

Разработка системы компьютерного зрения: детектирования лиц.

Научный руководитель и консультант:
Кузнецов А.С.

Выполнил: Кабаева А.С.

Постановка задачи

Цель:

Разработка и тестирование программно-аппаратной системы поиска лица в видео потоке данных. Применяемый алгоритм детектирования должен обладать частичной инвариантностью к условиям съемки и искажениям изображений.

Методы исследования:

методы теории обработки сигналов, теории распознавания образов, теории вероятностей и компьютерного зрения. Реализация этих методов осуществлена в виде модели в системе MatLab 7.04 программного обеспечения на языке программирования C++.

Актуальность:

Обработка изображений и видео потока данных с применением математических алгоритмов детектирования используется при решении ряда важных задач с точки зрения обеспечения безопасности и повышения качества обслуживания. Вариантами применения подобных систем на текущий момент являются внедрение их в:

- системы безопасности повышенной надежности,
- системы контроля работоспособности,
- устройства бытовой техники.

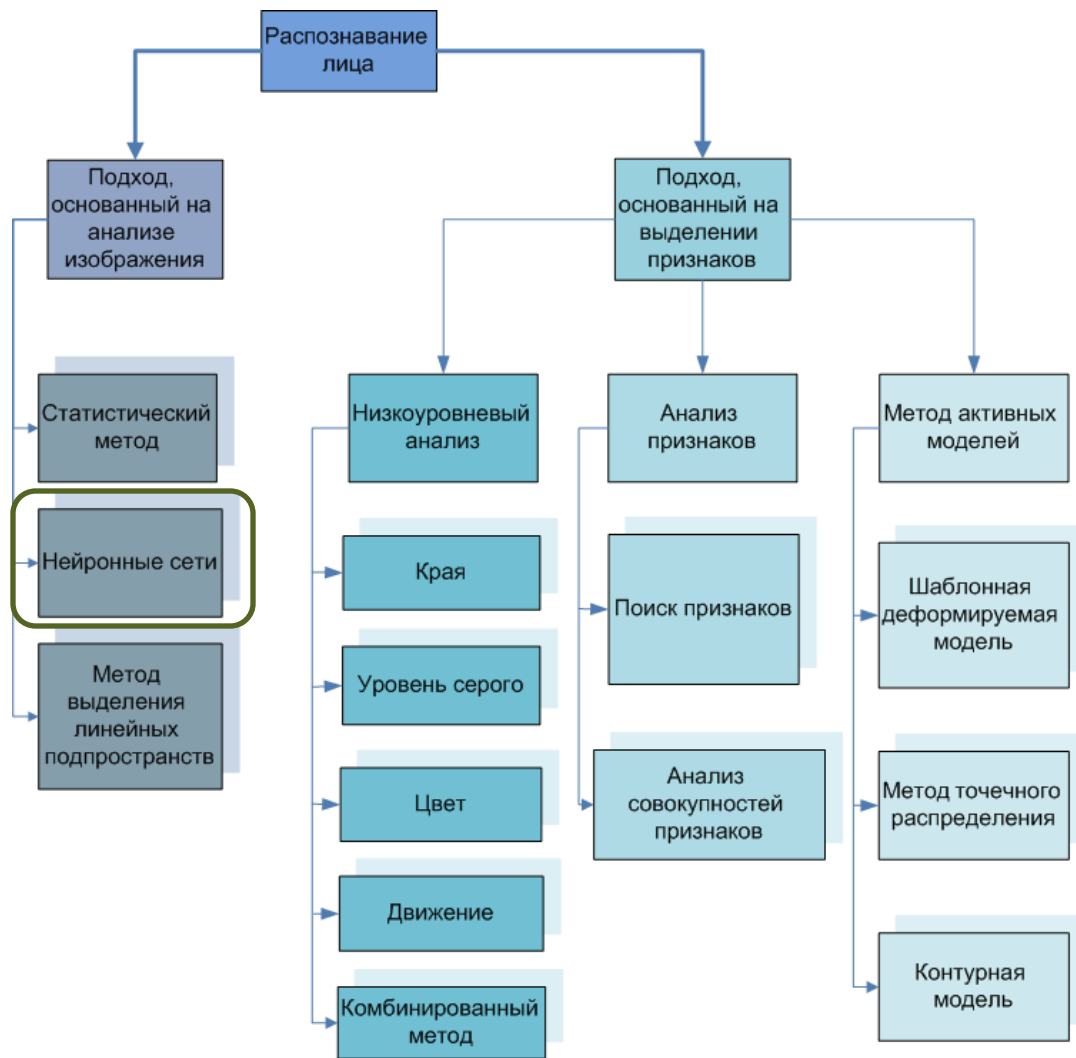
Подходы к решению задачи

Ограничения:

Система должна обладать инвариантностью к изменениям условий съемки (освещение, цветовой баланс камеры, искажение изображения, привносимые оптикой системы, качество изображения).

Решение:

Использование сверточных нейронных сетей.



Сверточные нейронные сети

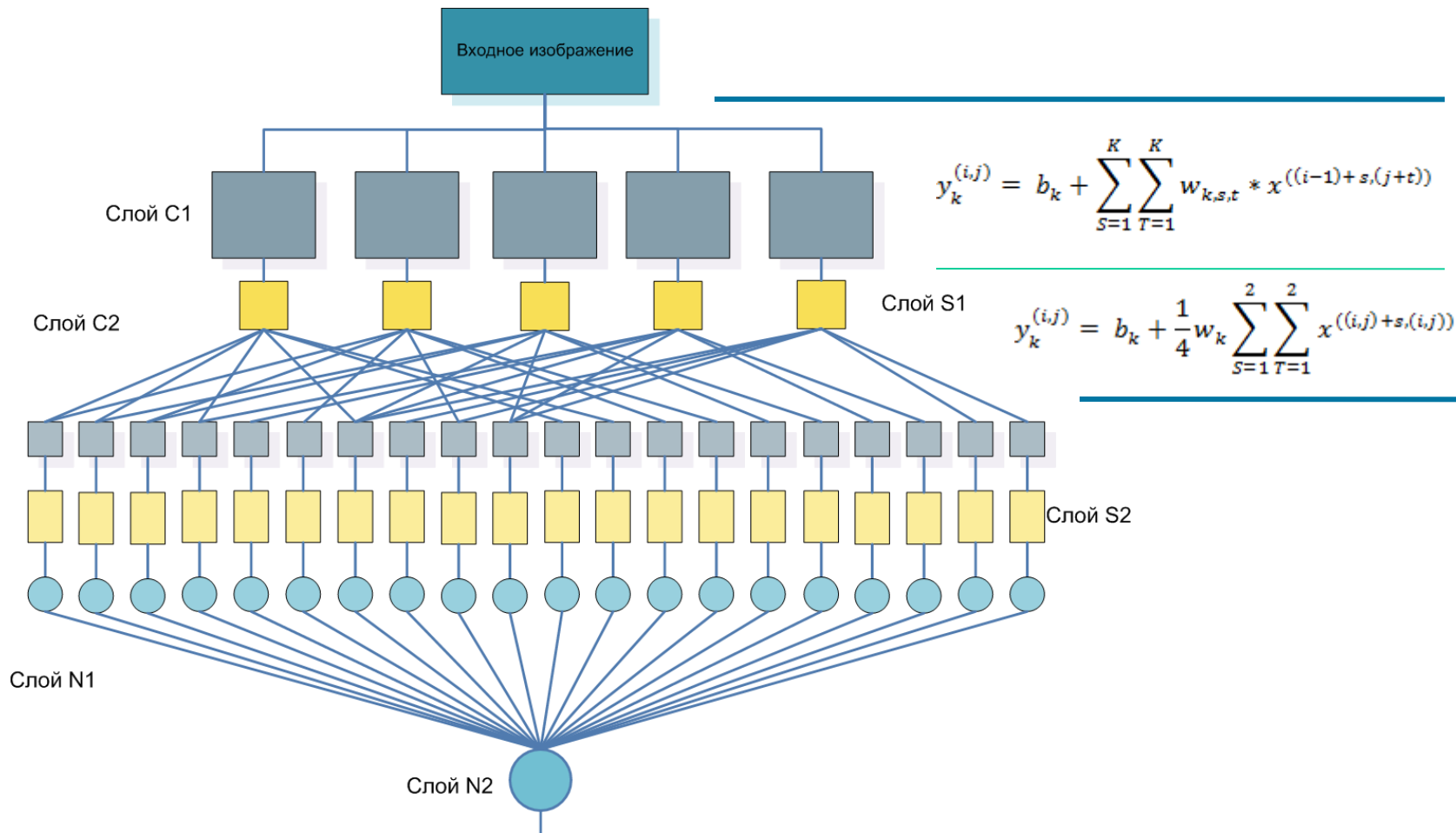
Разрабатываемая НС должна отвечать следующим основным параметрам:

