

Augmentation

Алексей Латышев, группа М4139

кафедра Компьютерных технологий
факультет Информационных Технологий и Программирования
Университет ИТМО

27 июня 2018



УНИВЕРСИТЕТ ИТМО

Что это? И зачем?

Augmentation — с английского увеличение, прирост. В рамках машинного обучения термин означает «раздувание» датасета для достижения одной из двух целей:

Что это? И зачем?

Augmentation — с английского увеличение, прирост. В рамках машинного обучения термин означает «раздувание» датасета для достижения одной из двух целей:

- Увеличить тренировочную выборку в случае, когда данных очень мало. Например, это используется в медицине в распознавании по картинке какой-нибудь сложной опухоли. Здесь фотографий распознаваемого объекта очень мало, приходится как-то решать эту проблему.



Что это? И зачем?

Augmentation — с английского увеличение, прирост. В рамках машинного обучения термин означает «раздувание» датасета для достижения одной из двух целей:

- Увеличить тренировочную выборку в случае, когда данных очень мало. Например, это используется в медицине в распознавании по картинке какой-нибудь сложной опухоли. Здесь фотографий распознаваемого объекта очень мало, приходится как-то решать эту проблему.
- Увеличить «разнообразность» представленных в тренировочном датасете картинок. В основном, как средство борьбы с переобучением. В работе исследовалась именно эта составляющая.



Что это? И зачем?

Augmentation — с английского увеличение, прирост. В рамках машинного обучения термин означает «раздувание» датасета для достижения одной из двух целей:

- Увеличить тренировочную выборку в случае, когда данных очень мало. Например, это используется в медицине в распознавании по картинке какой-нибудь сложной опухоли. Здесь фотографий распознаваемого объекта очень мало, приходится как-то решать эту проблему.
- Увеличить «разнообразность» представленных в тренировочном датасете картинок. В основном, как средство борьбы с переобучением. В работе исследовалась именно эта составляющая.



Что это? И зачем?

Augmentation — с английского увеличение, прирост. В рамках машинного обучения термин означает «раздувание» датасета для достижения одной из двух целей:

- Увеличить тренировочную выборку в случае, когда данных очень мало. Например, это используется в медицине в распознавании по картинке какой-нибудь сложной опухоли. Здесь фотографий распознаваемого объекта очень мало, приходится как-то решать эту проблему.
- Увеличить «разнообразность» представленных в тренировочном датасете картинок. В основном, как средство борьбы с переобучением. В работе исследовалась именно эта составляющая.

Далее мы будем рассматривать только задачи, входными данными которых являются картинки.



Один из способов это делать — различные «геометрические» преобразования над картинкой:

- 1 Различные аффинные преобразования



Один из способов это делать — различные «геометрические» преобразования над картинкой:

- ❶ Различные аффинные преобразования
 - Повороты



Один из способов это делать — различные «геометрические» преобразования над картинкой:

- 1 Различные аффинные преобразования
 - Повороты
 - Отражения



Один из способов это делать — различные «геометрические» преобразования над картинкой:

- 1 Различные аффинные преобразования
 - Повороты
 - Отражения
 - Сжатия к прямой



Один из способов это делать — различные «геометрические» преобразования над картинкой:

❶ Различные аффинные преобразования

- Повороты
- Отражения
- Сжатия к прямой
- Масштабирование



Один из способов это делать — различные «геометрические» преобразования над картинкой:

- ❶ Различные аффинные преобразования
 - Повороты
 - Отражения
 - Сжатия к прямой
 - Масштабирование
- ❷ Различные проективные преобразования



Один из способов это делать — различные «геометрические» преобразования над картинкой:

- ❶ Различные аффинные преобразования
 - Повороты
 - Отражения
 - Сжатия к прямой
 - Масштабирование
- ❷ Различные проективные преобразования
 - Наклоны



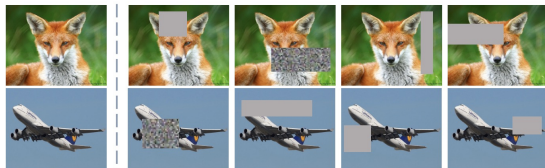
Один из способов это делать — различные «геометрические» преобразования над картинкой:

- ❶ Различные аффинные преобразования
 - Повороты
 - Отражения
 - Сжатия к прямой
 - Масштабирование
- ❷ Различные проективные преобразования
 - Наклоны
- ❸ Обрезка



Как? Random erasing.

