

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра компьютерных технологий и систем

АНАЛИЗ МИКРОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ

Курсовой проект

Пьянкова Егора Денисовича

студента 3 курса

специальность «прикладная
информатика»

Научный руководитель:

доктор технических наук,

заведующий кафедрой

компьютерных технологий и систем

Недзьведь Александр Михайлович

Оглавление

ВВЕДЕНИЕ	3
АКТУАЛЬНОСТЬ ПРОБЛЕМЫ	4
ГЛАВА 1. ПОЛУЧЕНИЕ И ОБРАБОТКА МИКРОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ	5
1.1 Микроскопические изображения.....	5
1.2 Компьютерное зрение.....	8
1.3 Видеопоследовательность и качество изображения.....	10
1.4 Стереопара.....	12
1.5 Оценка глубины изображения при помощи обычной камеры.	17
ГЛАВА 2. ЗАДАЧА РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЯ	18
2.1 Теория распознавания образов.....	18
2.2 Распознавание образа при помощи свёрточной нейронной сети.....	19
ГЛАВА 3. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ПОСТАВЛЕННЫХ ЗАДАЧ	20
3.1 Оценка глубины изображения при помощи обычной камеры.	20
3.2 Решение задачи распознавания изображения.	22
ЗАКЛЮЧЕНИЕ	25
ПРИЛОЖЕНИЕ.....	26
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	31

ВВЕДЕНИЕ

Достижения математики и информационных наук даёт существенный толчок прогрессу во многих других областях знаний. Как итог, все современные открытия происходят на стыке наук, исключением не стали биология и медицина.

Благодаря широкому использованию в медицине электронно-вычислительных комплексов значительно облегчается решение многих практических задач. Например, существует множество математических алгоритмов, применяющихся при обработке и анализе медицинских изображений и микроскопических изображений.

Современное развитие медицинских приложений часто опирается на анализ и обработку изображений. Существует целый ряд приложений по анализу трехмерных изображений и структур. Однако большинство оборудования для получения данных изображений основано на одном видеодатчике, то есть исходным для анализа является видеопоследовательность или же вовсе одно изображение.

Исходя из вышеизложенного, мы можем сформулировать конкретную цель, которую постараемся решить в рамках данного проекта: разработать приложения, позволяющие по минимальному набору входных данных получить информацию о клетке и оценить её 3D-форму.

Для достижения данной цели нам предстоит решить следующий набор задач:

1. Получить представление о современных направлениях в микроскопии.
2. Выбрать подходящий алгоритм для анализа входных данных.
3. Решить проблему распознавания вида клетки по фотографии.
4. Разработать прототип приложения.

АКТУАЛЬНОСТЬ ПРОБЛЕМЫ

Изображения анатомического и гистологического строения различных органов, тканей и клеток организма и их патологических изменений являются фундаментальными для медицинской науки. Диагностика заболеваний, лечение и управление терапевтическими процедурами опираются на данные, получаемые методами медицинской визуализации.

Общепатологические процессы протекают на различных уровнях: организменном, органном, тканевом, клеточном, ультраструктурном, молекулярном. Два первых уровня соответствуют макроскопическим изменениям, видимым невооруженным глазом, и предполагают изменения внешнего вида пораженных органов. На тканевом и клеточном уровнях патологические процессы выявляются с помощью светового микроскопа.

Приготовление препаратов для микроскопических исследований – сложный процесс, который влечёт за собой ухудшение качества изображений гистологических препаратов. Для преодоления трудностей, связанных с искажением изображения, требуются высокий профессионализм исследователя и выполнение сложной рутинной работы по корректировке результатов. Применение вычислительной техники значительно упрощает решение этих задач, облегчая работу с медицинскими изображениями.

Исходя из этого и формируется актуальность проблемы: достижение поставленной цели позволило бы снизить стоимость, увеличить качество и простоту анализа данных, полученных при помощи микроскопа.

ГЛАВА 1. ПОЛУЧЕНИЕ И ОБРАБОТКА МИКРОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ

1.1 Микроскопические изображения.

Для начала стоит разобраться, что же такое микроскопические изображения и как их получают. Ответы на эти вопросы может дать термин «микроскопия»

Микроскопия — это техническая область использования микроскопов для просмотра объектов и областей объектов, которые невозможно увидеть невооруженным глазом (объекты, которые не входят в диапазон разрешения нормального глаза).



Рисунок 1. Современный оптический тринокулярный микроскоп.

Современная микроскопия подразделяется на следующие виды:

1. Оптическая микроскопия (в свою очередь можно выделить ближнепольную оптическую микроскопию и инфракрасную микроскопию) – используется обыкновенный оптический микроскоп, знакомый многим ещё со школы.
2. Электронная микроскопия (включает в себя сканирующую (растровую) электронную микроскопию, просвечивающую электронную микроскопию и отражательную микроскопию) – применяется более усовершенствованная модель по сравнению с оптическим микроскопом, используется пучок электронов вместо светового потока.
3. Флуоресцентная микроскопия (выделяют подвид – двухфотонную лазерную микроскопию) – метод получения увеличенного изображения с использованием люминесценции возбуждённых атомов и молекул наблюдаемого образца.

4. Рентгеновская микроскопия (лазерная рентгеновская микроскопия) – метод, в котором применяется устройство для исследования очень малых объектов, размеры которых сопоставимы с длиной рентгеновской волны.
5. Сканирующая зондовая микроскопия (подразделяется на сканирующую туннельную микроскопию, атомносиловую микроскопию, ближнепольную оптическую микроскопию, магнитносиловую микроскопию, электросиловую микроскопию) – метод, в котором используется специальный класс микроскопов, предназначенных для получения изображения поверхности и её локальных характеристик. Процесс построения изображения основан на сканировании поверхности зондом.

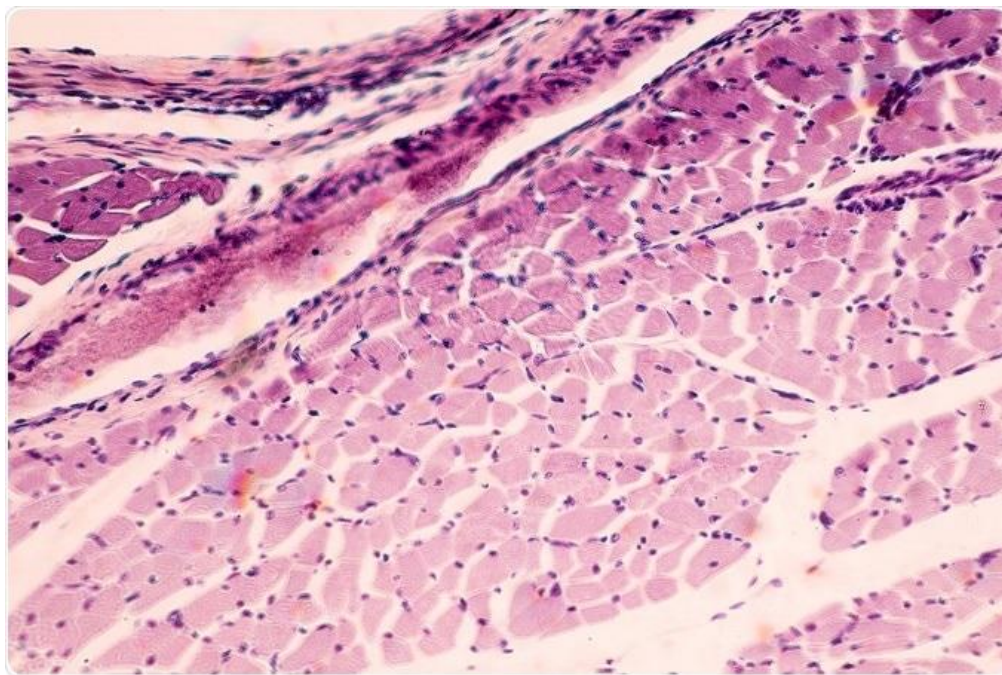


Рисунок 2. Поперечное сечение мышечного скелета под световым микроскопом.

