

Реализация алгоритма генерации изображений по сигналам головного мозга

Егор Мирошниченко
группа 16.Б10

Научный руководитель: д.ф.- м.н., проф. А.Н. Терехов

Научный консультант: ст.преп. А.Р.Ханов

Рецензент: к.т.н. А.А.Воробьева

спбгу

16 июня 2020 г.

Добрый день. Меня зовут Егор Мирошниченко, и я расскажу о своей работе "Реализация алгоритма генерации изображений по сигналам головного мозга".

- ▶ Активное развитие в области изучения работы головного мозга
- ▶ Улучшение и распространение устройств для сканирования мозга
- ▶ Появление удачных генеративных алгоритмов машинного обучения
- ▶ Успешные результаты авторов статьи Brain2Image

Изучение работы мозга является сложной задачей, которая привлекает внимание исследователей из разных областей. Последние исследования в области нейрофизиологии доказывают, что процессы мозговой активности могут быть выявлены из данных, полученных с помощью функциональной магнитно-резонансной томографии, *электроэнцефалографии (ЭЭГ)*, магнитоэнцефалографии.

А распространение и удешевление устройств для считывания этих сигналов за пределами стен исследовательских и медицинских центров способствует привлечению энтузиастов, что ускоряет развитие области. Одним из таких устройств является нейроинтерфейс Emotiv EPOC+, позволяющий снимать изменения электрических потенциалов на поверхности головы, которые называют электроэнцефалограммами.

Введение

- ▶ Активное развитие в области изучения работы головного мозга
- ▶ Улучшение и распространение устройств для сканирования мозга
- ▶ Появление удачных генеративных алгоритмов машинного обучения
- ▶ Успешные результаты авторов статьи Brain2Image

Уже существуют исследования о том, что сигналы головного мозга были использованы для обучения интеллектуальных систем для классификации эмоций, медицинских задач и т. п. И если эти задачи подразумевают выделение признаков активностей некоторых зон мозга, то задача визуального взаимодействия сильно сложнее и менее изучена. Однако существуют исследования доказывающие, что активность головного мозга содержит в себе конкретные образы, в зависимости от классов объектов, на которые смотрит человек.

- ▶ Активное развитие в области изучения работы головного мозга
- ▶ Улучшение и распространение устройств для сканирования мозга
- ▶ Появление удачных генеративных алгоритмов машинного обучения
- ▶ Успешные результаты авторов статьи Brain2Image ¹

Параллельно с этим быстро развивается область алгоритмов машинного обучения. В последние несколько лет появилось много новых подходов, в частности, появились алгоритмы, генерирующие изображения из некоторого скрытого пространства признаков. Хорошие результаты в этом направлении демонстрируют разнообразные автоэнкодеры и генеративно состязательные сети (GAN), которые находят применение в разных местах: от переноса стиля с одного изображения на другое, до подмены лиц на фото и видео.

Более того, исследователи из Италии разработали алгоритм Brain2Image, позволяющий по электроэнцефалограмме генерировать изображение объекта, на который смотрел человек во время считывания.

¹Kavasidis et al., Brain2Image: Converting Brain Signals into Images, 2017

Цель

Реализовать алгоритм генерации изображений по электроэнцефалограмме, полученной с помощью устройства Emotiv EPOC+.



Целью данной работы является реализация алгоритма генерации изображений из ЭЭГ, считанных с помощью нейроинтерфейса Emotiv EPOC+.

