

◇◇◇◇ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ◇◇◇◇

УДК 004.89; 616-002

Использование нейросетевых подходов в диагностировании заболеваний

Ахмед С.Х.*

Московский авиационный институт

(национальный исследовательский университет),

г. Москва, Российская Федерация

ORCID: <https://orcid.org/0000-0001-5057-4510>

e-mail: untronix@outlook.com

Скородумов С.В.**

Московский авиационный институт

(национальный исследовательский университет),

г. Москва, Российская Федерация

ORCID: <https://orcid.org/0000-0002-0442-2206>

e-mail: skorodum@gmail.com

Работа направлена на демонстрацию возможности применения нейросетевых подходов к диагностированию заболеваний на примере диагностики коронавируса. Проводится исследование тенденции распространения коронавируса, на основании данных, предоставляемых ВОЗ. Делаются выводы о возможности применения данных подходов.

Ключевые слова: нейронные сети, многослойные сети, сверточные нейронные сети, матрица ошибок, ВОЗ, коронавирус.

Для цитаты:

Ахмед С.Х., Скородумов С.В. Использование нейросетевых подходов в диагностировании заболеваний // Моделирование и анализ данных. 2020. Том 10. № 2. С. 49–61.

DOI: [10.17759/mda.2020100204](https://doi.org/10.17759/mda.2020100204)

**Ахмед Самир Халид*, бакалавр Московского Авиационного Института (национального исследовательского университета), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0001-5057-4510>, e-mail: untronix@outlook.com

***Скородумов Станислав Владимирович*, доцент кафедры 806 Московского Авиационного Института (национального исследовательского университета), г. Москва, Российская Федерация, e-mail: skorodum@gmail.com



1. ВВЕДЕНИЕ

В данной работе затрагивается актуальная, на сегодняшний день, задача обнаружения у пациента опасного заболевания – коронавируса, при помощи нейросетевых подходов. Вообще, применение искусственного интеллекта в медицине уже не ново. Стоит вспомнить одну из канонических задач машинного обучения – обнаружение у пациента рака, по размеру опухоли, или, к примеру, обнаружение болезни Паркинсона на основе сверточных нейронных сетей [1]. Также невозможно не упомянуть решение, представленное корпорацией IBM – IBM Watson [2]. Все они уже используются на практике и демонстрируют хорошие результаты в своих задачах.

Начало 2020 года ознаменовало собой возникновение новой угрозы – COVID-19, более известного как коронавирусная инфекция. Она проявила себя еще в ноябре 2019 года, в провинции Ухань, в Китае, и с тех пор успела охватить весь мир. На Рис. 1 показана тенденция распространения данного заболевания в некоторых странах с 20 марта по 1 апреля 2020 года. Более подробный анализ можно найти в [3].

Сравнение распространения вируса

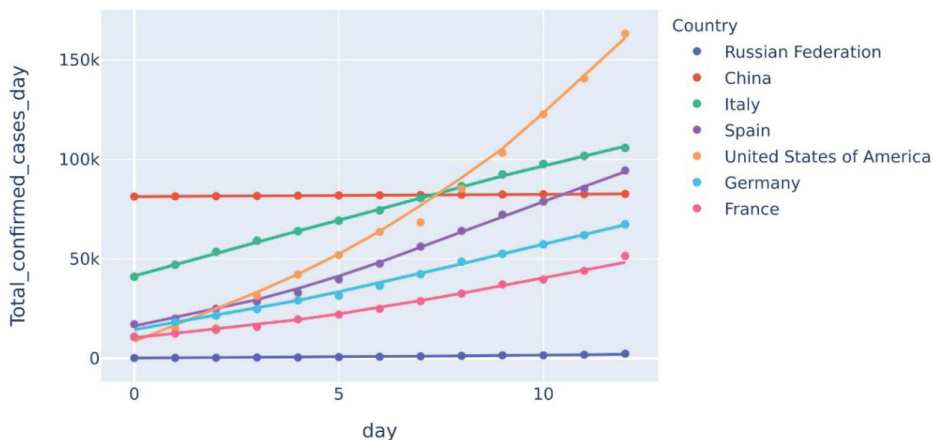


Рис. 1. Тенденция распространения коронавируса в России, Китае, Италии, США и других странах