В рамках данного проекта нас интересует только оптическая микроскопия и полученные при помощи оптического микроскопа изображения или видеопоследовательности, поэтому остановимся подробнее на ней.

Человеческий глаз представляет собой естественную оптическую систему, характеризующуюся определённым разрешением, то есть наименьшим расстоянием между элементами наблюдаемого объекта (воспринимаемыми как точки или линии), при котором они ещё могут быть отличены один от другого. Для нормального глаза при удалении от объекта на так называемом расстоянии наилучшего видения ($D = 250$ мм), среднестатистическое нормальное разрешения составляет $0,176$ мм. Размеры микроорганизмов, большинства растительных и животных клеток, мелких кристаллов, деталей микроструктуры металлов и сплавов и т. п. значительно меньше этой величины. Для наблюдения и изучения подобных объектов и предназначены оптические микроскопы различных типов.

В оптической микроскопии в настоящее время сделан прорыв, в результате которого преодолен фундаментальный рэлеевский критерий, заключающийся в том, что минимальный размер различимого объекта несколько меньше длины волны используемого света и принципиально ограничен дифракцией излучения. Это был предел

возможному в оптической микроскопии. До недавнего времени нельзя было преодолеть барьер, позволяющий различать структуры с расстоянием между элементами до 0,20 мкм.

Тем не менее выдающаяся последняя разработка оптической системы наноскопа с оптическим разрешением 10 нм расширила диапазон оптической микроскопии — наноскопии до десятков нанометров, что по сравнению с 0,20 мкм в 20 раз сократило расстояние между различаемыми элементами. (Например, размер белковых молекул, из которых состоит наш организм, колеблется от 3 до 10 нм).

Развитие видеотехники оказало существенное влияние на оптические микроскопы. Помимо упрощения документирования наблюдений электроника позволяет автоматизировать рутинные операции. А при отказе от непосредственного наблюдения глазом отпадает необходимость в классическом окуляре. В простейшем случае при модернизации микроскопа вместо окуляра устанавливается специальная оптическая конструкция для проецирования изображения на матричный фотоприёмник. Изображение фотоприёмника передаётся в ЭВМ и/или на дисплей.

Существуют также комбинированные профессиональные микроскопы, оснащённые третьим оптическим портом для установки фотоаппаратуры. В некоторых современных устройствах возможность прямого наблюдения глазом может отсутствовать полностью, что позволяет создавать простые и удобные в работе приборы компактного дизайна. Использование многоэлементных фотоприемников позволяет вести наблюдения не только в видимом, но и примыкающем к нему участках спектра.

Исходя из изложенного выше, будем считать задачу получения изображения решённой и сосредоточимся непосредственно на анализе полученных данных.

1.2 Компьютерное зрение.

Перед тем как приступить к анализу изображений при помощи компьютера, нам нужен какой-то набор методов, позволяющих компьютерам «видеть». Таким набором методов является компьютерное зрение – теория и технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов.

Компьютерное зрение также может быть описано как дополнение (но не обязательно противоположность) биологическому зрению. В биологии изучается зрительное восприятие человека и различных животных, в результате чего создаются модели работы таких систем в терминах физиологических процессов. Компьютерное зрение, с другой стороны, изучает и описывает системы компьютерного зрения, которые выполнены аппаратно или программно. Междисциплинарный обмен между биологическим и компьютерным зрением оказался весьма продуктивным для обеих научных областей. Подразделы компьютерного зрения включают воспроизведение действий, обнаружение событий, слежение, распознавание образов, восстановление изображений и некоторые другие.

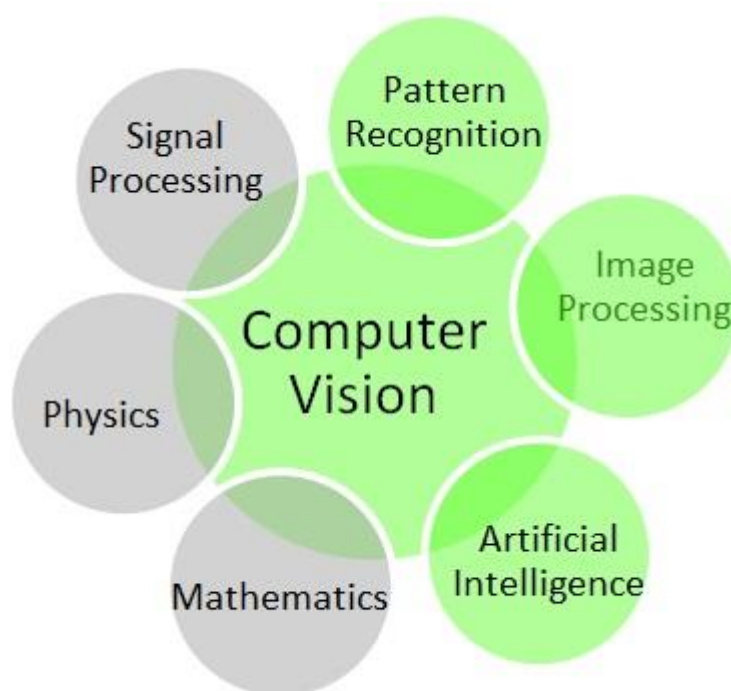


Рисунок 3. Основные компоненты, из которых состоит компьютерное зрение.

Область компьютерного зрения может быть охарактеризована как молодая, разнообразная и динамично развивающаяся. И хотя существуют более ранние работы, можно сказать, что только с конца 1970-х началось интенсивное изучение этой проблемы, когда компьютеры смогли управлять обработкой больших наборов данных, таких как изображения. Однако эти исследования обычно начинались с других областей, и, следовательно, нет стандартной формулировки проблемы компьютерного зрения. Также, и это даже более важно, нет стандартной формулировки того, как должна решаться проблема компьютерного зрения. Вместо этого, существует масса методов для решения

различных строго определённых задач компьютерного зрения, где методы часто зависят от задач и редко могут быть обобщены для широкого круга применения. Многие из методов и приложений все ещё находятся в стадии фундаментальных исследований, но всё большее число методов находит применение в коммерческих продуктах, где они часто составляют часть большей системы, которая может решать сложные задачи (например, в области медицинских изображений или измерения и контроля качества в процессах изготовления). В большинстве практических применений компьютерного зрения компьютеры предварительно запрограммированы для решения отдельных задач, но методы, основанные на знаниях, становятся всё более общими.

Важную часть в области искусственного интеллекта занимает автоматическое планирование или принятие решений в системах, которые могут выполнять механические действия, такие как перемещение робота через некоторую среду. Этот тип обработки обычно нуждается во входных данных, предоставляемых системами компьютерного зрения, действующими как видеосенсор и предоставляющими высокоуровневую информацию о среде и роботе. Другие области, которые иногда описываются как принадлежащие к искусственному интеллекту и которые используются относительно компьютерного зрения, это распознавание образов и обучающие методы. В результате, компьютерное зрение иногда рассматривается как часть области искусственного интеллекта или области компьютерных наук вообще.

