

Yandex Cup 2021 ML Challenge

CV & ASR

Kirill Brodt

Week (try) 1. CV Classification without classes

В этой задаче предлагаем вам обучить такую универсальную модель, которая способна решать задачу классификации изображений в режиме zero-shot. Например, по одному только названию класса (без дообучения) модель должна уметь отличать изображения варежек от изображений перчаток, фотографии Москвы от фотографий Казани, фотографии

Week (try) 1. CV Classification without classes

В этой задаче предлагаем вам обучить такую универсальную модель, которая способна решать задачу классификации изображений в режиме **zero-shot**. Например, по одному только названию класса (без дообучения) модель должна уметь отличать изображения варежек от изображений перчаток, фотографии Москвы от фотографий Казани, фотографии

Week (try) 1. CV Classification ~~without classes~~

В этой задаче предлагаем вам обучить такую универсальную модель, которая способна решать задачу классификации изображений в режиме **zero-shot**. Например, по одному только названию класса (без дообучения) модель должна уметь отличать изображения варежек от изображений перчаток, фотографии Москвы от фотографий Казани, фотографии

Week (try) 2. CV Baseline

Week (try) 2. CV Baseline

You do the research. Lightning will do everything else.

The ultimate PyTorch research framework. Scale your models, without the boilerplate.



Week (try) 2. CV Baseline

PyTorch Lightning

You do the research. Lightning will do everything else.

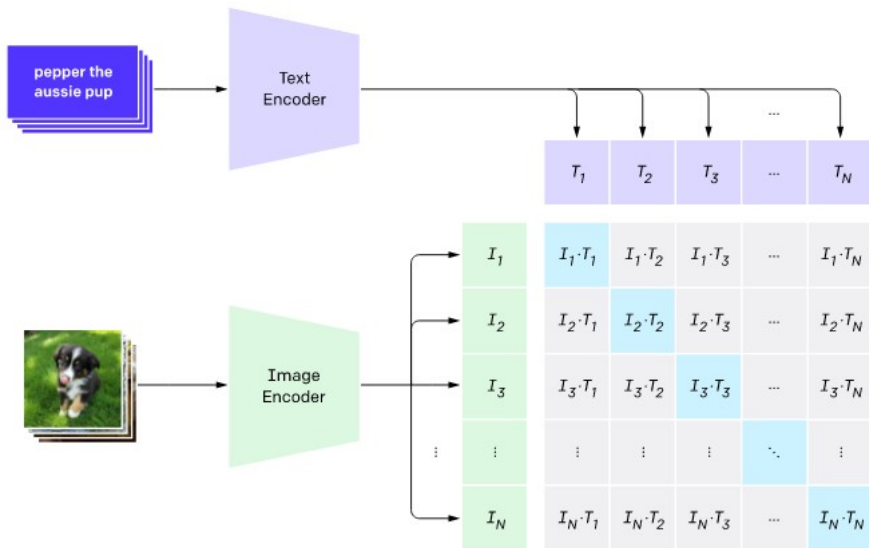
The ultimate PyTorch research framework. Scale your models without the boilerplate.

The logo for PyTorch Lightning, which consists of a white hexagon containing a stylized blue lightning bolt, set against a purple background.

