

УДК 004.93'11

ПОИСК СХОЖИХ АНАТОМИЧЕСКИХ ОБЛАСТЕЙ НА КТ-ИЗОБРАЖЕНИЯХ ЛЁГКИХ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ



Косарева А.А.
ассистент
кафедры электронн
ой техники и
технологии БГУИР,
аспирант БГУИР



Камлач П.В.
доцент кафедры
электронной
техники и
технологии
БГУИР,
кандидат
технических
наук, доцент



Ковалёв В.А.
заведующий
лабораторией
анализа
биомедицински
х изображений
ОИПИ НАН
Беларуси,
Доцент
кафедры
биомедицинско
й
информатики
ФПМИ
Кандидат
технических
наук



Снежко Э.В.
кандидат
технических
наук, ведущий
научный
сотрудник
лаборатории
анализа
биомедицинских
изображений
ОИПИ НАН
Беларуси



Раджабов А.Г.
младший
научный
сотрудник
лаборатории
анализа
биомедицинских
изображений
ОИПИ НАН
Беларуси,
аспирант
ОИПИ НАН
Беларуси

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Email: kosareva@bsuir.by

Косарева А.А.

Окончила Самарский Национальный Исследовательский Университет им С.П. Королёва. Аспирант БГУИР. Работает в БГУИР в должности ассистента. Область научного интереса – обработка данных КТ.

Камлач П.В.

Доцент кафедры электронной техники и технологии, кан. техн. наук, доцент. Область профессиональных интересов/исследований: медицинская электроника; - Информационные технологии в медицине.

Ковалёв В.А.

Доцент кафедры биомедицинской информатики ФПМИ, кандидат технических наук, Заведующий лабораторией анализа биомедицинских изображений ОИПИ НАН Беларуси.

Снежко Э.В.

Кандидат технических наук, ведущий научный сотрудник лаборатории анализа биомедицинских изображений ОИПИ НАН Беларуси.

Раджабов А.Г.

Младший научный сотрудник лаборатории анализа биомедицинских изображений ОИПИ НАН Беларуси, аспирант ОИПИ НАН Беларуси.

Аннотация: в данной работе исследуется эффективность использования моделей сверточных нейронных сетей класса EfficientNet при решении задачи классификации слоя изображения компьютерной томографии лёгких. Приводятся результаты отнесения слоя КТ изображения к одному из классов в соответствии с содержимым анатомической области грудной полости на слое: верхние доли лёгких, сердце, печень. Оцениваются результаты обучения нейронной сети разных версий, определяется оптимальная модель из класса EfficientNet для решения поставленной задачи. Формулируются выводы о пригодности данного класса моделей к извлечению нейросетевых признаков изображений компьютерной томографии лёгких.

Ключевые слова: компьютерная томография (КТ) лёгких, EfficientNet, поиск похожего изображения, задача классификации, дескриптор изображения, нейросетевые признаки

Введение.

Классификация медицинских изображений является важной задачей, решение которой в ряде приложений позволяет реализовать и автоматизировать процесс поддержки диагностики патологий [1]. Одним из основных этапов в решении задач такого типа является подготовка размеченного набора данных, что требует не только наличия изображений определенной модальности и анатомической области, но и участия высококвалифицированного медицинского специалиста. При этом доля «человеческой ошибки» сохраняется и возникает необходимость создания инструмента автоматической разметки медицинских изображений по определенным признакам.

Исследование, представленное в данной статье, является частью более общей работы, цель которой - разработка программного комплекса для поддержки автоматизированной диагностики патологий органов грудной полости [2]. Данная система позволит обеспечить поиск схожего диагностического случая среди базы. Один из способов реализации разрабатываемой системы – использование нейросетевых признаков, полученных при обучении сверточных нейронных сетей на размеченном наборе данных.

В данной статье исследуется эффективность использования нейросетевых признаков КТ-изображений грудной клетки, полученных при обучении моделей класса EfficientNet, для решения задачи поиска схожих анатомических участков грудной полости.

Материалы и методы.

