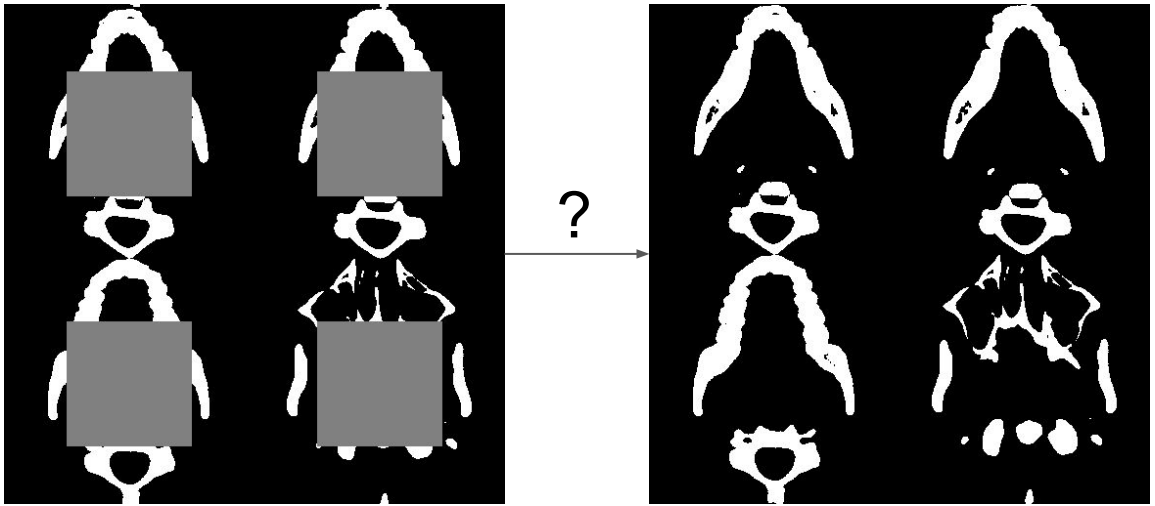
Классификация и восстановление изображений с применением глубоких генеративных сверточных состязательных сетей (DCGAN)

Ожегов Г. А.
Опрышко А. В.

Научный руководитель: Терехов В. И.

Постановка задачи

Создание **методики** для восстановления части изображения на основе его неповрежденной части (на примере восстановления срезов компьютерной томографии черепа).



Актуальность:

Проблема лечения тяжелой черепно-мозговой травмы является часто встречающиеся в современной медицине и имеет большое социально-экономическое значение

Практическая ценность:

Сейчас специалист тратит около 4 часов для создания одного импланта (заменителя поврежденной части черепа).
Методика позволит создавать импланты на 60% быстрее, что позволит увеличить период реабилитации пациентов после ЧМТ.

Решаемые задачи:

- Изучение различных подходов и алгоритмов для решения задачи классификации и восстановления изображений.
- Изучение возможности применения модели Generative adversarial network (генеративно-состязательная сеть, GAN) для задачи классификации и восстановления изображений.
- Разработка упрощенной модели GAN для проверки корректности выбранного подхода к решению задачи.
- Создание и исследование различных архитектур GAN в задаче классификации и восстановления изображений.
- Разработка методики классификации и восстановления изображений с помощью GAN.
- Анализ полученных результатов.
- Разработка программной реализации GAN.
- Улучшение метода восстановления путем предобработки обучающей выборки.
- Проведение вычислительных экспериментов на тестовых данных.

Общая задача классификации

Пусть: $X = \{x_1, ..., x_n\}$ – множество различных изображений
$Y = \{0, 1, ..., m\}$ – пространство возможных меток классов.

Существует неизвестная целевая зависимость — отображение $y^*: X \to Y$, значения которой известны только на объектах конечной обучающей выборки.

Требуется построить алгоритм $a: X \to Y$, способный классифицировать произвольный объект $x \in X$.

