

Методика разработки нейросетевых алгоритмов для решения задач обработки изображений

(часть 2)

Наталья Томашевич (Москва)

В продолжении статьи рассматриваются практические задачи на базе методики построения нейронной сети: 1) распознавание туберкулезных бацилл и 2) выделение лиц на цифровых изображениях. Обсуждаются результаты вычислительных экспериментов.

РЕШЕНИЕ ЗАДАЧИ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ТУБЕРКУЛЕЗНЫХ БАЦИЛЛ, ПОЛУЧЕННЫХ С ФЛЮОРЕСЦЕНТНОГО МИКРОСКОПА, С ПОМОЩЬЮ АНСАМБЛЯ НЕЙРОННЫХ СЕТЕЙ

Был разработан флюоресцентный микроскоп, получивший название микровидеотестер МВТ-ЛН (см. рис. 7). Он предназначен для компьютерного анализа патогенных микроорганизмов с использованием нейросетевого программного обеспечения.



Рис. 7. Флюоресцентный микроскоп, подставка для сканирования объектов, видеокамера и компьютер с нейросетевым программным обеспечением

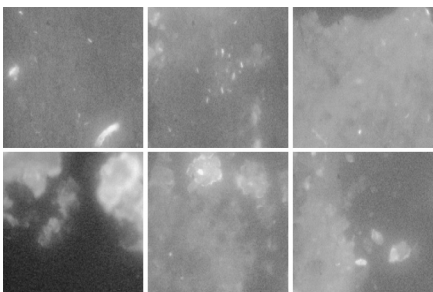


Рис. 8. Примеры анализа на туберкулез цифровых изображений, полученных с флюоресцентного микроскопа

Рассмотрим программное обеспечение этой системы, которое решает задачу выделения и подсчёта числа патогенных микроорганизмов на цифровом изображении, полученном с флюоресцентного микроскопа МВТ-ЛН.

1. Постановка задачи

Дано: Серое 256-уровневое изображение размера 244×260 пикселей, полученное с флюоресцентного микроскопа. Изображение получается в результате раскрашивания анализа слюны пациента флюоресцентным красителем, реагирующим на наличие туберкулезных микроорганизмов, и подаётся в растровом формате на компьютер. Изображение содержит различные объекты, которые можно разделить на два класса:

1. Туберкулезная палочка (имеет форму вытянутого эллипса, её длина может варьироваться примерно от 6 до 15 пикселей);
2. Другие объекты (имеют форму, отличную от эллиптической).

Требуется: Определить число палочек, указав их на изображении. К системе распознавания предъявляются следующие требования:

- инвариантность к сдвигу, повороту и незначительному изменению масштаба туберкулезной палочки (т.е. инвариантность к аффинным преобразованиям плоскости);
- инвариантность к яркости палочки. Палочка всегда светлее (ярче) фона, но не существует такого порога яркости, выше которого все точки изображения принадлежат палочкам;

- инвариантность к изгибу палочки.

Палочка представляет собой эллипсообразный объект, но, поскольку является грибом, может произвольно изгибаться.

Кроме того, система распознавания должна выполнять обработку исходного изображения за время не более 1,5 с на персональном компьютере со 128 Мб оперативной памяти и частотой процессора 400 МГц.

Примеры исходных изображений приведены на рисунке 8. На рисунке 9 приведены те же изображения, но с выделенными на них палочками (квадратики) и сомнительными объектами (кружочки). Является ли объект туберкулезной палочкой или нет, оператор определяет не по данному цифровому изображению, а глядя непосредственно в микроскоп. В отличие от серого и дискретного цифрового изображения, картинка в микроскопе цветная и непрерывная, что облегчает работу оператора.

3. Методика решения задачи и результаты

Как видно из примеров (см. рис. 8 и 9), помимо самих палочек, на исходном изображении присутствует большое число мелких и крупных ложных объектов. К крупным ложным объектам, как правило, относятся слизь или слюна пациента, а к мелким – кусочки красителя, выделяющего палочки, и различные непатогенные микроорганизмы.

