True` to use new step

API. This will be the default behaviour in future.

deprecation(

```
[4]: env.observation_space.shape # A frame from the game
```

```
[4]: (240, 256, 3)
```

```
[5]: env.action_space # All possible actions
```

```
[5]: Discrete(7)
```

```
[7]: # Creating a flag - restart or not
done = True
for step in range(100000): # This loop iterates through every frame of the game.
    if done:
        # Start the game
        env.reset()

        # perform any random action from our action space
        step_result = env.step(env.action_space.sample())
        state, reward, done, info = step_result[:4]
        env.render()
    # Close the game
    env.close()
```

```
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:174: UserWarning: WARN: Future
gym versions will require that `Env.reset` can be passed a `seed` instead of
using `Env.seed` for resetting the environment random number generator.
  logger.warn(
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:190: UserWarning: WARN: Future
gym versions will require that `Env.reset` can be passed `return_info` to return
information from the environment resetting.
  logger.warn(
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:195: UserWarning: WARN: Future
gym versions will require that `Env.reset` can be passed `options` to allow the
environment initialisation to be passed additional information.
  logger.warn(
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:227: DeprecationWarning: WARN:
Core environment is written in old step API which returns one bool instead of
two. It is recommended to rewrite the environment with new step API.
  logger.deprecation(
/home/joy/.local/lib/python3.10/site-packages/gym/core.py:57:
```

```

DeprecationWarning: WARN: You are calling render method, but you didn't
specified the argument render_mode at environment initialization. To maintain
backward compatibility, the environment will render in human mode.
If you want to render in human mode, initialize the environment in this way:
gym.make('EnvName', render_mode='human') and don't call the render method.
See here for more information: https://www.gymnasium.ml/content/api/
deprecation(
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:280: UserWarning: WARN: No render
modes was declared in the environment (env.metadata['render_modes'] is None or
not defined), you may have trouble when calling `.render()``.
logger.warn(
/home/joy/.local/lib/python3.10/site-
packages/gym_super_mario_bros/smb_env.py:148: RuntimeWarning: overflow
encountered in ubyte_scalars
return (self.ram[0x86] - self.ram[0x071c]) % 256

```

```

-----
KeyboardInterrupt                                Traceback (most recent call last)
Cell In [7], line 11
      9     step_result = env.step(env.action_space.sample())
     10     state, reward, done, info = step_result[:4]
--> 11     env.render()
     12 # Close the game
     13 env.close()

File ~/local/lib/python3.10/site-packages/gym/core.py:66, in _EnvDecorator.
    _deprecate_mode.<locals>.render(self, *args, **kwargs)
    56 elif self.spec is not None and "render_mode" not in self.spec.kwargs.
    keys(): # type: ignore
    57     deprecation(
    58         "You are calling render method, "
    59         "but you didn't specified the argument render_mode at_
    environment initialization. "
    (...)
    63         "See here for more information: https://www.gymnasium.ml/
    content/api/"
    64     )
--> 66 return render_func(self, *args, **kwargs)

File ~/local/lib/python3.10/site-packages/gym/core.py:433, in Wrapper.
    render(self, *args, **kwargs)
    429 def render(
    430     self, *args, **kwargs

```

```

431 ) -> Optional[Union[RenderFrame, List[RenderFrame]]]:
432     """Renders the environment."""
--> 433     return self.env.render(*args, **kwargs)

File ~/.local/lib/python3.10/site-packages/gym/core.py:66, in _EnvDecorator.
↳ _deprecate_mode.<locals>.render(self, *args, **kwargs)
    56 elif self.spec is not None and "render_mode" not in self.spec.kwargs.
↳ keys(): # type: ignore
    57     deprecation(
    58         "You are calling render method, "
    59         "but you didn't specified the argument render_mode at
↳ environment initialization. "
    (...
    63         "See here for more information: https://www.gymlibrary.ml/
↳ content/api/"
    64     )
---> 66 return render_func(self, *args, **kwargs)

```

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:433, in Wrapper.
↳ render(self, *args, **kwargs)
    429 def render(
    430     self, *args, **kwargs
    431 ) -> Optional[Union[RenderFrame, List[RenderFrame]]]:
    432     """Renders the environment."""
--> 433     return self.env.render(*args, **kwargs)

```

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:66, in _EnvDecorator.
↳ _deprecate_mode.<locals>.render(self, *args, **kwargs)
    56 elif self.spec is not None and "render_mode" not in self.spec.kwargs.
↳ keys(): # type: ignore
    57     deprecation(
    58         "You are calling render method, "
    59         "but you didn't specified the argument render_mode at
↳ environment initialization. "
    (...
    63         "See here for more information: https://www.gymlibrary.ml/
↳ content/api/"
    64     )
---> 66 return render_func(self, *args, **kwargs)

```

```

File ~/.local/lib/python3.10/site-packages/gym/wrappers/order_enforcing.py:51,
↳ in OrderEnforcing.render(self, *args, **kwargs)
    46 if not self._disable_render_order_enforcing and not self._has_reset:
    47     raise ResetNeeded(
    48         "Cannot call `env.render()` before calling `env.reset()`, if
↳ this is a intended action, "
    49         "set `disable_render_order_enforcing=True` on the OrderEnforcer,
↳ wrapper."

```

```

50     )
---> 51 return self.env.render(*args, **kwargs)

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:66, in _EnvDecorator.

```

↳ _deprecate_mode.<locals>.render(self, *args, **kwargs)
    56 elif self.spec is not None and "render_mode" not in self.spec.kwargs.
↳ keys(): # type: ignore
    57     deprecation(
    58         "You are calling render method, "
    59         "but you didn't specified the argument render_mode at
↳ environment initialization. "
    (...)
    63         "See here for more information: https://www.gymlibrary.ml/
↳ content/api/"
    64     )
---> 66 return render_func(self, *args, **kwargs)

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:433, in Wrapper.

```

↳ render(self, *args, **kwargs)
    429 def render(
    430     self, *args, **kwargs
    431 ) -> Optional[Union[RenderFrame, List[RenderFrame]]]:
    432     """Renders the environment."""
--> 433     return self.env.render(*args, **kwargs)

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:66, in _EnvDecorator.

```

↳ _deprecate_mode.<locals>.render(self, *args, **kwargs)
    56 elif self.spec is not None and "render_mode" not in self.spec.kwargs.
↳ keys(): # type: ignore
    57     deprecation(
    58         "You are calling render method, "
    59         "but you didn't specified the argument render_mode at
↳ environment initialization. "
    (...)
    63         "See here for more information: https://www.gymlibrary.ml/
↳ content/api/"
    64     )
---> 66 return render_func(self, *args, **kwargs)

```

File ~/.local/lib/python3.10/site-packages/gym/wrappers/env_checker.py:55, in

```

↳ PassiveEnvChecker.render(self, *args, **kwargs)
    53     return env_render_passive_checker(self.env, *args, **kwargs)
    54 else:
---> 55     return self.env.render(*args, **kwargs)

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:66, in _EnvDecorator.

```

↳ _deprecate_mode.<locals>.render(self, *args, **kwargs)

```

```

    56 elif self.spec is not None and "render_mode" not in self.spec.kwargs.
↳keys(): # type: ignore
    57     deprecation(
    58         "You are calling render method, "
    59         "but you didn't specified the argument render_mode at
↳environment initialization. "
    60     )
    61     "See here for more information: https://www.gymlibrary.ml/
↳content/api/"
    62 )
---> 66 return render_func(self, *args, **kwargs)

```

File ~/.local/lib/python3.10/site-packages/nes_py/nes_env.py:379, in NESEnv.

```

↳render(self, mode)
    373     self.viewer = ImageViewer(
    374         caption=caption,
    375         height=SCREEN_HEIGHT,
    376         width=SCREEN_WIDTH,
    377     )
    378     # show the screen on the image viewer
--> 379     self.viewer.show(self.screen)
    380 elif mode == 'rgb_array':
    381     return self.screen

```

File ~/.local/lib/python3.10/site-packages/nes_py/_image_viewer.py:144, in

```

↳ImageViewer.show(self, frame)
    138 self._window.dispatch_events()
    139 # create an image data object
    140 image = self.pyglet.image.ImageData(
    141     frame.shape[1],
    142     frame.shape[0],
    143     'RGB',
--> 144     frame.tobytes(),
    145     pitch=frame.shape[1]*-3
    146 )
    147 # send the image to the window
    148 image.blit(0, 0, width=self._window.width, height=self._window.height)

```

KeyboardInterrupt:

```
[8]: env.close()
```

```
[6]: state = env.reset() # the frames of the game
```

/home/joy/.local/lib/python3.10/site-

```

packages/gym/utils/passive_env_checker.py:174: UserWarning: WARN: Future
gym versions will require that `Env.reset` can be passed a `seed` instead of
using `Env.seed` for resetting the environment random number generator.
    logger.warn(
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:190: UserWarning: WARN: Future
gym versions will require that `Env.reset` can be passed `return_info` to return
information from the environment resetting.
    logger.warn(
/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:195: UserWarning: WARN: Future
gym versions will require that `Env.reset` can be passed `options` to allow the
environment initialisation to be passed additional information.
    logger.warn(

```

```

[7]: state.shape # the frame from the game that is 240px wide, 256 px high, 3
    ↪ channels.

```

```

[7]: (240, 256, 3)

```

```

[8]: env.step(1)[3]

```

```

/home/joy/.local/lib/python3.10/site-
packages/gym/utils/passive_env_checker.py:227: DeprecationWarning: WARN:
Core environment is written in old step API which returns one bool instead of
two. It is recommended to rewrite the environment with new step API.
    logger.deprecation(

```

```

[8]: {'coins': 0,
      'flag_get': False,
      'life': 2,
      'score': 0,
      'stage': 1,
      'status': 'small',
      'time': 400,
      'world': 1,
      'x_pos': 40,
      'x_pos_screen': 40,
      'y_pos': 79}

```

```

[9]: # Import FrameStacker wrapper and Grayscale wrapper
    from gym.wrappers import GrayScaleObservation, FrameStack
    # Importing Vectorization Wrappers

```

```

from stable_baselines3.common.vec_env import VecFrameStack, DummyVecEnv
from stable_baselines3.common.vec_env.base_vec_env import VecEnvWrapper
from stable_baselines3.common.vec_env import VecNormalize
from matplotlib import pyplot as plt

```

2023-10-21 12:13:19.813879: I tensorflow/tsl/cuda/cudart_stub.cc:28] Could not find cuda drivers on your machine, GPU will not be used.

2023-10-21 12:13:19.975323: I tensorflow/core/platform/cpu_feature_guard.cc:182] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.

To enable the following instructions: AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.

