

УДК 004.67

**ГЕНЕРАЦИЯ ИЗОБРАЖЕНИЙ ПО ТЕКСТОВОМУ ОПИСАНИЮ
ATTRIBUTES2IMAGE И DCGAN**



**GENERATION OF IMAGES ACCORDING TO TEXT DESCRIPTION
ATTRIBUTES2IMAGE AND DCGAN**

Тотухов Константин Евгеньевич

кандидат технических наук,
доцент кафедры «Информационных систем
и программирования»,
Кубанский государственный технологический университет
101ke@mail.ru

Климов Михаил Сергеевич

студент,
Кубанский государственный технологический университет
mklim2090@gmail.com

Аннотация. В данной статье представлен обзор некоторых моделей генерации изображений по текстовому описанию. В статье рассматриваются их архитектуры и краткое описание работы.

Ключевые слова: нейронная сеть, сверточные нейронные сети, атрибут, GAN.

Totuhov Konstantin Evgenievich

Candidate of Technical Sciences,
Associate Professor of the Department
of Information Systems and Programming,
Kuban State Technological University
101ke@mail.ru

Klimov Mikhail Sergeevich

Student,
Kuban State Technological University
mklim2090@gmail.com

Annotation. This article provides reviews of some models that create images from a text description. The article discusses their architecture and a brief description of the work.

Keywords: neural network, convolutional neural networks, attributes, GAN.

Нейронная сеть – это математическая модель, построенная по принципу биологических нейронных сетей – живых клеток организма, представляющая собой систему объединенных и взаимодействующих между собой искусственных нейронов [1]. Объединенные в большие сети, данные простые нейроны способны выполнять очень сложные для традиционных вычислительных систем задачи, такие как:

- Психодиагностика [2].
- Распознавание образов [3].
- Широкий спектр экономических задач [4].
- Предугадывание погодных условий [5].

Преобразование описания изображения на естественном языке в изображение является хорошей демонстрацией глубокого обучения. Задачи по классификации текста, такие как смысловой анализ текста успешно решаются глубокими рекуррентными нейросетями, которые способны изучать векторные представления текстовой строки.

Условная генерация изображений из визуальных атрибутов (англ. Conditional Image Generation from Visual Attributes, Attribute2Image [6]) – этот вид генеративной нейросети представляет генерируемое изображение как сочетание двух составляющих – переднего и заднего планов, тем самым получая генеративную модель со скрытыми переменными со множеством уровней. На рисунке 1 изображена архитектура Attribute2Image. Атрибут, в данном случае, это текстовое описание качеств генерируемого объекта, такие как: цвет, форма, черты лица и др.

Таким образом, данная генеративная модель способна синтезировать убедительные и реалистичные вариации исходного изображения (см. рис. 2).

Глубокая сверточная генеративная состязательная сеть (англ. Deep Convolutional Generative Adversarial Network, DCGAN) – одна из вариаций GAN – алгоритма машинного обучения на основе исходных данных, которая на их основе способна генерировать новые данные. GAN был создан Яном Гудфеллоу в 2014 году [7]. Сеть называется «состязательной» потому, что она состоит из двух частей – той которая отвечает за генерацию изображений и той, которая отбирает наиболее правдоподобные изображения и отбрасывает неудачные.