Ещё одной областью, связанной с компьютерным зрением, является обработка сигналов. Многие методы обработки одномерных сигналов, обычно временных сигналов, могут быть естественным путём расширены для обработки двумерных или многомерных сигналов в компьютерном зрении. Однако из-за своеобразной природы изображений существует много методов, разработанных в области компьютерного зрения и не имеющих аналогов в области обработки одномерных сигналов. Особым свойством этих методов является их нелинейность, что, вместе с многомерностью сигнала, делает соответствующую подобласть в обработке сигналов частью области компьютерного зрения.

Помимо упомянутых подходов к проблеме компьютерного зрения, многие из исследуемых вопросов могут быть изучены с чисто математической точки зрения. Например, многие методы основываются на статистике, методах оптимизации или геометрии. Наконец, большие работы ведутся в области практического применения компьютерного зрения — того, как существующие методы могут быть реализованы программно и аппаратно или как они могут быть изменены с тем, чтобы достичь высокой скорости работы без существенного увеличения потребляемых ресурсов.

1.3 Видеопоследовательность и качество изображения.

В компьютерном зрении данные могут быть представлены могут быть представлены множеством форм, таких как видеопоследовательность, изображения с различных камер или трехмерными данными, например, с устройства Kinect или медицинского сканера.

Нас интересует именно первая форма, потому что её проще всего получить и легче всего с ней работать. Очевидно, что видеопоследовательность – это набор однотипных фотографий, сделанных через равный промежуток времени, представленный в виде цифрового потока.

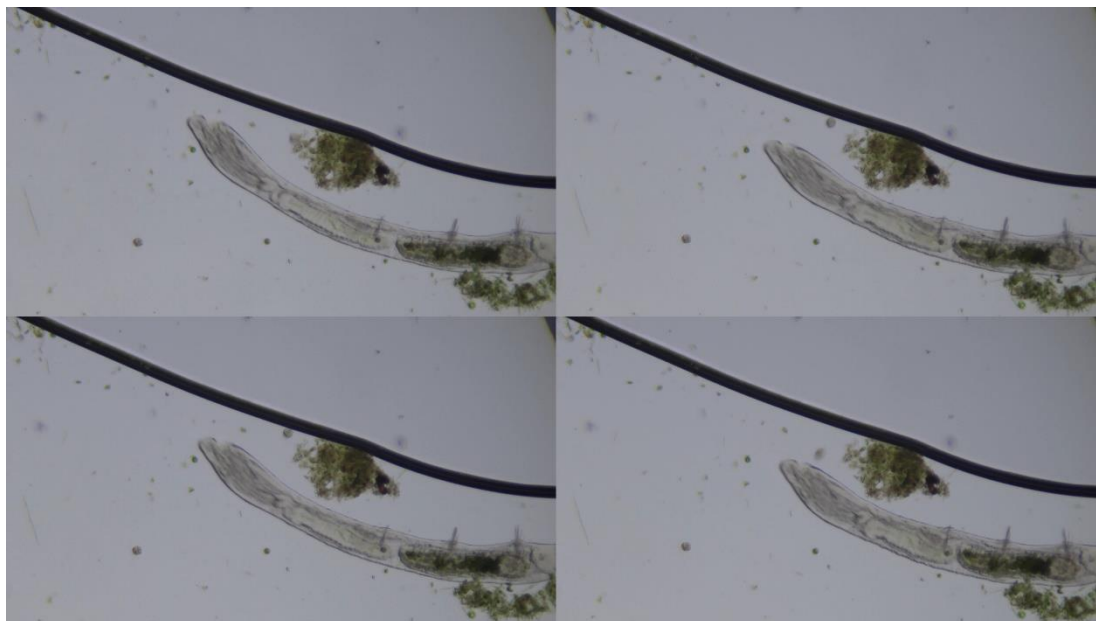


Рисунок 4. Пример выделения четырёх кадров из видеопоследовательности.

Теперь остановимся конкретнее на качестве получаемой фотографии. Качество изображения определяется разрешающей способностью микроскопа, т.е. минимальным расстоянием, на котором оптика микроскопа может различить отдельно две близко расположенные точки. разрешающая способность зависит от числовой апертуры объектива, конденсора и длины волны света, которым освещается препарат. Числовая апертура (раскрытие) зависит от угловой апертуры и показателя преломления среды, находящейся между фронтальной линзой объектива и конденсора и препаратом.

Угловая апертура объектива - это максимальный угол (АОВ), под которым могут попадать в объектив лучи, прошедшие через препарат. Числовая апертура объектива равна произведению синуса половины угловой апертуры на показатель преломления среды, находящейся между предметным стеклом и фронтальной линзой объектива. $N.A. = n \cdot \sin \alpha$ где, N.A. - числовая апертура; n - показатель преломления среды между препаратом и объективом; $\sin \alpha$ - синус угла α равного половине угла АОВ на схеме.

Таким образом, апертура сухих систем (между фронтальной линзой объектива и препаратом-воздух) не может быть более 1 (обычно не более 0,95). Среда, помещаемая между препаратом и объективом, называется иммерсионной жидкостью или иммерсией, а

объектив, рассчитанный для работы с иммерсионной жидкостью, называют иммерсионным. Благодаря иммерсии с более высоким показателем преломления чем у воздуха, можно повысить числовую апертуру объектива и, следовательно, разрешающую способность.

Разрешающая способность микроскопа зависит также от апертуры конденсора. Если считать апертуру конденсора равной апертуре объектива, то формула разрешающей способности имеет вид $R = \lambda / 2NA$, где R - предел разрешения; λ - длина волны; NA - числовая апертура. Из этой формулы видно, что при наблюдении в видимом свете (зеленый участок спектра – $\lambda = 550\text{нм}$), разрешающая способность (предел разрешения) микроскопа не может быть больше $0,2\text{мкм}$

Однако существует два способа повышения качества получаемого изображения:

Выбор большого угла светового конуса, как со стороны объектива, так и со стороны источника освещения. Благодаря этому, возможно, собрать в объективе более преломленные лучи света от очень тонких структур. Таким образом, первый путь повышения разрешения - это использование конденсора, числовая апертура которого соответствует числовой апертуре объектива.

Второй способ - использование иммерсионной жидкости между фронтальной линзой объектива и покровным стеклом. Так мы воздействуем на показатель преломления среды n , описанный в формуле выше. Его оптимальное значение, рекомендуемое для иммерсионных жидкостей, составляет 1.51.

1.4 Стереопара.

Итак, для начала следует разобраться, что же такое стереопара. Согласно Википедии, стереопара — пара плоских изображений одного и того же объекта (сюжета), имеющая различия между изображениями, призванные создать эффект объёма.

Данный эффект возникает в силу того, что расположенные на разном удалении от наблюдателя части сюжета при просмотре с разных точек имеют различное угловое смещение (параллакс).

Объекты с нулевым параллаксом наблюдаются с углом конвергенции ϕ_0 воспринимаются расположенными на удалении носителя изображения. Объекты отрицательным параллаксом наблюдаются с углом конвергенции ϕ_1 и кажутся выступающими вперед; объекты с положительным параллаксом наблюдаются углом конвергенции ϕ_2 и как находящиеся за плоскостью изображения.

Природа этого явления заключается в способности системы «глаза-мозг» обрабатывать зрительную информацию в некотором диапазоне $\Delta\phi$ ($\Delta\phi = \phi_1 - \phi_2$) угла конвергенции и объединять изображения в единый объемный образ. Максимальный диапазон $\Delta\phi$ составляет несколько градусов. Если параллакс изображений на стереофотографии превышает способности мозга по слиянию изображений, то комфортность просмотра нарушается, наблюдается двоение изображений. Поэтому при выполнении стереофотосъемки важно правильно выбирать величину стереобазиса.