2023-10-21 12:13:20.921044: W tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT

```

[10]: # 1. Create the base environment
env = gym_super_mario_bros.make("SuperMarioBros-v0")
# 2. Simplify the controls
env = JoypadSpace(env, SIMPLE_MOVEMENT)
# 3. Grayscale
env = GrayScaleObservation(env, keep_dim = True) # Grayscale the frames.
# 4. Wrap inside the dummy environment

env = DummyVecEnv([lambda: env]) #(Different from og code, getting error)

# 5. Stack the frames
env = VecFrameStack(env, 4, channels_order='last') # (Different from og code,
↳getting error)

print(env.observation_space)

```

/home/joy/.local/lib/python3.10/site-packages/gym/envs/registration.py:593:

UserWarning: WARN: The environment SuperMarioBros-v0 is out of date. You

should consider upgrading to version `v3`.

logger.warn(

Box(0, 255, (240, 256, 4), uint8)

/home/joy/.local/lib/python3.10/site-packages/gym/core.py:329:

DeprecationWarning: WARN: Initializing wrapper in old step API which

returns one bool instead of two. It is recommended to set `new_step_api=True` to use new step API. This will be the default behaviour in future.

deprecation(

/home/joy/.local/lib/python3.10/site-

packages/gym/wrappers/step_api_compatibility.py:39: DeprecationWarning:

WARN: Initializing environment in old step API which returns one bool instead of two. It is recommended to set `new_step_api=True` to use new step API. This will be the default behaviour in future.

```
deprecation(  
/home/joy/.local/lib/python3.10/site-  
packages/stable_baselines3/common/vec_env/patch_gym.py:49: UserWarning: You  
provided an OpenAI Gym environment. We strongly recommend transitioning to  
Gymnasium environments. Stable-Baselines3 is automatically wrapping your  
environments in a compatibility layer, which could potentially cause issues.  
warnings.warn(  

```

```
[11]: state = env.reset()  
state.shape # Now, there will only be one channel instead of 3 because we have  
↳grayscaled the frames.
```

```
/home/joy/.local/lib/python3.10/site-  
packages/gym/utils/passive_env_checker.py:174: UserWarning: WARN: Future  
gym versions will require that `Env.reset` can be passed a `seed` instead of  
using `Env.seed` for resetting the environment random number generator.  
logger.warn(  
/home/joy/.local/lib/python3.10/site-  
packages/gym/utils/passive_env_checker.py:190: UserWarning: WARN: Future  
gym versions will require that `Env.reset` can be passed `return_info` to return  
information from the environment resetting.  
logger.warn(  
/home/joy/.local/lib/python3.10/site-  
packages/gym/utils/passive_env_checker.py:195: UserWarning: WARN: Future  
gym versions will require that `Env.reset` can be passed `options` to allow the  
environment initialisation to be passed additional information.  
logger.warn(  

```

```
[11]: (1, 240, 256, 4)
```

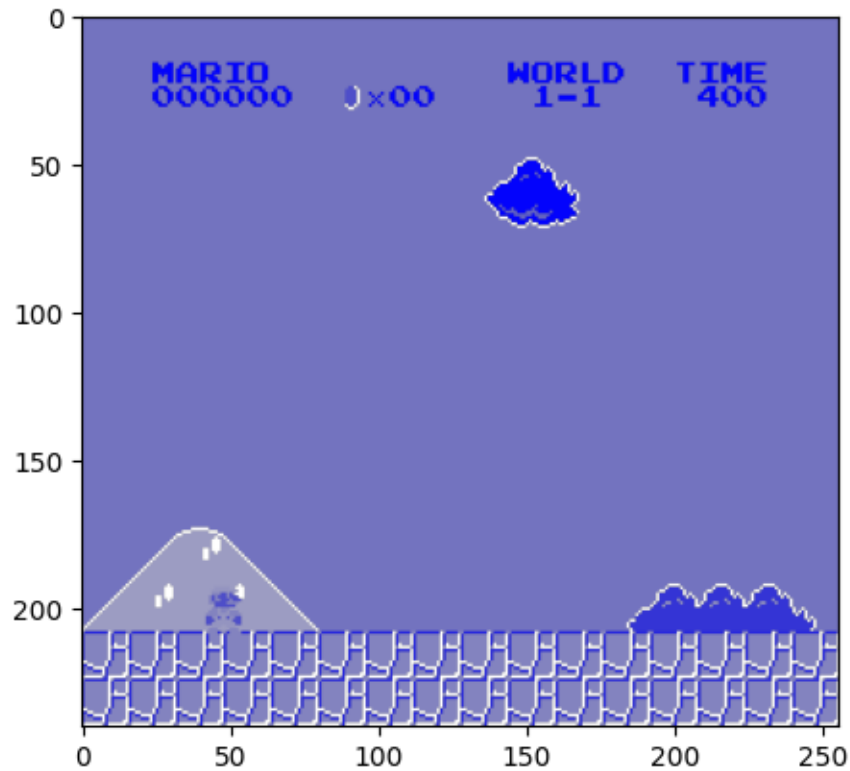
```
[12]: state, reward, done, info = env.step([5])
```

```
/home/joy/.local/lib/python3.10/site-  
packages/gym/utils/passive_env_checker.py:227: DeprecationWarning: WARN:  
Core environment is written in old step API which returns one bool instead of  
two. It is recommended to rewrite the environment with new step API.  
logger.deprecation(  

```

```
[13]: plt.imshow(state[0]) # Below is the grayscaled version of our Mario env's frame.
```

[13]: <matplotlib.image.AxesImage at 0x7f6d643f4220>



1.1 Reinforcement Learning :

The type of AI that we are using is called Reinforcement Learning. It follows four key elements: (A.R.E.A) 1) Agent : Here, mario acts as our agent.

- 2) Reward : Here, the reward function aims at moving left as far as possible as quickly as possible.
- 3) Environment : Entire game env that we have setup.
- 4) Action : Moving right, jumping, moving left are all the actions as defined in the action space, we have 7 actions.

1.2 Algorithm

The specific RL algorithm that we are using is called PPO: Proximal Policy Optimization.

2 Training the RL Model

```
[14]: # Importing os for file management
import os
# Importing the PPO algorithm
from stable_baselines3 import PPO
# Importing base_callbacks for saving the models
from stable_baselines3.common.callbacks import BaseCallback

[15]: CHECKPOINT_DIR = './train/'
LOG_DIR = './logs/'

[16]: # For saving our model locally every few steps
class TrainAndLoggingCallback(BaseCallback):

    def __init__(self, check_freq, save_path, verbose=1):
        super(TrainAndLoggingCallback, self).__init__(verbose)
        self.check_freq = check_freq
        self.save_path = save_path

    def _init_callback(self):
        if self.save_path is not None:
            os.makedirs(self.save_path, exist_ok=True)

    def _on_step(self):
        if self.n_calls % self.check_freq == 0:
            model_path = os.path.join(self.save_path, 'best_model_{}'.
                ↪format(self.n_calls))
            self.model.save(model_path)

        return True

[17]: # Setup model saving callback
callback = TrainAndLoggingCallback(check_freq=10000, save_path=CHECKPOINT_DIR)

[18]: # This is the AI model started
model = PPO('CnnPolicy', env, verbose=1, tensorboard_log=LOG_DIR, ↪
    ↪learning_rate=0.000001, n_steps=512)

# 1. 'CnnPolicy' is the neural network that we are using, consider it as the ↪
    ↪brain of our AI.
# 2. 'env' is loading our game environment into the AI model.
# 3. 'verbose = 1' provides us the below information about the training process.
# 4. 'tensorboard_log' stores the history of all the training in the path ↪
    ↪specified (log folder).
# 5. 'learning_rate' defines how fast our model will learn.
```

```
# 6. 'n_steps' defines how many frames we are going to wait per game before we
    ↳ update our neural network with new info. regarding new frames.
```

Using cpu device

Wrapping the env in a VecTransposeImage.

```
[33]: model.learn(total_timesteps=100000, callback=callback)
      # Imagining that every single move is a frame, total-timesteps is the number of
      ↳ moves that our AI will make.
```

Logging to ./logs/PP0_2

```
-----
| time/          |      |
|   fps          | 100  |
|  iterations    | 1    |
| time_elapsed   | 5    |
| total_timesteps | 512  |
-----
```

```
-----
| time/          |      |
|   fps          | 21   |
|  iterations    | 2    |
| time_elapsed   | 48   |
| total_timesteps | 1024 |
| train/         |      |
|  approx_kl     | 2.1567801e-05 |
|  clip_fraction | 0     |
|  clip_range    | 0.2   |
|  entropy_loss  | -1.95 |
| explained_variance | 0.000306 |
| learning_rate  | 1e-06 |
| loss           | 117   |
| n_updates      | 10    |
| policy_gradient_loss | -6.45e-05 |
| value_loss     | 324   |
-----
```

```
-----
| time/          |      |
|   fps          | 16   |
|  iterations    | 3    |
| time_elapsed   | 90   |
| total_timesteps | 1536 |
| train/         |      |
|  approx_kl     | 4.4594286e-05 |
|  clip_fraction | 0     |
|  clip_range    | 0.2   |
|  entropy_loss  | -1.95 |
| explained_variance | -0.0151 |
-----
```

	learning_rate		1e-06	
	loss		0.235	
	n_updates		20	
	policy_gradient_loss		-0.000449	
	value_loss		1.63	

	time/			
	fps		15	
	iterations		4	
	time_elapsed		133	
	total_timesteps		2048	
	train/			
	approx_kl		2.4720794e-06	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.95	
	explained_variance		-0.00316	
	learning_rate		1e-06	
	loss		0.113	
	n_updates		30	
	policy_gradient_loss		1.55e-06	
	value_loss		0.759	

	time/			
	fps		14	
	iterations		5	
	time_elapsed		175	
	total_timesteps		2560	
	train/			
	approx_kl		1.2435019e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.95	
	explained_variance		0.0261	
	learning_rate		1e-06	
	loss		0.202	
	n_updates		40	
	policy_gradient_loss		-0.000242	
	value_loss		0.697	

	time/			
	fps		14	
	iterations		6	
	time_elapsed		217	
	total_timesteps		3072	

train/	
approx_kl	8.227304e-06
clip_fraction	0
clip_range	0.2
entropy_loss	-1.95
explained_variance	0.00499
learning_rate	1e-06
loss	0.226
n_updates	50
policy_gradient_loss	-0.000157
value_loss	0.568

time/	
fps	13
iterations	7
time_elapsed	259
total_timesteps	3584
train/	
approx_kl	6.2121544e-06
clip_fraction	0
clip_range	0.2
entropy_loss	-1.95
explained_variance	0.0199
learning_rate	1e-06
loss	0.114
n_updates	60
policy_gradient_loss	-9.12e-05
value_loss	0.308

time/	
fps	13
iterations	8
time_elapsed	301
total_timesteps	4096
train/	
approx_kl	1.9338448e-05
clip_fraction	0
clip_range	0.2
entropy_loss	-1.95
explained_variance	-0.0127
learning_rate	1e-06
loss	0.109
n_updates	70
policy_gradient_loss	-0.000241
value_loss	0.283

time/		
fps	13	
iterations	9	
time_elapsed	342	
total_timesteps	4608	
train/		
approx_kl	6.529619e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	0.0117	
learning_rate	1e-06	
loss	0.119	
n_updates	80	
policy_gradient_loss	-0.000121	
value_loss	0.291	

