

Метод эволюционного согласования в задаче распознавания лиц с помощью нейронных сетей

В.В. Плякин, В.И. Протасов

Vladplyakin@yandex.ru | protvlad@gmail.com

Московский физико-технический институт (НИУ), Москва, Россия

Проблема формализации и автоматизации процесса распознавания человеческих лиц была затронута еще на самых ранних стадиях развития систем распознавания образов и остается актуальной до сих пор. Более того в течение последних десяти лет количество научных исследований и публикаций по данной тематике выросло в несколько раз, что свидетельствует о возрастании актуальности этой проблемы. Это можно объяснить тем, что современная вычислительная техника открывает новые возможности ее применения в самых различных сферах, и соответственно появилось множество прикладных задач, требующих своего скорейшего разрешения. Одним из практических применений теории распознавания образов является распознавание лиц, в задачу которого входит автоматическая локализация лица на изображении и идентификация человека по лицу. Интерес к процедурам, лежащим в основе процесса локализации и распознавания лиц, довольно значителен в связи с разнообразием их практического применения в таких областях, как охранные системы, верификация, криминалистическая экспертиза, телеконференции, компьютерные игры и т.д. Например, система распознавания лиц, разработанная в пекинском университете Циньхуа (Tsinghua University), была сертифицирована китайским Министерством общественной безопасности для использования в общественных местах. Японское отделение компании Omron, специализирующееся на технологиях распознавания, автоматизации и управления, разработало систему распознавания лица человека для мобильных телефонов. Компания Riya, основанная группой специалистов в области алгоритмов распознавания черт лица из Стэнфордского университета, начала открытое тестирование Web сервиса контекстного поиска изображений лиц в цифровых фотоальбомах. Обилие подобных примеров свидетельствует о практической значимости и востребованности методов распознавания лиц.

Ключевые слова: эволюционное согласование, распознавание лиц, нейронные сети.

Evolutionary matching method for face recognition using neural networks

V.V. Plyakin, V.I. Protasov

Moscow Institute of Physics and Technology (NRU), Moscow, Russia

The problem of formalizing and automating the process of recognizing human faces was touched upon at the earliest stages of the development of image recognition systems and remains relevant to this day. Moreover, over the past ten years, the number of scientific studies and publications on this topic has increased several times, which indicates an increase in the urgency of this problem. This can be explained by the fact that modern computing technology opens up new possibilities for its application in various fields, and, accordingly, a lot of applied problems have appeared that require their speedy resolution. One of the practical applications of the pattern recognition theory is face recognition, the task of which is to automatically localize a face in an image and identify a person by face. The interest in the procedures underlying the process of localization and face recognition is quite significant due to the variety of their practical applications in areas such as security systems, verification, forensic examination, teleconferences, computer games, etc. For example, the face recognition system developed at Beijing Tsinghua University has been certified by the Chinese Ministry of Public Security for use in public places. Omron Japan, which specializes in recognition, automation and control technologies, has developed a human face recognition system for mobile phones. Riya, founded by a group of specialists in facial recognition algorithms from Stanford University, has begun open testing of a Web service for contextual search of facial images in digital photo albums. The abundance of such examples indicates the practical importance and relevance of face recognition methods.

Ключевые слова: evolutionary matching, face recognition, neural networks.

1. Введение

В настоящее время все более широкое распространение получают биометрические системы идентификации. Такие системы основываются на учете уникальных биологических характеристик человека, которые однозначно определяют его образ и являются трудно подделываемыми. К основным биометрическим характеристикам относятся отпечатки пальцев, форма ладони, узор радужной оболочки, изображение лица. Распознавание человека по изображению лица наиболее распространено в биометрических системах благодаря тому, что:

1. Не требуется специальное или дорогостоящее оборудование.
2. Не нужен физический контакт с какими-либо устройствами ввода данных.

Сегодня технология распознавания лиц применяется в различных сферах:

- государственная безопасность (выявление нарушителей по имеющейся базе, пропуск на режимный объект и т. д.);
- маркетинг (анализ эмоций пользователя для выявления заинтересованности в той или иной рекламе);
- мобильные приложения (гримировка под различных персонажей);
- и некоторых других

В некоторых случаях необходимо обеспечить высокую надежность идентификации, и для этого требуется применять системы, использующие различные биометрические характеристики (так называемые мультимодальные системы). При построении подобных биометрических систем целесообразно применять нейросетевые методы распознавания лица.

Под распознаванием будем понимать установление тождественности неизвестного лица известной персоне на основании совпадения признаков. То есть понятие “распознавание” может быть определено как отнесение исследуемого объекта (изображения лица), задаваемого в виде совокупности наблюдений, к одному из взаимоисключающих классов (персоны), или заключение о том, что этот объект не относится к

известным классам.

На данный момент существует множество различных методов реализации данной технологии, но чаще всего в их основе лежит обобщенный процесс распознавания.

Для начала следует выделить общий процесс обработки лица при распознавании [1] (рис. 1).