Каждая палочка и каждый ложный объект исходного изображения имеют более или менее выраженные границы. Поэтому было бы хорошо, если бы из исходного серого 256-уровневого изображения мы получили чёрно-белое изображение, содержащее только палочки и небольшие ложные объекты, размеры которых сопоставимы с размера-

ми палочек. В этом случае задача распознавания палочек сводится к задаче выделения эллипсообразных объектов и вычисления отношения их длины к ширине, что несложно реализуется обычными методами без использования нейронных сетей. При отношении длины к ширине более 1,5...1,8 можно уверенно отнести объект к числу туберкулезных палочек.

Итак, рассмотрим методику построения нейронной сети, решающей задачу сегментации и фильтрации исходного серого изображения, в результате чего будет получено бинарное изображение, содержащие только палочки и небольшие ложные объекты. Эта нейронная сеть будет обеспечивать инвариантность к сдвигу, повороту, незначительному изменению масштаба и яркости объектов исходного изображения.

В каждой точке (x, y) исходного изображения будем рассматривать окно размером 15×15 точек (15 – это максимальная длина палочки), где (x, y) – центр этого окна; это обеспечит инвариантность к сдвигу. Для обеспечения инвариантности к повороту преобразуем квадратную пиксельную решётку рассматриваемого окна в круглую, как показано на рисунке 10. Затем каждый из получившихся восьми секторов круговой матрицы подаём на вход нейронной сети (НСДГ на рис. 10), выполняющей задачу определения

границ. Эта нейронная сеть также является инвариантной к яркости изображения.

Выходы восьми нейронных сетей НСДГ подаются на вход следующей нейронной сети (НСК на рис. 10), на выходе которой будет +1, если в рассматриваемом окне содержится палочка или небольшой ложный объект, и -1 – если другой объект. Таким образом, мы получаем бинарное изображение, на котором имеются только палочки и мелкие ложные объекты (на рисунке 11 в правом столбце они отмечены белым цветом, а фон – чёрным).

Затем остаётся классифицировать объекты получившегося в результате фильтрации и сегментации изображения. Для этого мы рассматриваем на полученном бинарном изображении только белые точки. Вокруг каждой белой точки сформируем окно и преобразуем это окно в круговую матрицу, так же как на рис. 10. Из восьми секторов круговой матрицы находим два не соседних сектора, содержащих максимальное число белых точек, которые расположены последовательно друг за другом при движении от центра кругового окна к его краям; это будет длина объекта, находящегося в окне. Аналогично находим ширину объекта и, вычислив отношение длины к ширине, классифицируем этот объект как палочку либо как ложный объект. Такая методика классификации позволяет

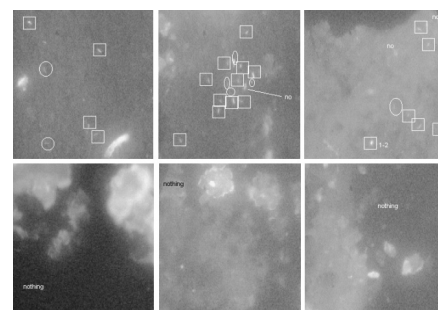


Рис. 9. Примеры цифровых изображений анализа на туберкулез

Квадратиками помечены туберкулезные палочки, кружочками – сомнительные объекты, на трёх нижних изображениях туберкулезные палочки отсутствуют

реализовать инвариантность к незначительному изменению масштаба палочки и к её изгибу.

Общее время решения задачи на компьютере с указанными выше характеристиками, в зависимости от количества палочек и мелких ложных объектов, составляет 1,1...1,3 с. В частности, для реализации нейросетевой части (фильтрации и сегментации исходного изображения) требуется 0,7 с.

Что касается качества решения задачи (см. рис. 11), то все палочки, различимые человеком на исходном изображении, были классифицированы правильно. Также были правильно классифицированы более 95% ложных объектов. Остальные объекты, которые на исходном изображении визуально не отличаются от палочек, распознаны не были¹.