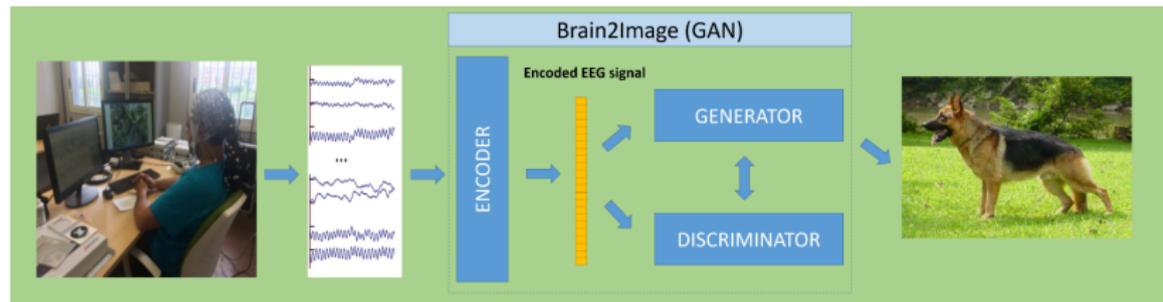
Задачи

- ▶ Проанализировать алгоритм Brain2Image
- ▶ Собрать необходимые данные для обучения алгоритма с помощью шлема Emotiv EPOC+
- ▶ Реализовать вариант алгоритма, использующий GAN

Для достижения этой цели были поставлены следующие задачи:

- Проанализировать алгоритм Brain2Image
- Собрать необходимые данные для обучения алгоритма с помощью шлема Emotiv EPOC+.
- Реализовать вариант алгоритма, использующий генеративно-состязательную сеть

Обзор решения



Общая схема алгоритма (из оригинальной статьи)

- ▶ Большое количество данных: 2000 изображений из 20 классов
- ▶ Высокоточный шлем для снятия ЭЭГ, 128 датчиков, частота дискретизации 1000 Гц

Общая схема алгоритма Brain2Image представлена на рисунке. С человека во время просмотра определенных изображений считывают ээг, далее по этим записям алгоритм пробует восстановить исходное изображение. Алгоритм разделяется на две части, первая энкодер выделяет из сырых данных ээг вектор признаков, в котором отражена информация о визуальных стимулах человека. Затем этот вектор передается в во вторую часть алгоритма — декодер, который из этих признаков пробует восстановить исходное изображение.

Стоит заметить, что авторами использовались большое количество данных 2000 изображений из 20 классов и высокоточный шлем с 128 датчиками и частотой дискретизации 1000 гц.

Обзор решения



(b) Jack-o'-Lantern

Из оригинальной статьи

- ▶ Сгенерированные изображения из отрезка ЭЭГ

Как можно видеть, на примере полученного ими изображения тыквы, результаты получились весьма многообещающими.

В связи с этим и была поставлена цель реализовать подобный алгоритм на более простом оборудовании.

Сбор данных

- ▶ Видеоряд из 350 изображений в 5 классах из базы ImageNet, 70 на класс
- ▶ Изображения из одного класса по 0.6 сек непрерывно, между классами нейтральное изображение 10 сек
- ▶ Считывание ЭЭГ во время просмотра изображений



Самолет



Рыбка клоун



Лошадь



Панда

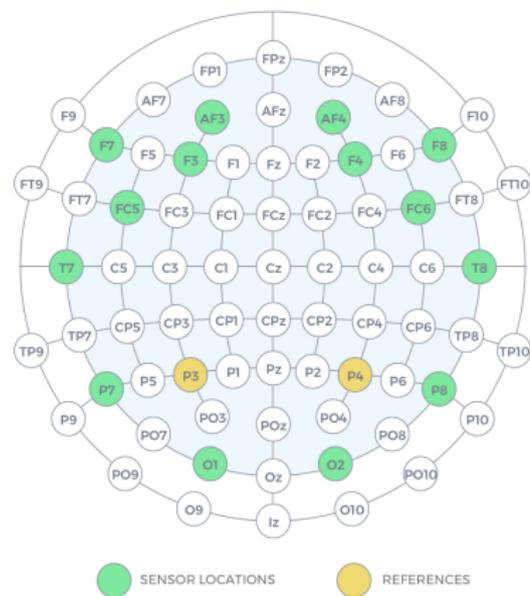


Тыква

Начинать подобную работу стоило со сбора необходимых данных. Для этого в базе ImageNet были отобраны по 70 изображений для 5 классов: самолет, рыба клоун, лошадь, панда, тыква, всего 350 изображений. из которых был составлен следующий видеоряд: все изображения одного класса показывались непрерывно по 0.6 сек на изображение, между классами показывалось нейтральное изображение в течение 10 секунд для отвлечения внимания от прошлого класса.