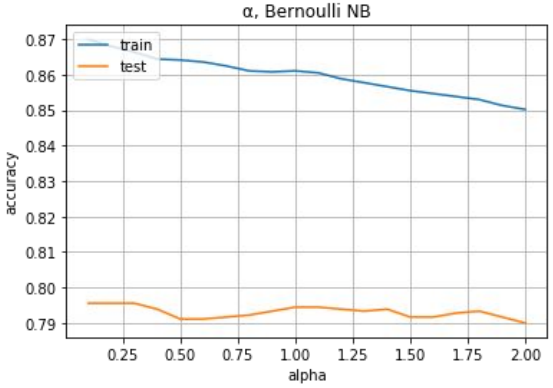
Постановка задачи классификации

Снимки срезов компьютерной томографии (КТ) не совпадают по номерам слоев в связи с различными размерами головы у разных людей.
Кроме того параметры съемки (шаг, снимаемый диапазон) у разных снимков различаются, что не дает однозначно сопоставлять изображения.

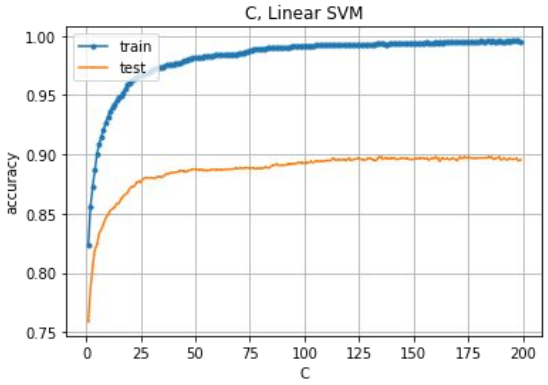
Необходимо классифицировать изображения срезов КТ для их последующего восстановления.

Результаты классификации снимков КТ (1/2)

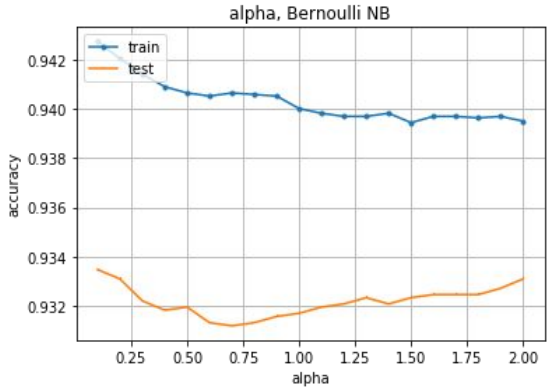
графики зависимости точности классификации от гиперпараметра



Naïve Bayes, визуальный «мешок слов» (BOVW)

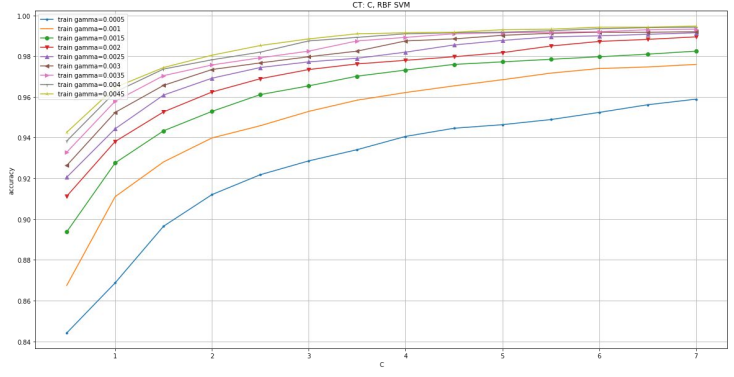


SVM с линейным ядром, визуальный «мешок слов» (BOVW)