Стереобазис – расстояния между точками съемки левого и правого кадра. Эта величина зависит от расстояния до фотографируемого объекта.

Существует 6 основных вариантов стереопар:

1. Горизонтальная стереопара – классическая стереопара, представляющая собой два изображения, расположенные горизонтально рядом друг с другом на расстоянии, обычно соответствующем межзрачковому расстоянию человека, хотя может и изменяться, в зависимости от рассматриваемого объекта. Например, для фотографии клетки, расстояние, очевидно, будет гораздо меньше. В свою очередь горизонтальная стереопара подразделяется на:
 - Параллельная стереопара - левое изображение предназначено для левого глаза, а правое изображение — для правого. Для рассматривания такой стереопары приходится направления взгляда глаз располагать параллельно, как если бы мы наблюдали очень удалённый объект, что требует специальных усилий при рассматривании стереопар.

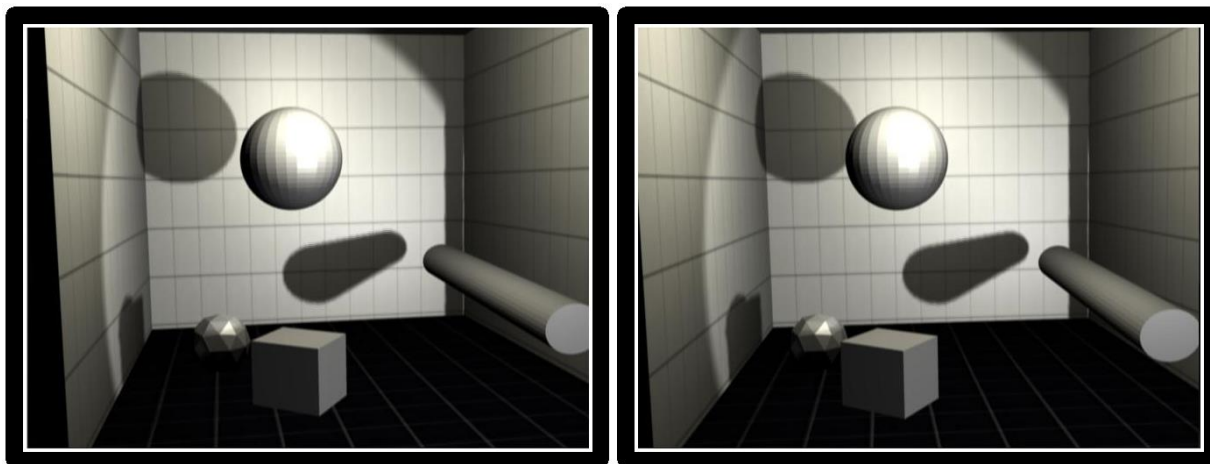


Рисунок 5. Параллельная стереопара.

- Перекрестная стереопара - левое изображение предназначено для правого глаза, а правое изображение — для левого. Для рассматривания такой стереопары приходится скрещивать глаза, как если бы мы наблюдали очень близкий объект, что проще, чем при рассматривании параллельных стереопар.

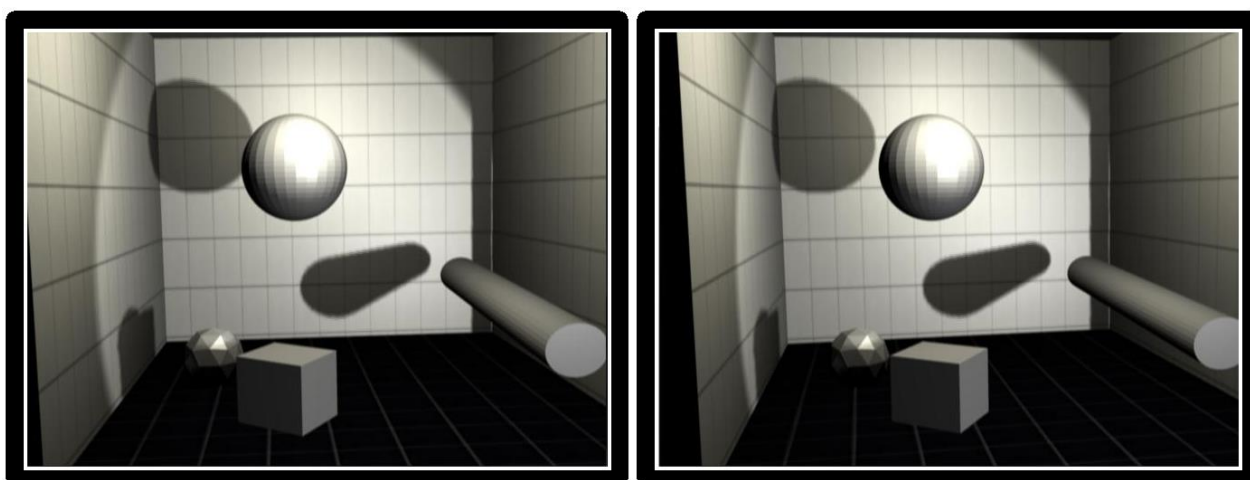


Рисунок 6. Перекрестная стереопара.

2. Вертикальная стереопара – представляет собой два изображения, расположенные друг над другом. Просмотр возможен в специальном плеере, преобразующем изображение для просмотра с любым типом стереочков.

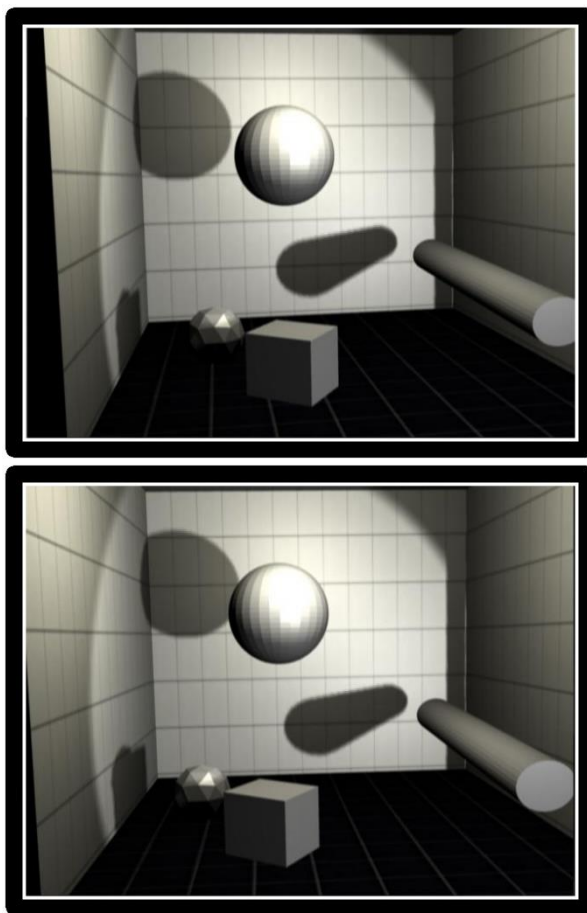


Рисунок 7. Вертикальная стереопара.

3. Чересстрочная стереопара – в четные строки вертикальной развертки записывается изображение одного ракурса, а в нечетные — другого. При таком методе пропадает половина вертикального разрешения у каждого ракурса, то есть разрешение фильма меняется, к примеру, с 720×480 до 720×240 точек.



Рисунок 8. Чересстрочная стереопара.

4. Анаглиф – изображение одного ракурса пропускается через один светофильтр, а изображение другого ракурса — через другой. Просмотр стереоизображения возможен через стереочки с цветными стеклами, соответствующими по цвету светофильтрам. Поскольку перспектива создаётся величиной отступа между версиями изображения, в глазу имитируется объёмная картинка.