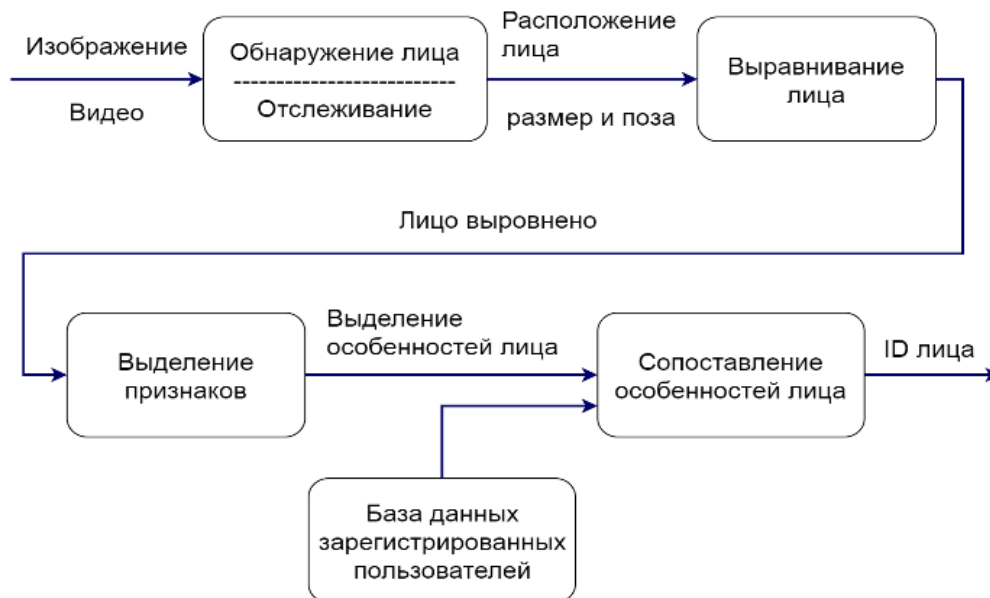


Рис.1. Общий процесс обработки лица при распознавании

На рис. 1 видно, что сначала производится обнаружение лица, определяются его расположение и размер. Далее происходит выравнивание лица, определение его признаков и выделение особенностей. На последнем этапе происходит сравнение особенностей лица с базой данных известных изображений.

К недостаткам распознавания человека по изображению лица можно отнести:

1. Сама по себе такая система не обеспечивает надежность идентификации на 100%. Система может ошибаться из-за таких факторов как: признаки старения, проявлений мимики, особенностей освещения и угла съемки.
2. Доля ошибок будет также расти с увеличением объема базы изображений лиц. Настройки для минимизации ошибок для каждой базы индивидуальны.

При построении подобных биометрических систем целесообразно применять нейросетевые методы распознавания лица.

Нейронная сеть (далее - НС) состоит из элементов, называемых формальными нейронами, каждый из которых элементарен по структуре и связан с другими нейронами. Каждый нейрон преобразует совокупность сигналов, поступающих к нему на вход, в выходной сигнал. Именно связи между нейронами, кодируемые весовыми коэффициентами, играют ключевую роль. Одно из основных преимуществ НС заключается в возможности параллельного функционирования ее элементов, что существенно повышает эффективность

решения задачи.

На сегодняшний день разработано множество видов нейронных сетей.

Многослойные нейронные сети (МНС) состоят из последовательно соединенных слоев, нейрон каждого из которых своими входами связан со всеми нейронами предыдущего слоя, а выходами - с нейронами последующего (рис. 2). Для активации таких нейронов служат разновидности линейных, пороговых и сигмоидных функций [2]. На рис. 2 представлена архитектура многослойной нейронной сети для распознавания изображений. Нейрон с максимальной активностью (цифра 1) указывает на принадлежность к распознанному классу. НС с одним решающим слоем способна формировать линейные разделяющие поверхности, что значительно сужает круг решаемых задач, в частности, такая сеть не сможет решить задачу типа „исключающее или“. НС с нелинейной функцией активации и двумя решающими слоями позволяет формировать любые выпуклые области в пространстве решений, а с тремя решающими слоями - области любой сложности, в том числе и невыпуклой. Обучение МНС осуществляется с помощью алгоритма обратного распространения ошибки. Такой алгоритм является разновидностью градиентного спуска в пространстве весов и обеспечивает минимизацию суммарной ошибки сети:

$$\Delta W = -\alpha \frac{dE}{dW}, E = \frac{1}{2} \sum_j (y_j - t_j)^2,$$

где y_j - выходное значение j -го нейрона сети, t_j - эталонное значение выходов сети.

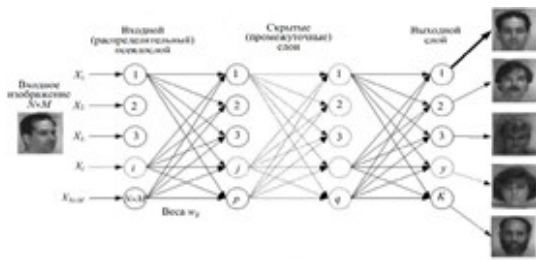


Рис.2. Архитектура многослойной нейронной сети для распознавания изображений

Скорректированные значения весов передаются от входов к выходам. Алгоритм обратного распространения является NP-трудным, поэтому время обучения сети увеличивается экспоненциально с ростом размерности данных.

Поскольку эталонные значения выходов известны, такой алгоритм относится к классу методов обучения с учителем. Применительно к извлечению ключевых признаков, когда происходит обучение сети реконструкции поданного на вход изображения, на скрытых нейронах сети формируется сжатое представление такого изображения, что может быть отнесено к классу методов самообучения. Инициализация МНС перед началом обучения производится случайным выбором весовых коэффициентов. Поэтому две разные обученные нейронные сети, обеспечивающие одинаковые значения ошибки, часто могут быть представлены различными разделяющими поверхностями, не сводимыми друг к другу. На этом тезисе основан метод коллективов (ансамблей) нейронных сетей, часто применяемый при распознавании по изображению лица: создается набор (коллектив) сетей, обученных решать одну и ту же задачу различными способами. Обобщенное, полученное таким методом решение точнее и надежнее, чем решение единственной нейронной сети.

