

ПРОСТРАНСТВЕННОЕ УВЕЛИЧЕНИЕ РАЗРЕШЕНИЯ ИЗОБРАЖЕНИЙ ПРИ ПОМОЩИ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНОЙ СЕТИ

Монгуш Т.В.¹, Куулар Ч.Т.²

¹ Томский политехнический университет, ИШИТР, студент гр. 8ВМ13, e-mail: tvn11@tpu.ru,

² Томский политехнический университет, ИШИТР, студент гр. 8ВМ13, e-mail: ctk1@tpu.ru

Введение

Целью работы является реализация генеративно-состязательной сети по реализации пространственного увеличения разрешения изображения.

Проблема увеличения разрешающей способности изображений из изображения более низкого качества изучается в рамках задачи суперразрешения. Это сложная обратная задача, так как существует большое количество возможных отображений изображения высокого качества в низкое.

К настоящему времени было предложено большое количество методов, позволяющих получить изображение высокого разрешения из изображения с низким разрешением. Данная проблема активно развивается с момента появления основополагающей работы Цая и Хуанга в 1984 году [1], за последние три десятилетия появилось большое количество методов, для повышения разрешения, начиная от методов обработки сигналов, заканчивая методами машинного обучения.

Описание алгоритма

В данной работе рассматривается увеличение разрешение изображения при помощи генеративно-состязательной сети (GAN – Generative adversarial network).

GAN – алгоритм машинного обучения без учителя, построенный на комбинации их двух нейронных сетей, одна из которых (G-generator) генерирует образцы, а вторая – (D-discriminator) старается отличить правильные (исходные) образцы от образцов генератора.

Модель генератора

Работа сети генератора заключается в том, что на вход подаётся изображение низкого разрешения, после чего он учится генерировать изображение высокого разрешения.

Генератор – свёрточная нейронная сеть (рис. 1). Генеративная сеть состоит из слоя свёртки с ядром 3×3 , 64-мя картами признаков, и с функцией активации PReLU, далее находятся 16 остаточных блоков с идентичной компоновкой. В каждом блоке находятся два свёрточных слоя с небольшими ядрами 3×3 и 64-мя картами признаков, за которыми расположены слои пакетной нормализации. В качестве функции активации используется PReLU. Далее идёт снова слой свёртки, после чего изображение увеличивается попиксельно с помощью двух свёрточных слоёв. Последним идёт слой свёртки с тангенциальной функцией активации.

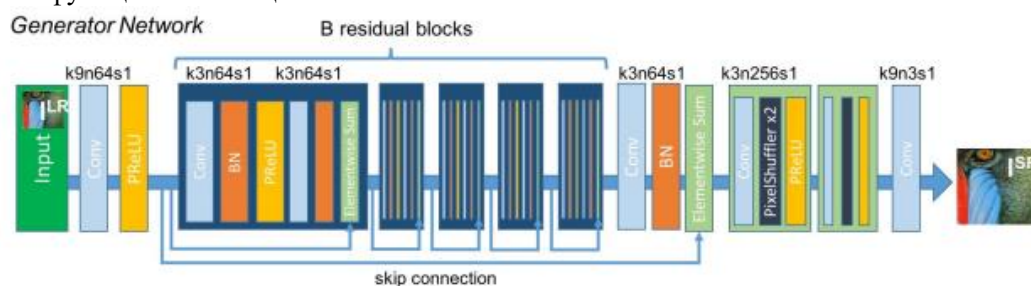


Рис. 1. Модель генератора

Модель дискриминатора

На вход дискриминатора подаётся изображение и метка, фейковое или реальное изображение. После чего дискриминатор при помощи свёртки пытается определить: реальное ли это изображение или сгенерированное.

Дискриминатор – свёрточная нейронная сеть, которая учится распознавать исходное и сгенерированное изображения. Сеть-дискриминатор содержит 8 свёрточных слоёв с возрастающим числом ядер фильтра 3×3 (с 64 до 512, каждый раз увеличиваясь в 2 раза, как в VGG сети). Пошаговые свёртки

используются для уменьшения разрешения изображения каждый раз, когда число признаков удваивается.