Стоит отметить экспоненциальный характер тенденции распространения заболевания для стран Европы и Америки, что в действительности совпадает с текущим положением дел. Также немаловажно отметить, что Китай после бурного роста в начале, что, к сожалению, не попало в отмеченный период, показывает малый, едва заметный, прирост, что означает, что в Китае прошел пик заболевания. При этом на фоне остальных стран, показатели Российской Федерации не кажутся столь устрашающими, даже несмотря на заметный прирост заболевших. Однако при бо-



лее детальном рассмотрении, можно сделать вывод, что процесс распространения в России набирает обороты. Это начинает становиться заметным уже на Рис. 2. При первом взгляде бросается в глаза тенденция роста заболевших у США, что наводит на мысль о том, что меры по предотвращению данной ситуации были приняты, к сожалению, поздно. Положительным моментом является тренд Италии и Германии, который дает понять, что на текущий момент у них намечен спад. По темпам возникновения новых случаев заболевания Российская Федерация уже перегнала Китай, что является тревожным звонком.

Новые случаи

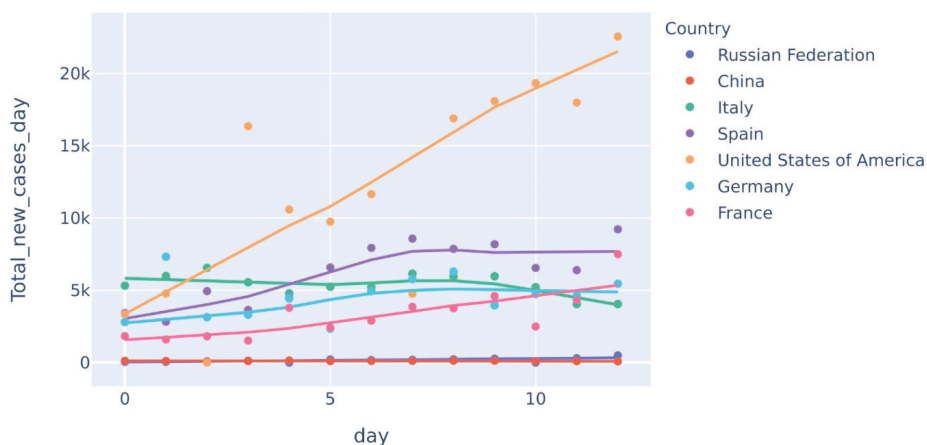


Рис. 2. Темпы обнаружения новых заболевших

Рис. 3 демонстрирует график действительного прироста числа, заболевших за двухнедельный период времени. На нем заметен характер изменения числа объектов с рассматриваемым диагнозом. Вырисовывается некоторый шаблон поведения исследуемой величины – после бурного роста следует однодневный период относительной стагнации, после которого следует дальнейший бурный, почти экспоненциальный рост. При этом период между этими паттернами увеличивается. Но двухнедельный период слишком мал для построения точной математической модели тенденции вируса в нашей стране.

Большинство государств для решения проблемы пандемии COVID-19 используют стратегию карантина и режима самоизоляции, а на врачей ложится огромная нагрузка по обеспечению необходимых мероприятий по поддержанию здоровья людей. На наш взгляд данная ситуация заставляет задуматься о внедрении новых систем обеспечения медицинских услуг, которые будут способны решить следующие проблемы:



Итоговое кол-во подтвержденных случаев

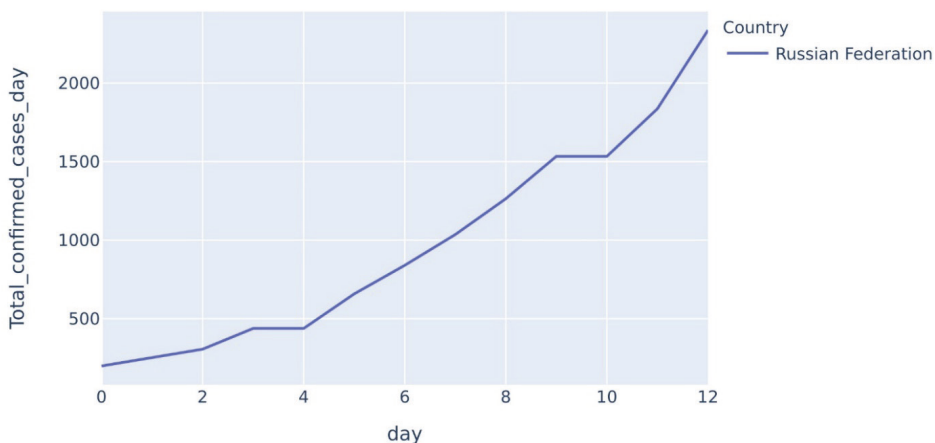


Рис. 3. Реальное положение дел в России в период с 20 марта по 1 апреля 2020 г.