Нейронные сети высокого порядка (НСВП) отличаются от МНС наличием одного слоя, на входы нейронов поступают также совокупности сигналов, которые могут рассматриваться как термы высокого порядка, являющиеся произведением двух или более компонентов входного вектора. Например, для сетей второго порядка такой вектор обеспечивает функционирование в соответствии с выражением:

$$S = \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j - T$$

Сети высокого порядка могут формировать сложные разделяющие поверхности, при этом добавлении компонент входного вектора в произведение можно получить класс полиномиальных разделяющих поверхностей. Сети также можно обучать по методу обратного распространения. Использование МНС в общем случае эффективнее, но существует ряд приложений, в которых сети высокого порядка проявляют себя лучше. Особенность НСВП заключается в том, что для распознавания некоторому классу достаточно предъявить обобщенный образ изображения без вариаций масштабов и поворотов. После обучения сеть будет распознавать известные классы инвариантно к масштабу и поворотам изображения. Такая сеть не

является полносвязной, она характеризуется высокой обучаемостью и быстродействием. Точность классификации такой сетью различающихся масштабом и углом поворота изображений выше по сравнению с МНС.

Радиально-базисные нейронные сети (РБНС) состоят из двух слоев (рис. 3). Первый слой описывается радиально-базисной активационной функцией:

$$y = \exp\left(\frac{-S^2}{2\sigma^2}\right),$$

где σ - среднеквадратичное отклонение, определяющее размер кластера, S - расстояние между вектором входных сигналов и сформированным вектором весовых коэффициентов W :

$$S^2 = |X - W|^2 = \sum_i (x_i - w_i)^2.$$

Значение S определяет расстояние до центра кластера от исходного изображения на входе конкретного нейрона. Второй (скрытый) слой представляет собой набор кластеров в пространстве образов и реализует первый этап кластеризации входного образа - значение активационной функции каждого нейрона быстро уменьшается с удалением от центра кластера. Последний слой нейронов может быть описан линейной активационной функцией, он реализует второй этап кластеризации - распределяет кластеры по классам (рис. 3).

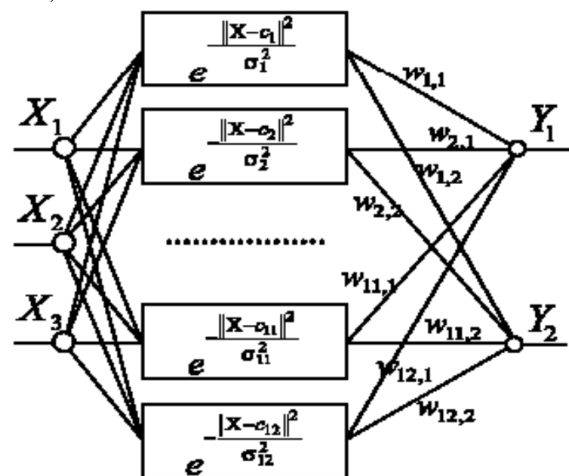


Рис.3. Распределение кластеров по классам

Основной особенностью нейронных сетей является их обучаемость. Обучение нейронной сети происходит с помощью готовых примеров. В процессе обучения извлекаются ключевые признаки и строится взаимосвязь между ними. После этого обученная нейронная сеть может применить полученный в процессе обучения опыт, для того чтобы распознать ранее неизвестный объект.

РБНС позволяют строить плохо разделяющиеся области и аппроксимировать многомерные функции. По сравнению с многослойной нейронной сетью, РБНС обучается на порядок быстрее, однако обладает намного худшей экстраполирующей способностью, т.е. не способна распознавать образы, значительно отличающиеся от образов-эталонов. Размерность РБНС больше, чем МНС, предназначенных для решения

аналогичных задач, поэтому эффективность РБНС уменьшается с ростом размерности входных данных. Обучение такой сети происходит в два этапа. На первом происходит обучение без учителя: первый слой выделяет компактно расположенные группы кластеров, при этом происходит корректировка центров кластеров. На втором этапе второй слой учится распределять по классам входные образы, пропущенные через первый слой. Если известны эталонные значения выходов, обучение обеспечивается матричными методами или алгоритмом обратного распространения ошибки. Рассмотренные типы нейронных сетей - МНС и РБНС - позволяют учесть топологию пространства изображения. Их принципы работы основываются на разбиении изображения на локальные области и иерархическом сопоставлении как их взаимного расположения, так и содержания. Такие сети наиболее

перспективны для распознавания изображений.

Когнитрон. В основу функционирования когнитрона (рис. 4) положена модель зрительной коры мозга. Каждый слой мозга реализует различные уровни обобщения - входной слой чувствителен к простым образам, таким как линии различной ориентации в плоскости, в то время как другие слои позволяют формировать более сложные, абстрактные и не зависящие от положения образы. Каждый нейрон когнитрона связан только с локальной областью предыдущего слоя, причем области могут взаимно перекрываться. Слоев в когнитроне обычно больше, чем в сетях других типов, таким образом достигается иерархическая организация. Высшие слои когнитрона реагируют на выделенные абстрактные признаки, поэтому на распознавание в меньшей степени влияют смещение и искажение исходного образа.