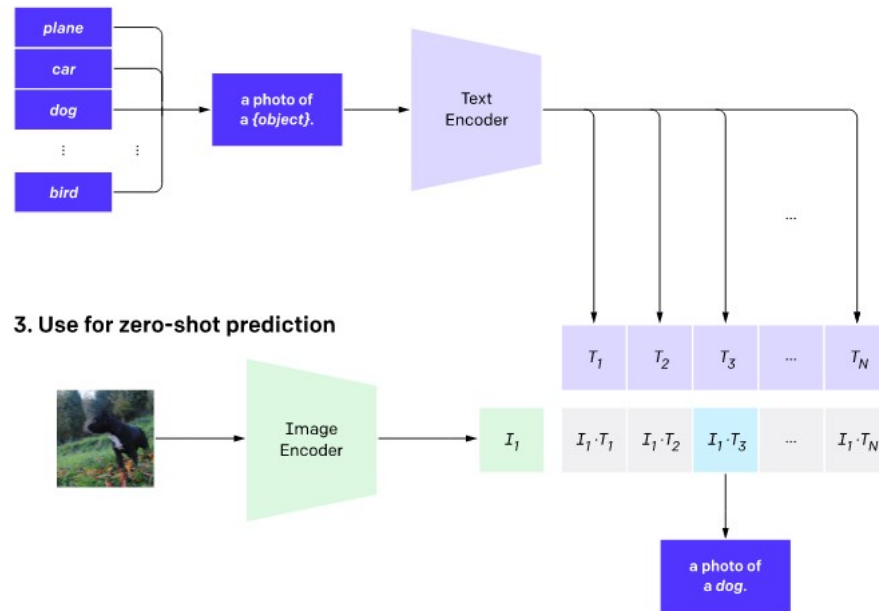
Week (try) 3. CV CLIP & Dataset

Week (try) 3. CV CLIP & Dataset

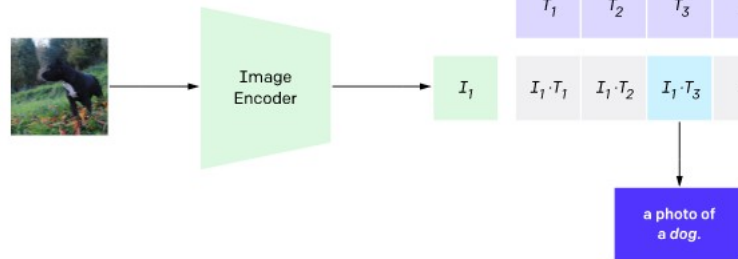
1. Contrastive pre-training



2. Create dataset classifier from label text

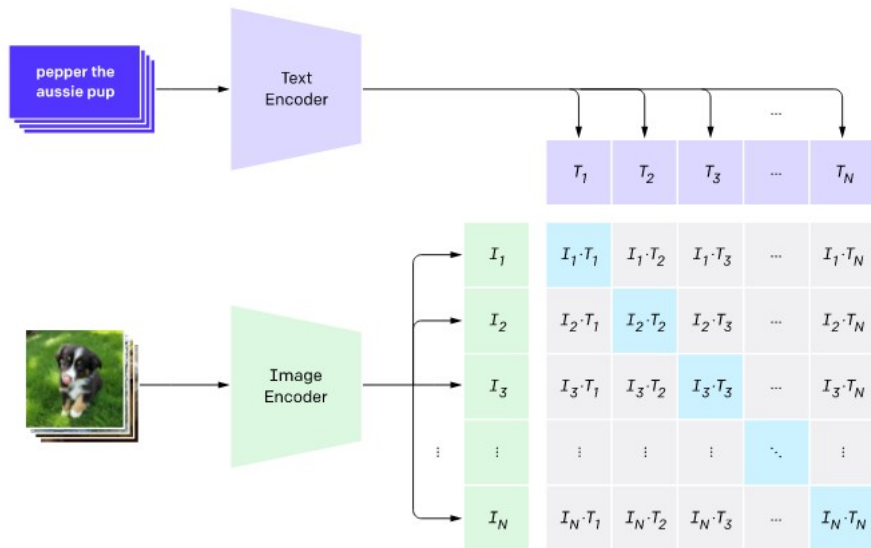


3. Use for zero-shot prediction

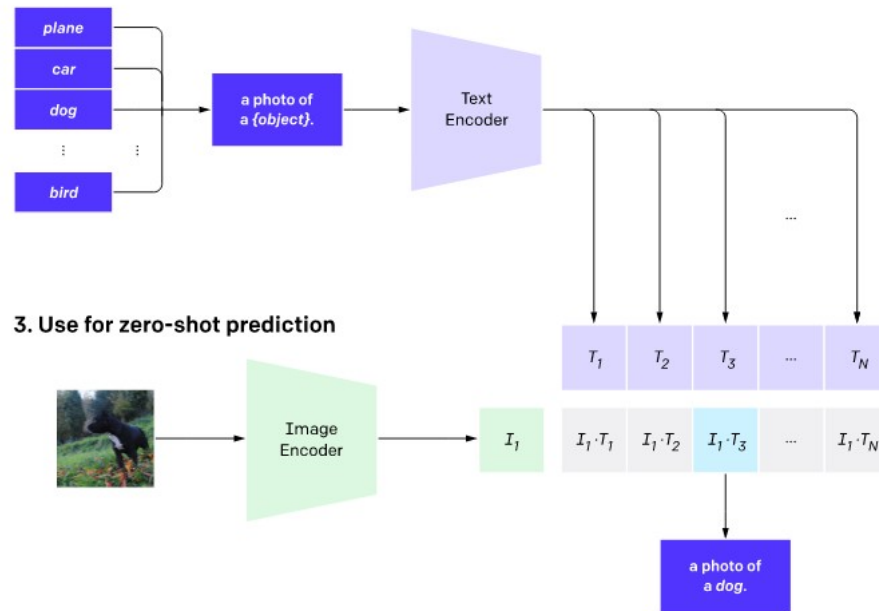


Week (try) 3. CV CLIP & Dataset

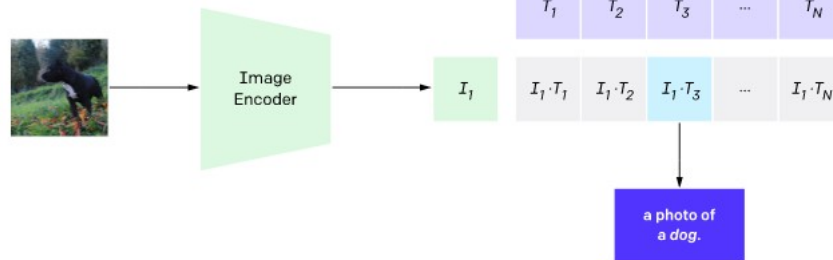
1. Contrastive pre-training



2. Create dataset classifier from label text



3. Use for zero-shot prediction



4.8kk images

Week 4. CV DGX

Week 4. CV DGX

- Baseline+

Week 4. CV DGX

- Baseline+
- Image encoder `ModifiedResnet`

Week 4. CV DGX

- Baseline+
- Image encoder `ModifiedResnet`
- Text encoder 12 layer Transformer

Week 4. CV DGX

- Baseline+
- Image encoder `ModifiedResnet`
- Text encoder 12 layer Transformer
- 1 week on 4 GPUs V100 32Gb from scratch (256 bs)

Week 4. CV DGX

- Baseline+
- Image encoder `ModifiedResnet`
- Text encoder 12 layer Transformer
- 1 week on 4 GPUs V100 32Gb from scratch (256 bs)

| # kbrodt at ltsun in ~ | | | | | | | | | | | | | |
|-------------------------------|------------|-----|---|--|--|--|--|--|--|----------|----------|----------|----------|
| > timew sum :all YandexCup.cv | | | | | | | | | | | | | |
| Wk | Date | Day | Tags | | | | | | | Start | End | Time | Total |