1. Снизить нагрузку на врачей, за счет смены канала общения и перекладывания некоторых задач на механизмы искусственного интеллекта
2. Уменьшить время пребывания людей в поликлиниках и больницах
3. Сделать систему прозрачной и централизованной, для обеспечения единого информационного пространства
4. Обеспечение сохранности данных, за счет современных технологий хранения данных (например, blockchain)

Для этого необходимо доказать возможность применения нейросетевых подходов в медицинских задачах, на что и нацелена данная статья. Это будет продемонстрировано на примере решения задачи диагностирования коронавируса по флюорографии пациента. Данное решение не претендует быть исчерпывающим, являясь своего рода лишь академическим примером возможности построения такого вида систем. Прежде чем переходить к описанию самого решения, необходимо дать несколько важных определений, а также описать используемые архитектуры нейросетей.

2. СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ

Дадим краткое описание архитектуры, следуя работе [4]: основной идеей сверточной архитектуры сетей состоит в том, что обработка участка изображения очень часто должна проходить независимо от конкретного расположения этого участка. Сеть делает предположение: покроем вход небольшими окнами и будем выделять признаки в каждом окне небольшой нейронной сетью, причем признаки в каждом окне будут выделяться те же. Затем эти окна мы будем заменять центральным пикселем и на ней можно применить второй слой и т.д. Иначе говоря мы сворачиваем наше

изображение, каждый раз уменьшая размерности полученных карт признаков. Это операция, по понятным причинам носит название свертка.

$$y_{i,j}^l = \sum_{-d \leq a,b \leq d} W_{a,b} * x_{i+a,j+b}^l \quad (1.1)$$

где $y_{i,j}^l$ – результат свертки на уровне l ;

x_i^l – ее вход, выход всего предыдущего слоя;

$W_{a,d}$ матрица весов $(2d+1)*(2d+1)$.

Формально, наше окно скалярно умножается на тензор свертки. Пример операции свертки показан на Рис. 4. Начиная с Рис. 4 и проходя по разделу, связанному с описанием работы архитектур, рисунки были заимствованы у источника [4].

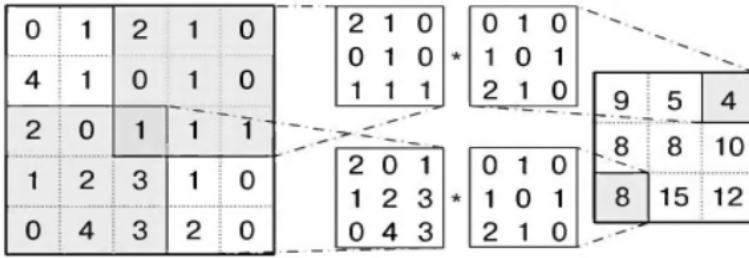


Рис. 4. Пример подсчета результат свертки [4].

Это преобразование обладает следующими свойствами:

- Сохраняет структуру входа.
- Разреженность.
- Использует один и тот же набор весов.
- Естественный регуляризатор

Помимо линейной свертки и следующей за ней нелинейности есть еще одна очень важная операция: субдискретизация. Ее смысл прост: в сверточных сетях исходят из предположения, что наличие или отсутствие того или иного признака гораздо важнее, чем его точные координаты (см. Рис. 5). Обычно на практике используется операция max pooling [4].

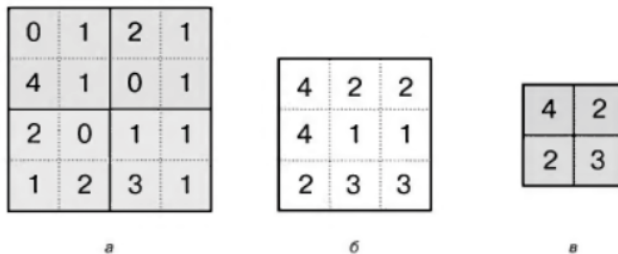


Рис. 5. Пример субдискретизации с окном размера 2x2 [4]



Стандартный слой сверточной сети состоит из трех компонентов:

1. Свертка в виде линейного отображения, выделяющая локальные признаки
2. Нелинейная функция, примененная покомпонентно к результатам свертки
3. Субдискретизация, которая обычно сокращает геометрический размер получившихся тензоров.