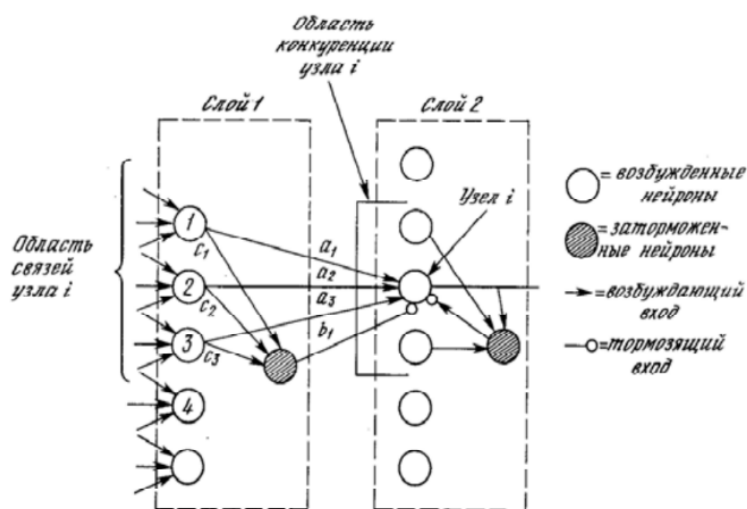


Рис.4. Когнитрон

Неокогнитрон. В зрительной коре мозга человека были обнаружены группы нейронов (узлы), реагирующие на такие элементы, как линии и углы определенной ориентации. На более высоком уровне узлы реагируют на более сложные и абстрактные образы - окружности, треугольники и прямоугольники. С увеличением уровня степень абстракции возрастает до тех пор, пока не сформируются узлы, реагирующие на лица и другие сложные объекты. В общем случае узлы последующих уровней получают на вход результаты обработки группы низкоуровневых узлов и, следовательно, реагируют на более широкую область визуального поля. Реакции высокоуровневых узлов более устойчивы к искажениям исходного образа. Неокогнитрон более точно, по сравнению с когнитроном, отражает строение зрительной коры и позволяет распознавать образы независимо от их преобразований: смещения, вращения, изменения масштаба и искажения [1]. Неокогнитрон может самообучаться, так и обучаться с учителем. На вход неокогнитрона поступают двумерные образы, аналогичные изображениям, сформированным сетчатой оболочкой глаза, и обрабатываются аналогично зрительной коре мозга человека. Главное отличие неокогнитрона от когнитрона -

двумерная организация локальных участков в виде иерархической структуры, состоящей из плоскостей (рис. 5). Слои состоят из простых и сложных плоскостей. Каждый нейрон простой плоскости связан с локальным двумерным участком плоскостей предыдущего слоя, значения весовых коэффициентов всех нейронов в пределах одной плоскости одинаковы, и таким образом плоскость реагирует на определенный образ, находящийся в участке изображения (на рис. 4 плоскости реагируют на букву С вне зависимости от угла поворота). Местоположение активированного нейрона в простой плоскости определяет участок, в котором найден этот образ, независимо от его искажения. Классический неокогнитрон является мощным средством распознавания изображений, однако требует труднодостижимых на сегодняшний день вычислительных затрат.

Сверточные нейронные сети (СНС). В классической многослойной нейронной сети межслойные нейронные соединения являются полностью связанными, изображение представлено в виде n -мерного вектора, не учитывающего ни двумерной локальной организации пикселей, ни возможностей деформации образа. Архитектура сверточной СНС (рис. 6) позволяет

преодолеть эти недостатки, в ней реализованы принципы архитектуры неокогнитрона, упрощенного и

дополненного алгоритмом обучения с обратным распространением ошибки.