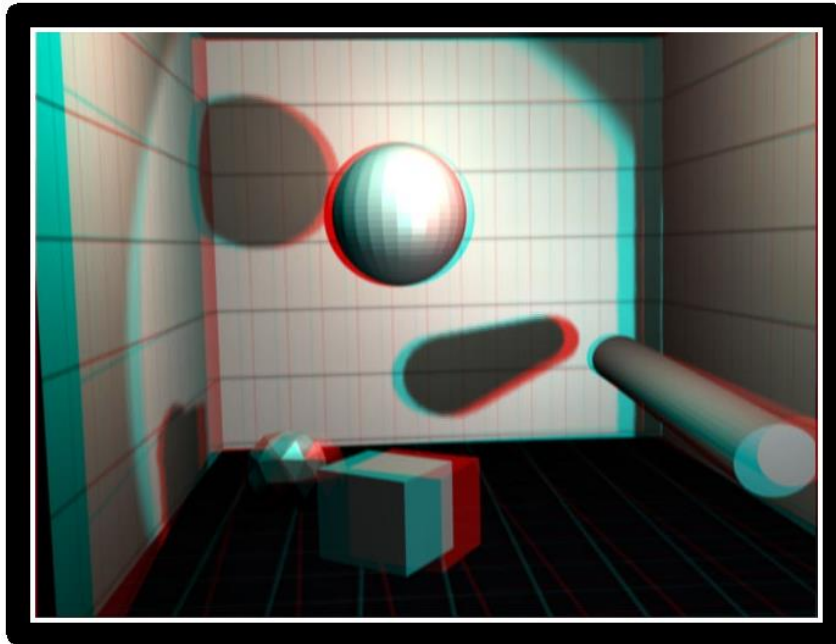


Рисунок 9. Анаглиф.

5. Поляризационный метод – использует пассивную поляризацию (поляризационные фильтры на источнике изображения и в очках), нужен стереомонитор либо два проектора и металлизированный экран.
6. Пэйдж-флип – изображения одного ракурса выводятся в четных кадрах, а изображение другого ракурса — в нечетных. Просмотр стереоизображения возможен через стерео-очки с ЖК-затворами закрывающимися попеременно для левого и правого глаза в такт кадрам.

Существуют два основных метода получения стереопары – параллельный и направленный методы фотосъемки. При параллельном способе направление оптической оси объектива не меняется, однако полученным при этом фотографиям требуется дополнительная обработка, называемая компенсационным сдвигом.

Компенсационным сдвигом называется относительное смещение изображений стереопары с целью достижения нулевого параллакса на одном из объектов съемки. При этом стереопарой является только та часть изображения, которая находится в зоне пересечения двух исходных изображений. Параллельный метод благодаря своей простоте является наиболее часто используемым. При получении стереопар таким способом далеко расположенные объекты не смещаются и на обоих кадрах находятся на одном и том же месте. Смещаются только объекты, расположенные на переднем плане, поэтому параллакс наблюдается только для переднего плана.

Второй способ получения стереоизображения – направленный – требует поворот камеры, однако полученные изображения обычно не требуют дополнительной обработки. Данный способ практически не реализуем для систем с одним датчиком видеозахвата.

Стереопара является исчерпывающим условием метрической реконструкции самых непредвиденных сложных объектов. Для этого необходимо знать фокусные расстояния, базу съемки и обеспечить параллельность осей, совпадение плоскостей кадров и синхронность работы затворов системы видеозахвата.

Стереопару для неподвижных объектов можно получать без специальных устройств, с помощью одного фотоаппарата или мобильного телефона с камерой. Для этого нужно сделать 2 кадра, сместившись на некоторое расстояние. Затем нужно оба кадра соединить на компьютере в один графический файл, так чтобы правое изображение было слева, а левое — справа. Тогда просмотр возможен без специальных очков на экране, это называется перекрёстная стереопара (см. выше). Для объектов, удалённых более чем на 400 м эффект нужно усиливать, если снимать разные фото не на расстоянии 6 см, что соответствует расстоянию между глазами, а сделать разницу больше. Тогда глубина пространства будет более ощутимой, особенно для пейзажей, они станут выглядеть как миниатюрные макеты. Для макросъёмки соответствующая база должна быть тем меньше, чем объект ближе.

Существуют программы, позволяющие создавать стереопары в разных форматах и обрабатывать их. Например, StereoPhoto Maker. Однако мы воспользуемся библиотекой OpenCV, потому что одновременно с составлением стереопары мы сможем решить задачи анализа изображений.

1.5 Оценка глубины изображения при помощи обычной камеры.

Существует специальная камера (давайте назовём её «глубинной»), предназначенная для оценки расстояния от фотографируемого объекта до самой камеры. Сразу возникает два логичных вопроса:

- 1) Каким образом данная камера получает информацию о глубине изображения?
- 2) Возможно ли получить подобный эффект при помощи обычной камеры?

Хороший пример «глубинной» камеры в повседневной жизни – Microsoft Kinect. В нём используется традиционная камера, совмещённая с инфракрасным датчиком, помогающим камере различать похожие объекты и вычислять их расстояние до объектива.

Однако не у всех есть доступ к Microsoft Kinect или другому дорогостоящему оборудованию. Именно здесь на помощь и приходит знаменитая библиотека OpenCV, позволяющая решить задачу определения глубины изображения при помощи обычной камеры. Стоит сделать замечание, что под обычной камерой мы подразумеваем камеру, не способную самостоятельно определить глубину изображения.

Предположим, что у нас есть простая камера, например, в нашем смартфоне. В этом случае нам придёт на помощь геометрия, а именно – эппипольная геометрия. Это особый тип геометрии, применяемый для решения проблем «стерео-зрения» - отвлечения компьютерного зрения, которое извлекает трёхмерную информацию из двух разных изображений одного и того же объекта. Как итог, мы получаем стереопару.

Мы ранее уже рассмотрели механизм получения стереопары, схематически он выглядит примерно следующим образом:

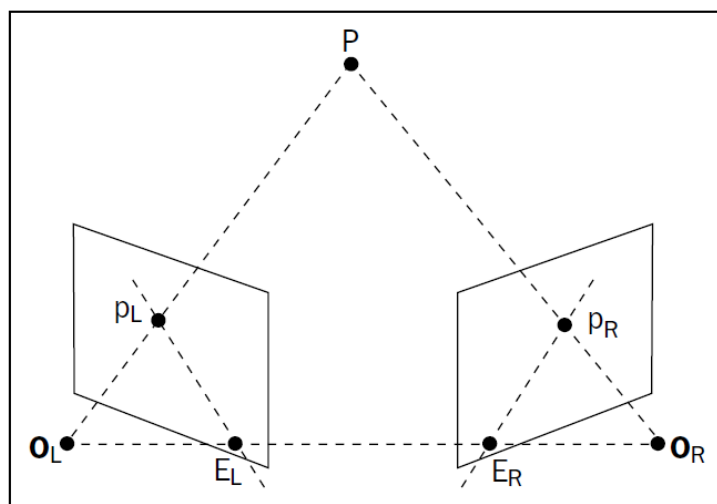


Рисунок 10. Визуальное представление механизма получения стереопары.

В библиотеке OpenCV уже реализованы методы эппипольной геометрии, позволяющие вычислить так называемую «карту несоответствий» - она представляет собой изображение, совокупность различных глубин, обнаруженных в двух изображениях. На этом теоретическая часть заканчивается, дело лишь за применением функций.

ГЛАВА 2. ЗАДАЧА РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЯ

2.1 Теория распознавания образов.