| W39 | 2021-10-02 | Sat | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 20:57:36 | 0:00:00 | 3:02:24 | 3:02:24 |
| W39 | 2021-10-03 | Sun | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 0:00:00 | 1:22:47 | 1:22:47 | |
| | | | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 21:19:22 | 0:00:00 | 2:40:38 | 4:03:25 |
| W40 | 2021-10-04 | Mon | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 0:00:00 | 0:48:19 | 0:48:19 | |
| | | | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 19:53:06 | 0:00:00 | 4:06:54 | 4:55:13 |
| W40 | 2021-10-05 | Tue | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 0:00:00 | 15:18:46 | 15:18:46 | |
| | | | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 20:59:29 | 22:06:31 | 1:07:02 | 16:25:48 |
| W40 | 2021-10-07 | Thu | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 19:45:30 | 20:55:40 | 1:10:10 | 1:10:10 |
| W40 | 2021-10-08 | Fri | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 22:31:24 | 0:00:00 | 1:28:36 | 1:28:36 |
| W40 | 2021-10-09 | Sat | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 0:00:00 | 9:38:33 | 9:38:33 | 9:38:33 |
| W41 | 2021-10-14 | Thu | Train a baseline model for CV, YandexCup.cv, cv, kaggle | | | | | | | 9:38:02 | 10:38:21 | 1:00:19 | 1:00:19 |
| | | | | | | | | | | | | | 41:44:28 |