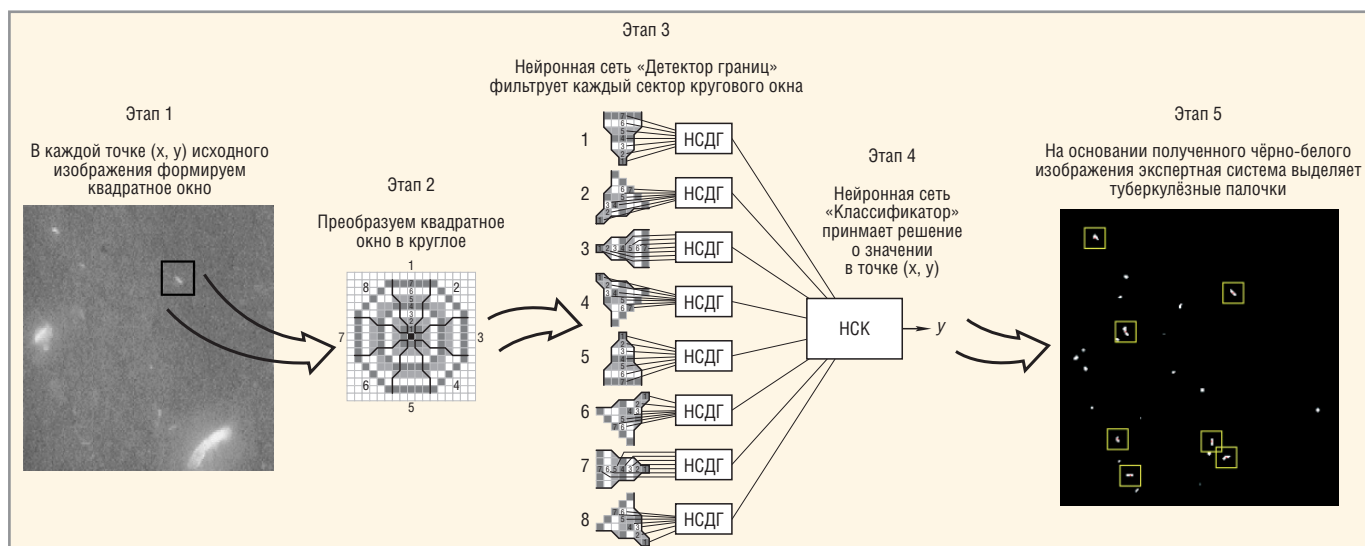


Рис. 10. Иллюстрация методики фильтрации и сегментации изображения на нейронной сети с реализацией инвариантности к сдвигу, повороту и яркости изображения

¹При преобразовании цветного непрерывного изображения на входе микроскопа в серое дискретное происходит потеря информации, поэтому по серой дискретной картинке на выходе микроскопа даже опытный специалист не всегда может классифицировать мелкий объект как туберкулезную бациллу или как ложный объект. – Прим. автора

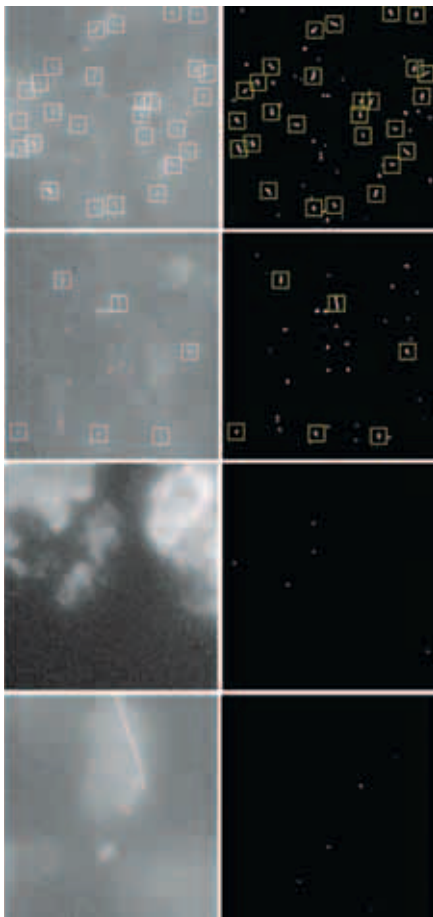


Рис. 11. Примеры решения задачи

Левый столбец содержит исходные изображения, на которых белыми квадратами помечены туберкулезные палочки. В правом столбце представлены решения системы для изображений левого столбца