Схема такого слоя представлена на Рис. 6:

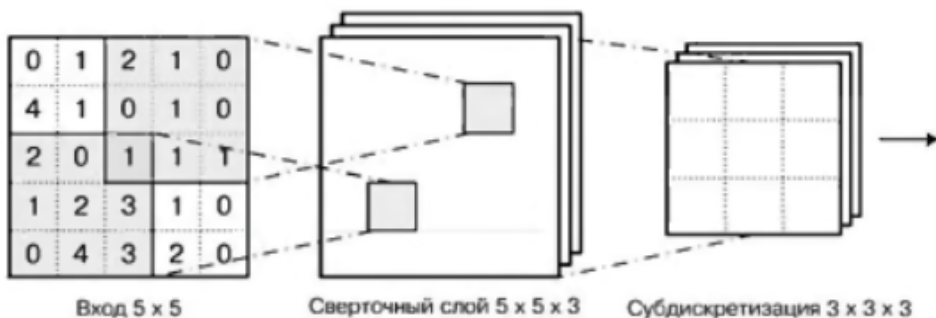


Рисунок 6. Пример сверточного слоя [4]

Следует обратить внимание на Рис 6. По сравнению с картинкой, на входе размерность тензора увеличилась: сверточная сеть обычно обучает сразу несколько карт признаков на каждом слое (на Рис. 6 три). Более подробно о процессе обучения этих сетей можно узнать в [4].

3. VGG

VGG является разработкой Группы визуальной геометрии Оксфордского университета. По факту это общее название двух конфигураций сетей на 16 и 19 слоев. В этой статье основной упор будет делаться на описании архитектуре с 16 слоями.

Что примечательного в данной архитектуре? Для этого следует сказать о том, как выстраивались архитектуры до VGG-16.

В качестве сверток использовались свертки размером 7x7 с шагом 2 или же 11x11 с шагом 3. Это было сделано с целью увеличить рецепторную способность сети, охватывая большую часть изображения. При этом соответственно увеличивалось число весов, что усложняло процесс обучения и сходимости сети. Ученые из Оксфордского Университета решили эту проблему, разбив сетку 7x7 на две сетки 3x3, в совокупности, имеющие тоже рецептивное поле, но при этом количество весов гораздо меньше 27 против 49. То есть при увеличении глубины сети, мы будем иметь меньшее количество весов. Как следствие разбиения большой свертки на меньшие дала внести нелинейность между такими слоями, увеличивая разрешающую способность сети.

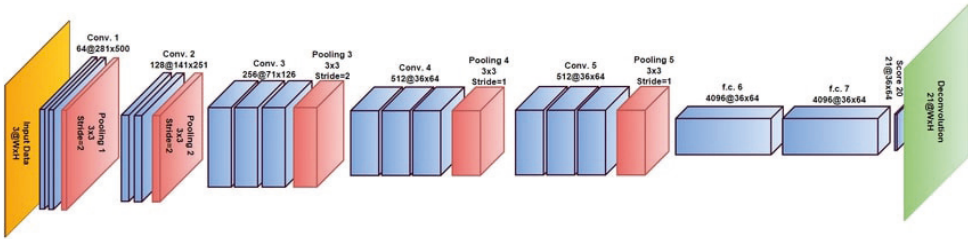


Рис. 7. Архитектура VGG-16

Синими блоками обозначены слои свертки, красными – слои субдискретизации
Стоит обратить внимание на то, что

- В сети идут по две-три свертки друг за другом без слоя субдискретизации
- Число карт признаков постепенно растет на более глубоких слоях

Это все ведет к тому, что мы сильно уменьшаем размерность исходного входа / изображения, параллельно порождая большее число объектов меньшего размера. В конце сети мы имеем практически сплюсненный вход, который мы растягиваем полносвязной сеткой.

Данная архитектура в 2014 году одержала победу в одной из номинаций ImageNet Large Scale Visual Recognition Competition [5].

4. ОСТАТОЧНОЕ ОБУЧЕНИЕ

Как показала практика, очень глубокие сети дают результат хуже чем менее глубокие. Увеличение ошибки наблюдается и на тестовом и даже на тренировочном наборе. Это частично обусловлено проблемой затухания градиентов при обратном распространении ошибок. Для решения проблемы деградации, Microsoft Research предложила идею: глубокое остаточное обучение [6]. Основное отличие от базовой модели в том, что в остаточном блоке существует прямая связь между выходом предыдущего слоя и следующим слоем, обходя некоторое преобразование. Примеры таких слоев продемонстрированы на Рис. 8.