1 - Топология нейронной сети должна обеспечивать вероятность обнаружения образа лица на уровне 0,95;

2 - Алгоритм обучения нейронной сети должен позволять обеспечивать эффективную работу нейронной сети при изменении углов поворота до $\pm 15^\circ$, уровня освещения в ~ 3 раза и изменении масштаба изображения в ~ 2 раза.



Топология сверточной нейронной сети



Топология СНС. Карты характеристик

C1- S1:

	1	2	3	4	5
1	█				
2		█			
3			█		
4				█	
5					█

S1- C2:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	█	█									█	█	█	█						
2			█	█							█				█	█	█			
3					█	█						█			█			█	█	
4							█	█					█			█		█		█
5									█	█				█			█		█	█

C2 - S2:

	1	2	...	19	20
1	█				
2		█			
...			█		
19				█	
20					█

S2 - N1

	1	2	...	19	20
1	█				
2		█			
...			█		
19				█	
20					█

N1- N2:

	1
1	█
2	█
...	█
19	█
20	█

Общее количество
настраиваемых синаптических
коэффициентов = 1351

Репрезентативность выборки:

Тестирование на базе данных ORL, содержащей изображения лиц с небольшими изменениями освещения, масштаба, пространственных поворотов, положения и различными эмоциями, показало приблизительно 98% точность распознавания, причём для известных лиц, предъявлялись варианты их изображений, отсутствующие в обучающем наборе.

Метод обучения:

Метод обратного распространения ошибки (обучение с учителем). Критерий – минимизация эмпирической ошибки. Выражения для коррекции синаптических коэффициентов имеет вид:

$$w_{i,j}(t+1) = w_{i,j}(t) + \eta \frac{dE}{dw_{i,j}}$$

Обучение НС

В качестве активационной функции необходимо использовалась гиперболический тангенс:

$$y = 1.7159 \tanh\left(\frac{2}{3}x\right)$$

Значения весовых коэффициентов были выбраны случайным образом из нормального распределения с нулевым средним и стандартным отклонением, где m – число связей, входящих в нейрон:

$$\sigma_w = \sqrt{m}$$

Обучение НС

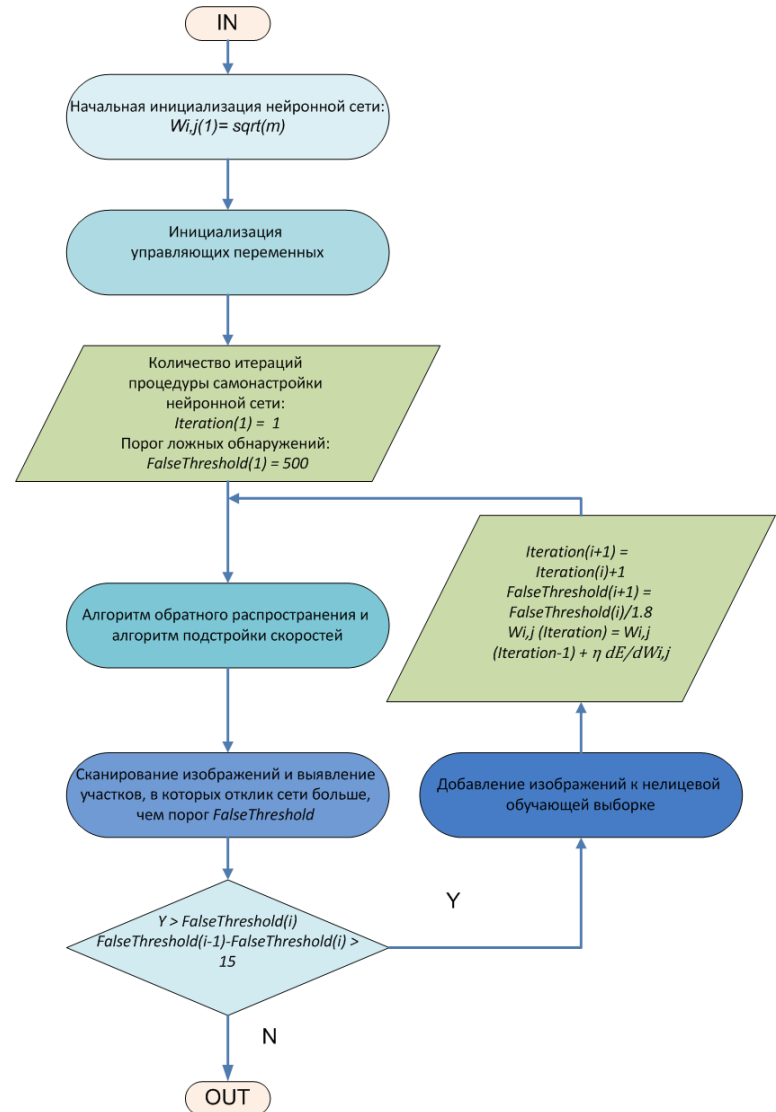
Шаг 1. Начальная инициализация нейронной сети. Создаются начальные обучающие выборки.

Шаг 2. Происходят начальная инициализация нейронной сети и инициализация управляющих переменных.

Шаг 3. Алгоритм обратного распространения с алгоритмом подстройки скоростей.

Шаг 4. Сканирование изображений и выявление участков, в которых отклик сети больше, чем порог.

Шаг 5. Полученные на шаге 4 изображения добавляются к нелицевой обучающей выборке. Уменьшение порога. Этими действиями к классифицирующим способностям нейронной сети предъявляются все более жесткие требования.



Обучение НС

Процесс обучения останавливается после пяти итераций, когда замечена сходимость процесса обучения, т.е. число ложных обнаружений остается примерно постоянным. Количество ложных лиц перед первой итерацией алгоритма самонастройки очень велико, но уже после первой итерации алгоритма самонастройки число ложных обнаружений уменьшается, а после 2 итераций существенно падает. Таким образом, нейронная сеть учится разделять образы на два класса.

Итоги обучения:



Количество лицевых образов	1200
Итоговое количество не лицевых образов	2100
Итоговая ошибка обучения	0,02