Немного отойдём от темы построения 3D-модели при помощи стереопары и вернёмся непосредственно к анализу самого изображения. Под анализом микроскопического изображения можно понимать множество вещей. В рамках данного проекта, хотелось бы решить, наверное, самую очевидную проблему – задачу распознавания вида клетки.

Возможно, эту проблему уже можно назвать классической в современной информатике, решением которой занимается теория распознавания образа – раздел информатики и смежных дисциплин, развивающий основы и методы классификации и идентификации предметов, явлений, процессов, сигналов, ситуаций и т. п. объектов, которые характеризуются конечным набором некоторых свойств и признаков.

Проблема распознавания образа приобрела выдающееся значение в условиях информационных перегрузок, когда человек не справляется с линейно-последовательным пониманием поступающих к нему сообщений, в результате чего его мозг переключается на режим одновременности восприятия и мышления, которому такое распознавание свойственно.

Неслучайно, таким образом, проблема распознавания образа оказалась в поле междисциплинарных исследований - в том числе в связи с работой по созданию искусственного интеллекта, а создание технических систем распознавания образа привлекает к себе всё большее внимание.

Рассмотрим некоторые методы распознавания графических образов, решаемых в рамках данной теории.

Для оптического распознавания образов можно применить метод перебора вида объекта под различными углами, масштабами, смещениями и т. д. Для букв нужно перебирать шрифт, свойства шрифта и т. д.

Второй подход — найти контур объекта и исследовать его свойства (связность, наличие углов и т. д.)

Ещё один подход — использовать искусственные нейронные сети. Этот метод требует либо большого количества примеров задачи распознавания (с правильными ответами), либо специальной структуры нейронной сети, учитывающей специфику данной задачи.

Именно на последнем подходе мы остановимся и рассмотрим его подробнее.

2.2 Распознавание образа при помощи свёрточной нейронной сети.

В рамках данного проекта нам интересно не столько теория искусственных нейронных сетей, сколько их практическое применение и решение поставленных задач, поэтому дадим определение свёрточной нейронной сети и причислим преимущества, которые делают данный вид нейронных сетей лучшим для распознавания образов.

Свёрточная нейронная сеть – специальная архитектура искусственных нейронных сетей, нацеленная на эффективное распознавание изображений, входит в состав технологий глубокого обучения (deep learning).

Основные преимущества:

1. Один из лучших алгоритмов по распознаванию и классификации изображений.
2. Меньшее количество настраиваемых весов, так как одно ядро весов используется целиком для всего изображения. Это подталкивает нейросеть при обучении к обобщению демонстрируемой информации, а не попиксельному запоминанию каждой показанной картинке.
3. Удобное распараллеливание вычислений, а, следовательно, возможность реализации алгоритмов работы и обучения сети на графических процессорах.
4. Относительная устойчивость к повороту и сдвигу распознаваемого изображения.
5. Обучение при помощи классического метода обратного распространения ошибки.

Однако самым главным недостатком свёрточной нейронной сети является большое количество варьируемых параметров сети и для новой задачи не существует какого-либо алгоритма, по которым нужно подбирать гиперпараметры обучения нашей нейронной сети.

Учёными из Оксфордского университета была разработана свёрточная нейронная сеть VGG16 (Visual Geometry Group) для распознавания объектов на изображениях, данная сеть состоит из 16 слоев. Именно её мы и будем использовать в практической реализации.

ГЛАВА 3. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ ПОСТАВЛЕННЫХ ЗАДАЧ

Перед нами стоит две задачи:

1. Оценить глубину изображения по стереопаре
2. Распознать клетку на фотографии

Обе эти задачи я постараюсь решить на языке программирования Python, так как для него существует огромное количество библиотек, которые позволяют без глубоких знаний в программировании решать множество

3.1 Оценка глубины изображения при помощи обычной камеры.

Для анализа стереопары мы будем использовать наиболее популярную библиотеку для компьютерного зрения – OpenCV. В ней есть специальный класс, предназначенный для работы со стереопарами, которым мы и воспользуемся.

Предположим, у нас есть два изображения, изъятых из видеопоследовательности, на которой запечатлено движение клетки. Иными словами, нам дана стереопара.

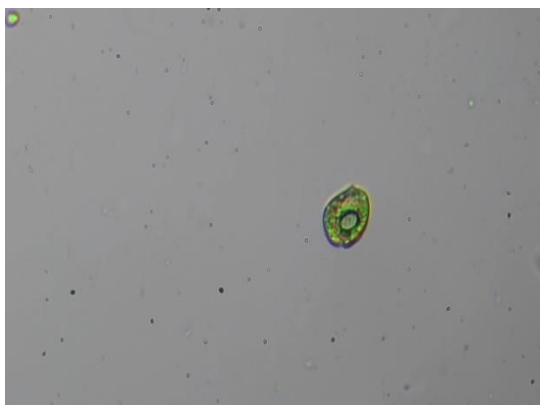


Рисунок 12. Левое изображение.

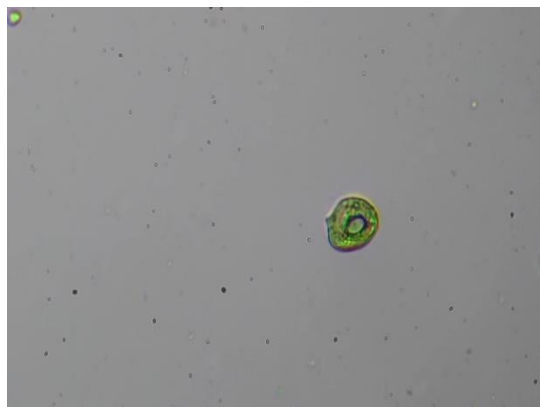


Рисунок 11. Правое изображение.

Если выполнить код из листинга 1, то мы получим трёхмерную характеристику стереопары – «карту» несоответствий и трёхмерное облако точек, дающее информацию о расположении объекта в пространстве. Например, оценка глубины для нашей пары изображений будет иметь вид:



Рисунок 13. Оценка глубины изображения.

К сожалению, из-за плохого качества фотографии на карте несоответствий появляются шумы, однако стоит отметить и то, что мы правильно смогли определить клетки и её подвижные части.

3.2 Решение задачи распознавания изображения.

Основным инструментом для решения нашей задачи является библиотека Keras – открытая нейросетевая библиотека, написанная на языке Python. В качестве бэк-энда для нашей библиотеки можно использовать фреймворки Theano или TensorFlow – они очень похожи, а разницы в работе программы не будет.

Как уже отмечалось ранее, мы воспользуемся свёрточной нейронной сетью VGG16. Данная нейронная сеть состоит из двух частей:

Первая часть сети выделяет характерные признаки в изображении. Состоит из чередующихся каскадов свертки и подвыборки. Сначала идут два каскада свертка-свертка-подвыборка, а затем три каскада свертка-свертка-свертка-подвыборка. Размер узла свертки во всех слоях 3x3. При подвыборке производится выбор максимального значения (Max Pooling) из квадрата размером 2x2.

Вторая часть отвечает за классификацию объекта на изображении по выделенным на предыдущем этапе признакам. Эта часть содержит 3 полносвязных слоя. В первых двух слоях 4096 нейронов, а в последнем - 1000 нейронов (по количеству классов объектов).