Формально мы строим отображение. Следуя определению данному выше, а также [4,6]:

$$x^{k+1} = F(x^k) + x^k \quad (1.2)$$

где x^{k+1} – вход следующего слоя $k+1$; x^k – вход слоя k ; $F(x^k)$ – исходная нелинейность.

Нейронные сети зачастую аппроксимируют некоторую функцию $R(x)$. Используя данное отображение, мы требуем, чтобы $F(x)$ аппроксимировала не функцию $R(x)$, а остаток $R(x) - x$, что зачастую проще осуществить. В таком случае остаточные связи отвечают за воспроизведение изменения входных аргументов.

Но как это отображение помогает решить проблему затухания градиентов? Давайте формально возьмем производную от данного слоя, следуя [4]:



$$\frac{\partial x^{k+1}}{\partial x^k} = 1 + \frac{\partial F(x^k)}{\partial x^k} \tag{1.3}$$

За счет возникновения остаточной связи, у нас к градиенту добавилась 1. То есть теперь полностью обученный слой не будет влиять на процесс обучения, так как несмотря на то, что его градиенты близки к нулю, за счет остаточной связи, они не будут нулями.

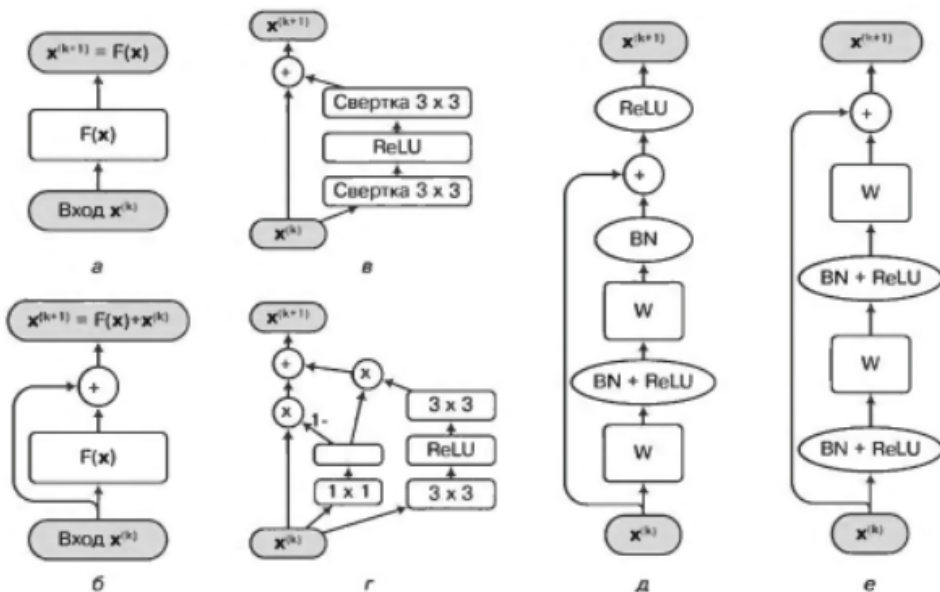


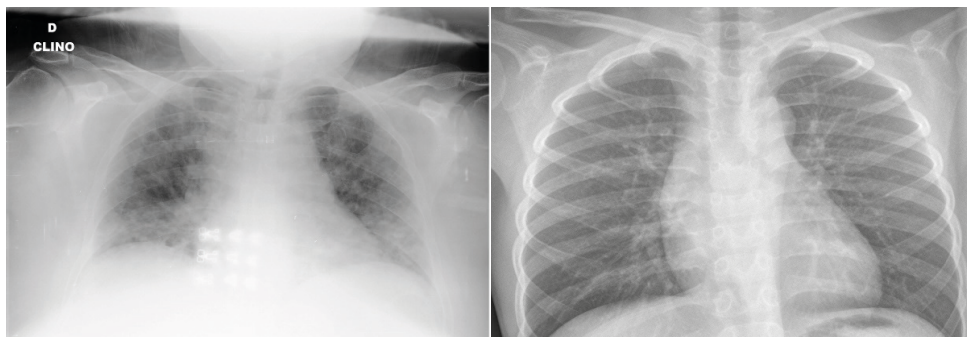
Рис. 8. Вариации остаточного блока и варианты остаточных сетей [4]

После того как даны необходимые определения пора перейти к рассмотрению самого эксперимента. Более детальное изложение материала можно увидеть в [4,6].

5. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА

Для диагностирования заболевания с помощью нейросетевых подходов был сформирован малый набор данных, содержащий в себе снимки флюорографии здоровых и больных пациентов. В нем содержатся 25 снимков, подтверждающих наличие заболевания, а также 25 снимков с пометкой здоров и 18 фотографий для тестирования. Данный набор, к сожалению, несет в себе лишь академический интерес, так как он слишком мал для полного формирования мнения по данной проблеме, но открывает иной взгляд на решение задачи.

На Рис. 9 приведен пример флюорографии больного и здорового пациента. Данный набор доступен по [3,8].



а) б)
 Рис. 9. Примеры флюорографии больного человека а) и здорового б)

В качестве базовых сетей было решено использовать предобученные сети VGG-16 и ResNet-50. Набор весов был получен путем обучения архитектур на наборе данных ImageNet, после этого, используя подход Transfer Learning [7], мы заменили последний слой на свой, что минимизировало время необходимое на обучение сети, так как вместо обучения всех весов, мы обучали лишь последние слои. Также за счет предобученности нейросетей мы имеем сети с большой обобщающей и выделяющей способностями.

На рис. 10 приведены матрицы ошибок для сравнения качества обученных моделей VGG и ResNet. После эмпирических проверок, многочисленных прогонов, было выявлено, что оптимальным числом эпох для обучения модели является 20 (рассматривались варианты с 20 эпохами, 40 эпохами и 100 эпохами). Для Resnet более оптимальным было число 40 эпох, но так как на этом числе эпох VGG начал явно переобучаться предпочтение отдалось 20 эпохам

Как заметно, данные модели хорошо обучились и смогли правильно классифицировать объекты в 16–17 случаях из 18. Но данные результаты, возможно обусловлены малочисленностью данных, или разделимостью данных (например, при добавлении объектов с простой пневмонией результаты падают до 80 процентов). Но это не показывает, какой же алгоритм лучше решает поставленную задачу. Посмотрим, как они себя вели на этапе обучения и валидации. Для этого взглянем на рис. 11 и 12

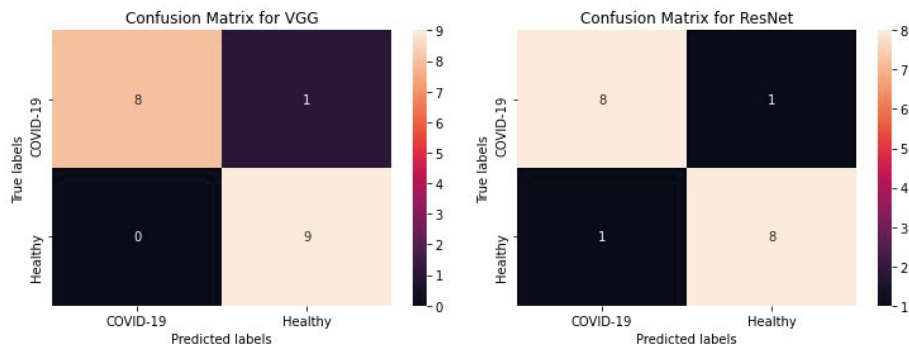
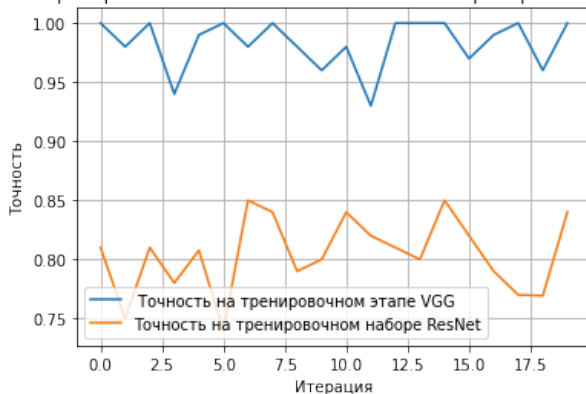