Week 4. CV DGX

- Baseline+
- Image encoder `ModifiedResnet`
- Text encoder 12 layer Transformer
- 1 week on 4 GPUs V100 32Gb from scratch (256 bs)

[illegible]














Week 4. CV DGX

- Baseline+
- Image encoder `ModifiedResnet`
- Text encoder 12 layer Transformer
- 1 week on 4 GPUs V100 32Gb from scratch (256 bs)

[illegible]

TODOs & Conclusions

- Train on all 5.6kk images
- Use pre-trained ruCLIP models

| № | Участник  я | Private score | Public score |
|---|--|---------------|--------------|
| 1 | Robert Kim  | 91,49 | 93,71 |
|  2 | kirill.brodt  | 87,13 | 91,87 |
| 3 | Андрей Чернов  | 77,79 | 84,77 |
| 4 | madcod  | 72,75 | 84,66 |
| 5 | slavaRED92  | 67,85 | 79,60 |
| 6 | Михаил Едуков  | 58,04 | 68,97 |
| 7 | Anton Chkin  | 57,90 | 70,78 |
| 8 | Артур  | 57,77 | 69,31 |
| 9 | denilv  | 56,49 | 70,29 |
| 10 | alexey.subach  | 56,44 | 68,09 |
| 11 | lacemaker  | 56,35 | 67,18 |

ASR

Задача голосовой активации – распознать фиксированный набор ключевых фраз в аудиопотоке. Основные отличия такой задачи от обычной задачи распознавания реч

ASR. Baseline+

ASR. Baseline+

- MelSpectrogram
(mels=128, fft=2048, hop=251)

ASR. Baseline+

- **MelSpectrogram** (mels=128, fft=2048, hop=251)
- **SpecAug** (40 freq, 60 time) & **MixUP** (uniform mixing)

ASR. Baseline+

- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- efficientnet-b{2..4} & Mixed Precision

ASR. Baseline+

- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- AdamW (lr=1e-3, wd=1e-5)
- efficientnet-b{2..4} & Mixed Precision

ASR. Baseline+

- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- efficientnet-b{2..4} & Mixed Precision
- AdamW (lr=1e-3, wd=1e-5)
- CosineAnnealingLR (T=25)

ASR. Baseline+

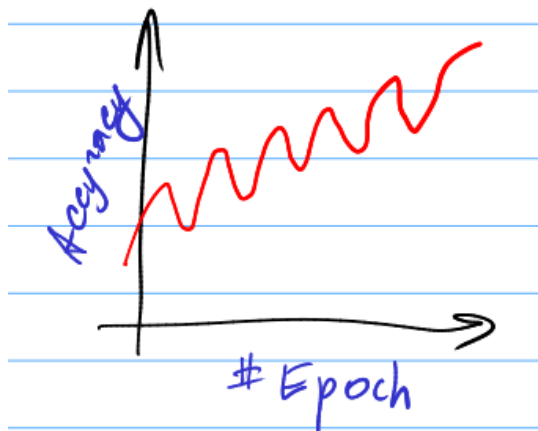
- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- efficientnet-b{2..4} & Mixed Precision
- AdamW (lr=1e-3, wd=1e-5)
- CosineAnnealingLR (T=25)
- 32 batch size / GPU (4 GPUs V100 32Gb)

ASR. Baseline+

- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- efficientnet-b{2..4} & Mixed Precision
- AdamW (lr=1e-3, wd=1e-5)
- CosineAnnealingLR (T=25)
- 32 batch size / GPU (4 GPUs V100 32Gb)
- 2k+ epochs...

