

Ray Wrapper: Vision Transformer

Ray Wrapper: Vision Transformer

Vision Transformer (ViT) на CIFAR-10 и CIFAR-100 — самые тяжёлые примеры пакета `ray_wrapper`, обучаются "с нуля" (без предобученных весов).

CIFAR-10 —

`ray_wrapper/examples/vit/cifar_10/train.py`

```
scale_config = {"num_workers": 2, "use_gpu": False}
train_config = {
    "num_layers": 6, "num_heads": 4, "hidden_dim": 192, "mlp_dim": 384,
    "lr": 1e-3, "epochs": 1, "batch_size_per_worker": 16,
}
data_config = {"data_root": "/tmp/mlumina_data/cifar10"}

train_vit_cifar10(scale_config=scale_config, train_config=train_config,
                  data_config=data_config)
```

Компактная конфигурация трансформера: 6 слоёв, 4 головы внимания, скрытая размерность 192, MLP-размерность 384 — заметно меньше "классического" ViT-Base (12 слоёв, 768/3072), рассчитана на демонстрацию на CPU/малом кластере, а не на конкурентоспособное качество.

`ray_wrapper/examples/vit/cifar_10/tune.py` подбирает гиперпараметры поверх той же архитектуры аналогично примерам FCNN/CNN (см. предыдущую страницу).

CIFAR-100 —

ray_wrapper/examples/vit/cifar_100/ train.py

Идентичная конфигурация модели (`num_layers=6`, `num_heads=4`, `hidden_dim=192`, `mlp_dim=384`), но `data_config={"data_root": "/tmp/mlumina_data/cifar100"}` и, соответственно, `train_vit_cifar100` — 100 классов вместо 10. Учитывая ту же компактную архитектуру, что и для CIFAR-10, стоит ожидать заметно более низкое качество на CIFAR-100 (задача сложнее, ёмкость модели та же) — это нормально для демонстрационного примера, не повод считать код сломанным.

??? ????????

Изнутри кластера (`head`/`worker`):

```
python ray_wrapper/examples/vit/cifar_10/train.py
python ray_wrapper/examples/vit/cifar_10/tune.py
python ray_wrapper/examples/vit/cifar_100/train.py
python ray_wrapper/examples/vit/cifar_100/tune.py
```

Как и в FCNN/CNN-примерах, при `use_gpu: False` обучение на CPU при `epochs=1` — это в первую очередь проверка того, что пайплайн работает end-to-end, а не полноценное обучение модели.

Revision #1

Created 2026-09-23 06:54:49 UTC by Маслов Алексей

Updated 2026-09-23 06:54:50 UTC by Маслов Алексей