

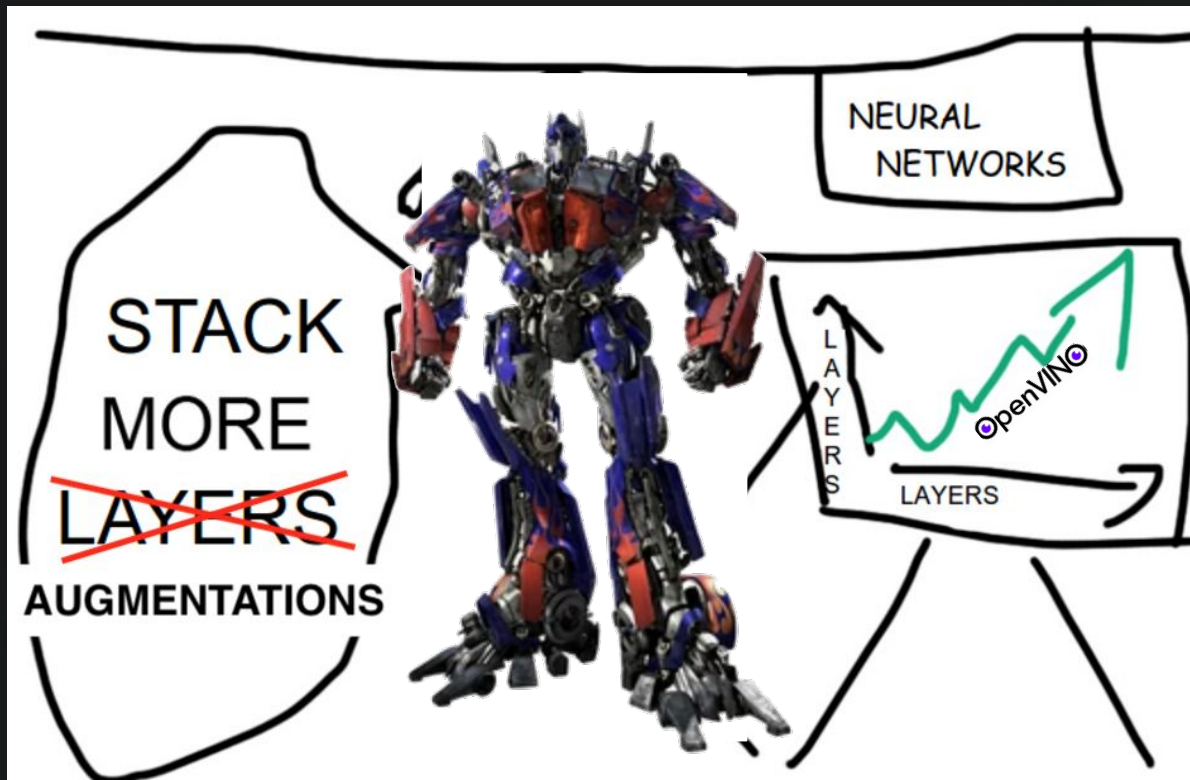
Андрей Попов

SIBUR CHALLENGE 2023//ГЕНЕРАЛИЗАЦИЯ

ПРЕЗЕНТАЦИЯ РЕШЕНИЯ
ПО ЗАДАЧЕ №1 «ВИДЕОАНАЛИТИКА»



Основная идея



Baseline - One image classification

- Учим классификатор на одном кадре с видео
- Предсказываем класс на кадрах с видео и берем самый частый класс как ответ
- Аугментации:
 - resize 224x224
 - horizontal flip
 - normalize

Public score: 0.62246

- Проблема: не учитываем темпоральную информацию

Video Classification models

- Обучаем модель из torchvision.models.video
- В моем случае - это Swin3D Tiny 🐷
- 4 кадра равномерно распределенных по видео (numpy.linspace)
- Аугментации:
 - resize 224x224
 - horizontal flip
 - normalize

Public score: 0.98028

<https://pytorch.org/vision/stable/models.html#video-classification>

Final augmentations

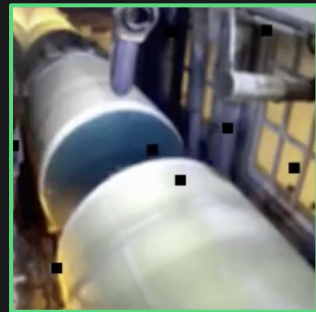
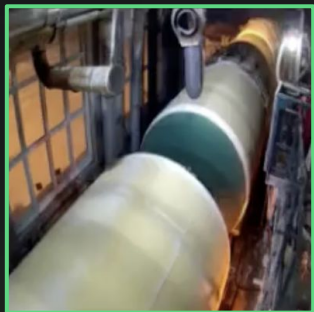
Training stage:

```
OneOf([Upsample(1.3), Downsample(0.7)]),  
OneOf([Pepper(), Salt()]),  
RandomShear(),  
RandomResizedCrop(224, 224, scale=(0.4, 1.0)),  
HorizontalFlip(),  
ShiftScaleRotate(),  
RGBShift(),  
RandomBrightnessContrast(),  
OneOf([Blur(), GaussNoise()]),  
HueSaturationValue(),  
Cutout(p=0.15),  
Normalize(),  
MixUp(alpha=0.2)
```

Inference stage:

```
Resize(256, 256),  
CenterCrop(224, 224),  
Normalize(),
```

Augmentation examples



Augmentation examples



Хаки во время обучения и инференса

Train:

- Label Smoothing (=0.1)
- Long training
- Train on full dataset
- Scheduler: CosineAnnealingLR (LR $1e-4 \rightarrow 0$)
- Optimizer: AdamW
- Exponential Moving Average (EMA)

Inference:

- PyTorch \rightarrow ONNX \rightarrow OpenVino (fp16=True) [121M \rightarrow 62Mb]
- Resize 256x256 \rightarrow CenterCrop 224x224

Public score: 98308

Хаки во время обучения и инференса

Train:

- Label Smoothing (=0.1)
- Long training
- Train on full dataset
- Scheduler: CosineAnnealingLR (LR $1e-4 \rightarrow 0$)
- Optimizer: AdamW
- ~~• Exponential Moving Average (EMA)~~

Inference:

- PyTorch \rightarrow ONNX \rightarrow OpenVino (fp16=True) [121M \rightarrow 62Mb]
- Resize 256x256 \rightarrow CenterCrop 224x224

Public score: 989044

Private score: 991651

Контакты



Андрей Попов

ML Engineer at NapoleonIT

telegram: *@and_pop_off*

github: *@traptrip*

kaggle: *@traptrip*

https://github.com/traptrip/sibur_2023_video_recognition

Сибур
диджитал

AI COMMUNITY

AI TODAY