Для проведения исследования был использован предварительно размеченный на три класса набор данных (аксиальные срезы КТ изображения, содержащие верхние доли лёгких, сердце, печень), включающий в себя 413 изображения компьютерной томографии лёгких [3]. Каждая КТ серия состояла из около 150 аксиальных срезов размерами 512x512 пикселей. Была выполнена конвертация аксиальных слоёв в формат PNG (*.png). Предварительно было выполнено отсечение значений вокселей, не попадающих в диапазон между нижней и верхней границами по шкале Хаунсфилда, равными -125 и +255 соответственно (Рисунок 1) и приведено к диапазону от 0 до 255. Данный диапазон значений подходит для визуализации мягких тканей в области грудной полости [4], что является оптимальным вариантом при решении задачи классификации изображений внутренних органов.

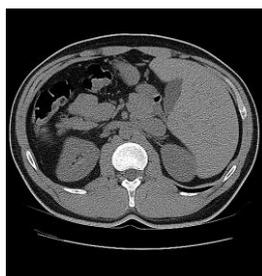


Рисунок 1. Изображение, значения которого находятся в диапазоне от -125 до +225

Ну

В качестве модели для обучения были использованы модели класса EfficientNet. Модели этого класса были разработаны в ходе исследования нового метода масштабирования нейронных сетей [5]. Масштабирование сети может происходить по трём направлениям: преобразование глубины (d), ширины (w) и разрешения (r). Каждый из этих параметров влияет на точность работы обученной сети. В классе моделей EfficientNet предлагается оценить эти параметры с помощью коэффициента смешивания (φ), описанного формулой:

$$\begin{aligned} \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2, \\ \alpha^\varphi &= d, \\ \beta^\varphi &= w, \\ \gamma^\varphi &= r, \\ \alpha, \beta, \gamma &\geq 1, \end{aligned} \quad (1)$$

где α, β, γ – промежуточные переменные, описывающие взаимосвязь между искомыми параметрами масштаба модели при установке определенного коэффициента смешивания.

Данное преобразование называется гибридным. Именно оно легло в основу разработки архитектуры моделей класса EfficientNet.

Так как разные модели EfficientNet имеют разное разрешение входного слоя (Таблица 1), предварительная подготовка данных, подаваемых на вход нейронной сети, отличается.

Таблица 1. Различие входного слоя моделей класса EfficientNet

Модель	Разрешение входного слоя
EfficientNetB0	224x224
EfficientNetB1	240x240
EfficientNetB2	260x260
EfficientNetB3	300x300
EfficientNetB4	380x380
EfficientNetB5	456x456
EfficientNetB6	528x528
EfficientNetB7	600x600

В данной работе были исследованы архитектуры EfficientNetB0 – EfficientNetB5. Это обусловлено тем, что исходное изображение имеет разрешение 512x512 пикселей, и приведение к большему разрешению не даст функционально новых признаков, при этом только увеличит время процесса обучения.

Результаты.

На рисунках 1-2 представлены графики изменения значений функций точности и потерь во время обучения на подготовленном наборе данных. Обучение проходило в течение 100 эпох. Параметры обучения оставались постоянными, менялись только размеры входного пакета данных при обучении для EfficientNetB4 и EfficientNetB5 (размер требуемого объема памяти графического ускорителя для обучения моделей с таким большим количеством параметров приводит к необходимости уменьшения размера входного пакета во время обучения сети).