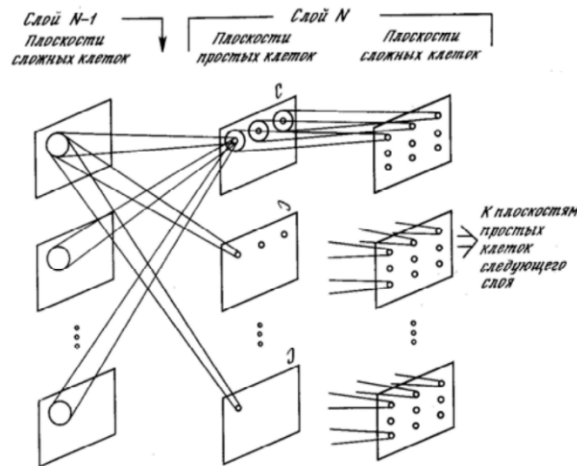


Рис.5. Иерархическая структура

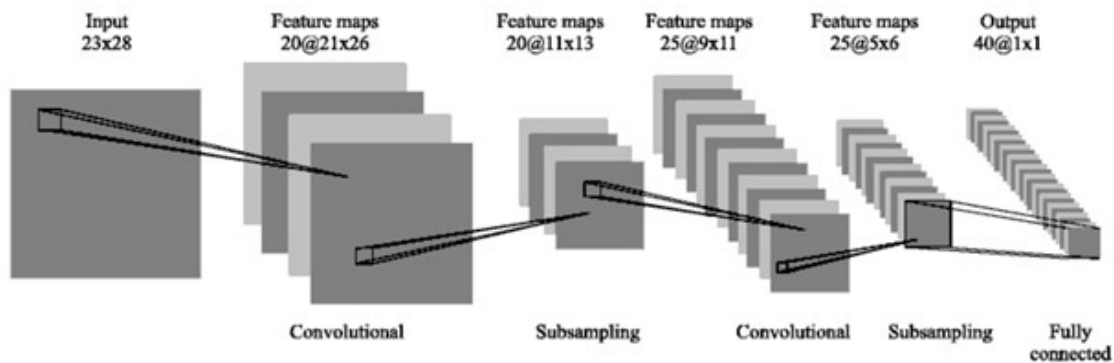


Рис.6. Архитектура сверточной НС

В СНС используются локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов), общие весовые коэффициенты (обеспечивают детектирование отдельных черт лица, находящихся в любом фрагменте изображения) и иерархическая организация с пространственными подвыборками (Spatial subsampling). СНС обеспечивает частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям. Архитектура СНС многослойна. Слои подразделяются на два типа: сверточные (Convolutional) и подвыборочные (Subsampling), чередующиеся друг с другом. В каждом слое имеется набор из нескольких плоскостей, причем нейроны одной плоскости имеют одинаковые весовые коэффициенты, поступающие ко всем локальным участкам предыдущего слоя (как в зрительной коре человека), изображение предыдущего слоя „сканируется“ небольшим окном и „взвешивается“ набором весовых коэффициентов, а результат отображается на соответствующий нейрон текущего слоя. Таким образом, плоскости называются картами характеристик (feature maps), каждая из них выделяет “свои” участки изображения в любом месте предыдущего слоя. Следующий за сверточным подвыборочный слой уменьшает масштаб плоскостей за счет локального усреднения значений реакции слоя на

выходах нейронов, таким образом достигается иерархическая организация СНС. Последующие слои извлекают более общие характеристики, меньше зависящие от искажений изображения. Обучается СНС стандартным методом обратного распространения ошибки. Сравнение МНС и СНС показало существенные преимущества последней как по скорости, так и по надежности классификации. Полезным свойством СНС является и то, что характеристики, формируемые на выходах верхних слоев структуры, могут применяться для классификации по методу ближайшего соседа (например, при вычислении евклидова расстояния), причем СНС может успешно извлекать такие характеристики и для образов, отсутствующих в обучающем наборе. Для СНС характерны высокая скорость обучения и быстрое действие.

Глубинная нейронная сеть (ГНС, англ. DNN - Deep neural network) - это искусственная нейронная сеть (ИНС) с несколькими слоями между входным и выходным слоями. ГНС подбирает корректный метод математических преобразований, позволяющий преобразовывать входные данные в выходные, независимо от линейной или нелинейной. Сеть продвигается по слоям, рассчитывая вероятность каждого выхода. Пользователь может просмотреть результаты и выбрать вероятности, которые должна отображать сеть

(выше определенного порога, например), и вернуть в сеть предложенную метку. Каждое математическое преобразование считается слоем, а сложные ГНС имеют много слоев.

В первую очередь к глубинному обучению относятся следующие методы и их вариации:

- определенные системы обучения без учителя, такие как ограниченная машина Больцмана для предварительного обучения, автокодировщик, глубокая сеть доверия, генеративно-состязательная сеть;
- определенные системы обучения с учителем, такие как свёрточная нейронная сеть, которая вывела на новый уровень технологии распознавания образов;
- рекуррентные нейронные сети, позволяющие обучаться на процессах во времени;
- рекурсивные нейронные сети, позволяющие включать обратную связь между элементами схемы и цепочками.

Глубокое обучение является апробированной выборкой из широкого семейства методов машинного обучения для представлений данных, наиболее соответствующих характеру задачи.

2. Минимизация вероятности ошибок при распознавании объектов

Перед группой свёрточных нейронных сетей (СНС), обученных на реальных базах, поставлена задача совместного использования всех вычислительных мощностей для надежной идентификации ряда объектов. Известно, что использование нейросетей в задачах распознавания может привести к трем типам ответов - «правильный ответ», «неправильный» и «объект не распознан».

Основным требованием, предъявляемым к группе нейронных сетей, является требование минимизации вероятности ошибочного распознавания, которая для единичной сети достаточно велика. При этом результатами распознавания должны быть либо «правильный» ответ, либо ответ «объект не распознан» с вероятностью неправильного ответа, не превышающей наперед заданной малой величины.

До сих пор в известной нам литературе эта проблема почти не обсуждалась, хотя потребность в ее решении возникает во многих прикладных задачах, таких, например, как принятие решений в технических системах, в медицинской диагностике, при чрезвычайных ситуациях и многих других случаях, в которых цена ошибки велика.

Действительно, как показывает опыт, в ряду равномерно расположенных по степени трудности заданий, начиная с некоторого критического ее значения, существует зона, в которой нейросеть дает неправильное решение с вероятностью, превышающей наперед заданную величину.

В этой связи возникает ряд проблем. Во-первых, проблема доказательства возможности создания группы нейронных сетей, дающей неправильные решения задач любой степени трудности с

вероятностью, не превышающей наперед заданную малую величину. Во-вторых, проблема определения предельной степени трудности решения задачи распознавания с заданной наперед вероятностью правильного ответа группой нейронных сетей с определяемыми заранее их характеристиками.

3. Метод эволюционного согласования

Задачу предлагается решать методом эволюционного согласования решений, координирующего совместную работу группы нейронных сетей и потенциально удовлетворяющего сформулированным выше требованиям. Этот метод объединяет две стадии. На первой стадии - стадии формирования популяции решений - нейронные сети, работая поодиночке, генерируют возможные варианты ответов, причем им предоставляется возможность отказаться от принятия решения, в случае если они не уверены в ответе. На второй и последующих стадиях - стадиях обмена вариантами - нейронные сети, которые отказались от ответа, выбирают из предложенных ответов правильные на их взгляд или вновь отказываются от них. Процесс итераций продолжается до тех пор, пока у большинства нейронных сетей не совпадут ответы.