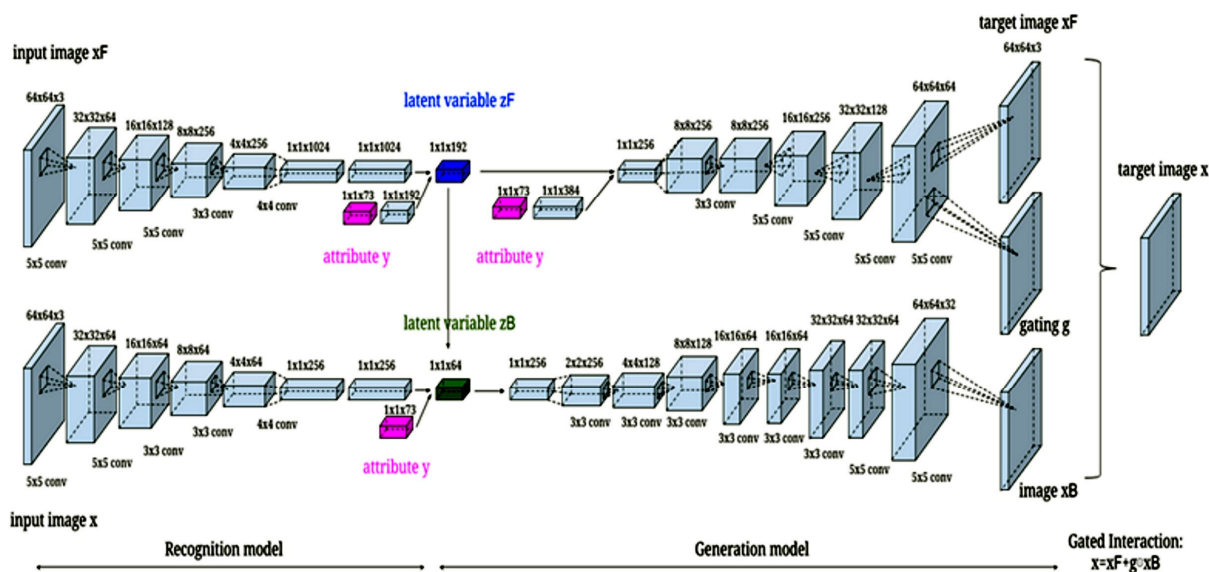


Рисунок 1 – Архитектура Attribute2Image

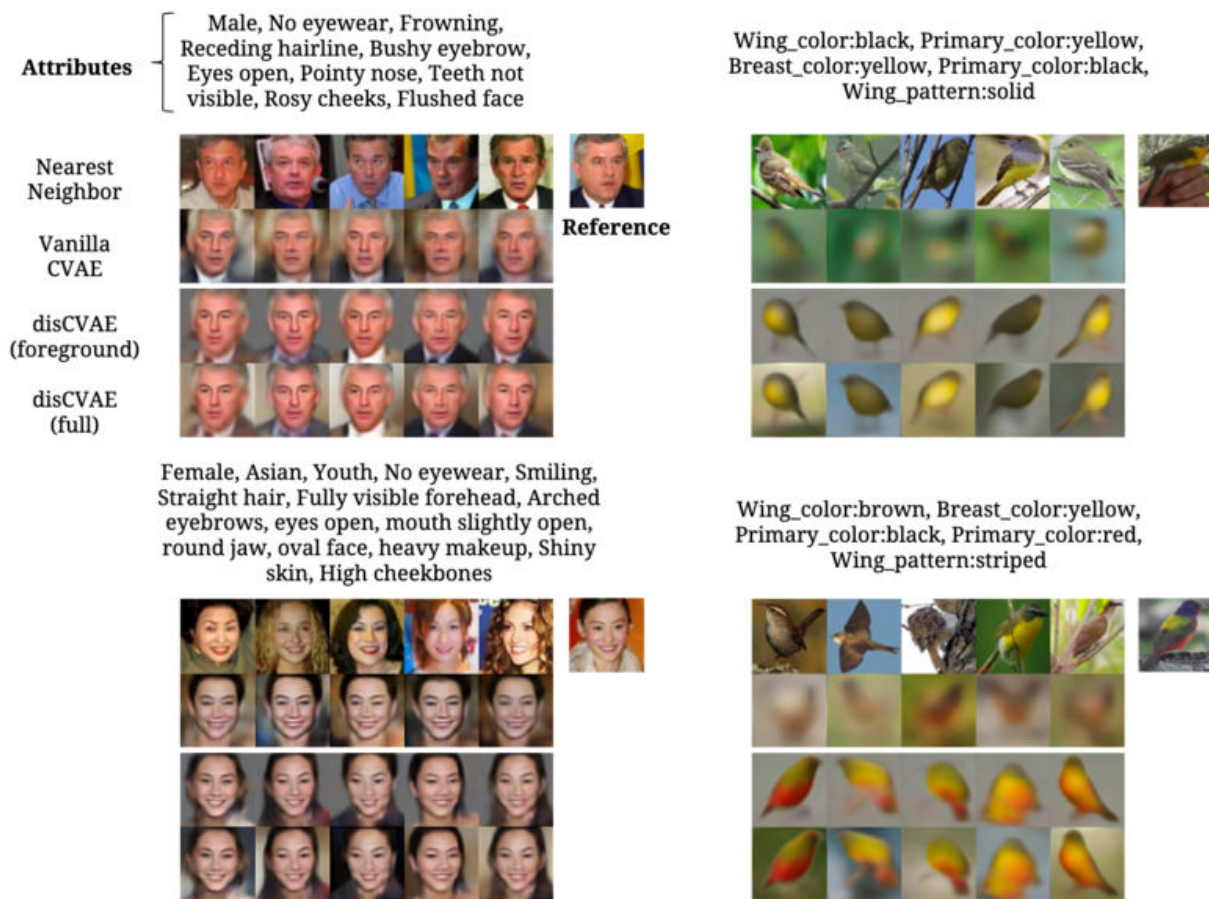


Рисунок 2 – Пример результата работы Attribute2Image

DCGAN, в свою очередь, обуславливается текстовым описанием, кодируемым гибридной сверточно-рекуррентной нейронной сетью. Во многих случаях эта сеть способна генерировать изображения размером 64 на 64 пикселя. Архитектура DCGAN изображена на рисунке 3.