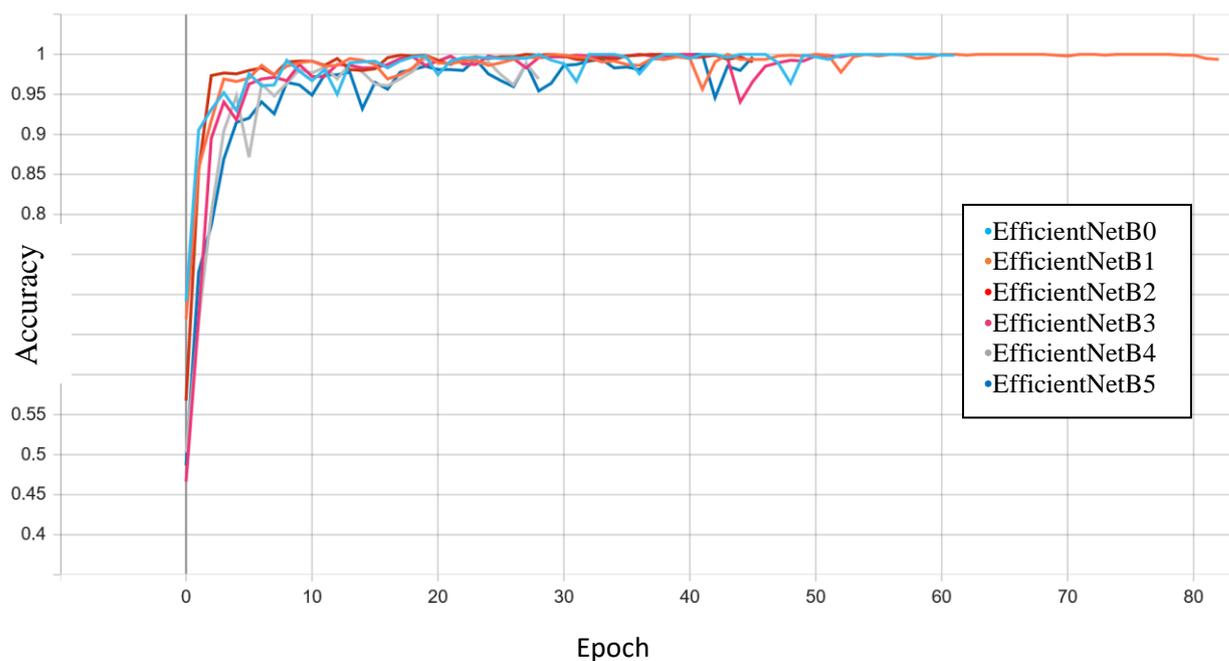


Рисунок 2. Кривая значений функции точности на обучающем наборе данных для моделей класса EfficientNet

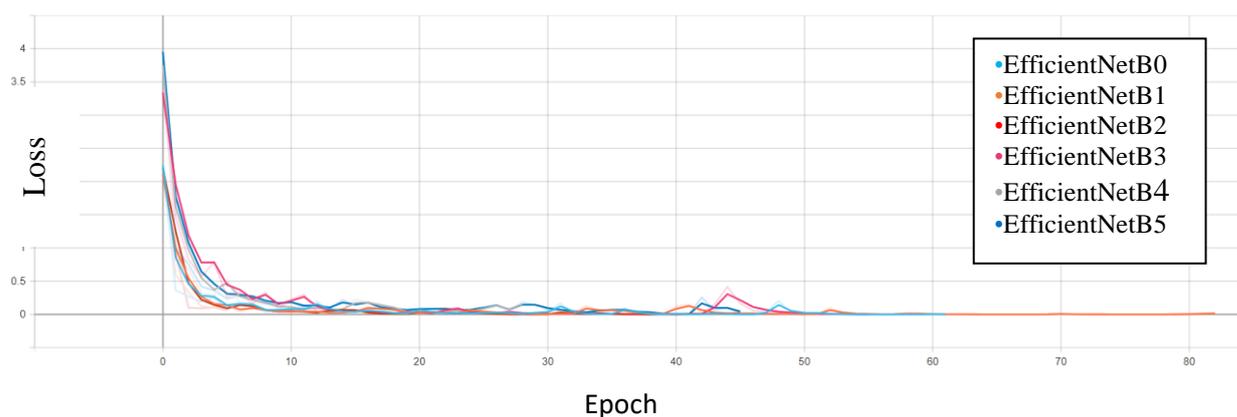


Рисунок 3. Кривая значений функции потерь на обучающем наборе данных для моделей класса EfficientNet

В таблице 2 приведены: количество параметров (весов) обучаемой нейронной сети, максимальные значения функций точности и потерь во время обучения. Все модели показали устойчивый результат при обучении. Наименьшее значение функции потерь на контрольной выборке имели модели EfficientNetB1 и EfficientNetB4. Однако остальные модели также показывают хороший результат, поэтому было необходимо ввести дополнительные параметры для выбора оптимальной модели в рамках решения нашей задачи.

Для оценки способности моделей распознать схожесть слоёв изображения компьютерной томографии лёгких, относящихся к одному классу, был проведён дополнительный эксперимент.

Таблица 2. Результаты обучения моделей класса EfficientNet на исследуемом наборе данных

Модель	Общее количество обучаемых параметров	Val_loss	Val_accuracy
EfficientNetB0	4,011,391	0.00009	1.
EfficientNetB1	6,517,027	0.00001	1.
EfficientNetB2	7,705,221	0.00040	1.
EfficientNetB3	10,700,843	0.00007	1.
EfficientNetB4	17,553,995	0.00003	1.
EfficientNetB5	28,346,931	0.00114	1.

