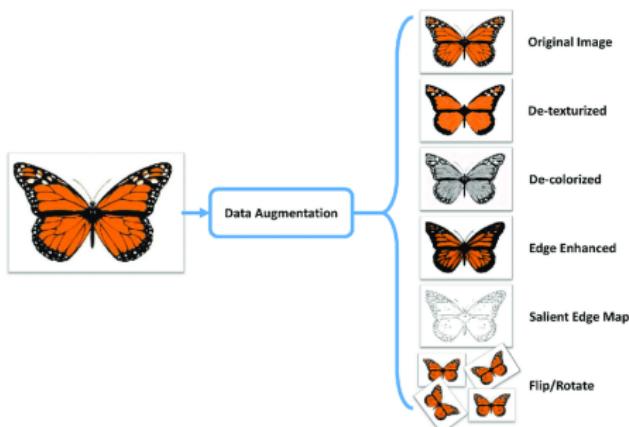


# Расширение обучающей выборки

Виктор Китов  
[victorkitov.github.io](http://victorkitov.github.io)



# Расширение обучающей выборки

Расширение обучающей выборки (data augmentation):

$$\{(x_n, y_n)\}_{n=1}^N \longrightarrow \{(f_\theta(x_m), y_m)\}_{m=1}^M$$

- $f_\theta(x)$  преобразование, сохраняющее отклик
  - может быть комбинация таких преобразований
- $\theta$  случайные параметры преобразования
- $M \gg N$  - расширение #обучающих объектов.

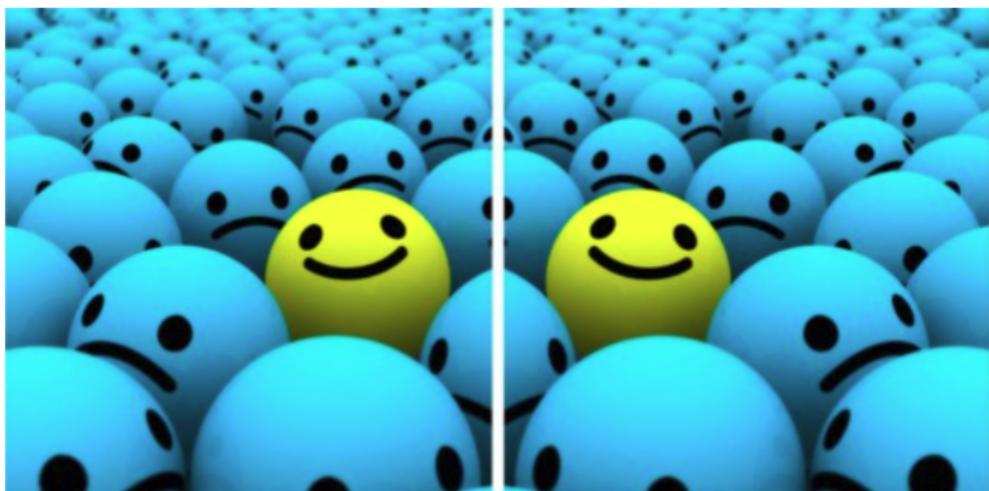
Применения:

- более точная настройка модели
- прогнозы модели будут устойчивы к заданным преобразованиям
- трансферное обучение: например, есть дневные фото, а хотим классифицировать ночные.