ЭКСПЕРИМЕНТЫ ПО ОБУЧЕНИЮ НЕЙРОННОЙ СЕТИ ЗАДАЧЕ ВЫДЕЛЕНИЯ ЛИЦА НА ДВУМЕРНОМ ИЗОБРАЖЕНИИ

1. Нейросетевая постановка задачи выделения лиц

Традиционно задача выделения лица с помощью нейронной сети решается следующим образом (см. рис. 12). В каждой точке (x, y) исходного изображения (как правило, серой

фотографии) формируется окно размера $N \times N$. Из содержимого окна формируется вектор признаков, который подаётся на вход нейронной сети. Для формирования вектора признаков используются самые разные методы: от моментов и дискриминантного анализа до тех же самых нейронных сетей. До недавнего времени выделение вектора признаков из окна было необходимым этапом, позволяющим существенно уменьшить число входов нейронной сети. В наши дни, благодаря высокой производительности вычислительной техники, стало возможным использовать в качестве входа нейронной сети всё окно размера $N \times N$ целиком, что позволяет избежать потерь информации при преобразовании изображения окна в вектор признаков.

Итак, сформированный из окна вектор признаков подаётся на вход нейронной сети. На выходе нейронная сеть должна выдать +1, если окно содержит лицо человека, и -1 – если не содержит. То есть задача выделения лица сводится к задаче классификации с двумя классами образов. Поскольку в реальных задачах размер и ориентация искомого лица человека, как правило, неизвестны, то в каждой точке (x, y) формируется не одно, а множество окон разного размера и ориентации, которые и подаются на вход сети.

4.2. Эксперименты

В данной работе при построении нейронной сети преследуются такие цели:

1. Нейронная сеть должна обучаться всем примерам, предъявленным ей в обучающей выборке (процент ошибки на обучающей выборке должен быть равен нулю);

2. Обученная нейронная сеть должна обладать способностью к обобщению. То есть она должна правильно классифицировать новые примеры, в частности, новые изображения лиц людей, незначительно отличающиеся от изображений лиц, присутствующих в обучающей выборке.

Другими словами, если в обучающей выборке представлены примеры изображения чернокожих людей en face, то сеть должна правильно выделять и других чернокожих людей en face. Такие требования, предъявляемые к нейронной сети, позволят в будущем сформировать репрезентативную обучающую выборку и обучить нейронную сеть задаче выделения лиц со 100-% вероятностью.

Инвариантность к мимике лица и повороту головы достигается за счёт включения в обучающую выборку примеров лиц с разной мимикой и поворотами головы в 3D-пространстве. Инвариантность к масштабу и поворотам изображения на углы, кратные 90° , достигается за счёт формирования в каждой точке исходного изображения нескольких окон различного масштаба, повернутых на углы, кратные 90° . Поскольку внутри окна лица также могут иметь разный масштаб, то инвариантность к незначительному изменению масштаба и поворотам плоскости на углы, меньшие 45° , также обеспечивается нейронной сетью. Инвариантность к яркости изображения, как было сказано выше, включена в алгоритм построения нейронной сети. Инвариантность к различным источникам света реализуется за счёт включения в обучающую выборку соответствующих примеров.