Нейросетевыми признаками изображений назывался вектор значений на выходе слоя обученной нейронной сети, предшествующего полносвязному слою. Таким образом, при подаче входного изображения (слоя КТ изображения) получался вектор-дескриптор данного изображения, содержащий 1280 значений. Расстояние между дескрипторами оценивалось как $L1$ -норма разности двух дескрипторов. Чем меньше расстояние, тем ближе друг к другу изображения в пространстве признаков.

В качестве тестовой выборки для оценки расстояний между изображениями были выбраны соседние слои каждого класса (верх лёгкого, печень, сердце) одного пациента. (Рисунок 4). Данные слои анатомически схожи между собой и в рамках данного исследования предполагалось, что они являются наиболее близкими друг к другу из всего набора данных. Соответственно, чем меньше расстояние между дескрипторами этих изображений, тем более точна модель, позволяющая получить данный дескриптор.

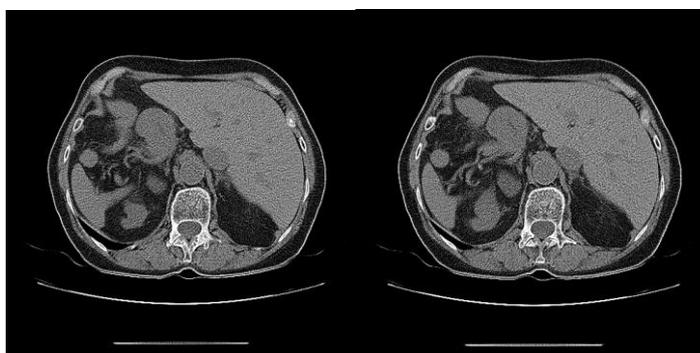


Рисунок 4. Соседние слои класса печень

В таблице 3 приведены значения расстояний между нейросетевыми дескрипторами разных моделей в разных классах.

Таблица 3 – Расстояние между соседними слоями одного пациента

Модель	Верх лёгкого	Печень	Сердце	Среднее
EfficientNetB0	0.039283723	0.01314193	0.025917768	0.026114
EfficientNetB1	0.071293	0.059221447	0.05568123	0.062065
EfficientNetB2	0.05748436	0.33201528	0.017679334	0.135726
EfficientNetB3	0.07199988	0.039694548	0.021171093	0.044289
EfficientNetB4	0.03687218	0.0048017204	0.021511555	0.021062
EfficientNetB5	0.0031098798	0.005650103	0.005241023	0.004667
Среднее по классу	0.046674	0.075754	0.024534	

Наилучший результат показала модель EfficientNetB5. Усредненное расстояние между изображениями каждого класса равно 0.004667. На втором и третьем месте находятся EfficientNetB4 (0.021062) и EfficientNetB0 (0.026114). Лучше всего определялась «схожесть» изображений класса «сердце» (0.024534), а хуже всего – класса «печень» (0.046674). При этом обученная модель EfficientNetB5 занимает 343,6 Мб, а EfficientNetB0 49,4Мб, что примерно в 7 раз меньше. По этой причине использование модели EfficientNetB0 было принято приемлемым для практического использования в составе разрабатываемого программного комплекса.

На последнем этапе выполнялась апробация обученной модели EfficientNetB0 на другом наборе данных [2], определялись двадцать ближайших слоёв среди всех изображений компьютерной томографии всех пациентов (~20000 слоёв) (Рисунок 5).

Все найденные ближайшие изображения относятся к одному классу, следовательно, нейросетевой дескриптор, полученный с использованием обученной модели EfficientNetB0 может быть использован при разработке системы поиска схожих анатомических областей на изображениях компьютерной томографии.