time/		
fps	13	
iterations	10	
time_elapsed	384	
total_timesteps	5120	
train/		
approx_kl	2.38996e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	-0.000897	
learning_rate	1e-06	
loss	0.164	
n_updates	90	
policy_gradient_loss	-0.000476	
value_loss	0.311	

time/		
fps	13	
iterations	11	
time_elapsed	426	
total_timesteps	5632	
train/		
approx_kl	6.2802574e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	-0.00626	

	learning_rate		1e-06	
	loss		0.15	
	n_updates		100	
	policy_gradient_loss		-9.08e-05	
	value_loss		0.296	

	time/			
	fps		13	
	iterations		12	
	time_elapsed		468	
	total_timesteps		6144	
	train/			
	approx_kl		1.7475104e-06	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.95	
	explained_variance		0.00767	
	learning_rate		1e-06	
	loss		0.0702	
	n_updates		110	
	policy_gradient_loss		-2.33e-05	
	value_loss		0.183	

	time/			
	fps		13	
	iterations		13	
	time_elapsed		510	
	total_timesteps		6656	
	train/			
	approx_kl		1.7445302e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.95	
	explained_variance		0.00677	
	learning_rate		1e-06	
	loss		0.0883	
	n_updates		120	
	policy_gradient_loss		-0.000302	
	value_loss		0.198	

	time/			
	fps		12	
	iterations		14	
	time_elapsed		552	
	total_timesteps		7168	

train/		
approx_kl	8.637784e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	0.00922	
learning_rate	1e-06	
loss	0.0572	
n_updates	130	
policy_gradient_loss	-0.000108	
value_loss	0.142	

time/		
fps	12	
iterations	15	
time_elapsed	594	
total_timesteps	7680	
train/		
approx_kl	7.1949326e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	-0.00887	
learning_rate	1e-06	
loss	0.0924	
n_updates	140	
policy_gradient_loss	-0.000114	
value_loss	0.175	

time/		
fps	12	
iterations	16	
time_elapsed	636	
total_timesteps	8192	
train/		
approx_kl	1.2647128e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	0.00416	
learning_rate	1e-06	
loss	0.0979	
n_updates	150	
policy_gradient_loss	-0.000187	
value_loss	0.17	

time/		
fps	12	
iterations	17	
time_elapsed	678	
total_timesteps	8704	
train/		
approx_kl	2.3511122e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	0.026	
learning_rate	1e-06	
loss	149	
n_updates	160	
policy_gradient_loss	-0.000228	
value_loss	294	

time/		
fps	12	
iterations	18	
time_elapsed	720	
total_timesteps	9216	
train/		
approx_kl	4.059635e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.95	
explained_variance	0.0113	
learning_rate	1e-06	
loss	44.3	
n_updates	170	
policy_gradient_loss	-3.89e-05	
value_loss	255	

time/		
fps	12	
iterations	19	
time_elapsed	762	
total_timesteps	9728	
train/		
approx_kl	3.8114144e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.209	

	learning_rate	1e-06	
	loss	0.187	
	n_updates	180	
	policy_gradient_loss	-0.00023	
	value_loss	1.35	

	time/		
	fps	12	
	iterations	20	
	time_elapsed	805	
	total_timesteps	10240	
	train/		
	approx_kl	7.752562e-06	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.94	
	explained_variance	0.0338	
	learning_rate	1e-06	
	loss	0.129	
	n_updates	190	
	policy_gradient_loss	-6.25e-05	
	value_loss	1.05	

	time/		
	fps	12	
	iterations	21	
	time_elapsed	847	
	total_timesteps	10752	
	train/		
	approx_kl	2.3921253e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.94	
	explained_variance	0.00107	
	learning_rate	1e-06	
	loss	0.15	
	n_updates	200	
	policy_gradient_loss	-0.000273	
	value_loss	0.774	

	time/		
	fps	12	
	iterations	22	
	time_elapsed	889	
	total_timesteps	11264	

train/		
approx_kl	2.8160284e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.0178	
learning_rate	1e-06	
loss	0.196	
n_updates	210	
policy_gradient_loss	-0.000395	
value_loss	0.694	

time/		
fps	12	
iterations	23	
time_elapsed	931	
total_timesteps	11776	
train/		
approx_kl	2.5177607e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	-0.0108	
learning_rate	1e-06	
loss	0.144	
n_updates	220	
policy_gradient_loss	-0.000354	
value_loss	0.368	

time/		
fps	12	
iterations	24	
time_elapsed	973	
total_timesteps	12288	
train/		
approx_kl	1.3188925e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.00112	
learning_rate	1e-06	
loss	0.122	
n_updates	230	
policy_gradient_loss	-0.00022	
value_loss	0.324	

time/		
fps	12	
iterations	25	
time_elapsed	1015	
total_timesteps	12800	
train/		
approx_kl	5.4466072e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.0046	
learning_rate	1e-06	
loss	0.169	
n_updates	240	
policy_gradient_loss	-4.21e-07	
value_loss	0.313	

time/		
fps	12	
iterations	26	
time_elapsed	1057	
total_timesteps	13312	
train/		
approx_kl	1.3685902e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.0115	
learning_rate	1e-06	
loss	0.197	
n_updates	250	
policy_gradient_loss	-0.000164	
value_loss	0.381	

time/		
fps	12	
iterations	27	
time_elapsed	1099	
total_timesteps	13824	
train/		
approx_kl	3.358815e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.00298	

	learning_rate		1e-06	
	loss		0.151	
	n_updates		260	
	policy_gradient_loss		-0.000433	
	value_loss		0.249	

	time/			
	fps		12	
	iterations		28	
	time_elapsed		1141	
	total_timesteps		14336	
	train/			
	approx_kl		3.1557167e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.94	
	explained_variance		-0.000942	
	learning_rate		1e-06	
	loss		0.0738	
	n_updates		270	
	policy_gradient_loss		-0.000382	
	value_loss		0.171	

	time/			
	fps		12	
	iterations		29	
	time_elapsed		1183	
	total_timesteps		14848	
	train/			
	approx_kl		6.210094e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.94	
	explained_variance		0.00577	
	learning_rate		1e-06	
	loss		0.0811	
	n_updates		280	
	policy_gradient_loss		-0.000602	
	value_loss		0.179	

	time/			
	fps		12	
	iterations		30	
	time_elapsed		1225	
	total_timesteps		15360	

train/	
approx_kl	2.9693823e-05
clip_fraction	0
clip_range	0.2
entropy_loss	-1.94
explained_variance	0.00632
learning_rate	1e-06
loss	0.0614
n_updates	290
policy_gradient_loss	-0.000399
value_loss	0.181

time/	
fps	12
iterations	31
time_elapsed	1267
total_timesteps	15872
train/	
approx_kl	2.2504362e-05
clip_fraction	0
clip_range	0.2
entropy_loss	-1.94
explained_variance	0.00391
learning_rate	1e-06
loss	0.166
n_updates	300
policy_gradient_loss	-0.00026
value_loss	0.248

time/	
fps	12
iterations	32
time_elapsed	1309
total_timesteps	16384
train/	
approx_kl	7.920957e-05
clip_fraction	0
clip_range	0.2
entropy_loss	-1.94
explained_variance	-0.000856
learning_rate	1e-06
loss	0.159
n_updates	310
policy_gradient_loss	-0.000849
value_loss	0.235

time/		
fps	12	
iterations	33	
time_elapsed	1352	
total_timesteps	16896	
train/		
approx_kl	0.00010355236	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.102	
learning_rate	1e-06	
loss	109	
n_updates	320	
policy_gradient_loss	0.000438	
value_loss	303	

time/		
fps	12	
iterations	34	
time_elapsed	1393	
total_timesteps	17408	
train/		
approx_kl	3.835419e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.102	
learning_rate	1e-06	
loss	8.42	
n_updates	330	
policy_gradient_loss	-8.63e-05	
value_loss	22.3	

time/		
fps	12	
iterations	35	
time_elapsed	1433	
total_timesteps	17920	
train/		
approx_kl	2.6355847e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.0797	

	learning_rate		1e-06	
	loss		0.0664	
	n_updates		340	
	policy_gradient_loss		-5.31e-05	
	value_loss		0.572	

	time/			
	fps		12	
	iterations		36	
	time_elapsed		1474	
	total_timesteps		18432	
	train/			
	approx_kl		2.5877496e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.94	
	explained_variance		-0.254	
	learning_rate		1e-06	
	loss		0.123	
	n_updates		350	
	policy_gradient_loss		-0.000222	
	value_loss		0.432	

	time/			
	fps		12	
	iterations		37	
	time_elapsed		1515	
	total_timesteps		18944	
	train/			
	approx_kl		1.991808e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.94	
	explained_variance		0.0379	
	learning_rate		1e-06	
	loss		0.136	
	n_updates		360	
	policy_gradient_loss		-0.000188	
	value_loss		0.54	

	time/			
	fps		12	
	iterations		38	
	time_elapsed		1557	
	total_timesteps		19456	

train/		
approx_kl	1.781329e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.00989	
learning_rate	1e-06	
loss	0.0694	
n_updates	370	
policy_gradient_loss	-0.000199	
value_loss	0.289	

time/		
fps	12	
iterations	39	
time_elapsed	1599	
total_timesteps	19968	
train/		
approx_kl	1.7243321e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.135	
learning_rate	1e-06	
loss	0.107	
n_updates	380	
policy_gradient_loss	-0.000186	
value_loss	0.368	

time/		
fps	12	
iterations	40	
time_elapsed	1641	
total_timesteps	20480	
train/		
approx_kl	2.8004986e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	-0.0108	
learning_rate	1e-06	
loss	0.0614	
n_updates	390	
policy_gradient_loss	-0.000183	
value_loss	0.158	

time/		
fps	12	
iterations	41	
time_elapsed	1684	
total_timesteps	20992	
train/		
approx_kl	6.583112e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	-0.00657	
learning_rate	1e-06	
loss	0.109	
n_updates	400	
policy_gradient_loss	-0.000746	
value_loss	0.235	

time/		
fps	12	
iterations	42	
time_elapsed	1726	
total_timesteps	21504	
train/		
approx_kl	1.8963008e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	-0.00452	
learning_rate	1e-06	
loss	0.057	
n_updates	410	
policy_gradient_loss	-0.000209	
value_loss	0.16	

time/		
fps	12	
iterations	43	
time_elapsed	1768	
total_timesteps	22016	
train/		
approx_kl	3.2766606e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.94	
explained_variance	0.00805	

	learning_rate	1e-06	
	loss	0.0626	
	n_updates	420	
	policy_gradient_loss	-0.000261	
	value_loss	0.185	

	time/		
	fps	12	
	iterations	44	
	time_elapsed	1810	
	total_timesteps	22528	
	train/		
	approx_kl	6.722158e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.94	
	explained_variance	-0.0288	
	learning_rate	1e-06	
	loss	0.051	
	n_updates	430	
	policy_gradient_loss	-0.000712	
	value_loss	0.117	