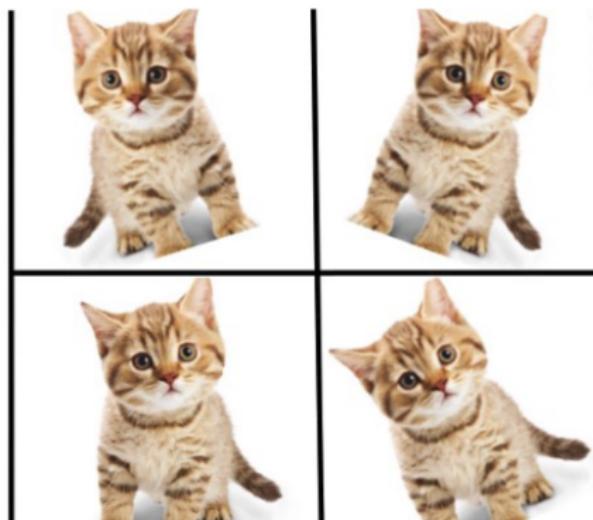
## #объектов и #параметров

|            | VGGNet   | DeepVideo  | GNMT  |
|------------|--|--|---|
| Used For   | Identifying Image Category   | Identifying Video Category   | Translation   |
| Input      | Image<br> | Video<br> | English Text<br> |
| Output     | 1000 Categories  | 47 Categories  | French Text   |
| Parameters | 140M   | ~100M  | 380M  |
| Data Size  | 1.2M Images with assigned Category   | 1.1M Videos with assigned Category   | 6M Sentence Pairs, 340M Words   |
| Dataset    | ILSVRC-2012  | Sports-1M  | WMT'14  |