Помимо реальных изображений и поддельных, на вход нейросети подаются реальные изображения, не совпадающие с текстовым описанием, благодаря которым сеть обучается отличать подходящие изображения от неподходящих [8]. На рисунке 4 изображен результат работы нейросети.

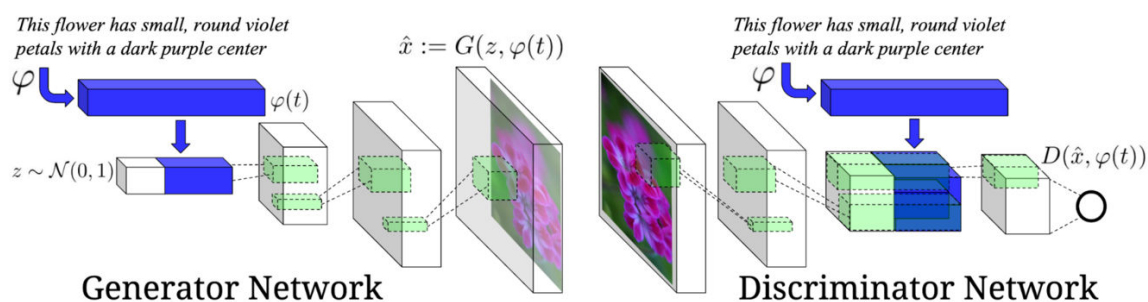


Рисунок 3 – Архитектура DCGAN



Рисунок 4 – Результат работы DCGAN

Литература

1. Нейронная сеть. Большая российская энциклопедия: [в 35 т.] / гл. ред. Ю.С. Осипов. – М. : Большая российская энциклопедия, 2004–2017.
2. Горбан А.Н., Россиев Д.А., Доррер М.Г. МультиНейрон – Нейросетевая симуляция для медицинского и применения в психологии // Wcnn'95, Washington, D.C.: World Congress on Neural Networks, 1995.
3. Лекан Я., Бенгио Я. Сверточные нейронные сети для изображений, речи и временных рядов. The Handbook of Brain Theory and Neural Networks, MIT Press, 1995
4. Калацкая Л.В., Новиков В.А., Садов В.С. Организация и обучение искусственных нейронных сетей: Экспериментальное учеб. пособие. – Минск : Изд-во БГУ, 2003.
5. Яндекс анонсирует собственную технологию прогнозирования погоды Метеум. С точностью до дома. – URL : habr.com
6. Attribute2Image: Условная генерация изображений из визуальных атрибутов / Я. Синьчен [и др.].
7. Генеративные состязательные сети / Г. Ян [и др.]. – Университет Монреаль.
8. Скотт Р. Генеративно-состязательный синтез текста в изображения.

References

1. Neural network. Big Russian Encyclopedia: [in 35 vols] / ed. by Yu.S. Osipov. – M. : Big Russian Encyclopedia, 2004-2017.
2. Gorban A.N., Rossiev D.A., Dorrer M.G. Multi-Neuron – Neural Network Simulation for Medical and Psychological Applications // Wcnn'95, Washington, D.C. : World Congress on Neural Networks, 1995.
3. Lecan J., Bengio J. Convergent neural networks for images, speech, and time series. The Handbook of Brain Theory and Neural Networks, MIT Press, 1995.
4. Kalatskaya L.V., Novikov V.A., Sadov V.S. Organization and training of artificial neural networks: Experimental tutorial. – Minsk : Publishing house of Belarusian State University, 2003.
5. Yandex announces its own weather forecasting technology Meteum. With an accuracy to the house. – URL : habr.com.
6. Attribute2Image: Conditional image generation from visual attributes / Y. Xincheng [et al.]
7. Generative adversarial networks / G. Yang [et al.]. – University of Montreal.
8. Scott R. Generative-adversarial synthesis of text into images.