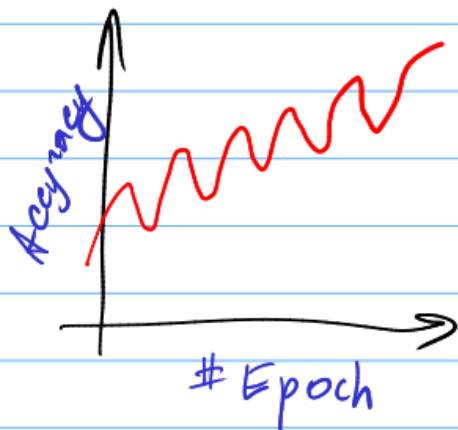
ASR. Baseline+

- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- efficientnet-b{2..4} & Mixed Precision
- AdamW (lr=1e-3, wd=1e-5)
- CosineAnnealingLR (T=25)
- 32 batch size / GPU (4 GPUs V100 32Gb)
- 2k+ epochs...



ASR. Baseline+

- MelSpectrogram (mels=128, fft=2048, hop=251)
- SpecAug (40 freq, 60 time) & MixUP (uniform mixing)
- efficientnet-b{2..4} & Mixed Precision
- AdamW (lr=1e-3, wd=1e-5)
- CosineAnnealingLR (T=25)
- 32 batch size / GPU (4 GPUs V100 32Gb)
- 2k+ epochs...



```
# kbrodrt at ltsun in ~  
> timew sum :all YandexCup.asr
```

| Wk | Date | Day | Tags | | Start | End | Time | Total |
|-----|------------|-----|--|--|----------|----------|---------|---------|
| W39 | 2021-09-27 | Mon | Train a baseline model for ASR, YandexCup.asr, asr, kaggle | | 21:31:04 | 23:37:26 | 2:06:22 | 2:06:22 |
| W39 | 2021-10-01 | Fri | Train a baseline model for ASR, YandexCup.asr, asr, kaggle | | 19:00:43 | 21:04:40 | 2:03:57 | 2:03:57 |
| W40 | 2021-10-09 | Sat | Train a baseline model for ASR, YandexCup.asr, asr, kaggle | | 11:41:17 | 13:29:48 | 1:48:31 | 1:48:31 |
| W40 | 2021-10-10 | Sun | Train a baseline model for ASR, YandexCup.asr, asr, kaggle | | 9:19:18 | 11:32:41 | 2:13:23 | 2:13:23 |
| W41 | 2021-10-13 | Wed | Train a baseline model for ASR, YandexCup.asr, asr, kaggle | | 11:01:55 | 12:01:55 | 1:00:00 | 1:00:00 |
| W42 | 2021-10-20 | Wed | Train a baseline model for ASR, YandexCup.asr, asr, kaggle | | 17:28:44 | 18:01:01 | 0:32:17 | 0:32:17 |

9:44:30

| № | Участник  я | Private score | Public score |
|-----|--|---------------|--------------|
| 1 | slavaRED92  | 96,350 | 96,350 |
| 2 | kirill.brodt  | 95,910 | 95,800 |
| 3 | parkov33@gmail.com  | 92,010 | 92,070 |
| 4 | lyghter  | 91,350 | 91,430 |
| 5 | Александр Мамаев  | 89,270 | 89,120 |
| 6-7 | rcnegby  | 89,180 | 89,430 |
| 6-7 | Andrey P.  | 89,180 | 89,080 |
| 8 | SazerLife  | 88,900 | 88,950 |
| 9 | Tonymit  | 88,660 | 88,730 |
| 10 | Сергей Злобин  | 87,620 | 87,420 |
| 11 | arefiev.mc@gmail.com  | 87,280 | 87,130 |
| 12 | digital.qubit  | 87,150 | 87,220 |
| 13 | dmitriy.weezi  | 86,920 | 86,750 |
| 14 | yar-panda2  | 86,800 | 86,920 |
| 15 | kovtun5  | 83,550 | 83,640 |
| 16 | Truff4ut  | 83,510 | 83,320 |
| 17 | Valery Khanaev  | 83,460 | 83,460 |
| 18 | Victor-Kras2008  | 83,050 | 83,240 |
| 19 | DrVitos  | 82,980 | 83,140 |
| 20 | Yourgospodin  | 82,650 | 83,140 |