РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММНЫХ СРЕДСТВ ПОИСКА «ЦИФРОВЫХ ДВОЙНИКОВ» ПО РЕНТГЕНОВСКИМ ИЗОБРАЖЕНИЯМ

Д. В. Сизова

Белорусский государственный университет, г. Минск;
darya.sizova31@gmail.com;
науч. рук. – В. А. Ковалев, канд. техн. наук

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: рентгеновские изображения; сиамские нейронные сети; сопоставление изображений; мешок визуальных слов; ключевые точки.

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 первого алгоритма сопоставления изображений легла сиамская нейронная сеть [1]. Она представляет собой две идентичные подсети, выходами которых являются вектора признаков, за которыми следует слой вычисления Евклидова расстояния между полученными векторами. Для изображений базы данных хранятся заранее найденные вектора признаков, которые попарно сравниваются с вектором признаков входного изображения. Обучение сети происходит посредством минимизации функции контрастных потерь.

Вначале модель была обучена на положительных и случайно составленных отрицательных парах изображений. В силу того что изображения негативных пар были слишком различными, сеть не различала похожие изображения, принадлежащие разным людям. Таким образом, появилась необходимость составления более сложных негативных пар изображений [2]. Получить такие пары удалось итеративно: с помощью уже обученной модели для снимков находятся ближайшие к ним среди тех, что принадлежат другим людям, и из них составляется пара. Затем модель обучается на новых парах, и по тому же принципу составляются еще более сложные пары. Также были применены аугментации к тренировочному набору.

Для сравнения моделей был произведен поиск пары для 250 тестовых изображений среди 15000 снимков базы данных, результаты представлены на рисунке 1.

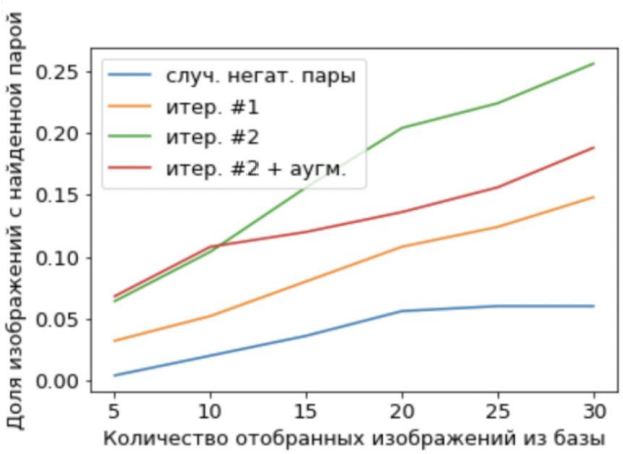

Рис 1. Сравнение нейросетевых моделей

Следующим подходом являлся «мешок визуальных слов», истоки которого лежат в области обработки текстов. Метод основан на поиске ключевых точек и их дескрипторов, в данной работе был выбран алгоритм SIFT (англ. Scale-Invariant Feature Transform) [3].

Дескрипторы различных точек могут описывать похожие ключевые точки, поэтому их можно кластеризовать каким-либо алгоритмом кластеризации получив словарь, элементами которого являются часто повторяющиеся элементы изображения. т.н. «визуальные слова». Для кластеризации в данной работе был выбран мини-пакетный метод *k*-средних (англ. mini batch *k*-means), который заключается в кластеризации векторов, случайно выбираемых на каждой итерации в заданном количестве [4]. Каждое изображение может быть представлено вектором-гистограммой частот визуальных слов.

Поиск изображений производился по алгоритму, аналогичному тому, что использовался в нейросетевом подходе. В качестве векторов признаков были взяты полученные гистограммы. Было проанализировано влияние количества кластеров на точность модели. Как видно на рисунке 2, точность оказалась не выше точности, полученной с нейросетевой моделью.

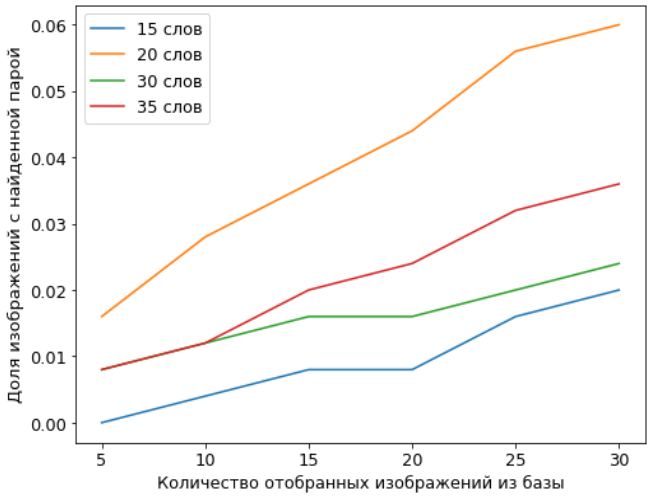

Рис 2. Сравнение моделей метода «мешка визуальных слов»