Naïve Bayes, попиксельно

Алгоритм классификации и векторизации изображений	Гиперпараметр	F-score
Naïve Bayes + BOVW (Harris)	alpha	81%
Linear SVM + BOVW (Harris)	alpha	88%
Naïve Bayes	C	93,3%
SVM с RBF-ядром, попиксельно	C	94,8%



SVM с RBF-ядром, попиксельно

Результаты классификации снимков КТ (2/2)

значения F-меры в зависимости от конфигурации нейронных сетей

Алгоритм классификации и векторизации изображений	Гиперпараметр — кол-во слоев	F-score
Нейронные сети, попиксельно	2 слоя	35,8%
	3 слоя	44,8%
Свёрточные нейронные сети, попиксельно	2 пары	96,76%
	3 пары	97,76%

Восстановление с помощью GAN

Для создания восстановленного изображения к исходным пикселям нужно добавить завершенные пиксели

$$x_r = M \circ y + (1 - M) \circ G(\hat{z})$$

Для количественного определения «ближайшего» отображения определяем функцию, состоящую из контекстной потери и перцептивной потери

$$L_{contextual} = ||M \circ G(z) - M \circ y||_1 \qquad L_{perceptual}(z) = \log(1 - D(G(z))$$

$$L(z) = L_{contextual} + \lambda L_{perceptual}(z)$$

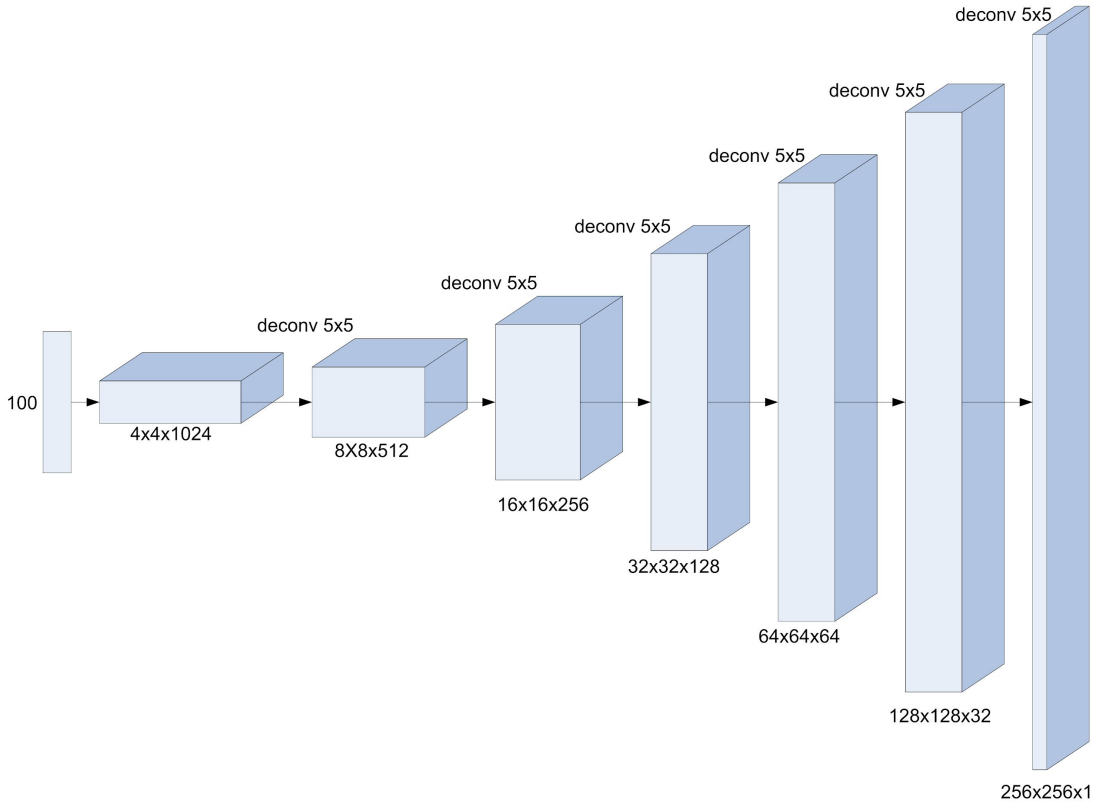
Для нахождение неоходимо минимизировать функцию потерь

$$\hat{z} = argmin_z(L(z))$$

Проектирование GAN на основе глубоких генеративных сверточных состязательных сетей (DCGAN) для восстановления изображений