Ясно, что этот ответ может быть трех видов: «правильный», «неправильный» или «неопределенно». Вероятности возникновения таких ответов зависят от числа нейронных сетей M в группе, трудности задачи β , от исходных вероятностей генерации правильных и неправильных ответов у одиночных нейронных сетей и, соответственно, от вероятностей правильного или неправильного выбора «чужих» ответов на стадиях согласования решений.

В [1] доказано, что при использовании этого метода и соблюдении ряда условий вероятность принятия правильного решения группой нейронных сетей по сравнению с вероятностью принятия такого решения одиночной нейронной сетью может достигать единицы, а вероятность принятия неправильного решения - снизиться до нуля. Результаты компьютерного моделирования искомых вероятностей при использовании метода эволюционного согласования приведены на рис. 7.

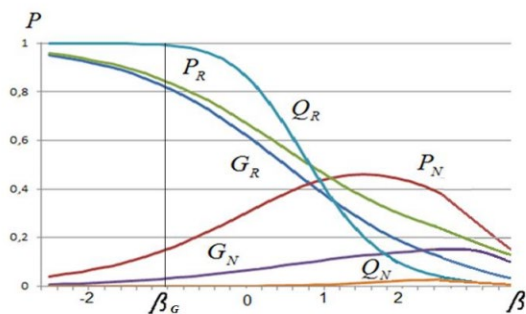


Рис. 7. Результаты компьютерного моделирования искомых вероятностей при использовании метода эволюционного согласования

На рис. 7 представлена диаграмма зависимости вероятностей правильного и неправильного решения

одиночными нейронными сетями и группой нейронных сетей от трудности задачи. Здесь P_R и P_N - вероятности “правильного” и “неправильного” решения задачи трудности одиночными нейронными сетями в составе группы по окончании итераций согласования; Q_R и Q_N - соответствующие вероятности решения задачи группой нейронных сетей.

Использование результатов модели Раша [2] и теоремы Кондорсе о “присяжных” [3] позволило найти методом эволюционного согласования аналитическое решение задач определения условий ошибочного принятия решений с вероятностью, не превышающей наперед заданной малой величины в предположении об их произвольной сложности.

В основу алгоритмов ГНС были заложены следующие теоремы [4]:

Теорема 1. Пусть групповая нейронная сеть состоит из M одиночных нейронных сетей с подготовленностями к решению задач θ_{GR} , θ_{GS} , θ_{ER} и θ_{ES} , численные значения которых лежат в интервале $[\theta - \delta, \theta + \delta]$, $\theta \in \mathbf{R}$ и $\delta \in \mathbf{R}$. Здесь θ_{GR} - подготовленность одиночной нейронной сети к генерации правильных решений, которая измеряется по всем правильным и неправильным решениям, θ_{ER} и θ_{ES} - соответствующие подготовленности при оценивании чужих решений.

Утверждается, что при $M \rightarrow \infty$ и $\delta < \ln 2$ вероятность P_N неправильного решения задачи произвольной трудности β группой нейронных сетей стремится к нулю.

Теорема 2. Для произвольной группы из M нейронных сетей существует задача с пределом трудности β_G , при превышении которого эта группа уже не в состоянии правильно решить задачу с заданной вероятностью q .

Доказательства данных теорем приведены в работе [2], мы же воспользуемся готовыми выкладками.

3.1. Постановка задачи

Для сравнения эффективности работы одиночной нейронной сети и групп нейронных сетей с использованием метода эволюционного согласования, описанного в работе [4].

Проведя сравнительный анализ существующих видов нейронных сетей, для серии экспериментов была выбрана глубокая нейронная сеть, как обладающая наилучшей разрешающей способностью в области распознавания объектов, в данном случае лица человека.

Программа, используемая для серии следующих экспериментов, моделирует идентификацию людей по лицу с помощью нейронных сетей глубокого обучения (ГНС).

Сначала ГНС была обучена на БД человеческих лиц в различных ракурсах. Т.е. сеть могла отличать фото 3D лица произвольного человека, изображения которого ранее не подавались на вход НС, от других предметов, не являющихся лицами.

Результатом работы такой ГНС является вектор в 128-мерном пространстве, значениями его компонент являются числа в промежутке от -0,38 до 0,38. Стоит

отметить, для одного и того же человека в разных ракурсах компоненты этого вектора меняются незначительно - менее 10%. При фотографировании лица в разных ракурсах (где-то в пределах ± 20 градусов от фронтального) в 128-мерном пространстве формируется кластер близко расположенных точек, характеризующий данного человека. Его идентификатором можно считать центр этого кластера.

Предположим, что наша БД будет содержать миллион людей, тогда в 128-мерном пространстве будет миллион точек, причем можно найти такие две точки, что пифагорово расстояние между ними будет равно самому минимальному значению из всех попарных расстояний. Такое расстояние далее мы будем называть **R1**. Это расстояние зависит от разрешения системы и количества людей в БД.

Другой характеристикой системы является расстояние самой удаленной точки кластера какого-либо человека от центра. Такой расстояние далее мы будем обозначать как **R2**.

Ясно, что если $R2 < R1$, то наша система будет давать идеальное распознавание человека. Но в реальных случаях распознавания это недостижимый результат. Вследствие этого, для данного разрешения системы при увеличении числа людей и соответственно числа кластеров, R1 и R2 начнут пересекаться и появятся **ошибки первого и второго рода**.