## Отражение по горизонтали



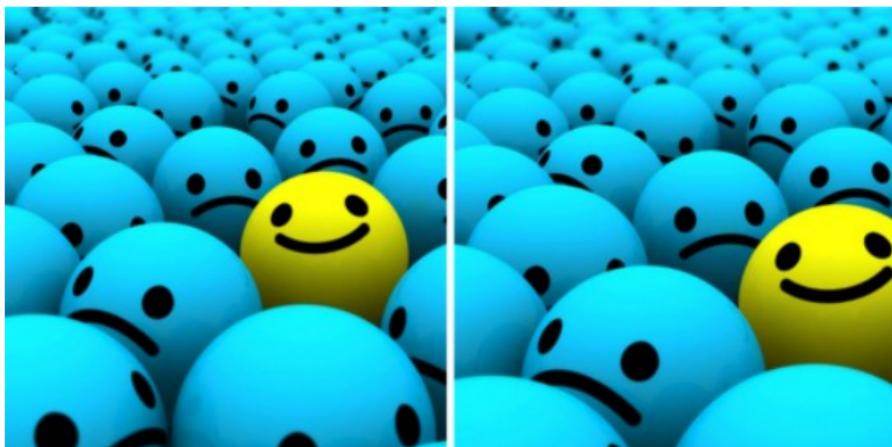
## Повороты



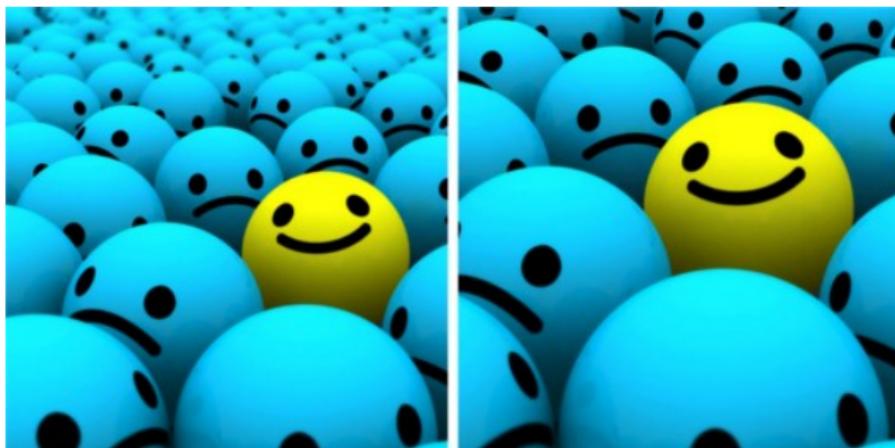
## Смещения



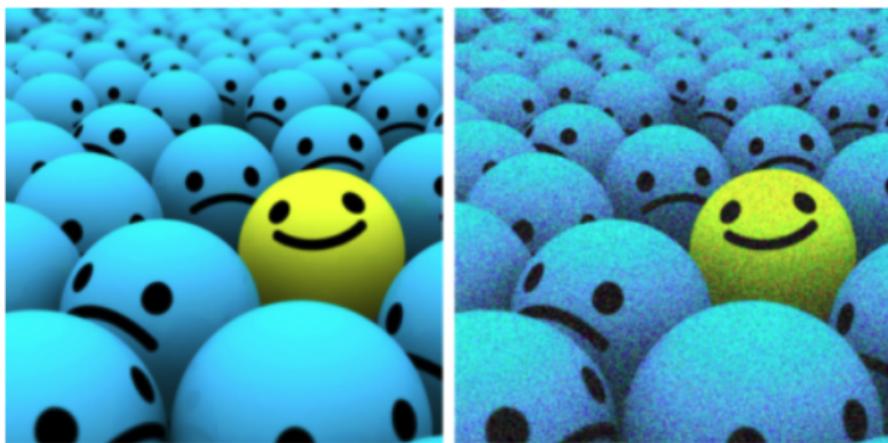
## Перемасштабирование



## Обрезка фрагментов



## Добавление шума (Гуссов, соль&перец)



## Изменение яркости и контраста (brightness, contrast)



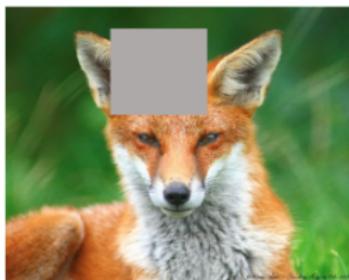
## Изменение насыщенности (saturation)



## Случайные сдвиги RGB каналов (hue)

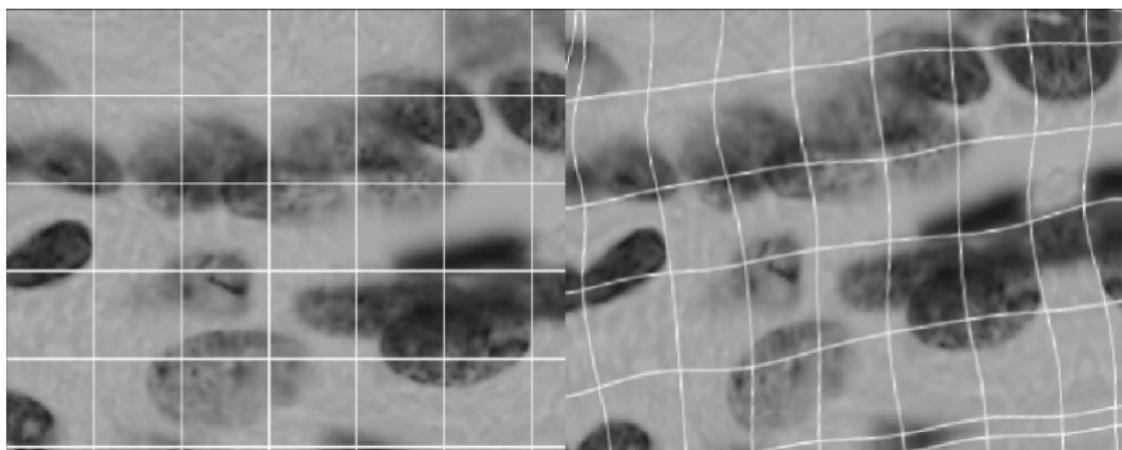


## Затирание фрагмента (константой, шумом)



## Эластичная деформация

- Elastic transform. Часто используется в задачах сегментации в медицине.



## Заключение

- Расширение обучающей выборки (data augmentation) позволяет:
  - точнее настроить модель
  - придать прогнозам модели устойчивость к заданным преобразованиям
  - выполнять трансферное обучение: например, есть дневные фото, а хотим классифицировать ночные.
- Больше вариантов, если комбинировать разные преобразования.
- Сила преобразования не должна быть слишком большой.
  - иначе модель привыкнет к неестественным изображениям
  - при слишком сильном повороте  $b \rightarrow 9$ ,  $p \rightarrow d$  и т.д.
- Существуют свои способы расширение выборки для текстов, звуков, графов и т.д.