	time/		
	fps	12	
	iterations	45	
	time_elapsed	1852	
	total_timesteps	23040	
	train/		
	approx_kl	5.105522e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.93	
	explained_variance	-0.00666	
	learning_rate	1e-06	
	loss	0.0537	
	n_updates	440	
	policy_gradient_loss	-0.000442	
	value_loss	0.117	

	time/		
	fps	12	
	iterations	46	
	time_elapsed	1894	
	total_timesteps	23552	

train/		
approx_kl	6.681564e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.93	
explained_variance	-0.00293	
learning_rate	1e-06	
loss	0.0606	
n_updates	450	
policy_gradient_loss	-0.000737	
value_loss	0.13	

time/		
fps	12	
iterations	47	
time_elapsed	1935	
total_timesteps	24064	
train/		
approx_kl	0.00018384983	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.93	
explained_variance	0.00859	
learning_rate	1e-06	
loss	0.111	
n_updates	460	
policy_gradient_loss	-0.00126	
value_loss	0.166	

time/		
fps	12	
iterations	48	
time_elapsed	1977	
total_timesteps	24576	
train/		
approx_kl	2.0184903e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.93	
explained_variance	0.00731	
learning_rate	1e-06	
loss	4.93	
n_updates	470	
policy_gradient_loss	-0.000298	
value_loss	10.2	

time/		
fps	12	
iterations	49	
time_elapsed	2018	
total_timesteps	25088	
train/		
approx_kl	0.00014711102	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.93	
explained_variance	0.0359	
learning_rate	1e-06	
loss	192	
n_updates	480	
policy_gradient_loss	-0.000319	
value_loss	490	

time/		
fps	12	
iterations	50	
time_elapsed	2059	
total_timesteps	25600	
train/		
approx_kl	3.4080935e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.93	
explained_variance	0.116	
learning_rate	1e-06	
loss	0.274	
n_updates	490	
policy_gradient_loss	-0.000184	
value_loss	1.06	

time/		
fps	12	
iterations	51	
time_elapsed	2099	
total_timesteps	26112	
train/		
approx_kl	3.7656282e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.93	
explained_variance	0.34	

	learning_rate		1e-06	
	loss		0.0528	
	n_updates		500	
	policy_gradient_loss		-0.000229	
	value_loss		0.346	

	time/			
	fps		12	
	iterations		52	
	time_elapsed		2140	
	total_timesteps		26624	
	train/			
	approx_kl		6.170559e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.93	
	explained_variance		0.0334	
	learning_rate		1e-06	
	loss		0.103	
	n_updates		510	
	policy_gradient_loss		-0.000401	
	value_loss		0.303	

	time/			
	fps		12	
	iterations		53	
	time_elapsed		2181	
	total_timesteps		27136	
	train/			
	approx_kl		8.158642e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.93	
	explained_variance		0.00653	
	learning_rate		1e-06	
	loss		0.0758	
	n_updates		520	
	policy_gradient_loss		-0.0006	
	value_loss		0.236	

	time/			
	fps		12	
	iterations		54	
	time_elapsed		2221	
	total_timesteps		27648	

train/		
approx_kl	0.00014260586	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	0.07	
learning_rate	1e-06	
loss	0.107	
n_updates	530	
policy_gradient_loss	-0.000893	
value_loss	0.257	

time/		
fps	12	
iterations	55	
time_elapsed	2262	
total_timesteps	28160	
train/		
approx_kl	8.463778e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	0.127	
learning_rate	1e-06	
loss	0.0566	
n_updates	540	
policy_gradient_loss	-0.000761	
value_loss	0.192	

time/		
fps	12	
iterations	56	
time_elapsed	2303	
total_timesteps	28672	
train/		
approx_kl	3.074319e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	-0.118	
learning_rate	1e-06	
loss	0.062	
n_updates	550	
policy_gradient_loss	-0.000286	
value_loss	0.18	

time/		
fps	12	
iterations	57	
time_elapsed	2344	
total_timesteps	29184	
train/		
approx_kl	4.205585e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	0.081	
learning_rate	1e-06	
loss	0.0612	
n_updates	560	
policy_gradient_loss	-0.000374	
value_loss	0.131	

time/		
fps	12	
iterations	58	
time_elapsed	2387	
total_timesteps	29696	
train/		
approx_kl	2.0968146e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	0.0188	
learning_rate	1e-06	
loss	0.043	
n_updates	570	
policy_gradient_loss	-0.000125	
value_loss	0.111	

time/		
fps	12	
iterations	59	
time_elapsed	2430	
total_timesteps	30208	
train/		
approx_kl	0.00015285402	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	0.0443	

	learning_rate		1e-06	
	loss		0.108	
	n_updates		580	
	policy_gradient_loss		-0.00114	
	value_loss		0.167	

	time/			
	fps		12	
	iterations		60	
	time_elapsed		2471	
	total_timesteps		30720	
	train/			
	approx_kl		2.3047556e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.92	
	explained_variance		0.476	
	learning_rate		1e-06	
	loss		0.0981	
	n_updates		590	
	policy_gradient_loss		-0.00024	
	value_loss		0.165	

	time/			
	fps		12	
	iterations		61	
	time_elapsed		2513	
	total_timesteps		31232	
	train/			
	approx_kl		0.00018868095	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.92	
	explained_variance		0.00325	
	learning_rate		1e-06	
	loss		0.0937	
	n_updates		600	
	policy_gradient_loss		-0.000857	
	value_loss		0.194	

	time/			
	fps		12	
	iterations		62	
	time_elapsed		2554	
	total_timesteps		31744	

train/		
approx_kl	4.6035042e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.92	
explained_variance	0.00195	
learning_rate	1e-06	
loss	0.102	
n_updates	610	
policy_gradient_loss	-0.0002	
value_loss	0.182	

time/		
fps	12	
iterations	63	
time_elapsed	2595	
total_timesteps	32256	
train/		
approx_kl	4.3555745e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.91	
explained_variance	-0.00449	
learning_rate	1e-06	
loss	0.0704	
n_updates	620	
policy_gradient_loss	-0.000253	
value_loss	0.184	

time/		
fps	12	
iterations	64	
time_elapsed	2637	
total_timesteps	32768	
train/		
approx_kl	0.0003685396	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.91	
explained_variance	0.1	
learning_rate	1e-06	
loss	107	
n_updates	630	
policy_gradient_loss	-0.000122	
value_loss	360	

time/		
fps	12	
iterations	65	
time_elapsed	2678	
total_timesteps	33280	
train/		
approx_kl	0.00031270366	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.167	
learning_rate	1e-06	
loss	58	
n_updates	640	
policy_gradient_loss	0.000458	
value_loss	96.1	

time/		
fps	12	
iterations	66	
time_elapsed	2720	
total_timesteps	33792	
train/		
approx_kl	9.151979e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	-0.117	
learning_rate	1e-06	
loss	0.154	
n_updates	650	
policy_gradient_loss	-0.000148	
value_loss	1.08	

time/		
fps	12	
iterations	67	
time_elapsed	2762	
total_timesteps	34304	
train/		
approx_kl	1.890643e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	-0.00557	

	learning_rate	1e-06	
	loss	0.148	
	n_updates	660	
	policy_gradient_loss	0.00027	
	value_loss	0.856	

	time/		
	fps	12	
	iterations	68	
	time_elapsed	2804	
	total_timesteps	34816	
	train/		
	approx_kl	6.219186e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.9	
	explained_variance	0.00117	
	learning_rate	1e-06	
	loss	0.2	
	n_updates	670	
	policy_gradient_loss	-0.000307	
	value_loss	0.555	

	time/		
	fps	12	
	iterations	69	
	time_elapsed	2846	
	total_timesteps	35328	
	train/		
	approx_kl	5.159527e-06	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.9	
	explained_variance	0.249	
	learning_rate	1e-06	
	loss	0.158	
	n_updates	680	
	policy_gradient_loss	0.000162	
	value_loss	0.524	

	time/		
	fps	12	
	iterations	70	
	time_elapsed	2888	
	total_timesteps	35840	

train/		
approx_kl	4.8471382e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.00595	
learning_rate	1e-06	
loss	0.144	
n_updates	690	
policy_gradient_loss	-0.000188	
value_loss	0.486	

time/		
fps	12	
iterations	71	
time_elapsed	2930	
total_timesteps	36352	
train/		
approx_kl	3.5290723e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.221	
learning_rate	1e-06	
loss	0.0965	
n_updates	700	
policy_gradient_loss	-0.000192	
value_loss	0.377	

time/		
fps	12	
iterations	72	
time_elapsed	2972	
total_timesteps	36864	
train/		
approx_kl	0.00012104842	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.0805	
learning_rate	1e-06	
loss	0.139	
n_updates	710	
policy_gradient_loss	-0.000881	
value_loss	0.384	

time/		
fps	12	
iterations	73	
time_elapsed	3013	
total_timesteps	37376	
train/		
approx_kl	6.950507e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.00241	
learning_rate	1e-06	
loss	0.108	
n_updates	720	
policy_gradient_loss	-0.000368	
value_loss	0.29	

time/		
fps	12	
iterations	74	
time_elapsed	3055	
total_timesteps	37888	
train/		
approx_kl	5.573081e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.00478	
learning_rate	1e-06	
loss	0.113	
n_updates	730	
policy_gradient_loss	-0.00037	
value_loss	0.266	

time/		
fps	12	
iterations	75	
time_elapsed	3097	
total_timesteps	38400	
train/		
approx_kl	0.0001605642	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	-0.0335	

	learning_rate		1e-06	
	loss		0.0408	
	n_updates		740	
	policy_gradient_loss		-0.000725	
	value_loss		0.117	

	time/			
	fps		12	
	iterations		76	
	time_elapsed		3139	
	total_timesteps		38912	
	train/			
	approx_kl		0.00011019211	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.9	
	explained_variance		-0.0152	
	learning_rate		1e-06	
	loss		0.0708	
	n_updates		750	
	policy_gradient_loss		-0.000577	
	value_loss		0.13	

	time/			
	fps		12	
	iterations		77	
	time_elapsed		3181	
	total_timesteps		39424	
	train/			
	approx_kl		8.609216e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.91	
	explained_variance		0.205	
	learning_rate		1e-06	
	loss		0.0405	
	n_updates		760	
	policy_gradient_loss		-0.000587	
	value_loss		0.107	

	time/			
	fps		12	
	iterations		78	
	time_elapsed		3223	
	total_timesteps		39936	

train/		
approx_kl	0.00016480172	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.91	
explained_variance	0.106	
learning_rate	1e-06	
loss	0.0472	
n_updates	770	
policy_gradient_loss	-0.000902	
value_loss	0.132	

time/		
fps	12	
iterations	79	
time_elapsed	3266	
total_timesteps	40448	
train/		
approx_kl	0.00023095077	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.191	
learning_rate	1e-06	
loss	0.0579	
n_updates	780	
policy_gradient_loss	-0.00107	
value_loss	0.132	