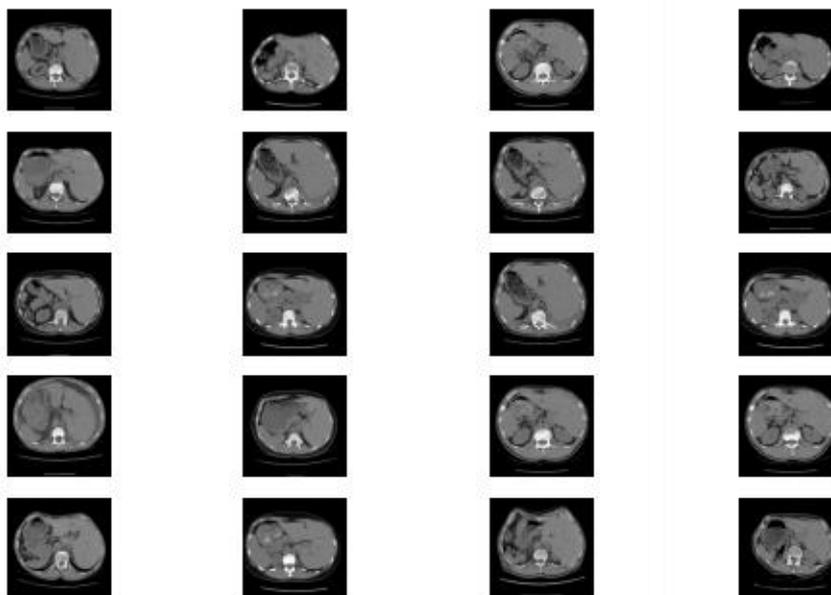


Рисунок 5. Окно работы программы с выводом двадцати ближайших изображений

Заключение.

Данное исследование позволяет сделать вывод о том, что нейросетевые дескрипторы, полученные при обучении моделей класса EfficientNet, пригодны к использованию при разработке системы поиска ряда анатомических областей на изображениях компьютерной томографии. Наилучший результат при этом показала модель EfficientNetB5, однако с точки зрения вычислительной эффективности приемлемо также для решения обозначенной задачи использовать модель EfficientNetB0.

Следующим этапом исследования является обзор других архитектур нейронных сетей и традиционных методов извлечения признаков изображений и введения параметра оценки «схожести», позволяющего проводить сравнительный анализ исследуемых методов.

Список использованных источников

- [1] Yadav, S.S., Jadhav, S.M. Deep convolutional neural network based medical image classification for disease diagnosis. J Big Data 6, 113 (2019). <https://doi.org/10.1186/s40537-019-0276-2>
- [2] Исследование и подготовка архива КТ-изображений патологий лёгких для системы автоматического поиска заданного анатомического участка // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник научных статей VII Международной научно-практической конференции, Минск, 19-20 мая 2021 года / редкол.: В. А. Богуш [и др.]. – Минск : Бестпринт, 2021. – С. 253–257.
- [3] Morozov, S. et al. MosMedData: chest CT scans with COVID-19 related findings. Preprint at <https://arxiv.org/abs/2005.06465> (2020).
- [4] Murphy, A., Baba, Y. Windowing (CT). Reference article, Radiopaedia.org. (accessed on 09 Feb 2022) <https://doi.org/10.53347/rID-52108>
- [5] Mingxing Tan and Quoc V Le. EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of International Conference on Machine Learning(ICML), 2019.

FINDING A SIMILAR ANATOMICAL AREA ON THE LUNG`S CT-IMAGE USING EFFICIENTNET

Kosareva A.A. <i>Assistant of Electronic Engineering and Technology Department, PhD student, BSUIR</i>	Kamlach P.V. <i>Associate professor of Electronic Engineering and Technology Department, PhD, BSUIR</i>	Kovalev V.A <i>PhD, Head of the Biomedical Image Analysis group, United Institute of Informatics Problems</i>	Snezhko E.V. <i>PhD, Leading Researcher, United Institute of Informatics Problems</i>	Radzhabov A.G. <i>PhD student, Junior Researcher United Institute of Informatics Problems</i>
--	---	---	---	---

BSUIR, Republic of Belarus
Email: kosareva@bsuir.by

Abstract: This paper explores the efficiency of using the EfficientNet network when solving the task of determining the similar layer of the computer tomography lung`s image. The results of searching for a similar layer based on pre-marked images are given. Three grades are used: the top of a lung, heart, liver. The results of learning the neural network of different versions are estimated, the optimal learning option is determined under the objective task. Formulate conclusions about the suitability of this class of models to extract neural network signs of images of computer tomography of lungs.

Key words: Computed tomography (CT) of lungs, EfficientNet, searching similar image, classification task, image descriptor, neural network features