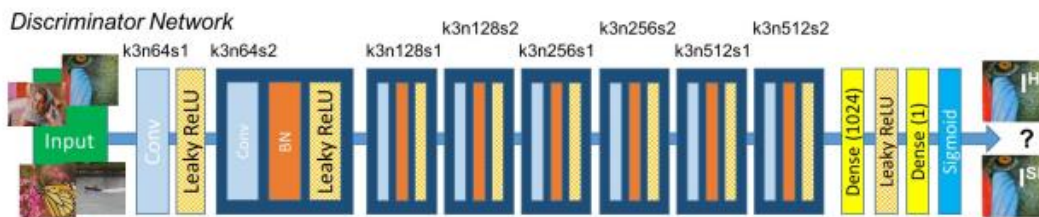


Рис. 2. Модель дискриминатора

Функция ошибки

Определение функции ошибки восприятия играет очень важную роль для производительности сети генератора, поэтому в данной работе функция потери восприятия определена как взвешенная сумма потерь содержания и состязательных потерь.

Потери содержания определяются на основе слоёв активации ReLU предварительно обученной VGG сети и определяются как евклидово расстояние между признаками восстановленного и опорного изображений.

Таким образом сеть отдаёт предпочтение более естественным изображениям, пытаясь обмануть дискриминатор [2].

Эксперименты

Эксперименты проводились на ноутбуке Lenovo Legion 5 15ach6. Характеристики устройства приведены в таблице ниже.

Таблица 1

Характеристики ноутбука Lenovo Legion 5 15ach6

Комплектующее	Модель	Характеристика
Видеокарта	Nvidia RTX 3060 Laptop	6 ГБ
Процессор	AMD Ryzen 7 5800H	8 ядер, 16 потоков
Оперативная память	Kingston ValueRAM	32 ГБ

В качестве данных были взяты данные DIV2K. DIV2K [3] — это популярный набор данных с одним изображением сверхвысокого разрешения, который содержит 1000 изображений с различными сценами и разделен на 800 для обучения.

Результаты

Обучение проводилось на 800 изображениях высокого разрешения, с параметром скорости обучения 0.0001, обучение проведено для 400 эпох, на каждую эпоху требуется 2 минуты (13-14 часов). На Рис. 3, 4 приведены результаты работы модели, real(hr) – исходное изображение высокого разрешения, lr – изображение низкого разрешения, gen(hr) – сгенерированное изображение.



Рис. 3. Результат работы сети. 300 эпоха [disc_loss=5.28e-6, gen_loss=0.802]



Рис. 4. Результат работы сети. 399 эпоха [$disc_loss=4.05e-6$, $gen_loss=0.79$]

По результатам сети можно сделать вывод, что генератор сумел научиться создавать изображения, но при этом качество изображения ещё далеко от оригинала.

Таблица 2

Результаты обучения модели генеративно состязательной сети

Эпоха	Ошибка дискриминатора	Ошибка генератора
0	0,0317	1,34
99	$6,98 \cdot 10(-5)$	0,972
199	$2,09 \cdot 10(-5)$	0,887
299	$9,66 \cdot 10(-6)$	0,82
399	$4,05 \cdot 10(-6)$	0,79

Заключение

В работе была реализована собственная модель для повышения разрешения изображения при помощи генеративно-состязательной сети. Но результаты показали, что, генеративно-состязательных сетей есть один существенный недостаток – их очень сложно обучать.

К тому же полученные результаты относительно исходных изображений низкого разрешения получились более детальными, но на сгенерированных изображениях прослеживаются различные артефакты (зернистость). Для устранения данных недостатков, возможно, требуется большее число эпох обучения, либо произвести процедуру сглаживания полученного изображения при помощи фильтров Гаусса.

Список использованных источников

1. R. Y. Tsai and T. S. Huang. Multipleframe image restoration and registration. In Advances in Computer Vision and Image Processing. – Greenwich, CT, 1984. – 317–339.
2. Повышение разрешения изображения с помощью AI. [Электронный ресурс]. – URL: <https://www.reg.ru/blog/increase-image-resolution-using-ai/>
3. DIV2K dataset: DIVerse 2K resolution high quality images as used for the challenges NTIRE. [Электронный ресурс]. – <https://data.vision.ee.ethz.ch/cvl/DIV2K/>