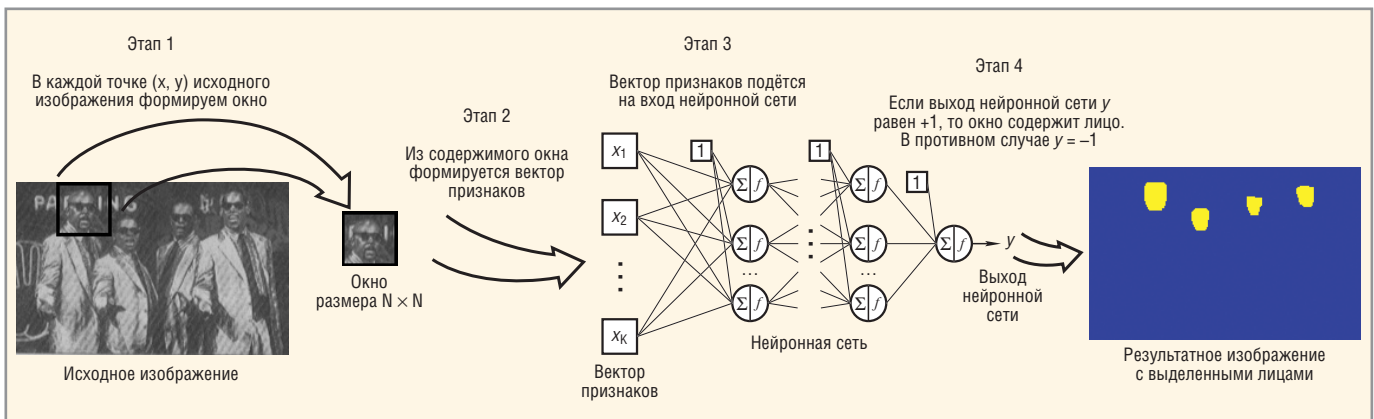


Рис. 12. Решение задачи выделения лица с помощью нейронной сети



Рис. 13. Результаты экспериментов

- (а) Исходное изображение. Обведены два лица, включенные в обучающую выборку;
 (б) Результат работы нейронной сети для рисунка (а). Чёрные точки соответствуют выходу сети -1 , т.е. классу «не лицо». Белые точки соответствуют выходу сети $+1$, т.е. классу «лицо»;
 (в) Исходное изображение. Обведены два лица, включенные в обучающую выборку;
 (г) Результат работы нейронной сети для рисунка (в). Чёрные точки соответствуют выходу сети -1 , т.е. классу «не лицо». Белые точки соответствуют выходу сети $+1$, т.е. классу «лицо».

В связи с недавним началом этой работы представлены результаты экспериментов на небольшой выборке. На рисунке 13а приведено исходное изображение, содержащее лица пяти разных белых людей. Два человека (их лица обведены) включены в обучающую выборку. На рисунке 13б приведён результат работы нейронной сети. Чёрные точки соответствуют выходу сети, равному -1 , т.е. клас-

су «не лицо». Белые точки соответствуют выходу сети, равному $+1$, т.е. классу «лицо». Видно, что нейронная сеть обучилась на всех примерах обучающей выборки. При этом она обладает хорошими обобщающими свойствами: при наличии в выборке изображений двух людей сеть правильно выделила трёх других людей.

На рис. 13в приведено другое исходное изображение, содержащее

лица четырёх различных темнокожих людей в солнцезащитных очках. В обучающую выборку также включены два обведённых лица. Как видно из рис. 13г, нейронная сеть, как и в предыдущем случае, обучена на всех примерах обучающей выборки и обладает хорошими обобщающими свойствами: двое людей, не включённых в выборку, правильно отнесены сетью к классу «люди». ©

**18-20 сентября
2007**



**Новосибирск
Россия**

**СИБИРСКИЙ ФОРУМ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ И ТЕЛЕКОММУНИКАЦИЙ**

СИБСВЯЗЬ. СИБКОМПЬЮТЕР. СИБИНТЕРНЕТ
XV международная специализированная выставка средств связи и телекоммуникаций,
коммуникационного оборудования, информационных технологий и компьютерной техники, сервисов Интернет

ЭЛЕКТРОНСИБ
Специализированная выставка электронных компонентов и технологического оборудования

СИБИРЬ-ТЕЛЕРАДИОВЕЩАНИЕ
Специализированная выставка кабельного и спутникового телевидения, радиовещания и
широкополосных телекоммуникаций

Выставочное Общество СИБИРСКАЯ ЯРМАРКА
 Россия, 630049, Новосибирск, Красный пр-т, 220/10. Телефон: (383) 210-62-90, факс: 225-98-45
 E-mail: ponkrat@sibfair.ru; www.sibfair.ru

Реклама