Для этого были найдены 5 добровольцев и организованы сессии записи, где они по очереди просматривали этот видеоряд во время которого, производилась запись данные ээг с нейрошлема.

Emotive EPOC+



- ▶ 14 датчиков
- ▶ Частота дискретизации 128/256 Гц

Emotiv EPOC+ представляет из себя устройство в виде шлема, с 14-ю датчиками прилегающими к коже головы, с которых с заданной частотой дискретизации 128 или 256 Гц считываются данные и передаются по беспроводному каналу приемнику. Режим работы был выбран 256 Гц для увеличения количества полученной информации.

Предобработка данных

- ▶ Сегментация данных по временным отметкам показываемых изображений
- ▶ Удаление 40 мс от начала каждого сегмента для уменьшения влияния предыдущего изображения
- ▶ Укорачивание сегментов по длине наименьшего — 120 записей
- ▶ После удаления поврежденных элементов осталось 2358 массива размером 120x14
- ▶ Стандартизация вдоль каналов

Непрерывные серии ЭЭГ были разделены на сегменты по временным отметкам показываемых изображений. Далее были удалены 40мс начала каждого сегмента для уменьшения влияния предыдущего изображения. Все сегменты выровнены по длине 120 записей. После удаления поврежденных частей получилось 2358 массивов размером 120 записей на 14 каналов. И вдоль каналов была проведена стандартизация значений — вычитание среднего и деление на стандартное отклонение.

Энкодер

- ▶ После апробации разных моделей, лучший результат был получен двуслойной LSTM моделью с 56 узлами в каждом
- ▶ Список моделей:

Структура	Accuracy
2xLSTM(56)	45.97%
3xLSTM(28)	39.86%
LSTM(56) + Dense(14)	43.01%
Модели без LSTM	<36.00%

Задача энкодера — выделение из входного отрезка ЭЭГ признаков, содержащих визуальную информацию. Для этого сперва строится классификатор, определяющий на какой класс изображений смотрел человек во время снятия ЭЭГ. Далее из него выделяется энкодер путем удаления последнего слоя.

Были опробованы разные модели и комбинации, включая слои LSTM узлов, полносвязные нейронные сети, SVM. Наилучшие результаты показали модели, использовавшие LSTM. Точность некоторых из опробованных классификаторов показана в таблице.

Энкодер

- ▶ Получившийся классификатор показал точность 45.97% на тестовом множестве. Точность по каждому классу приведена в таблице:

Класс	Ассурасу
Самолет	43.34%
Лошадь	45.19%
Тыква	48.52%
Панда	53.13%
Рыбка клоун	39.60%

Лучший результат показал классификатор состоящий из двух LSTM слоев, по 56 узлов в каждом, с добавленным обычным слоем из 5 нейронов для финальной классификации. Он показал точность 45.97% на тестовом множестве. Точность по классам приведена в таблице. Из него был удален последний слой и оставшаяся модель, которая на каждый кусок ЭЭГ выдает вектор признаков длины 56, взята как энкодер.