time/		
fps	12	
iterations	80	
time_elapsed	3308	
total_timesteps	40960	
train/		
approx_kl	0.0008784523	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.133	
learning_rate	1e-06	
loss	145	
n_updates	790	
policy_gradient_loss	1.18e-06	
value_loss	448	

time/		
fps	12	
iterations	81	
time_elapsed	3350	
total_timesteps	41472	
train/		
approx_kl	9.9674915e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.88	
explained_variance	0.011	
learning_rate	1e-06	
loss	0.0517	
n_updates	800	
policy_gradient_loss	-0.000349	
value_loss	1.25	

time/		
fps	12	
iterations	82	
time_elapsed	3392	
total_timesteps	41984	
train/		
approx_kl	9.992835e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	-0.00663	
learning_rate	1e-06	
loss	0.0526	
n_updates	810	
policy_gradient_loss	-0.000352	
value_loss	0.438	

time/		
fps	12	
iterations	83	
time_elapsed	3434	
total_timesteps	42496	
train/		
approx_kl	0.00022892875	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	-0.0138	

	learning_rate	1e-06	
	loss	0.047	
	n_updates	820	
	policy_gradient_loss	-0.00145	
	value_loss	0.371	

	time/		
	fps	12	
	iterations	84	
	time_elapsed	3475	
	total_timesteps	43008	
	train/		
	approx_kl	7.514702e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.9	
	explained_variance	0.0344	
	learning_rate	1e-06	
	loss	0.0932	
	n_updates	830	
	policy_gradient_loss	-0.000376	
	value_loss	0.301	

	time/		
	fps	12	
	iterations	85	
	time_elapsed	3517	
	total_timesteps	43520	
	train/		
	approx_kl	5.2896445e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.9	
	explained_variance	0.0758	
	learning_rate	1e-06	
	loss	0.0617	
	n_updates	840	
	policy_gradient_loss	-0.000259	
	value_loss	0.194	

	time/		
	fps	12	
	iterations	86	
	time_elapsed	3559	
	total_timesteps	44032	

train/		
approx_kl	0.00029565243	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.0653	
learning_rate	1e-06	
loss	0.0672	
n_updates	850	
policy_gradient_loss	-0.00136	
value_loss	0.186	

time/		
fps	12	
iterations	87	
time_elapsed	3601	
total_timesteps	44544	
train/		
approx_kl	0.00013438682	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.213	
learning_rate	1e-06	
loss	0.062	
n_updates	860	
policy_gradient_loss	-0.0008	
value_loss	0.157	

time/		
fps	12	
iterations	88	
time_elapsed	3642	
total_timesteps	45056	
train/		
approx_kl	7.899932e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.112	
learning_rate	1e-06	
loss	0.0495	
n_updates	870	
policy_gradient_loss	-0.000551	
value_loss	0.158	

time/		
fps	12	
iterations	89	
time_elapsed	3684	
total_timesteps	45568	
train/		
approx_kl	0.00015839946	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.192	
learning_rate	1e-06	
loss	0.0716	
n_updates	880	
policy_gradient_loss	-0.0011	
value_loss	0.166	

time/		
fps	12	
iterations	90	
time_elapsed	3725	
total_timesteps	46080	
train/		
approx_kl	0.00011557422	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.0869	
learning_rate	1e-06	
loss	0.0702	
n_updates	890	
policy_gradient_loss	-0.000903	
value_loss	0.141	

time/		
fps	12	
iterations	91	
time_elapsed	3767	
total_timesteps	46592	
train/		
approx_kl	0.00033916696	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.0258	

	learning_rate	1e-06	
	loss	0.059	
	n_updates	900	
	policy_gradient_loss	-0.00181	
	value_loss	0.14	

	time/		
	fps	12	
	iterations	92	
	time_elapsed	3809	
	total_timesteps	47104	
	train/		
	approx_kl	4.618184e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.89	
	explained_variance	-0.0109	
	learning_rate	1e-06	
	loss	0.0633	
	n_updates	910	
	policy_gradient_loss	-0.000169	
	value_loss	0.118	

	time/		
	fps	12	
	iterations	93	
	time_elapsed	3851	
	total_timesteps	47616	
	train/		
	approx_kl	5.8326637e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.89	
	explained_variance	0.0889	
	learning_rate	1e-06	
	loss	0.0452	
	n_updates	920	
	policy_gradient_loss	-0.000515	
	value_loss	0.115	

	time/		
	fps	12	
	iterations	94	
	time_elapsed	3893	
	total_timesteps	48128	

train/	
approx_kl	0.00016686879
clip_fraction	0
clip_range	0.2
entropy_loss	-1.89
explained_variance	0.00114
learning_rate	1e-06
loss	0.0512
n_updates	930
policy_gradient_loss	-0.000883
value_loss	0.114

time/	
fps	12
iterations	95
time_elapsed	3935
total_timesteps	48640
train/	
approx_kl	3.5886187e-06
clip_fraction	0
clip_range	0.2
entropy_loss	-1.9
explained_variance	0.00914
learning_rate	1e-06
loss	2.72
n_updates	940
policy_gradient_loss	0.000197
value_loss	10.3

time/	
fps	12
iterations	96
time_elapsed	3977
total_timesteps	49152
train/	
approx_kl	0.0006195295
clip_fraction	0
clip_range	0.2
entropy_loss	-1.89
explained_variance	0.058
learning_rate	1e-06
loss	123
n_updates	950
policy_gradient_loss	-0.000338
value_loss	374

time/		
fps	12	
iterations	97	
time_elapsed	4019	
total_timesteps	49664	
train/		
approx_kl	8.178968e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	-0.0242	
learning_rate	1e-06	
loss	0.0556	
n_updates	960	
policy_gradient_loss	-0.000102	
value_loss	1.19	

time/		
fps	12	
iterations	98	
time_elapsed	4062	
total_timesteps	50176	
train/		
approx_kl	9.635079e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.043	
learning_rate	1e-06	
loss	0.0638	
n_updates	970	
policy_gradient_loss	-0.000357	
value_loss	0.389	

time/		
fps	12	
iterations	99	
time_elapsed	4104	
total_timesteps	50688	
train/		
approx_kl	2.7743517e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.131	

	learning_rate		1e-06	
	loss		0.0651	
	n_updates		980	
	policy_gradient_loss		-3.53e-05	
	value_loss		0.287	

	time/			
	fps		12	
	iterations		100	
	time_elapsed		4146	
	total_timesteps		51200	
	train/			
	approx_kl		3.202283e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.89	
	explained_variance		-0.0121	
	learning_rate		1e-06	
	loss		0.0476	
	n_updates		990	
	policy_gradient_loss		-0.000121	
	value_loss		0.177	

	time/			
	fps		12	
	iterations		101	
	time_elapsed		4188	
	total_timesteps		51712	
	train/			
	approx_kl		5.1365234e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.89	
	explained_variance		0.0958	
	learning_rate		1e-06	
	loss		0.0453	
	n_updates		1000	
	policy_gradient_loss		-0.000387	
	value_loss		0.213	

	time/			
	fps		12	
	iterations		102	
	time_elapsed		4230	
	total_timesteps		52224	

train/		
approx_kl	1.7060665e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.000889	
learning_rate	1e-06	
loss	0.0583	
n_updates	1010	
policy_gradient_loss	-5.36e-05	
value_loss	0.132	

time/		
fps	12	
iterations	103	
time_elapsed	4272	
total_timesteps	52736	
train/		
approx_kl	6.0526538e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.9	
explained_variance	0.00357	
learning_rate	1e-06	
loss	0.0692	
n_updates	1020	
policy_gradient_loss	-0.000412	
value_loss	0.134	

time/		
fps	12	
iterations	104	
time_elapsed	4314	
total_timesteps	53248	
train/		
approx_kl	7.442944e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	-0.00206	
learning_rate	1e-06	
loss	0.0582	
n_updates	1030	
policy_gradient_loss	-0.000478	
value_loss	0.133	

time/		
fps	12	
iterations	105	
time_elapsed	4356	
total_timesteps	53760	
train/		
approx_kl	0.00012411259	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.504	
learning_rate	1e-06	
loss	0.0988	
n_updates	1040	
policy_gradient_loss	-0.000813	
value_loss	0.142	

time/		
fps	12	
iterations	106	
time_elapsed	4397	
total_timesteps	54272	
train/		
approx_kl	0.000108138425	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	-0.00908	
learning_rate	1e-06	
loss	0.0683	
n_updates	1050	
policy_gradient_loss	-0.000622	
value_loss	0.129	

time/		
fps	12	
iterations	107	
time_elapsed	4439	
total_timesteps	54784	
train/		
approx_kl	0.00041910843	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.89	
explained_variance	0.132	

	learning_rate	1e-06	
	loss	0.0638	
	n_updates	1060	
	policy_gradient_loss	-0.00185	
	value_loss	0.145	

	time/		
	fps	12	
	iterations	108	
	time_elapsed	4481	
	total_timesteps	55296	
	train/		
	approx_kl	3.0186027e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.88	
	explained_variance	0.0322	
	learning_rate	1e-06	
	loss	0.0914	
	n_updates	1070	
	policy_gradient_loss	-0.00017	
	value_loss	0.135	

	time/		
	fps	12	
	iterations	109	
	time_elapsed	4523	
	total_timesteps	55808	
	train/		
	approx_kl	0.0002167488	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.88	
	explained_variance	-0.000298	
	learning_rate	1e-06	
	loss	0.0386	
	n_updates	1080	
	policy_gradient_loss	-0.00139	
	value_loss	0.0976	

	time/		
	fps	12	
	iterations	110	
	time_elapsed	4565	
	total_timesteps	56320	

train/		
approx_kl	4.309381e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.88	
explained_variance	0.012	
learning_rate	1e-06	
loss	0.0479	
n_updates	1090	
policy_gradient_loss	-0.000209	
value_loss	0.0948	

time/		
fps	12	
iterations	111	
time_elapsed	4606	
total_timesteps	56832	
train/		
approx_kl	0.0007202084	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.87	
explained_variance	0.174	
learning_rate	1e-06	
loss	70.8	
n_updates	1100	
policy_gradient_loss	0.000407	
value_loss	197	

time/		
fps	12	
iterations	112	
time_elapsed	4649	
total_timesteps	57344	
train/		
approx_kl	0.0010185962	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.226	
learning_rate	1e-06	
loss	79.2	
n_updates	1110	
policy_gradient_loss	0.0014	
value_loss	185	

time/		
fps	12	
iterations	113	
time_elapsed	4691	
total_timesteps	57856	
train/		
approx_kl	3.3829827e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	-0.00883	
learning_rate	1e-06	
loss	0.085	
n_updates	1120	
policy_gradient_loss	0.000276	
value_loss	1.03	

time/		
fps	12	
iterations	114	
time_elapsed	4733	
total_timesteps	58368	
train/		
approx_kl	9.416242e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.0114	
learning_rate	1e-06	
loss	0.133	
n_updates	1130	
policy_gradient_loss	-0.000468	
value_loss	0.944	