Еще один пример Augmentation. С каждой картинкой перед тем как скормить ее на обучение модели с вероятностью p сделают следующее. Выберут случайный прямоугольник на картинке и заменят его на белый шум, либо закрасят серым цветом. Подробнее можно посмотреть в статье <https://arxiv.org/pdf/1708.04896.pdf>



Algorithm 1: Random Erasing Procedure

Input : Input image I ;

Image size W and H ;

Area of image S ;

Erasing probability p ;

Erasing area ratio range s_l and s_h ;

Erasing aspect ratio range r_1 and r_2 .

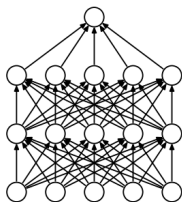
Output: Erased image I^* .

Initialization: $p_1 \leftarrow \text{Rand}(0, 1)$.

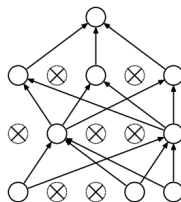
```
1 if  $p_1 \geq p$  then
2    $I^* \leftarrow I$ ;
3   return  $I^*$ .
4 else
5   while True do
6      $S_e \leftarrow \text{Rand}(s_l, s_h) \times S$ ;
7      $r_e \leftarrow \text{Rand}(r_1, r_2)$ ;
8      $H_e \leftarrow \sqrt{S_e \times r_e}$ ,  $W_e \leftarrow \sqrt{\frac{S_e}{r_e}}$ ;
9      $x_e \leftarrow \text{Rand}(0, W)$ ,  $y_e \leftarrow \text{Rand}(0, H)$ ;
10    if  $x_e + W_e \leq W$  and  $y_e + H_e \leq H$  then
11       $I_e \leftarrow (x_e, y_e, x_e + W_e, y_e + H_e)$ ;
12       $I(I_e) \leftarrow \text{Rand}(0, 255)$ ;
13       $I^* \leftarrow I$ ;
14      return  $I^*$ .
15    end
16  end
17 end
```

Как? Dropout.

Это не совсем Augmentation, но рядом. Dropout — рандомное выключение некоторых вершин в слое нейронной сети. А точнее каждая вершина слоя, к которой применен dropout с вероятностью p даст нулевой вклад в следующий слой, на данном этапе обучения. При этом во время теста все вершины уже включены. Подробнее <https://www.cs.toronto.edu/~hinton/absps/JMLRdropout.pdf>



(a) Standard Neural Net



(b) After applying dropout.



Как? Больше нейронных сетей ...

Еще Augmentation можно делать с помощью нейронных сетей и обучать вместе с моделью. Получается неплохо.

① GAN

② AugNet предложенный в

<http://cs231n.stanford.edu/reports/2017/pdfs/300.pdf>

Augmentation Network

1. Conv with 16 channels and 3x3 filters. Relu activations.
2. Conv with 16 channels and 3x3 filters. Relu activations.
3. Conv with 16 channels and 3x3 filters. Relu activations.
4. Conv with 16 channels and 3x3 filters. Relu activations.
5. Conv with 3 channels and 3x3 filters.

1



УНИВЕРСИТЕТ ИТМО

Что я делал?

Основная цель проверить, как влияет Augmentation на точность обучения. Для этого рассмотрим несколько датасетов и несколько архитектур нейронных сетей. Обучим, применяя и не применяя различные методы Augmentation. Сравним точность и итоговый средний loss.



- ① MNIST
- ② FashionMNIST
- ③ CIFAR10



Augmentation		Simple		Conv		LeNet	
MNIST	without	0.1140	96.80%	0.0490	98.39%	?	
	ROT(10)	0.1086	96.89%	0.0498	98.46%	?	
Fashion	without	0.3874	86.40%	0.4030	84.95%	?	
	ROT(10)	0.3741	86.60%	0.4308	83.74%	?	
CIFAR10	without	1.4939	47.85%	1.3542	52.98%	1.1091	61.86%
	CROP(32, pad=4) HFLIP	1.5120	46.24%	1.4643	48.03%	1.1627	58.99%