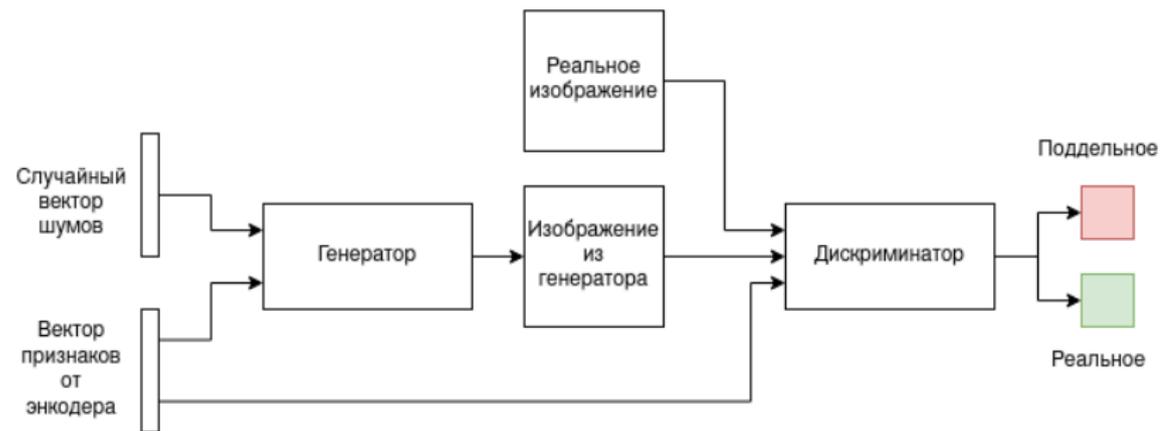
Декодер

- ▶ Генерирует изображение из пространства признаков, выделенных энкодером
- ▶ Для реализации выбрана генеративно-состязательная сеть (GAN)
- ▶ Сложно обучаются, требовательны к данным, поэтому:
 - ▶ Целевой размер изображений был выбран 32x32 пикселя
 - ▶ Было проведено увеличения числа изображений посредством модификации исходных

Следующим шагом стало создание декодера-генератора, который из векторов полученных от энкодера генерирует финальное изображение. Для этих целей была выбрана генеративно-состязательная сеть.

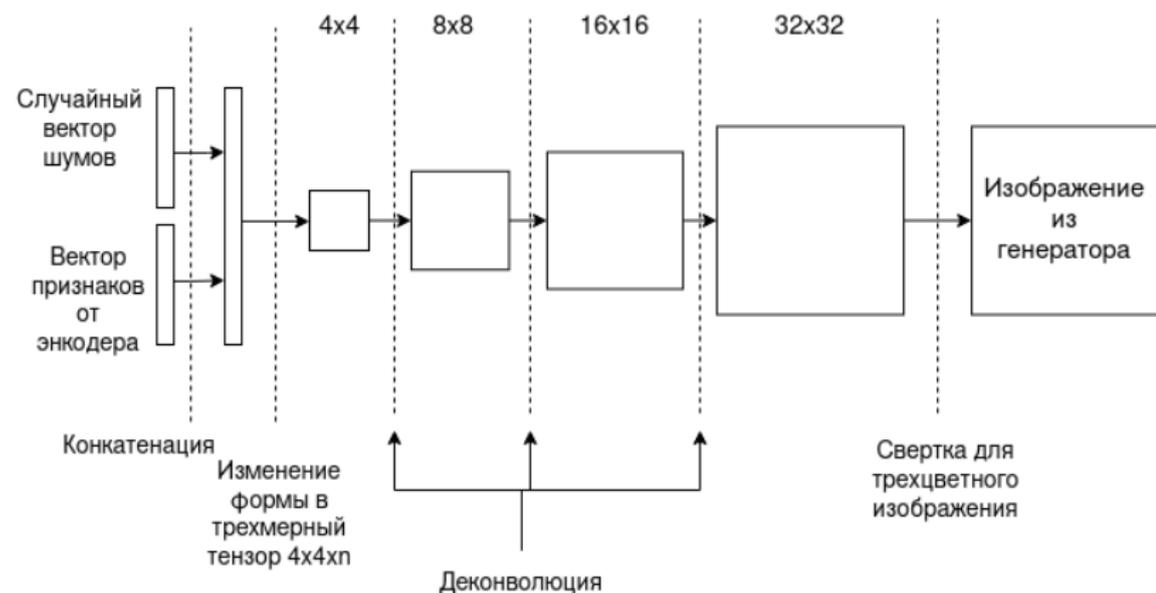
Она сложно обучается и требовательна к объему данных. Поэтому целевой размер изображений был выбран 32x32 пикселей, а база картинок была раздута с помощью разных искажений исходных.