В данной работе мы будем рассматривать модель предприятия, в котором находится ровно N идентифицированных работников. При этом каждому из N людей при входе необходимо присвоить пропуск, который дает допуск в определенное только для него помещение. Первоначально, каждый из работников заносится в специальную базу данных (далее БД) по следующему алгоритму. Делается несколько снимков каждого человека в разных ракурсах (количество ракурсов может быть различным). Данные ракурсы обрабатываются нейронной сетью и на выходе получается уникальный вектор в 128-ми пространстве, соответствующий данному человеку. Далее, все полученные векторы для конкретной персоны усредняются и заносятся в БД как эталонные значения. Данная операция совершается для всех работников.

После формирования БД происходит следующий этап проверки. Приходя на предприятие, человек (не обязательно работник) снова фотографируется в разных ракурсах, далее находится средний вектор. Этот вектор и сравнивается со всеми существующими N векторами и находится тот, у которого пифагорово расстояние (далее POR) меньше R1. Расстояние между двумя точками в таком случае будет определяться следующей формулой: $p = (p_1, \dots, p_n)$ и $q = (q_1, \dots, q_n)$, расстояние $d(p, q)$ между ними определяется следующим образом:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2},$$

где n - это мерность пространства (в нашем случае n=128).

При обнаружении такого ближайшего вектора в БД, перед пропускной системой ставятся следующие задачи:

1. Не допустить, чтобы незарегистрированный человек (“чужой”) не смог получить никакого пропуска при распознавании. Это соответствует минимизации ошибки 1-го рода.
2. Обеспечить, чтобы зарегистрированный в системе человек (“свой”) всегда получал только свой пропуск и не мог получить пропуск другого человека. Это соответствует минимизации ошибки 2-го рода.
3. Высокая вероятность, что зарегистрированному человеку будет выдан именно его пропуск $E \sim 0.99$. Следовательно, наша система может характеризоваться следующими параметрами:
 - N - число зарегистрированных человек в системе;
 - E - вероятность того, что зарегистрированному человеку будет выдан его пропуск;
 - E_1 - вероятность того, что зарегистрированному человеку будет выдан чужой пропуск;
 - E_2 - вероятность того, что чужому человеку будет выдан любой пропуск (ошибка I - го рода);
 - E_3 - вероятность того, что своему человеку не будет выдан пропуск при первом предъявлении (ошибка II рода);
 - E_4 - вероятность того, что чужому человеку не будет выдан никакой пропуск.
 Учитывая, что полная вероятность равна 1, получаем следующее равенство:

$$P_{\text{полн}} = E + E_1 + E_3 = E_2 + E_4 = 1$$

3.3. Описание программы

С помощью ГНС сформирована БД, состоящая из 1201 ракурса, что соответствует изображениям порядка 500 отдельно взятых человек. Данные ракурсы являются результатом обработки фотографий реальных людей.

Ракурс заносится в БД как уникальный вектор в 128-пространстве. БД для экономии памяти преобразована из диапазона чисел с плавающей запятой (-0.38;0.38) в диапазон целых чисел от 0 до 255 по формуле:

$$X_n = \max\{n \in Z \mid n \leq 318.78 * (x_n + 0.4)\}$$

База данных имеет следующий формат: 128 целочисленных компонент вектора и числовой идентификатор человека, всего 1201 строка по 129 значений в каждой.

Каждая строка как совокупность компонент векторов представляет собой определенный ракурс изображения человека. Каждый ракурс уникален. При этом у одного человека может быть несколько ракурсов съемки, т.е. значения последнего параметра - ID - в базе могут повторяться в рамках строки.

Описание программы приведено на рис. 8.

Параметр 1	Параметр 2	...	Параметр 128	ID человека
Значение атрибута	Значение атрибута	...	Значение атрибута	ID 1
...				
Значение атрибута	Значение атрибута	...	Значение атрибута	ID i
...				
Значение атрибута	Значение атрибута	...	Значение атрибута	ID K

i-й ракурс {

Рис.8. Описание программы

3.2. Сравнение эффективности одиночной ГНС и комитета ГНС на малой БД

Для начала предположим, что наше предприятие состоит из 100 человек. Для такого случая выберем следующие параметры: LV - число компонент вектора, LV=128. NITER - число своих людей, проходящих через пропускную систему и число чужих, тоже пытающихся пройти через систему, NITER = 10^2-10^8 (так как число входов влияет только на масштаб, сами

вероятности ошибок 1-го (E_2) и 2-го рода (E_3) от числа итераций не зависят). NB - число людей на предприятии, NB = 100. DEL - величина разброса точек в кластере от средней точки в относительных единицах, DEL = 0.111. Опытным путем было установлено, что данное значение является наилучшим для данной БД. POR - пифагорово расстояние, настраиваемое для получения оптимального соотношения величина ошибок I и II рода. В данном случае $0 < POR < 300$.

Интерфейс программы, содержащей одиночную нейронную сеть, представлен на рис. 9.