В листинге 2 содержится исходный код для создания и обучения нашей нейронной сети, запустим его и начинаем её обучать.

```
Using Theano backend.  
Found 2000 images belonging to 2 classes.  
Found 2000 images belonging to 2 classes.  
Epoch 1/10  
1/25 [>.....] - ETA: 227s - loss: 0.6982 - acc: 0.5000  
2/25 [=>.....] - ETA: 211s - loss: 1.6672 - acc: 0.5156  
3/25 [==>.....] - ETA: 196s - loss: 1.4385 - acc: 0.5312  
4/25 [===>.....] - ETA: 186s - loss: 1.4036 - acc: 0.5078  
5/25 [====>.....] - ETA: 176s - loss: 1.2646 - acc: 0.5375  
6/25 [=====>.....] - ETA: 167s - loss: 1.1496 - acc: 0.5573  
7/25 [=====>.....] - ETA: 157s - loss: 1.0531 - acc: 0.5848  
8/25 [=====>.....] - ETA: 148s - loss: 0.9756 - acc: 0.6172  
9/25 [=====>.....] - ETA: 140s - loss: 0.9189 - acc: 0.6354  
10/25 [=====>.....] - ETA: 130s - loss: 0.8745 - acc: 0.6469  
11/25 [=====>.....] - ETA: 122s - loss: 0.8601 - acc: 0.6335
```

Рисунок 14. Начало обучения нейронной сети.

Спустя 10 эпох мы получили нашу нейронную сеть, точность которой на данных для валидации достигла 87%, что является неплохим результатом.

```

Epoch 10/10
1/25 [>.....] - ETA: 257s - loss: 0.3473 - acc: 0.8438
2/25 [=>.....] - ETA: 240s - loss: 0.3534 - acc: 0.8438
3/25 [==>.....] - ETA: 229s - loss: 0.2767 - acc: 0.8854
4/25 [===>.....] - ETA: 219s - loss: 0.2528 - acc: 0.8906
5/25 [====>.....] - ETA: 218s - loss: 0.2623 - acc: 0.8875
6/25 [=====>.....] - ETA: 212s - loss: 0.2580 - acc: 0.8958
7/25 [=====>.....] - ETA: 199s - loss: 0.2650 - acc: 0.8929
8/25 [=====>.....] - ETA: 187s - loss: 0.2576 - acc: 0.8906
9/25 [=====>.....] - ETA: 175s - loss: 0.2773 - acc: 0.8750
10/25 [=====>.....] - ETA: 164s - loss: 0.2913 - acc: 0.8719
11/25 [=====>.....] - ETA: 152s - loss: 0.3035 - acc: 0.8636
12/25 [=====>.....] - ETA: 141s - loss: 0.2980 - acc: 0.8672
13/25 [=====>.....] - ETA: 131s - loss: 0.2978 - acc: 0.8630
14/25 [=====>.....] - ETA: 121s - loss: 0.2942 - acc: 0.8638
15/25 [=====>.....] - ETA: 109s - loss: 0.2838 - acc: 0.8688
16/25 [=====>.....] - ETA: 100s - loss: 0.2758 - acc: 0.8730
17/25 [=====>.....] - ETA: 89s - loss: 0.2716 - acc: 0.8768
18/25 [=====>.....] - ETA: 78s - loss: 0.2839 - acc: 0.8733
19/25 [=====>.....] - ETA: 66s - loss: 0.2828 - acc: 0.8734
20/25 [=====>.....] - ETA: 55s - loss: 0.2761 - acc: 0.8781
21/25 [=====>.....] - ETA: 43s - loss: 0.2820 - acc: 0.8765
22/25 [=====>.....] - ETA: 32s - loss: 0.2908 - acc: 0.8707
23/25 [=====>.....] - ETA: 21s - loss: 0.2974 - acc: 0.8682
24/25 [=====>.....] - ETA: 10s - loss: 0.2934 - acc: 0.8698
25/25 [=====] - 624s - loss: 0.2943 - acc: 0.8700 - val_loss: 0.3272 - val_acc: 0.8610

```

Рисунок 15. Завершение обучения нейронной сети и сохранение результата.

Теперь напишем ещё одну программу, которая бы загружала сохранённую нейронную сеть, получала бы изображение и говорила, что именно на ней изображено: миоцит или фибропласт. Данную программу можно найти в приложении, она находится в листинге 3.

Передадим в нашу нейронную сеть, например, следующую фотографию миоцита:

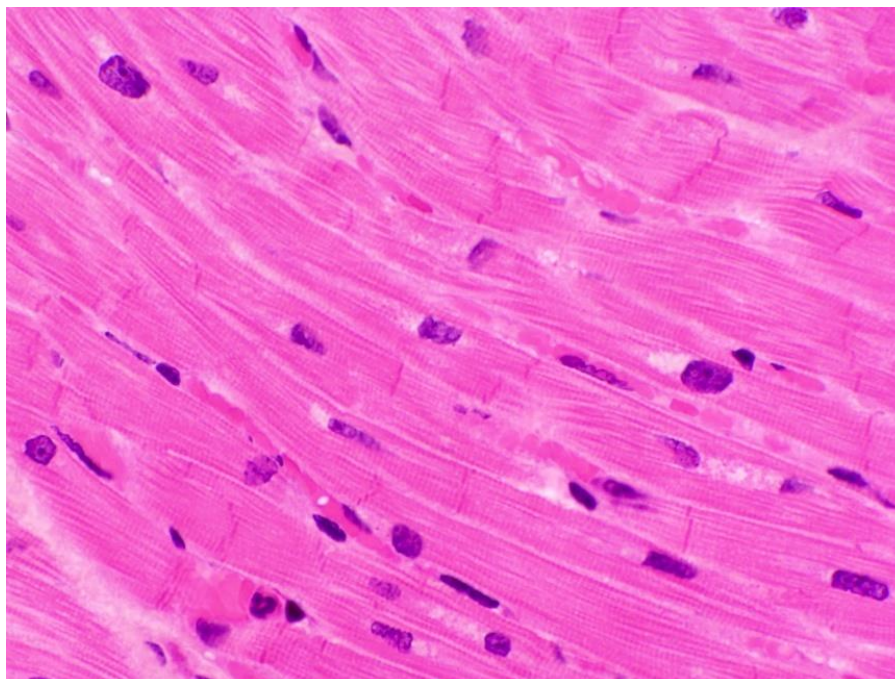


Рисунок 16. Изображение миоцита, которое нейронная сеть попытается распознать.

В результате выполнения программы мы получим следующее:

```
Using Theano backend.  
[[ 1.00000000e+00  4.05398448e-40]]  
миоцит  
  
Process finished with exit code 0
```

Рисунок 17. Результат распознавания изображения.

В квадратных скобках показаны вероятности того, что данная картинка относится к какой-либо из категорий. Первая – для миоцита, вторая – для фибропласта. Хотя мы и успешно смогли распознать изображение клетки, данная нейронная сеть ещё далека от идеала и может быть усовершенствована несколькими способами: увеличением числа классов распознаваемых клеток, расширением базы фотографий, увеличением количества фотографий для распознавания и выбором более оптимальных гиперпараметров обучения.

Что касается гиперпараметров обучения – это отдельная область знаний, которая позволяет выбрать наиболее подходящие значения для обучений нейронной сети (размер выборки, количество эпох обучения, функцию валидации и т.д.), однако увеличение гиперпараметров для обучения нейронной сети может негативно сказаться на скорости её обучения, которая может возрасти не на один порядок.

ЗАКЛЮЧЕНИЕ

В рамках данного курсового проекта, была сформулирована проблема анализа микроскопических изображений и подготовлена почва для создания полноценной системы для достижения поставленной цели. Методика создания стереопары в совокупности с алгоритмами сегментации динамических объектов, определения оптической оси и коррекции высот может позволить анализировать реальные изображения в медицине и биологии. Более того, было предложено решение проблемы распознавание изображения, которое позволит получить наиболее полную информацию о самой клетке и её состоянии.