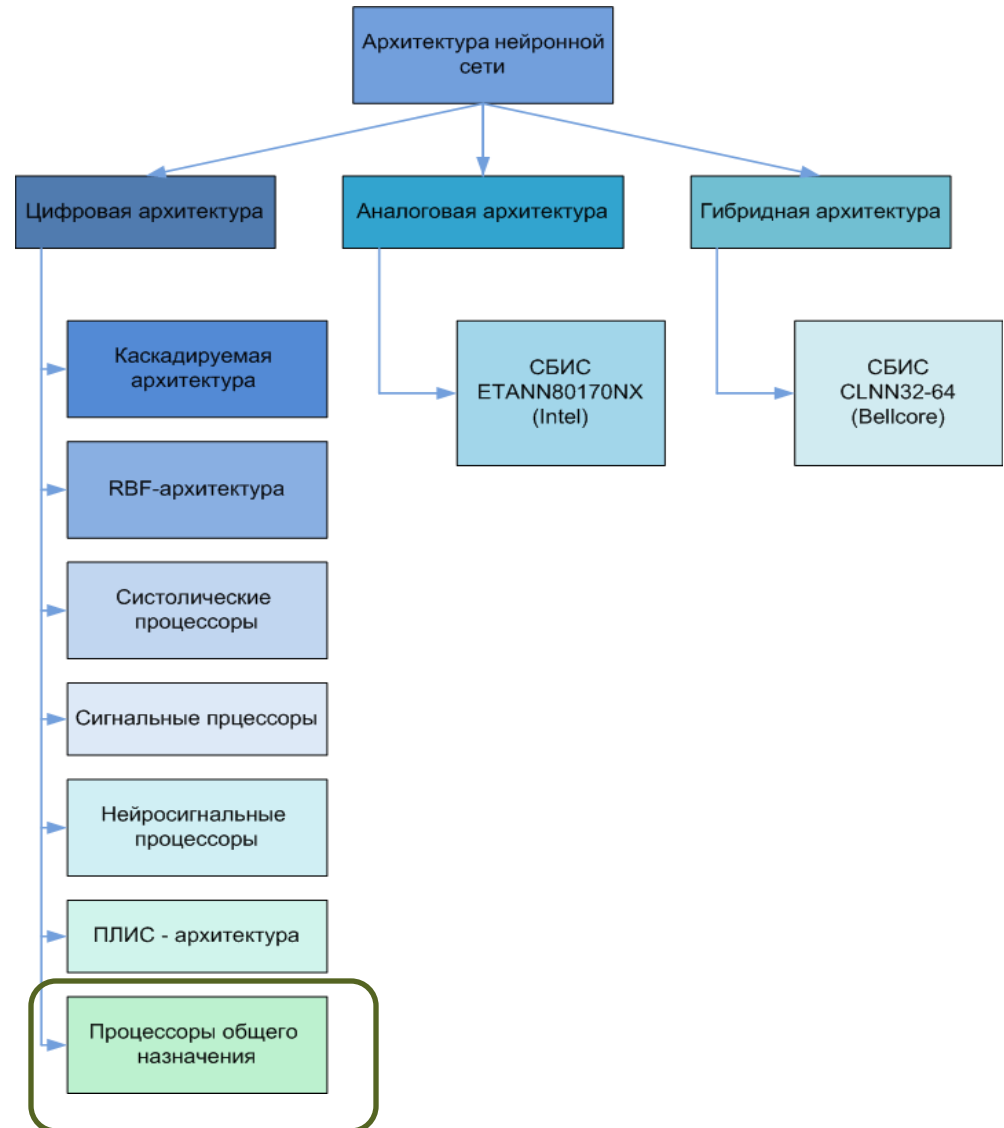
Аппаратная среда реализации

Ограничения:

- Возможность интеграции приложений с другими программными средствами
- Доступность
- Гибкость
- Масштабируемость
- Совместимость с другими процессорами
- Высокий ожидаемый экономический эффект

Решение:

PC - платформа.



Реализация

Алгоритм локализации

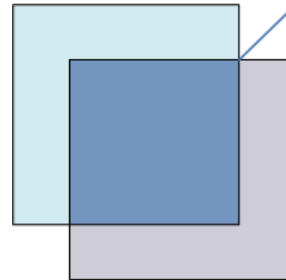
Разработанный алгоритм локализации лица можно представить в двух этапах:

1. сканирование различных масштабов входного изображения с большим постоянным шагом по ширине и по высоте участков-кандидатов, т.е. тех, в которых отклик нейронной сети больше некоторого порога;
2. повторное сканирование с малым шагом в некоторой окрестности выявленных участков и принятие решения о принадлежности их к классу лиц.

Оптимизация вычислений

Применение техники общих вычислений:

Общая область между двумя областями сканирования



Использование техники общих вычислений является важной особенностью сверточных нейронных сетей, достигается экономия вычисления порядка 40 раз.