Вместо вычисления частот встречаемости дескрипторов ключевых точек, можно сравнивать изображения путем сопоставления ключевых точек двух снимков. В нашем случае использовался детектор ORB (англ. oriented FAST and rotated BRIEF) [5] и SIFT. Сопоставление точек происходило путем попарного сравнения дескрипторов обоих изображений.

Расстояние между снимками вычислялось как среднее расстояние между сопоставленными парами точек. Для вычисления среднего выбирались не все пары, так как в таком случае было сильное влияние шума. Поэтому на тренировочном наборе был произведен подбор параметра n – количества самых близких пар точек. Наилучших результаты удалось достигнуть при $n = 10$. Результаты представлены на рисунке 4.

Несмотря на высокую точность, у данного подхода есть существенные недостатки: большой объем занимаемой памяти для хранения дескрипторов изображений и низкая скорость поиска. Для оптимизации было уменьшено количество хранимых дескрипторов, что позволило сократить объем занимаемой памяти в 2.5 раза и получить ускорение в 2 раза. Для этого из 10000 тренировочных изображений были составлены случайные пары, найдены ключевые точки соответствующих изображений, их дескрипторы, и произведено сопоставление точек. Для точек из топ-10 с помощью ранее обученной модели k-средних были найдены кластеры. Оказалось, что абсолютное большинство таких точек принадлежит 7 кластерам из 20. Таким образом, для изображений базы данных хранились только дескрипторы принадлежащие данным кластерам.

Как видно на графике 3, ценой меньшего потребления памяти и скорости поиска точность снизилась.

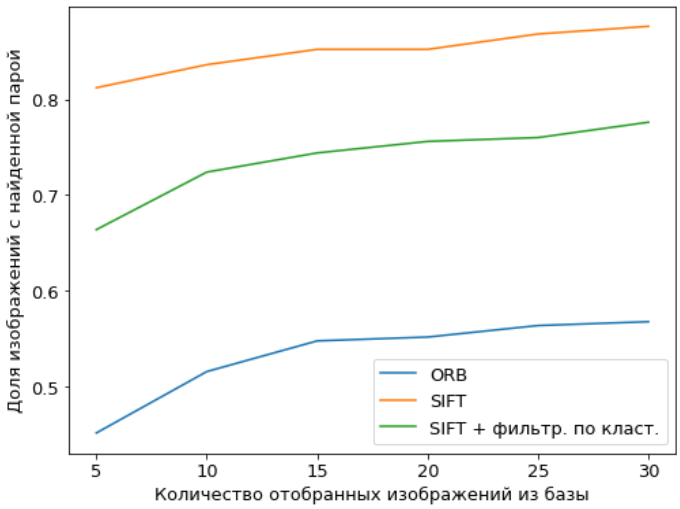

Рис 3. Сравнение моделей, основанных на сопоставлении ключевых точек

В данной работе были рассмотрены три подхода к сопоставлению изображений: нейросетевой, метод мешка визуальных слов и метод сопоставления ключевых точек. Также было произведено сравнение данных методов. Наилучших результатов удалось достигнуть путем сопоставления ключевых точек, такой метод показал точность 87,6% при поиске пары для 250 изображений среди базы данных в 15000 снимков при выборе топ-30 ближайших изображений. Полученные результаты можно использовать как основу для дальнейших исследований и улучшения алгоритма поиска «цифровых двойников».

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. – P. 378-383.
2. *Li M.* et al. Deep instance-level hard negative mining model for histopathology images // International Conference on Medical Image Computing and Computer-Assisted Intervention. – Springer, Cham, 2019. – P. 514-522.
3. *Xie B.* et al. An Image Retrieval Algorithm Based on Gist and Sift Features // Int. J. Netw. Secur. – 2018. – Vol. 20. – №. 4. – P. 609-616.
4. *Sculley D.* Web-scale *k*-means clustering // Proceedings of the 19th international conference on the World wide web. – 2010. – P. 1177-1178.
5. *Luo C.* et al. Overview of image matching based on ORB algorithm // Journal of Physics: Conference Series. – IOP Publishing, 2019. – Vol. 1237. – №. 3. – 032020.