time/		
fps	12	
iterations	115	
time_elapsed	4775	
total_timesteps	58880	
train/		
approx_kl	0.00012746197	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.0218	

	learning_rate		1e-06	
	loss		0.116	
	n_updates		1140	
	policy_gradient_loss		-0.000518	
	value_loss		0.609	

	time/			
	fps		12	
	iterations		116	
	time_elapsed		4817	
	total_timesteps		59392	
	train/			
	approx_kl		6.091944e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.86	
	explained_variance		0.0139	
	learning_rate		1e-06	
	loss		0.127	
	n_updates		1150	
	policy_gradient_loss		-0.000122	
	value_loss		0.455	

	time/			
	fps		12	
	iterations		117	
	time_elapsed		4858	
	total_timesteps		59904	
	train/			
	approx_kl		0.00010348763	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.86	
	explained_variance		0.00156	
	learning_rate		1e-06	
	loss		0.143	
	n_updates		1160	
	policy_gradient_loss		-0.000698	
	value_loss		0.397	

	time/			
	fps		12	
	iterations		118	
	time_elapsed		4900	
	total_timesteps		60416	

train/		
approx_kl	5.0334027e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.169	
learning_rate	1e-06	
loss	0.125	
n_updates	1170	
policy_gradient_loss	-0.000255	
value_loss	0.332	

time/		
fps	12	
iterations	119	
time_elapsed	4943	
total_timesteps	60928	
train/		
approx_kl	0.00016150868	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	-0.0014	
learning_rate	1e-06	
loss	0.149	
n_updates	1180	
policy_gradient_loss	-0.000954	
value_loss	0.347	

time/		
fps	12	
iterations	120	
time_elapsed	4984	
total_timesteps	61440	
train/		
approx_kl	0.00014481845	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.00844	
learning_rate	1e-06	
loss	0.16	
n_updates	1190	
policy_gradient_loss	-0.00082	
value_loss	0.319	

time/		
fps	12	
iterations	121	
time_elapsed	5026	
total_timesteps	61952	
train/		
approx_kl	9.388605e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.00355	
learning_rate	1e-06	
loss	0.22	
n_updates	1200	
policy_gradient_loss	-0.000503	
value_loss	0.311	

time/		
fps	12	
iterations	122	
time_elapsed	5067	
total_timesteps	62464	
train/		
approx_kl	7.499533e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	-0.0101	
learning_rate	1e-06	
loss	0.0589	
n_updates	1210	
policy_gradient_loss	-0.000444	
value_loss	0.155	

time/		
fps	12	
iterations	123	
time_elapsed	5109	
total_timesteps	62976	
train/		
approx_kl	3.6163605e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.87	
explained_variance	0.0257	

	learning_rate		1e-06	
	loss		0.0468	
	n_updates		1220	
	policy_gradient_loss		-0.000257	
	value_loss		0.113	

	time/			
	fps		12	
	iterations		124	
	time_elapsed		5151	
	total_timesteps		63488	
	train/			
	approx_kl		0.00013010972	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.86	
	explained_variance		-0.0306	
	learning_rate		1e-06	
	loss		0.041	
	n_updates		1230	
	policy_gradient_loss		-0.000859	
	value_loss		0.105	

	time/			
	fps		12	
	iterations		125	
	time_elapsed		5193	
	total_timesteps		64000	
	train/			
	approx_kl		5.649624e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.86	
	explained_variance		0.199	
	learning_rate		1e-06	
	loss		0.0491	
	n_updates		1240	
	policy_gradient_loss		-0.000298	
	value_loss		0.152	

	time/			
	fps		12	
	iterations		126	
	time_elapsed		5235	
	total_timesteps		64512	

train/		
approx_kl	4.4743996e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	0.000675	
learning_rate	1e-06	
loss	0.0581	
n_updates	1250	
policy_gradient_loss	-0.000183	
value_loss	0.114	

time/		
fps	12	
iterations	127	
time_elapsed	5276	
total_timesteps	65024	
train/		
approx_kl	0.0013513819	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	0.224	
learning_rate	1e-06	
loss	91.5	
n_updates	1260	
policy_gradient_loss	0.00126	
value_loss	324	

time/		
fps	12	
iterations	128	
time_elapsed	5319	
total_timesteps	65536	
train/		
approx_kl	1.8454739e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.84	
explained_variance	0.444	
learning_rate	1e-06	
loss	0.243	
n_updates	1270	
policy_gradient_loss	-1.64e-05	
value_loss	8.75	

time/		
fps	12	
iterations	129	
time_elapsed	5361	
total_timesteps	66048	
train/		
approx_kl	0.00017793779	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	0.263	
learning_rate	1e-06	
loss	0.0754	
n_updates	1280	
policy_gradient_loss	-0.000975	
value_loss	0.476	

time/		
fps	12	
iterations	130	
time_elapsed	5403	
total_timesteps	66560	
train/		
approx_kl	4.007644e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	-0.0297	
learning_rate	1e-06	
loss	0.0552	
n_updates	1290	
policy_gradient_loss	0.000132	
value_loss	0.405	

time/		
fps	12	
iterations	131	
time_elapsed	5445	
total_timesteps	67072	
train/		
approx_kl	3.6559533e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	-0.0215	

	learning_rate		1e-06	
	loss		0.0605	
	n_updates		1300	
	policy_gradient_loss		-0.000137	
	value_loss		0.29	

	time/			
	fps		12	
	iterations		132	
	time_elapsed		5487	
	total_timesteps		67584	
	train/			
	approx_kl		9.479339e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.85	
	explained_variance		-0.0413	
	learning_rate		1e-06	
	loss		0.0996	
	n_updates		1310	
	policy_gradient_loss		-0.000626	
	value_loss		0.228	

	time/			
	fps		12	
	iterations		133	
	time_elapsed		5529	
	total_timesteps		68096	
	train/			
	approx_kl		1.0531978e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.85	
	explained_variance		-0.0179	
	learning_rate		1e-06	
	loss		0.0852	
	n_updates		1320	
	policy_gradient_loss		4.67e-05	
	value_loss		0.201	

	time/			
	fps		12	
	iterations		134	
	time_elapsed		5571	
	total_timesteps		68608	

train/		
approx_kl	7.988722e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	-0.0255	
learning_rate	1e-06	
loss	0.0525	
n_updates	1330	
policy_gradient_loss	-0.000358	
value_loss	0.155	

time/		
fps	12	
iterations	135	
time_elapsed	5613	
total_timesteps	69120	
train/		
approx_kl	9.0383925e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	-0.0133	
learning_rate	1e-06	
loss	0.0509	
n_updates	1340	
policy_gradient_loss	-0.000411	
value_loss	0.128	

time/		
fps	12	
iterations	136	
time_elapsed	5654	
total_timesteps	69632	
train/		
approx_kl	5.265104e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.86	
explained_variance	-0.0864	
learning_rate	1e-06	
loss	0.0721	
n_updates	1350	
policy_gradient_loss	-0.000233	
value_loss	0.195	

time/		
fps	12	
iterations	137	
time_elapsed	5697	
total_timesteps	70144	
train/		
approx_kl	0.00031185616	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.85	
explained_variance	-0.02	
learning_rate	1e-06	
loss	0.0479	
n_updates	1360	
policy_gradient_loss	-0.00159	
value_loss	0.14	

time/		
fps	12	
iterations	138	
time_elapsed	5739	
total_timesteps	70656	
train/		
approx_kl	0.0015150894	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.83	
explained_variance	0.22	
learning_rate	1e-06	
loss	101	
n_updates	1370	
policy_gradient_loss	0.00261	
value_loss	268	

time/		
fps	12	
iterations	139	
time_elapsed	5781	
total_timesteps	71168	
train/		
approx_kl	0.0024212827	
clip_fraction	0.00371	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	0.283	

	learning_rate		1e-06	
	loss		160	
	n_updates		1380	
	policy_gradient_loss		0.000159	
	value_loss		459	

	time/			
	fps		12	
	iterations		140	
	time_elapsed		5824	
	total_timesteps		71680	
	train/			
	approx_kl		0.00020098616	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.79	
	explained_variance		-0.0353	
	learning_rate		1e-06	
	loss		0.231	
	n_updates		1390	
	policy_gradient_loss		-0.000559	
	value_loss		3.79	

	time/			
	fps		12	
	iterations		141	
	time_elapsed		5866	
	total_timesteps		72192	
	train/			
	approx_kl		0.0001736983	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.8	
	explained_variance		-0.0411	
	learning_rate		1e-06	
	loss		0.0655	
	n_updates		1400	
	policy_gradient_loss		0.000209	
	value_loss		1.88	

	time/			
	fps		12	
	iterations		142	
	time_elapsed		5907	
	total_timesteps		72704	

train/		
approx_kl	0.00031178736	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.642	
learning_rate	1e-06	
loss	0.0683	
n_updates	1410	
policy_gradient_loss	-0.00253	
value_loss	1.19	

time/		
fps	12	
iterations	143	
time_elapsed	5949	
total_timesteps	73216	
train/		
approx_kl	2.9282295e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.0847	
learning_rate	1e-06	
loss	12.1	
n_updates	1420	
policy_gradient_loss	-0.000263	
value_loss	26.6	

time/		
fps	12	
iterations	144	
time_elapsed	5991	
total_timesteps	73728	
train/		
approx_kl	4.287029e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.0507	
learning_rate	1e-06	
loss	13.8	
n_updates	1430	
policy_gradient_loss	-0.000528	
value_loss	30.8	

time/		
fps	12	
iterations	145	
time_elapsed	6033	
total_timesteps	74240	
train/		
approx_kl	0.00015299709	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.065	
learning_rate	1e-06	
loss	0.136	
n_updates	1440	
policy_gradient_loss	-0.000478	
value_loss	1.74	

time/		
fps	12	
iterations	146	
time_elapsed	6075	
total_timesteps	74752	
train/		
approx_kl	0.00015581073	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.0531	
learning_rate	1e-06	
loss	0.0454	
n_updates	1450	
policy_gradient_loss	-0.000297	
value_loss	0.978	

time/		
fps	12	
iterations	147	
time_elapsed	6117	
total_timesteps	75264	
train/		
approx_kl	0.00013859512	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.83	
explained_variance	-0.0419	

	learning_rate		1e-06	
	loss		0.0505	
	n_updates		1460	
	policy_gradient_loss		-0.000499	
	value_loss		0.766	

	time/			
	fps		12	
	iterations		148	
	time_elapsed		6159	
	total_timesteps		75776	
	train/			
	approx_kl		0.0001523419	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.84	
	explained_variance		0.675	
	learning_rate		1e-06	
	loss		0.143	
	n_updates		1470	
	policy_gradient_loss		-0.000455	
	value_loss		0.766	

	time/			
	fps		12	
	iterations		149	
	time_elapsed		6201	
	total_timesteps		76288	
	train/			
	approx_kl		0.00011460984	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.84	
	explained_variance		0.0351	
	learning_rate		1e-06	
	loss		0.061	
	n_updates		1480	
	policy_gradient_loss		-0.000342	
	value_loss		0.53	