Рис. 10. Матрица ошибок для VGG и ResNet



Точность на тренировочном этапе VGG vs Точность на тренировочном этапе ResNet



Потери во время тренировки VGG vs Потери во время тренировки ResNet

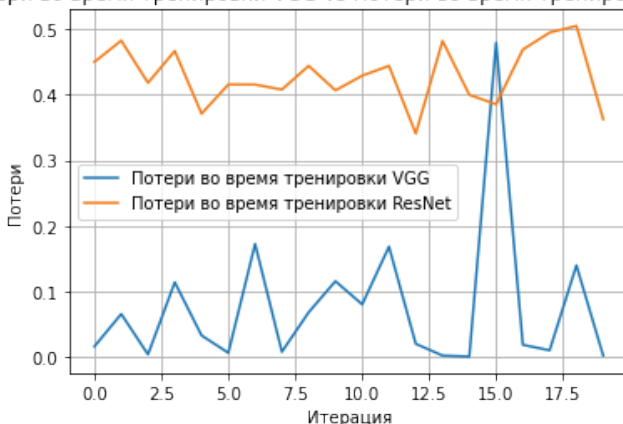


Рис. 11. Поведение моделей на этапе обучения

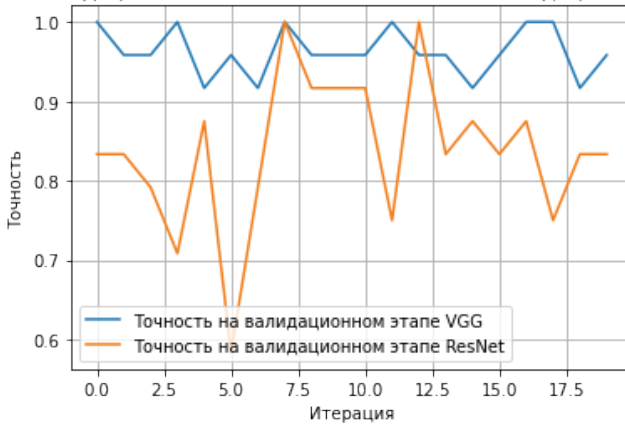
Первое что стоит отметить, это высокие показатели точности на этапе обучения и их процесс стабилизации и сходимости на данном числе итераций. К сожалению, это не столь очевидно при просмотре функции потерь на данном числе эпох: ошибка имеет большую дисперсию на рассматриваемой дистанции. Стоит также заметить, какой прирост точности и убыль функции потерь демонстрирует старенькая модель VGG, существенно превосходя ResNet по этим показателям, что может быть обусловлено количеством обучающих сэмплов или числом эпох. Необходимо сказать, что для воспроизведения сопоставимых результатов ResNet пришлось прогонять несколько раз, что показывает некоторую слабость этой модели при применении ее на малый набор данных или при малом числе итераций. Есть также предположение о некоторой избыточности модели. Данные графики показывают, что обобщение прошло хорошо, что вызвано скорее характером объектов, нежели работой самих моделей.



Стоит сделать замечание, анализ процесса обучения не говорит нам об обобщающей способности алгоритма, но дает нам некоторые инсайды, по поводу того, как оно протекало и насколько оно сошлось к решению.

Чтобы делать выводы о качестве модели, о ее обобщающей способности рассмотрим ее показатели на валидации.

Точность на валидационном этапе VGG vs Точность на валидационном этапе ResNet



Потери во время валидации VGG vs Потери во время валидации Loss ResNet

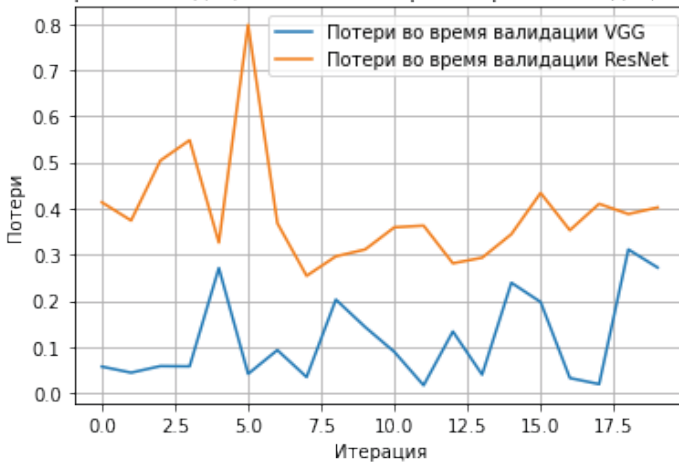


Рис. 12. Показатели моделей на валидации

Смотря на показатели при валидации моделей, стоит заметить, что Resnet на валидации показывает себя нестабильно, что свидетельствует о том, что Resnet не сошелся к окончательному решению за 20 итераций, вероятно, необходимо больше итераций, либо больше примеров, в тоже самое время на данном наборе данных VGG показывает себя стабильнее и в конечном итоге имеет меньшую дисперсию, что отображается на функции потерь.