Общая схема cGAN



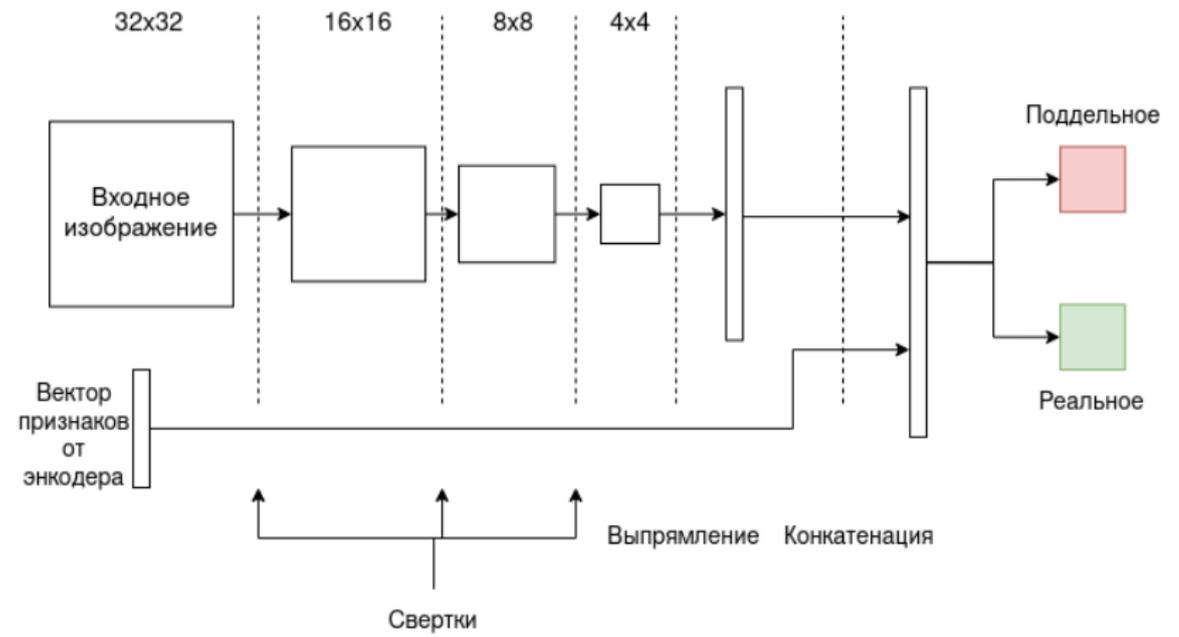
Общая схема GAN изображена на схеме. Она состоит из двух основных элементов: генератора и дискриминатора, которые тренируются параллельно. Учитывая входной вектор признаков энкодера, дискриминатор пытается предсказать реальное ли переданное ему изображение или подделка. А генератор учится создавать изображения, которые дискриминатор не отличит от реальных. В ходе обучения они улучшаются с помощью друг друга, пока генератор не достигнет нужного уровня.

Схема генератора



Модель генератора следующая: на вход подается вектор случайного шума и вектор признаков от энкодера, они конкатенируются и изменяется их форма до тензора с профилем 4x4. Далее он проходит через три деконволюционных слоя, каждый из которых удваивает профиль в два раза. И финальная свертка выдает трехцветное изображение 32x32

Схема дискриминатора



Модель дискриминатора похожа на отраженный генератор: входное изображение пройдя через три свертки преобразуется тензор с профилем 4x4. Он выпрямляется в вектор, к нему добавляется вектор признаков энкодера и финальный слой выдает предсказание.

Полученные изображения

- ▶ Малое количество исходных данных
- ▶ Большая часть изображений, порядка 80%, совсем не передает исходных цветов и форм
- ▶ Примеры удачных изображений:



Тыква, рыба, самолет, лошадь, панда

В связи с малым количеством исходных данных большая часть сгенерированных изображений не имеют ничего общего с исходными. Но несмотря на это некоторые изображения частично передают цвета и форму исходных. Примеры удачных можно увидеть на слайде, субъективно.

Результаты

- ▶ Организован сбор данных с помощью устройства Emotiv EPOC+
- ▶ Реализован классификатор для определения класса изображения на которое смотрит человек во время считывания ЭЭГ
- ▶ Реализован генератор изображений из пространства признаков энкодера на основе GAN

Таким образом в ходе работы получены следующие результаты:

- Организован сбор данных с помощью устройства Emotiv EPOC+
- Реализован классификатор для определения класса изображения на которое смотрит человек во время считывания ЭЭГ
- Реализован генератор изображений из пространства признаков энкодера на основе GAN

Спасибо за внимание.