	time/			
	fps		12	
	iterations		150	
	time_elapsed		6242	
	total_timesteps		76800	

train/		
approx_kl	0.00015744916	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.84	
explained_variance	0.0772	
learning_rate	1e-06	
loss	0.0649	
n_updates	1490	
policy_gradient_loss	-0.000292	
value_loss	0.359	

time/		
fps	12	
iterations	151	
time_elapsed	6285	
total_timesteps	77312	
train/		
approx_kl	5.3031603e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.83	
explained_variance	-0.0344	
learning_rate	1e-06	
loss	0.0437	
n_updates	1500	
policy_gradient_loss	-0.000188	
value_loss	0.253	

time/		
fps	12	
iterations	152	
time_elapsed	6327	
total_timesteps	77824	
train/		
approx_kl	0.0002339991	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.83	
explained_variance	-0.00921	
learning_rate	1e-06	
loss	0.0697	
n_updates	1510	
policy_gradient_loss	-0.00107	
value_loss	0.2	

time/		
fps	12	
iterations	153	
time_elapsed	6369	
total_timesteps	78336	
train/		
approx_kl	0.00031094323	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.84	
explained_variance	0.00636	
learning_rate	1e-06	
loss	0.0329	
n_updates	1520	
policy_gradient_loss	-0.00131	
value_loss	0.189	

time/		
fps	12	
iterations	154	
time_elapsed	6411	
total_timesteps	78848	
train/		
approx_kl	0.00054658076	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.83	
explained_variance	0.134	
learning_rate	1e-06	
loss	96.5	
n_updates	1530	
policy_gradient_loss	-0.000438	
value_loss	245	

time/		
fps	12	
iterations	155	
time_elapsed	6453	
total_timesteps	79360	
train/		
approx_kl	0.000433572	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.45	

	learning_rate		1e-06	
	loss		45.3	
	n_updates		1540	
	policy_gradient_loss		-0.00054	
	value_loss		100	

	time/			
	fps		12	
	iterations		156	
	time_elapsed		6495	
	total_timesteps		79872	
	train/			
	approx_kl		3.6942423e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.82	
	explained_variance		-0.0168	
	learning_rate		1e-06	
	loss		0.117	
	n_updates		1550	
	policy_gradient_loss		-6.5e-06	
	value_loss		0.697	

	time/			
	fps		12	
	iterations		157	
	time_elapsed		6537	
	total_timesteps		80384	
	train/			
	approx_kl		6.376114e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.82	
	explained_variance		-0.00162	
	learning_rate		1e-06	
	loss		0.139	
	n_updates		1560	
	policy_gradient_loss		-0.000214	
	value_loss		0.494	

	time/			
	fps		12	
	iterations		158	
	time_elapsed		6580	
	total_timesteps		80896	

train/		
approx_kl	0.00014303823	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.331	
learning_rate	1e-06	
loss	25.6	
n_updates	1570	
policy_gradient_loss	0.00024	
value_loss	62.7	

time/		
fps	12	
iterations	159	
time_elapsed	6622	
total_timesteps	81408	
train/		
approx_kl	6.7938585e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.233	
learning_rate	1e-06	
loss	1.56	
n_updates	1580	
policy_gradient_loss	0.000202	
value_loss	4.62	

time/		
fps	12	
iterations	160	
time_elapsed	6663	
total_timesteps	81920	
train/		
approx_kl	0.00017160829	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.00948	
learning_rate	1e-06	
loss	0.147	
n_updates	1590	
policy_gradient_loss	-0.000368	
value_loss	1.43	

time/		
fps	12	
iterations	161	
time_elapsed	6705	
total_timesteps	82432	
train/		
approx_kl	9.664684e-06	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.0052	
learning_rate	1e-06	
loss	0.11	
n_updates	1600	
policy_gradient_loss	0.000185	
value_loss	0.94	

time/		
fps	12	
iterations	162	
time_elapsed	6747	
total_timesteps	82944	
train/		
approx_kl	8.9262845e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	-0.0307	
learning_rate	1e-06	
loss	0.157	
n_updates	1610	
policy_gradient_loss	-0.000586	
value_loss	0.672	

time/		
fps	12	
iterations	163	
time_elapsed	6789	
total_timesteps	83456	
train/		
approx_kl	0.00020587805	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	-0.0109	

	learning_rate		1e-06	
	loss		0.113	
	n_updates		1620	
	policy_gradient_loss		-0.000652	
	value_loss		0.535	

	time/			
	fps		12	
	iterations		164	
	time_elapsed		6831	
	total_timesteps		83968	
	train/			
	approx_kl		0.00012078183	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.81	
	explained_variance		0.0713	
	learning_rate		1e-06	
	loss		0.124	
	n_updates		1630	
	policy_gradient_loss		-0.000511	
	value_loss		0.386	

	time/			
	fps		12	
	iterations		165	
	time_elapsed		6873	
	total_timesteps		84480	
	train/			
	approx_kl		0.00025412044	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.82	
	explained_variance		-0.00447	
	learning_rate		1e-06	
	loss		0.0569	
	n_updates		1640	
	policy_gradient_loss		-0.000846	
	value_loss		0.207	

	time/			
	fps		12	
	iterations		166	
	time_elapsed		6914	
	total_timesteps		84992	

train/		
approx_kl	0.00014541461	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.000424	
learning_rate	1e-06	
loss	0.053	
n_updates	1650	
policy_gradient_loss	-0.000425	
value_loss	0.175	

time/		
fps	12	
iterations	167	
time_elapsed	6957	
total_timesteps	85504	
train/		
approx_kl	0.00014247303	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.0167	
learning_rate	1e-06	
loss	0.0489	
n_updates	1660	
policy_gradient_loss	-0.000481	
value_loss	0.145	

time/		
fps	12	
iterations	168	
time_elapsed	6998	
total_timesteps	86016	
train/		
approx_kl	0.0002763929	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.00201	
learning_rate	1e-06	
loss	0.0385	
n_updates	1670	
policy_gradient_loss	-0.000878	
value_loss	0.143	

time/		
fps	12	
iterations	169	
time_elapsed	7040	
total_timesteps	86528	
train/		
approx_kl	0.00018240884	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.83	
explained_variance	-0.0455	
learning_rate	1e-06	
loss	0.0705	
n_updates	1680	
policy_gradient_loss	-0.000672	
value_loss	0.14	

time/		
fps	12	
iterations	170	
time_elapsed	7082	
total_timesteps	87040	
train/		
approx_kl	0.00097097503	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.82	
explained_variance	0.246	
learning_rate	1e-06	
loss	162	
n_updates	1690	
policy_gradient_loss	-0.000377	
value_loss	513	

time/		
fps	12	
iterations	171	
time_elapsed	7124	
total_timesteps	87552	
train/		
approx_kl	1.467939e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	-0.0259	

	learning_rate	1e-06	
	loss	0.0578	
	n_updates	1700	
	policy_gradient_loss	-6.09e-06	
	value_loss	0.684	

	time/		
	fps	12	
	iterations	172	
	time_elapsed	7165	
	total_timesteps	88064	
	train/		
	approx_kl	0.0001691794	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.82	
	explained_variance	0.0193	
	learning_rate	1e-06	
	loss	0.0599	
	n_updates	1710	
	policy_gradient_loss	-0.000949	
	value_loss	0.294	

	time/		
	fps	12	
	iterations	173	
	time_elapsed	7207	
	total_timesteps	88576	
	train/		
	approx_kl	0.0001343633	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.82	
	explained_variance	-0.0236	
	learning_rate	1e-06	
	loss	0.0684	
	n_updates	1720	
	policy_gradient_loss	-0.000197	
	value_loss	0.218	

	time/		
	fps	12	
	iterations	174	
	time_elapsed	7249	
	total_timesteps	89088	

train/		
approx_kl	0.00024553062	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	0.148	
learning_rate	1e-06	
loss	44.8	
n_updates	1730	
policy_gradient_loss	0.000637	
value_loss	110	

time/		
fps	12	
iterations	175	
time_elapsed	7291	
total_timesteps	89600	
train/		
approx_kl	3.8649072e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	0.0206	
learning_rate	1e-06	
loss	0.0931	
n_updates	1740	
policy_gradient_loss	1.62e-05	
value_loss	0.696	

time/		
fps	12	
iterations	176	
time_elapsed	7333	
total_timesteps	90112	
train/		
approx_kl	2.646586e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	-0.07	
learning_rate	1e-06	
loss	0.056	
n_updates	1750	
policy_gradient_loss	0.000156	
value_loss	0.601	

time/		
fps	12	
iterations	177	
time_elapsed	7375	
total_timesteps	90624	
train/		
approx_kl	0.00013761909	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	-0.0391	
learning_rate	1e-06	
loss	0.0511	
n_updates	1760	
policy_gradient_loss	-0.000553	
value_loss	0.388	

time/		
fps	12	
iterations	178	
time_elapsed	7417	
total_timesteps	91136	
train/		
approx_kl	6.215391e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	0.0214	
learning_rate	1e-06	
loss	0.0556	
n_updates	1770	
policy_gradient_loss	-0.000154	
value_loss	0.223	

time/		
fps	12	
iterations	179	
time_elapsed	7459	
total_timesteps	91648	
train/		
approx_kl	4.1826162e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	-0.0167	

	learning_rate		1e-06	
	loss		0.0849	
	n_updates		1780	
	policy_gradient_loss		-0.000139	
	value_loss		0.239	

	time/			
	fps		12	
	iterations		180	
	time_elapsed		7500	
	total_timesteps		92160	
	train/			
	approx_kl		8.934201e-05	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.81	
	explained_variance		-0.0457	
	learning_rate		1e-06	
	loss		0.0418	
	n_updates		1790	
	policy_gradient_loss		-0.000374	
	value_loss		0.163	

	time/			
	fps		12	
	iterations		181	
	time_elapsed		7542	
	total_timesteps		92672	
	train/			
	approx_kl		0.00015510048	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.81	
	explained_variance		-0.00581	
	learning_rate		1e-06	
	loss		0.0335	
	n_updates		1800	
	policy_gradient_loss		-0.000571	
	value_loss		0.137	

	time/			
	fps		12	
	iterations		182	
	time_elapsed		7584	
	total_timesteps		93184	

train/		
approx_kl	0.00021820643	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	-0.0335	
learning_rate	1e-06	
loss	0.06	
n_updates	1810	
policy_gradient_loss	-0.00083	
value_loss	0.137	

time/		
fps	12	
iterations	183	
time_elapsed	7625	
total_timesteps	93696	
train/		
approx_kl	0.00023878354	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	0.0242	
learning_rate	1e-06	
loss	0.0377	
n_updates	1820	
policy_gradient_loss	-0.000796	
value_loss	0.108	