6. ВЫВОДЫ

В данной работе были протестированы нейросетевые алгоритмы VGG-16 и ResNet-50 и подход к обучению Transfer Learning применительно к задачам детектирования патологий на малом наборе визуальных данных [снимках флюорографии]. Результаты показали перспективность нейросетевого подхода на данных задачах. Однако стоит отметить недостаточный размер искомого датасета для построения полноценного детектора. Данная работа прежде всего предназначалась как демонстрация возможности построения подобных систем.

Литература

1. *Andrés Ortiz, Jorge Munilla, Manuel Martínez-Ibañez, Juan M. Górriz, Javier Ramírez and Diego Salas-Gonzalez.* Parkinson's Disease Detection Using Isosurfaces-Based Features and Convolutional Neural Networks // *Front Neuroinform.* 2019;13:48. doi: 10.3389/fninf.2019.00048
2. *IBM Watson Website* [Электронный ресурс] // IBM Watson. URL: <https://www.ibm.com/watson-health>
3. *Github* [Электронный ресурс] // Github. <https://github.com/UntronixOnly/Articles>
4. *Николенко С., Кадурин А., Архангельская Е.* Глубокое обучение // Спб.: Питер, 2019—480 с: ил.
5. *Imagenet* [Электронный ресурс] // Imagenet competition. URL: <http://www.image-net.org/>
6. *Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun.* Deep Residual Learning for Image Recognition [Электронный ресурс] // ArXiv. URL: <https://arxiv.org/pdf/1512.03385.pdf>
7. *Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, Qing He.* A Comprehensive Survey on Transfer Learning [Электронный ресурс] // ArXiv. URL: <https://arxiv.org/pdf/1911.02685.pdf>



The Use of Neural Network Approaches in the Diagnosis of Diseases

Samir Kh. Akhmed*

Moscow Aviation Institute (National Research University),
Moscow, Russian Federation
ORCID: <https://orcid.org/0000-0001-5057-4510>
e-mail: untronix@outlook.com

Stanislav V. Skorodumov**

Moscow Aviation Institute (national research University),
Moscow, Russian Federation
ORCID: <https://orcid.org/0000-0002-0442-2206>
e-mail: skorodum@gmail.com

The work is aimed at demonstrating the possibility of using neural network approaches to diagnosing diseases using the example of the diagnosis of coronavirus. A study of the trend in the spread of coronavirus is being conducted, based on data provided by WHO. Conclusions are drawn about the possibility of applying these approaches.

Keywords: neural networks, multilayer networks, convolutional neural networks, error matrix, WHO, coronavirus.

For citation:

Ahmed S. Kh., Skorodumov S.V. The Use of Neural Network Approaches in the Diagnosis of Diseases. *Modelirovanie i analiz dannykh = Modelling and Data Analysis*, 2020. Vol. 10, no. 2, pp. 49–61. DOI:10.17759/mda.2020100204 (In Russ., abstr. in Engl.).

References

1. Andrés Ortiz, Jorge Munilla, Manuel Martínez-Ibañez, Juan M. Górriz, Javier Ramírez and Diego Salas-Gonzalez. Parkinson's Disease Detection Using Isosurfaces-Based Features and Convolutional Neural Networks. *Front Neuroinform.* 2019;13:48. doi: 10.3389/fninf.2019.00048
2. IBM Watson Website [Electronic resource]. IBM Watson. URL: <https://www.ibm.com/watson-health>
3. Github [Electronic resource]. Github. <https://github.com/Untronix0nly/Articles>
4. Nikolenko S., Kadurin A., Archangelskaya E. Deep learning // Saint Petersburg: Piter, 2019–480 s: II.
5. Imagenet [Electronic resource]. Imagenet competition. URL: <http://www.image-net.org/>
6. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition [Electronic resource]. ArXiv. URL: <https://arxiv.org/pdf/1512.03385.pdf>
7. Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, Qing He. A Comprehensive Survey on Transfer Learning [Electronic resource]. ArXiv. URL: <https://arxiv.org/pdf/1911.02685.pdf>

***Akhmed Samir Khalid**, Bachelor of the Moscow Aviation Institute (National Research University), Moscow. Moscow, Russian Federation, ORCID: <https://orcid.org/0000-0001-5057-4510>, e-mail: untronix@outlook.com

****Skorodumov Stanislav Vladimirovich**, Associate Professor, Chair 806, Moscow Aviation Institute (National Research University), Moscow, Russian Federation, e-mail: skorodum@gmail.com