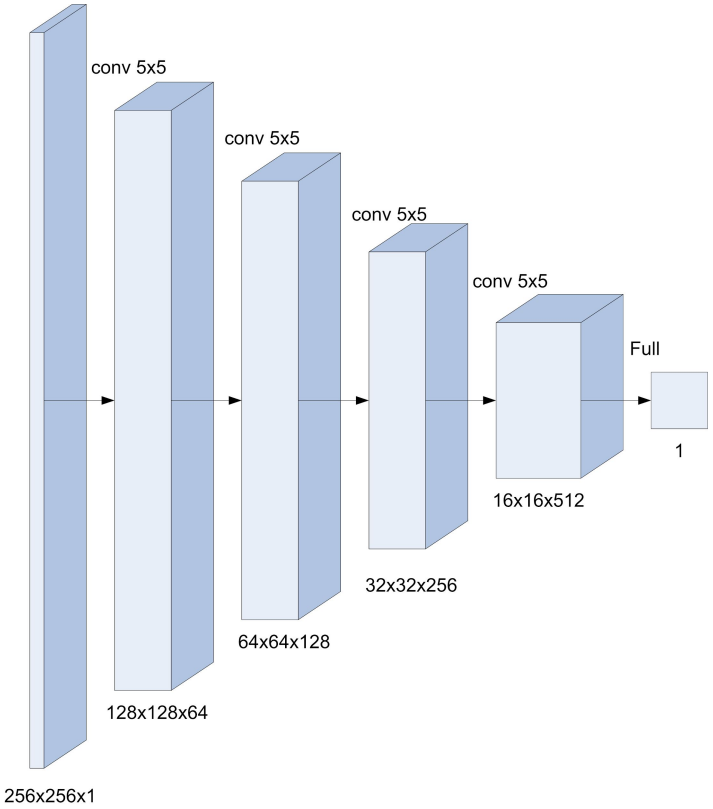
1) Проектирование различных архитектур сети DCGAN.
2) Подготовка обучающей выборки различными методами.

Архитектура генератора сети DCGAN



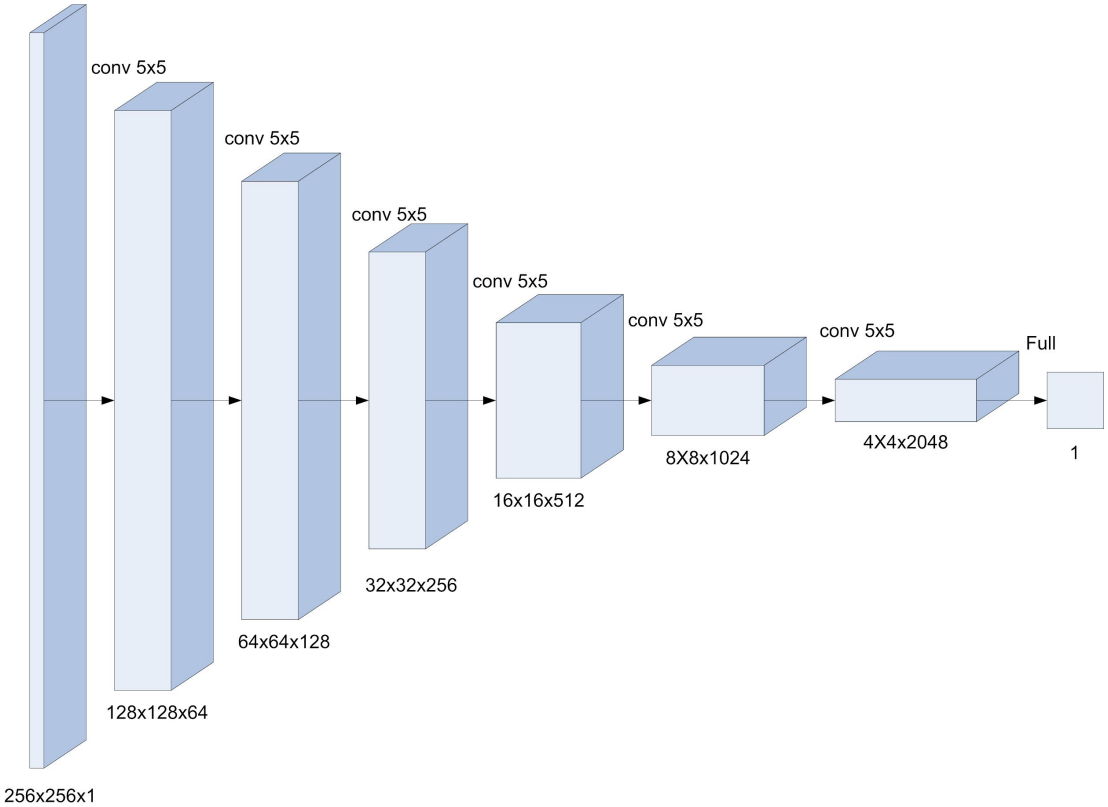
Генератор является общим для первой и второй архитектуры.
Вход сети представляет собой 100-мерный вектор случайного шума, полученный из равномерного распределения между [-1, 1], а затем следует полносвязанный слой, преобразованный в размеры $4 \times 4 \times 1024$.
Каждый из следующих слоев является разверточными слоями, где количество каналов уменьшается вдвое, а размер изображения удваивается из предыдущих уровней.
Всего в сети используется 6 разверточные слоев.
Выходной слой имеет размер $256 \times 256 \times 1$.

Первая архитектура дискриминатора сети DCGAN



Дискриминатор - сверточная нейросеть, которая на вход принимает некоторое изображение, а на выходе получает вероятность того, что изображение является оригиналом.
Входной слой представляет собой изображение размером 256 × 256 × 1, за которым следует серия сверточных слоев, где размер изображения уменьшается вдвое, а количество каналов удваивается от предыдущего уровня.
Всего в сети используется 4 сверточный слоя, а выходной уровень - это softmax двух классов (оригинал или фальшивка)

Вторая архитектура дискриминатора сети DCGAN



Дискриминатор - сверточная нейросеть, которая на вход принимает некоторое изображение, а на выходе получает вероятность того, что изображение является оригиналом.
Входной слой представляет собой изображение размером 256 × 256 × 1, за которым следует серия сверточных слоев, где размер изображения уменьшается вдвое, а количество каналов удваивается от предыдущего уровня.
Всего в сети используется 6 сверточный слоев.

Вычислительные эксперименты

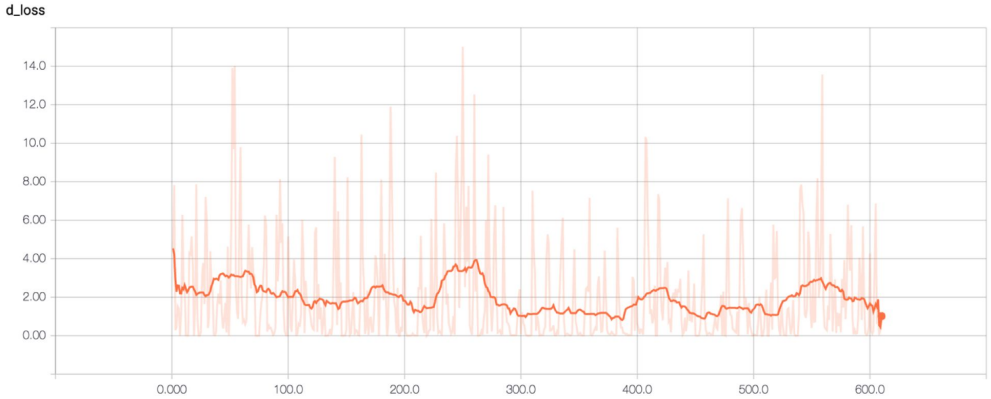
Проведем обучения 2х архитектур сетей с каждым способом подготовки исходных данных.