ПРИЛОЖЕНИЕ

В этом разделе располагается весь исходный код. Его получилось не очень много, поэтому создавать и выкладывать ссылку на репозиторий не имеет особого смысла, гораздо проще разместить его прямо здесь.

Листинг 1. Оценка глубины стереопары при помощи библиотеки OpenCV.

```
import numpy as np
import cv2

ply_header = """ply
format ascii 1.0
element vertex %(vert_num)d
property float x
property float y
property float z
property uchar red
property uchar green
property uchar blue
end_header
"""

def update(val=0):
    stereo.setBlockSize(cv2.getTrackbarPos('window_size', 'disparity'))
    stereo.setUniquenessRatio(cv2.getTrackbarPos('uniquenessRatio', 'disparity'))
    stereo.setSpeckleWindowSize(cv2.getTrackbarPos('speckleWindowSize', 'disparity'))
    stereo.setSpeckleRange(cv2.getTrackbarPos('speckleRange', 'disparity'))
    stereo.setDisp12MaxDiff(cv2.getTrackbarPos('disp12MaxDiff', 'disparity'))
    print('computing disparity...')
    disp = stereo.compute(imgL, imgR).astype(np.float32) / 16.0
    cv2.imshow('left', imgL)
    cv2.imshow('disparity', (disp - min_disp) / num_disp)
    print('generating 3d point cloud...')
    h, w = imgL.shape[:2]
    f = 0.8 * w # guess for focal length
    q = np.float32([[1, 0, 0, -0.5 * w],
                    [0, -1, 0, 0.5 * h],
                    [0, 0, 0, -f],
                    [0, 0, 1, 0]])
    points = cv2.reprojectImageTo3D(disp, q)
    colors = cv2.cvtColor(imgL, cv2.COLOR_BGR2RGB)
    mask = disp > disp.min()
    out_points = points[mask]
    out_colors = colors[mask]
    out_fn = 'out.ply'
    write_ply(out_fn, out_points, out_colors)
```

```

def write_ply(fn, verts, colors):
    verts = verts.reshape(-1, 3)
    colors = colors.reshape(-1, 3)
    verts = np.hstack([verts, colors])
    with open(fn, 'wb') as f:
        f.write((ply_header % dict(vert_num=len(verts))).encode('utf-8'))
        np.savetxt(f, verts, fmt='%f %f %f %d %d %d ')

if __name__ == "__main__":
    window_size = 5
    min_disp = 16
    num_disp = 192 - min_disp
    blockSize = window_size
    uniquenessRatio = 1
    speckleRange = 3
    speckleWindowSize = 3
    disp12MaxDiff = 200
    P1 = 600
    P2 = 2400
    imgL = cv2.imread('images/image_L.jpg')
    imgR = cv2.imread('images/image_R.jpg')
    cv2.namedWindow('disparity')
    cv2.createTrackbar('speckleRange', 'disparity', speckleRange, 50, update)
    cv2.createTrackbar('window_size', 'disparity', window_size, 21, update)
    cv2.createTrackbar('speckleWindowSize', 'disparity', speckleWindowSize, 200, update)
    cv2.createTrackbar('uniquenessRatio', 'disparity', uniquenessRatio, 50, update)
    cv2.createTrackbar('disp12MaxDiff', 'disparity', disp12MaxDiff, 250, update)
    stereo = cv2.StereoSGBM_create(
        minDisparity=min_disp,
        numDisparities=num_disp,
        blockSize=window_size,
        uniquenessRatio=uniquenessRatio,
        speckleRange=speckleRange,
        speckleWindowSize=speckleWindowSize,
        disp12MaxDiff=disp12MaxDiff,
        P1=P1,
        P2=P2)
    update()
    cv2.waitKey()

```

Листинг 2. Обучение нейронной сети для распознавания 2-х видов клеток – миоцитов и фибропласт.

```

from keras import applications
from keras.layers import Dense

```

```

from keras.layers import GlobalAveragePooling2D
from keras.models import Model
from keras.preprocessing.image import ImageDataGenerator

# Размер изображений
img_width, img_height = 150, 150

# Путь к каталогу с изображениями для обучения
train_data_dir = 'data/train'

# Путь к каталогу с изображениями для валидации
validation_data_dir = 'data/validation'

# Количество изображений для обучения
nb_train_samples = 25

# Количество изображений для валидации
nb_validation_samples = 25

# Количество эпох
epochs = 10

# Размер выборки
batch_size = 32

# Загружаем сеть VGG16 без части, которая отвечает за классификацию
base_model = applications.VGG16(weights='imagenet', include_top=False)

# Добавляем слои классификации
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation='relu')(x)

# Выходной слой с двумя нейронами для классов "миоцит" и "фибропласт"
predictions = Dense(2, activation='softmax')(x)

# Составляем сеть из двух частей
model = Model(inputs=base_model.input, outputs=predictions)

# "Замораживаем" сверточные уровни сети VGG16
# Обучаем только вновь добавленные слои
for layer in base_model.layers:
    layer.trainable = False

# Компилируем модель
model.compile(optimizer='rmsprop', loss='categorical_crossentropy', metrics=['accuracy'])

# Создаем генератор данных для обучения
datagen = ImageDataGenerator(rescale=1. / 255)
train_generator = datagen.flow_from_directory(
    train_data_dir,
    target_size=(img_width, img_height),
    batch_size=batch_size,
    class_mode='categorical')

```

```

# Создаем генератор данных для валидации
validation_generator = datagen.flow_from_directory(
    validation_data_dir,
    target_size=(img_width, img_height),
    batch_size=batch_size,
    class_mode='categorical')

# Обучаем модель с помощью генератора
model.fit_generator(
    train_generator,
    steps_per_epoch=nb_train_samples,
    epochs=epochs,
    validation_data=validation_generator,
    validation_steps=nb_validation_samples)

# Генерируем описание модели в формате json
model_json = model.to_json()

# Записываем модель в файл
json_file = open("network.json", "w")
json_file.write(model_json)
json_file.close()

# Сохраняем веса модели
model.save_weights("network.h5")

```

Листинг 3. Использование ранее созданной нейронной сети для распознавания вида клетки.

```

import numpy as np
from keras.applications.vgg16 import preprocess_input
from keras.models import model_from_json
from keras.preprocessing import image

# Список классов
classes = ['миоцит', 'фибропласт']

# Загружаем обученную модель
json_file = open("network.json", "r")
loaded_model_json = json_file.read()
json_file.close()
loaded_model = model_from_json(loaded_model_json)
loaded_model.load_weights("network.h5")

# Компилируем модель
loaded_model.compile(optimizer='rmsprop', loss='categorical_crossentropy',
    metrics=['accuracy'])

# Загружаем изображение для распознавания
img = image.load_img('test/cell.jpg', target_size=(224, 224))

```

```
x = image.img_to_array(img)
x = np.expand_dims(x, axis=0)
x = preprocess_input(x)

# Запускаем распознавание
prediction = loaded_model.predict(x)
print(prediction)
print(classes[np.argmax(prediction)])
```

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. С.В. Абламейко, А.М. Недзведь Обработка оптических изображений клеточных структур в медицине
2. Minichino J., Howse J. - Learning OpenCV 3 Computer Vision with Python, Second Edition – 2015
3. <https://github.com/opencv/opencv/tree/master/samples/python>
4. <https://www.asozykin.ru/courses/nnp/python>
5. <https://gggdomi.github.io/keras-workshop/notebook.html>
6. Л. Шапиро, Дж. Стокман. Компьютерное зрение — М.: Бином. Лаборатория знаний, 2006.
7. <https://triaxes.com/ru/articles/3dimension-photograph1>