time/		
fps	12	
iterations	184	
time_elapsed	7667	
total_timesteps	94208	
train/		
approx_kl	0.0001907614	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.81	
explained_variance	-0.0258	
learning_rate	1e-06	
loss	0.0414	
n_updates	1830	
policy_gradient_loss	-0.000709	
value_loss	0.111	

time/		
fps	12	
iterations	185	
time_elapsed	7709	
total_timesteps	94720	
train/		
approx_kl	0.00012336986	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.8	
explained_variance	-0.00253	
learning_rate	1e-06	
loss	0.0471	
n_updates	1840	
policy_gradient_loss	-0.000564	
value_loss	0.104	

time/		
fps	12	
iterations	186	
time_elapsed	7751	
total_timesteps	95232	
train/		
approx_kl	0.0018382153	
clip_fraction	0.00293	
clip_range	0.2	
entropy_loss	-1.78	
explained_variance	0.301	
learning_rate	1e-06	
loss	138	
n_updates	1850	
policy_gradient_loss	-0.000126	
value_loss	396	

time/		
fps	12	
iterations	187	
time_elapsed	7792	
total_timesteps	95744	
train/		
approx_kl	0.00015607162	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.77	
explained_variance	-0.025	

	learning_rate		1e-06	
	loss		0.126	
	n_updates		1860	
	policy_gradient_loss		-0.000225	
	value_loss		1.9	

	time/			
	fps		12	
	iterations		188	
	time_elapsed		7834	
	total_timesteps		96256	
	train/			
	approx_kl		0.0028756773	
	clip_fraction		0.00586	
	clip_range		0.2	
	entropy_loss		-1.75	
	explained_variance		0.153	
	learning_rate		1e-06	
	loss		139	
	n_updates		1870	
	policy_gradient_loss		-0.00146	
	value_loss		321	

	time/			
	fps		12	
	iterations		189	
	time_elapsed		7876	
	total_timesteps		96768	
	train/			
	approx_kl		0.00014718226	
	clip_fraction		0	
	clip_range		0.2	
	entropy_loss		-1.73	
	explained_variance		0.692	
	learning_rate		1e-06	
	loss		6.49	
	n_updates		1880	
	policy_gradient_loss		-0.000221	
	value_loss		21.5	

	time/			
	fps		12	
	iterations		190	
	time_elapsed		7917	
	total_timesteps		97280	

train/	
approx_kl	1.9481522e-05
clip_fraction	0
clip_range	0.2
entropy_loss	-1.73
explained_variance	-0.0226
learning_rate	1e-06
loss	0.277
n_updates	1890
policy_gradient_loss	0.000678
value_loss	1.77

time/	
fps	12
iterations	191
time_elapsed	7960
total_timesteps	97792
train/	
approx_kl	0.0002280568
clip_fraction	0
clip_range	0.2
entropy_loss	-1.74
explained_variance	-0.0435
learning_rate	1e-06
loss	0.189
n_updates	1900
policy_gradient_loss	-0.000742
value_loss	1.09

time/	
fps	12
iterations	192
time_elapsed	8001
total_timesteps	98304
train/	
approx_kl	0.00011534197
clip_fraction	0
clip_range	0.2
entropy_loss	-1.74
explained_variance	-0.0243
learning_rate	1e-06
loss	0.13
n_updates	1910
policy_gradient_loss	-0.000201
value_loss	1.07

time/		
fps	12	
iterations	193	
time_elapsed	8043	
total_timesteps	98816	
train/		
approx_kl	0.00021190918	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.75	
explained_variance	-0.0199	
learning_rate	1e-06	
loss	0.138	
n_updates	1920	
policy_gradient_loss	-0.000838	
value_loss	0.611	

time/		
fps	12	
iterations	194	
time_elapsed	8085	
total_timesteps	99328	
train/		
approx_kl	4.2675994e-05	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.76	
explained_variance	-0.00344	
learning_rate	1e-06	
loss	0.164	
n_updates	1930	
policy_gradient_loss	0.000275	
value_loss	0.51	

time/		
fps	12	
iterations	195	
time_elapsed	8127	
total_timesteps	99840	
train/		
approx_kl	0.00015855033	
clip_fraction	0	
clip_range	0.2	
entropy_loss	-1.75	
explained_variance	-0.0142	

	learning_rate	1e-06	
	loss	0.0709	
	n_updates	1940	
	policy_gradient_loss	-0.000825	
	value_loss	0.401	

	time/		
	fps	12	
	iterations	196	
	time_elapsed	8169	
	total_timesteps	100352	
	train/		
	approx_kl	6.7998655e-05	
	clip_fraction	0	
	clip_range	0.2	
	entropy_loss	-1.75	
	explained_variance	-0.0277	
	learning_rate	1e-06	
	loss	0.0675	
	n_updates	1950	
	policy_gradient_loss	-0.000441	
	value_loss	0.263	

[33]: <stable_baselines3.ppo.ppo.PPO at 0x7f649bc53370>

3 Test it out

```
[19]: # Load model
model = PPO.load('./train/best_model_100000')
```

```
[20]: state = env.reset()
```

```
[21]: # Start the game
state = env.reset()
# Loop through the game
while True:
    action, _ = model.predict(state)
    state, reward, done, info = env.step(action)
    env.render()
```

```
/home/joy/.local/lib/python3.10/site-
packages/stable_baselines3/common/vec_env/base_vec_env.py:234: UserWarning: You
tried to call render() but no `render_mode` was passed to the env constructor.
warnings.warn("You tried to call render() but no `render_mode` was passed to
the env constructor.")
```

```

/home/joy/.local/lib/python3.10/site-
packages/gym_super_mario_bros/smb_env.py:148: RuntimeWarning: overflow
encountered in ubyte_scalars
    return (self.ram[0x86] - self.ram[0x071c]) % 256

```

```

-----
KeyboardInterrupt                                Traceback (most recent call last)
Cell In [21], line 6
      4 while True:
      5     action, _ = model.predict(state)
----> 6     state, reward, done, info = env.step(action)
      7     env.render()

File ~/.local/lib/python3.10/site-packages/stable_baselines3/common/vec_env/
↳base_vec_env.py:197, in VecEnv.step(self, actions)
    190 """
    191 Step the environments with the given action
    192
    193 :param actions: the action
    194 :return: observation, reward, done, information
    195 """
    196 self.step_async(actions)
--> 197 return self.step_wait()

File ~/.local/lib/python3.10/site-packages/stable_baselines3/common/vec_env/
↳vec_frame_stack.py:33, in VecFrameStack.step_wait(self)
    30 def step_wait(
    31     self,
    32 ) -> Tuple[Union[np.ndarray, Dict[str, np.ndarray]], np.ndarray, np.
↳ndarray, List[Dict[str, Any]],]:
----> 33     observations, rewards, dones, infos = self.venv.step_wait()
    34     observations, infos = self.stacked_obs.update(observations, dones,
↳infos) # type: ignore[arg-type]
    35     return observations, rewards, dones, infos

File ~/.local/lib/python3.10/site-packages/stable_baselines3/common/vec_env/
↳dummy_vec_env.py:58, in DummyVecEnv.step_wait(self)
    55 def step_wait(self) -> VecEnvStepReturn:
    56     # Avoid circular imports
    57     for env_idx in range(self.num_envs):
----> 58         obs, self.buf_rews[env_idx], terminated, truncated, self.
↳buf_infos[env_idx] = self.envs[env_idx].step(
    59             self.actions[env_idx]
    60         )
    61         # convert to SB3 VecEnv api
    62         self.buf_dones[env_idx] = terminated or truncated

```

```

File ~/.local/lib/python3.10/site-packages/shimmy/openai_gym_compatibility.py:
↳257, in GymV21CompatibilityV0.step(self, action)
    248 def step(self, action: ActType) -> tuple[Any, float, bool, bool, dict]:
    249     """Steps through the environment.
    250
    251     Args:
    (...)
    255         (observation, reward, terminated, truncated, info)
    256     """
--> 257     obs, reward, done, info = self.gym_env.step(action)
    259     if self.render_mode is not None:
    260         self.render()

```

```

File ~/.local/lib/python3.10/site-packages/gym/core.py:495, in_
↳ObservationWrapper.step(self, action)
    493 def step(self, action):
    494     """Returns a modified observation using :meth:`self.observation`_
↳after calling :meth:`env.step`. """
--> 495     step_returns = self.env.step(action)
    496     if len(step_returns) == 5:
    497         observation, reward, terminated, truncated, info = step_returns

```

```

File ~/.local/lib/python3.10/site-packages/nes_py/wrappers/joypad_space.py:74,
↳in JoypadSpace.step(self, action)
    59 """
    60 Take a step using the given action.
    61
    (...)
    71
    72 """
    73 # take the step and record the output
---> 74 return self.env.step(self._action_map[action])

```

```

File ~/.local/lib/python3.10/site-packages/gym/wrappers/time_limit.py:60, in_
↳TimeLimit.step(self, action)
    48 def step(self, action):
    49     """Steps through the environment and if the number of steps elapsed
↳exceeds ``max_episode_steps`` then truncate.
    50
    51     Args:
    (...)
    57         "TimeLimit.truncated"=False if the environment terminated
    58     """
    59     observation, reward, terminated, truncated, info =_
↳step_api_compatibility(
---> 60         self.env.step(action),
    61         True,
    62     )

```

```

63     self._elapsed_steps += 1
65     if self._elapsed_steps >= self._max_episode_steps:

File ~/.local/lib/python3.10/site-packages/gym/wrappers/order_enforcing.py:37, in
↳ OrderEnforcing.step(self, action)
    35 if not self._has_reset:
    36     raise ResetNeeded("Cannot call env.step() before calling env.
↳ reset()")
---> 37 return self.env.step(action)

File ~/.local/lib/python3.10/site-packages/gym/wrappers/step_api_compatibility.
↳ py:52, in StepAPICompatibility.step(self, action)
    43 def step(self, action):
    44     """Steps through the environment, returning 5 or 4 items depending
↳ on `new_step_api`.
    45
    46     Args:
    (...)
    50         (observation, reward, terminated, truncated, info) or
↳ (observation, reward, done, info)
    51     """
---> 52     step_returns = self.env.step(action)
    53     if self.new_step_api:
    54         return step_to_new_api(step_returns)

File ~/.local/lib/python3.10/site-packages/gym/wrappers/env_checker.py:39, in
↳ PassiveEnvChecker.step(self, action)
    37     return env_step_passive_checker(self.env, action)
    38 else:
---> 39     return self.env.step(action)

File ~/.local/lib/python3.10/site-packages/nes_py/nes_env.py:293, in NESEnv.
↳ step(self, action)
    291 self.controllers[0][:] = action
    292 # pass the action to the emulator as an unsigned byte
--> 293 _LIB.Step(self._env)
    294 # get the reward for this step
    295 reward = self._get_reward()

KeyboardInterrupt:

```

```

[23]: # gym-super-mario-bros==7.3.0
      # nes-py==8.1.8
      # gym==0.25.1
      # stable-baselines3==2.1.0
      # opencv-python==4.8.0.76

```


[]: