

# РАЗРАБОТКА МОДЕЛЕЙ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВАРИАЦИОННОГО АВТОЭНКОДЕРА

*О.М. Гергет, д.т.н., проф. ОИТ ИШИТР,  
В.В. Лантев, магистрант гр. 8ВМ93  
Томский политехнический университет  
E-mail: vv139@tpu.ru*

## Введение

В настоящее время машинное обучение набирает все большую популярность. Наиболее востребованными являются алгоритмы машинного зрения способные решать такие популярные задачи, как: классификация изображений [1], детектирование объектов [2], распознавание текстов [3] и т.п. В основе каждого успешного обучения лежит датасет<sup>1</sup> с множеством уникальных примеров. Зачастую именно работа с данными, в частности, сбор и разметка, является самой сложной и трудоемкой. Разработчики нередко прибегают к алгоритмам аугментации и синтеза новых данных с целью расширения обучающей выборки.

Целью данной работы являлась разработка модели искусственной нейронной сети (ИНС) для вариационного автоэнкодера (Variational Autoencoder – VAE). VAE позволил не только сгладить шумы на реальных и синтетических данных, но и внести дополнительную уникальность.

## Описание алгоритма

VAE состоят из двух нейронных сетей: энкодера и декодера. Энкодер получает на вход данные и выполняет над ними ряд преобразований, приводя их к компактному и сжатому виду. Далее расположен особый слой скрытого пространства длиной  $n$ , разделяющийся на два подвектора длиной  $n/2$ : вектор средних значений  $\mu$  и вектор стандартного отклонения  $\sigma$ . Среднее значение вектора определяет точку, в окрестностях которой будет вершина вектора, а стандартное отклонение определяет насколько далеко может располагаться вершина от среднего значения. Таким образом, входному объекту соответствует не одна точка в скрытом пространстве, а непрерывная область. В завершении модели располагается декодер, задача которого восстановить изображение из скрытого пространства с максимальным приближением. В частном случае архитектура декодера является отражением энкодера, что позволяет отнести VAE к семейству моделей U-Net. При обучении VAE функционал ошибки базируется на 2-х метриках: средняя квадратичная ошибка (MSE) и расстояние Кульбака-Лейблера (KLD), которое показывает информационное расхождение (относительную энтропию) двух вероятностных распределений. В результате функционал ошибки  $L$  имел следующий вид:

$$L = K.mean(Reconstruction loss + KLD loss) \quad (1)$$

$$Reconstruction loss = Original dim * MSE \quad (2)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3)$$

$$KLD = \sum_{i=0}^N p(x_i) \cdot \log\left(\frac{p(x_i)}{q(x_i)}\right) \quad (4)$$

где  $K.mean$  – среднее значение тензора,  $Original dim$  – размерность исходного изображения,  $y_i$  – значение наблюдения,  $\hat{y}_i$  – предсказанное значение наблюдения,  $N$  – количество наблюдений в выборке,  $p(x_i)$  – исходное распределение,  $q(x_i)$  – аппроксимирующее распределение.

Такой функционал позволяет расположить области в скрытом пространстве как можно ближе друг к другу, но при этом оставить их различными в виде отдельных составляющих. В таком случае, будет возможно производить гладкую интерполяцию и генерировать уникальные данные.

Для разработки моделей искусственных нейронных сетей (ИНС) был использован язык программирования Python с использованием открытой программной библиотеки TensorFlow для машинного обучения, разработанной компанией Google. В основе модели используется предварительно обученная ИНС выделения признаков (feature vector) EfficientNet B1 [4]. Отличительной особенностью данной архитектуры является использование нового метода составного

<sup>1</sup> Датасет (от англ. Data set) — термин, используемый для файловой системы мейнфреймов от IBM; коллекция из логических записей, хранящихся в виде кортежа.

масштабирования (compound scaling method), который равномерно изменяет глубину/ширину/разрешение с фиксированными пропорциями между ними.

Разработка VAE производится главным образом для обработки изображений медицинского характера, в частности снимков МРТ и КТ. Отличительной особенностью данных изображений является одноканальность (градации серого), в следствии чего принято решение в качестве первого слоя сети установить сверточный слой с значением фильтра равное 3. Это позволит привести входное изображение к требуемому формату сети EfficientNet без увеличения вычислительных затрат. Следующей ключевой особенностью модели является архитектура декодера, представленная на рисунке 1, восстанавливающая изображение размерностью 224x224x3 из скрытого пространства размерностью 512.

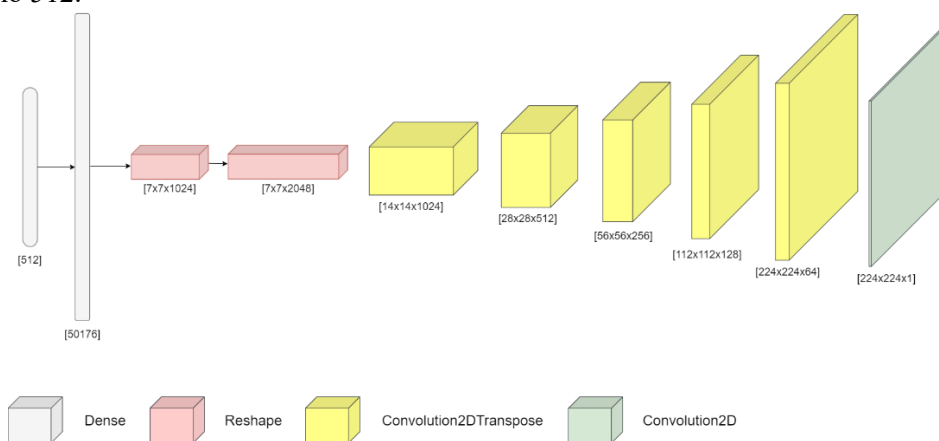


Рис.1. Архитектура «Декодер».

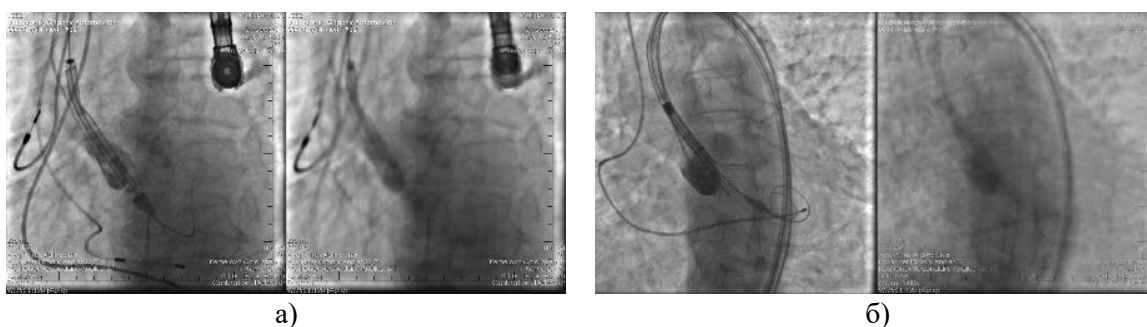


Рис.2. Результаты тестирования: а) входные изображения; б) результат.

## Заключение

В результате исследования разработана собственная архитектура VAE, а также проведена настройка гиперпараметров. Результаты работы модели представлены на рисунке 2 под буквами а и б, где изображения в левой части — это пример входного изображения, а правая — результат. Следует отметить, что при реконструкции изображения с наложением шумов, имея лишь 40% информации от исходного изображения, VAE, посредством собственной сверточной архитектуры, способен восстановить 75–80% исходной информации. Данный метод использован в качестве улучшения исходной выборки модели детектирования объектов.

## Список использованных источников

1. Сикорский О.С. Обзор свёрточных нейронных сетей для задачи классификации изображений // Новые информационные технологии в автоматизированных системах. 2017. №20. [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/obzor-svyortochnyh-neyronnyh-setey-dlya-zadachi-klassifikatsii-izobrazheniy> (дата обращения: 06.03.2021).
2. Ерохин Д. Ю., Ершов М. Д. Современные сверточные нейронные сети для обнаружения и распознавания объектов // Цифровая обработка сигналов. – 2018. – №. 3. – С. 64.
3. Зарипова Р. С., Кривоногова А. Е. Распознавание текстов с использованием нейронных сетей // Nova Um. Ru. – 2018. – №. 11. – С. 38.
4. Mingxing Tan and Quoc V. Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML 2019. [Электронный ресурс]. – URL: <https://arxiv.org/abs/1905.11946> (дата обращения: 06.03.2021)