

АЛГОРИТМ ПОИСКА «ЦИФРОВЫХ ДВОЙНИКОВ» В БАЗАХ ДАННЫХ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ

Д. В. Сизова, В. А. Ковалев

Белорусский государственный университет, Минск, Беларусь

E-mail: darya.sizova31@gmail.com

В последние годы остро стоит проблема защиты личных данных, к которым, в свою очередь, относятся медицинские изображения. Ввиду человеческого фактора, не исключены ошибки при внесении изображений врачом в базу данных, что влечет за собой необходимость разработки алгоритма, позволяющего извлекать из базы изображения, принадлежащие одному человеку. Целью данной работы является разработка и сравнение алгоритмов поиска похожих изображений в базе данных на примере рентгеновских снимков легких.

Ключевые слова: *биоинформатика, сиамская нейронная сеть, рентгеновские изображения.*

Введение. В настоящее время существует ряд алгоритмов для сопоставления изображений, различающихся по точности и производительности, начиная с поиска классических признаков изображения, заканчивая моделями глубокого обучения, на которых основано большое количество современных алгоритмов компьютерного зрения.

В основу нашего алгоритма сопоставления изображений легла сиамская нейронная сеть, которая для двух входных изображений определяет то, насколько они близки на основании расстояния между векторами признаков, соответствующими данным изображениям. Она представляет собой две идентичные подсети, выходами которых являются вектора признаков, за которыми следует слой вычисления Евклидова расстояния между полученными векторами. В качестве функции потерь была выбрана функция контрастных потерь, побуждающая пары снимков одного человека (положительные) находиться близко друг к другу в пространстве признаков, в то же время отдаляя пары снимков разных людей (негативные):

$$l(x_i, x_j, z_{ij}) = (1 - z_{ij}) \|h_i - h_j\|_2^2 + z_{ij} \max(0, \tau - \|h_i - h_j\|_2^2)$$

где x_i, x_j – входные изображения, h_i, h_j – соответствующие им вектора признаков, z_{ij} – метка пары (0, если положительная, и 1 в противном случае), τ – заранее заданный порог.

Для изображений базы данных хранятся заранее найденные вектора признаков, которые попарно сравниваются с вектором признаков входного изображения.

Эксперименты. В исходном наборе данных имелось порядка 94 тысяч рентгеновских снимков легких мужчин и женщин возрастом от 18 до

50 лет. Половина данных вошла в состав положительных пар, из другой половины случайным образом были составлены негативные пары. Все изображения были сжаты до размера 224 x 224 пикселя.

Выбор порога и архитектуры подсетей. Для выбора порога из промежутка $[0.1, 1.0]$ с шагом 0.1 была обучена сиамская сеть на основе предобученной сети ResNet-50, произведено сравнение значений площади под ROC-кривой для предсказаний модели на валидационном наборе данных. Наилучших результатов удалось достигнуть при значении $\tau = 0.5$.

Далее, было произведено сравнение сети, содержащей подсеть с ResNet-50, с сетью, основанной на архитектуре EfficientNet B4. В случае с ResNet-50 на тестовом наборе площадь под ROC-кривой была равна 0.955, а для модели с EfficientNet B4 она составила 0.94. Таким образом, ResNet-50 была взята за основу подсетей сиамской нейронной сети.

Несмотря на неплохие результаты классификации пар на негативные и позитивные, модель показала себя не лучшим образом при поиске пары для 250 тестовых изображений среди 15000 снимков базы данных (см. таблицу 1), что можно объяснить тем, что в тренировочном наборе изображения негативных пар были слишком различными и сеть не научилась различать похожие изображения, принадлежащие разным людям. Таким образом, появилась необходимость составления более сложных негативных пар изображений.

Составление негативных пар. Получить более сложные пары удалось итеративно, используя уже обученную сиамскую модель. На первой итерации были выбраны изображения в количестве равном количеству негативных пар, которое присутствовало ранее в наборе. Для них среди снимков других людей были найдены ближайшие по такому же принципу, как происходит поиск в базе данных. На второй итерации модель была обучена уже на новых парах и, как и ранее, составлены негативные пары. Для этой модели на рисунках 1 и 2 показан пример поиска похожих изображений, полученных при сравнении векторов признаков.

Сравнение моделей, обученных на парах с первой и второй итерации, можно увидеть в таблице 1.

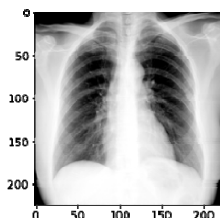


Рис. 1. Входное изображение



Рис. 2. Результат поиска пяти ближайших изображений в базе данных

Аугментации. Также, к тренировочным данным были применены аугментации, которые могут улучшить обобщающую способность модели, выступая как способ регуляризации [2]. Были выбраны следующие трансформации снимков: поворот на градус из промежутка $[-5, 5]$ и небольшое увеличение. Как показал эксперимент, прироста точности модели достигнуть не удалось. Результаты можно увидеть в таблице 1.

Таблица 1

Результаты поиска пары в зависимости от способа формирования тренировочных данных

Способ формирования пар	Доля изображений, для которых была найдена пара среди топ-п ближайших снимков					
	5	10	15	20	25	30
Случайно	0.004	0.02	0.036	0.056	0.06	0.06
Итеративно (1 ит.)	0.032	0.052	0.08	0.108	0.124	0.148
Итеративно (2 ит.)	0.064	0.104	0.156	0.204	0.224	0.256
Итеративно (2 ит.) + аугментации	0.068	0.108	0.12	0.136	0.156	0.188

Заключение. В данной работе была рассмотрена нейронная сеть для сопоставления изображений, обученная на различных наборах тренировочных данных, и произведена оценка качества поиска пары для изображений в базе данных. Полученные результаты указывают на недостаточную для использования в прикладных программах точность модели и необходимость дальнейших исследований для ее улучшения или выбора другого подхода получения векторов признаков.

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Melekhov I., Kannala J., Rahtu E. Siamese network features for image matching // 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE, 2016. С. 378-383.
2. Zoph B. et al. Learning data augmentation strategies for object detection // European Conference on Computer Vision. Springer, Cham, 2020. С. 566-583.