Реализация

Дополнительный модуль статистики:

1. Анализ количества обнаружений лица на нескольких последних кадрах.
2. Наличие счетчика и «списка лиц»

Индексирование по времени:

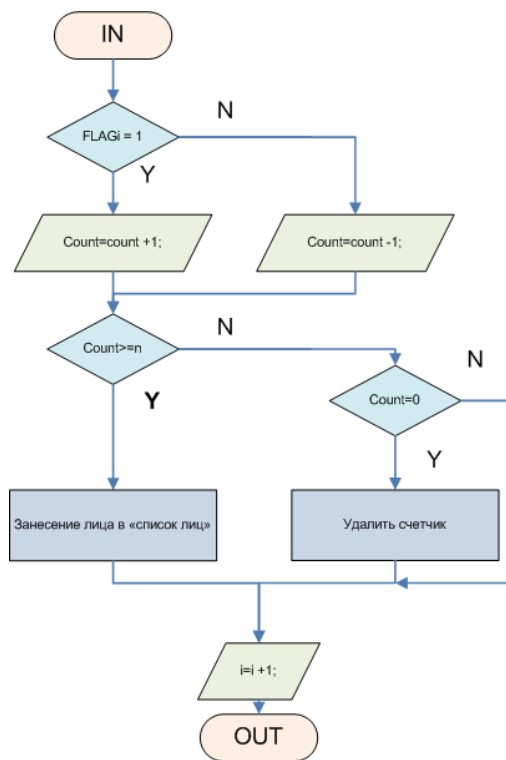
Временная сегментация основывается на проценте появившихся и исчезнувших лиц. Сравнение производится попарно между несколькими соседними кадрами :

$$TS = \frac{L}{N + L} + \frac{M}{N + M}$$

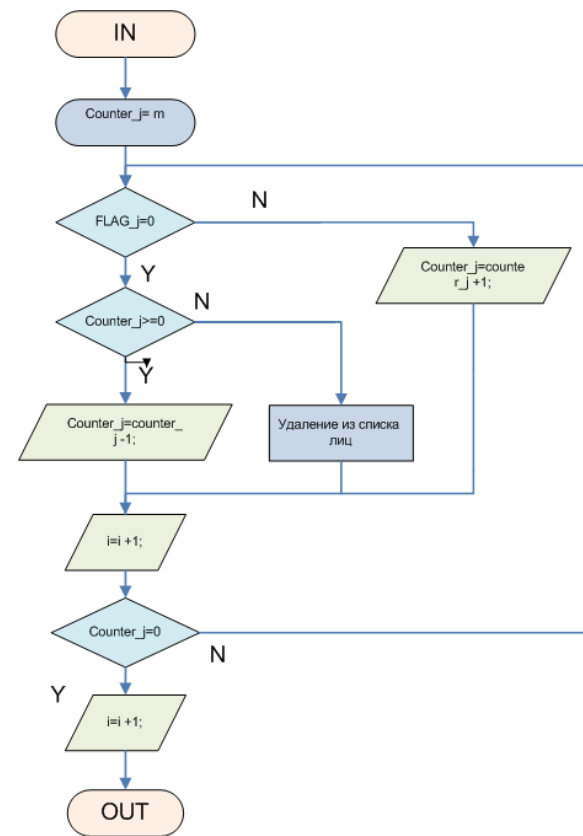
Где L/M – число пропавших/ появившихся лиц

N – число лиц, которые присутствовали в обоих кадрах.

Анализ обнаружений



Счетчик «списка лиц»



Реализация

Расчетные данные:

- Идентификатор индекса;
- Времена начала и конца показа кадра;
- Количество лиц и их координаты;
- Продолжительность сегмента с небольшим изменением состава присутствующих лиц в кадре.



```

D:\application\Indexes\test3\DEBATE\64x260_DEBATE.avi.xml
D:\application\Indexes\test3\DEBATE\64x260_DEBATE...
<FeatureData ID="0x12f0" StartTime="257933" StopTime="258000">0</FeatureData>
<FeatureData ID="0x12f0" StartTime="258000" StopTime="258066">0</FeatureData>
<FeatureData ID="0x12f0" StartTime="258066" StopTime="258133">1, 60x203 41</FeatureData>
<FeatureData ID="0x13f0" StartTime="258066">5400</FeatureData>
<FeatureData ID="0x12f0" StartTime="258133" StopTime="258200">1, 60x203 41</FeatureData>
<FeatureData ID="0x12f0" StartTime="258200" StopTime="258266">1, 60x203 41</FeatureData>
<FeatureData ID="0x12f0" StartTime="258266" StopTime="258333">1, 60x203 41</FeatureData>
<FeatureData ID="0x12f0" StartTime="258333" StopTime="258400">2, 60x203 41, 348x81 32</FeatureData>
<FeatureData ID="0x13f0" StartTime="258333">266</FeatureData>
<FeatureData ID="0x12f0" StartTime="258400" StopTime="258466">2, 61x204 40, 348x81 32</FeatureData>
<FeatureData ID="0x12f0" StartTime="258466" StopTime="258533">2, 62x202 38, 348x81 32</FeatureData>
<FeatureData ID="0x12f0" StartTime="258533" StopTime="258600">2, 64x202 40, 348x81 32</FeatureData>
<FeatureData ID="0x12f0" StartTime="258600" StopTime="258666">2, 348x81 32, 64x202 40</FeatureData>
<FeatureData ID="0x12f0" StartTime="258666" StopTime="258733">2, 346x79 32, 65x202 38</FeatureData>
<FeatureData ID="0x12f0" StartTime="258733" StopTime="258800">1, 70x200 37</FeatureData>
<FeatureData ID="0x13f0" StartTime="258733">400</FeatureData>
<FeatureData ID="0x12f0" StartTime="258800" StopTime="258866">2, 74x199 35, 258x214 33</FeatureData>
<FeatureData ID="0x12f0" StartTime="258866" StopTime="258933">2, 74x199 35, 258x214 33</FeatureData>
<FeatureData ID="0x12f0" StartTime="258933" StopTime="259000">2, 74x199 35, 258x214 33</FeatureData>
<FeatureData ID="0x12f0" StartTime="259000" StopTime="259066">2, 74x199 38, 258x214 33</FeatureData>
<FeatureData ID="0x12f0" StartTime="259066" StopTime="259133">2, 258x214 33, 81x198 36</FeatureData>
<FeatureData ID="0x12f0" StartTime="259133" StopTime="259200">0</FeatureData>
<FeatureData ID="0x13f0" StartTime="259133">400</FeatureData>
<FeatureData ID="0x12f0" StartTime="259200" StopTime="259266">0</FeatureData>
<FeatureData ID="0x12f0" StartTime="259266" StopTime="259333">0</FeatureData>
<FeatureData ID="0x12f0" StartTime="259333" StopTime="259400">0</FeatureData>
<FeatureData ID="0x12f0" StartTime="259400" StopTime="259466">0</FeatureData>
<FeatureData ID="0x12f0" StartTime="259466" StopTime="259533">0</FeatureData>
  
```

Тестирование системы

Методика эксперимента:

В результате проведения тестирования на 5 различных видео последовательностях, на которых в общей сложности находилось 21 лицо и с заданной надежностью на уровне 0,95.

	Истинное обнаружение	Ложное обнаружение	Пропуск лица
Вероятность	0.952	0.143	0
Точность	от 0,05 до 0,75	от 0,36 до 0,86	-

**Надежность полученных оценок при этом была на уровне 0,95.

Заключение

1. Разработанная система может использоваться как в качестве самостоятельного продукта, так и в качестве вспомогательного модуля для решения смежных задач детектирования лиц.
2. Программно реализованная топология нейронной сети гарантирует высокое качество работы детектора, а применяемые дополнительные алгоритмы позволяют существенно повысить скорость обработки данных.
3. Разработанная система прошла тестирование на реальных видео последовательных и подтвердила свое качество и надежность при эксплуатации.
4. В случае продвижения проекта на рынок прикладного программного обеспечения с привлечением заемных средств по ставке, не превышающей 20%, проект обеспечивает следующие экономические показатели:
 - Чистый доход составляет 609000 тыс. рублей;
 - Чистая недисконтируемая прибыль проекта составляет 25911 тыс. рублей;
 - Внутренняя норма рентабельности проекта составляет 37 %;
 - Недисконтируемый срок окупаемости проекта – 9 квартал (1-й квартал 2013 года).

Благодарю за внимание!