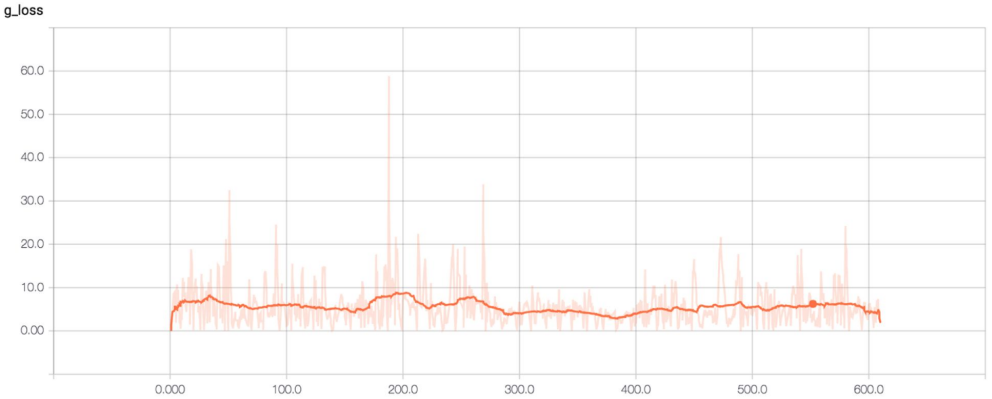
Способы подготовки исходных данных:

1) В естественном порядке съемки компьютерной томографии черепа - порядок изображений начинается с шейных отделов позвоночника и заканчивается черепной коробкой.
2) Кластеризация изображений.
3) Перемешивание изображений из исходного набора данных случайным образом.

Наилучшие результаты достигнуты при использовании 2 способа.

2 способ: сеть, обученная на основе классифицированных данных

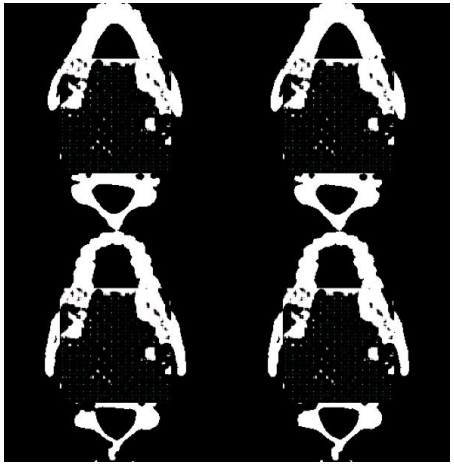




Изображение, порожденное обученным генератором

Восстановление поврежденных снимков КТ

Для сравнения бинарных изображений используется расстояние Хемминга



После восстановления поврежденного снимка КТ расстояние Хемминга d = 712

Направления дальнейших исследований

1. Большее количество экспериментов с «визуальными словами»: размер, предобработка, предобработка изображений перед извлечением ключевых точек.
2. Использование генетических алгоритмов для перебора бóльшего числа параметров алгоритмов классификации изображений КТ.
3. Эксперименты с индексами Хаунсфилда (количественной шкалой рентгеновской плотности): извлечение из снимков изображений КТ с различными индексами Хаунсфилда.