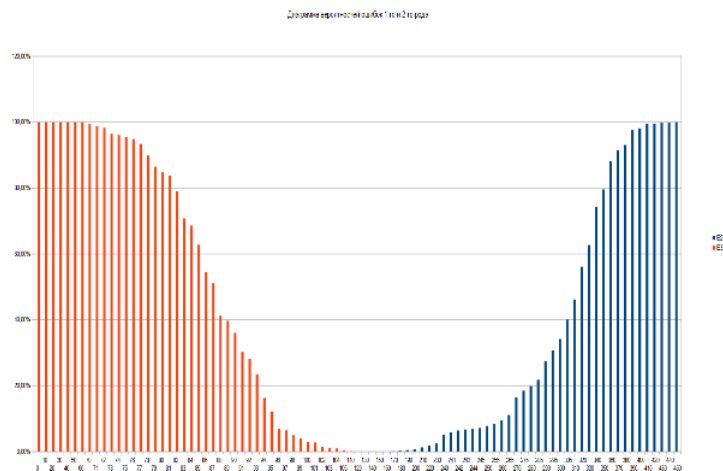


Рис.9. Интерфейс программы, содержащей одиночную нейронную сеть

Стоит отметить, что при одних и тех же начальных данных ГНС выдаёт результаты, которые могут отличаться друг от друга в пределах 5-10%. Чтобы избежать неточностей, для каждого значения проводилась

серия из 5 экспериментов, а затем полученные данные усреднялись.

Распределение вероятностей ошибок 1 и 2 рода для данного эксперимента приведено на рис. 10.

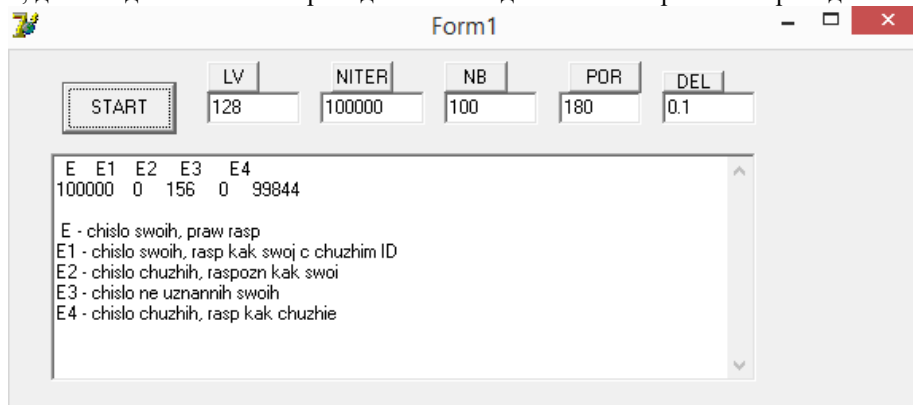


Рис.10. Распределение вероятностей ошибок 1 и 2 рода для данного эксперимента

Из диаграммы можно сделать вывод, что наилучшие настроечные параметры POR для данного сценария находятся в диапазоне от 110 до 160.

При увеличении числа NITER до 10^6 получаем более точные данные - рис.11.

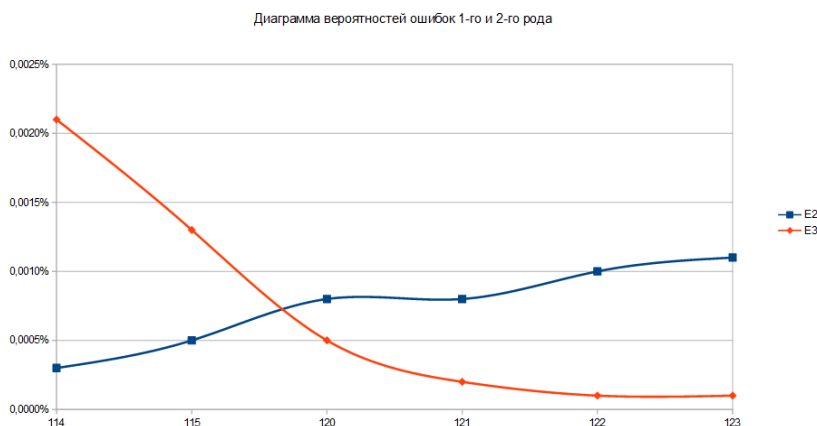


Рис.11. Более точные данные

Минимальные значения ошибок 1-го и 2-го рода (E2 и E3 соответственно) для предприятия в 100 человек достигаются при $POR = 121 \pm 0.5$

При этом порядок ошибок колеблется около 10^{-5} (при условии, что $E \rightarrow 1$).

Уменьшить данный порядок ошибки не

представляется возможным ввиду разрешающей способности одиночной нейронной сети.

Ни при каких других значениях POR ошибка 1-го рода и 2-го рода не становится меньше 10^{-6} (при условии, что $E \rightarrow 1$).

В следующей серии исследований будем

рассматривать уже не одиночную ГНС, а группу ГНС из 15 нейронных сетей.

Организованы они следующим образом: 5 групп по 3 ГНС в группе. На каждую такую ГНС будут подаваться не все 128 компонент вектора, а лишь его часть в виде задаваемых нами параметра RAZ (при этом каналы могут пересекаться).

Согласовываться между собой внутри одной группы и среди других групп нейронные сети будут при помощи метода эволюционного согласования, описанного выше.

Для предприятия в 100 человек были выбраны следующие параметры :

LV - число компонент вектора, LV=128.

RAZ - число каналов, подаваемых на одиночную ГНС; RAZ = 64.

NITER - число “своих” людей, проходящих через пропускную систему, и число чужих, также пытающихся пройти через систему. Данный параметр будет изменяться для получения более точных результатов и сравнения вероятностей ошибок 1-го и 2-го рода с аналогичными входными данными у одиночной ГНС. $NITER = 10^2 \cdot 10^8$.

NB - число людей на предприятии, NB =100.

DEL - величина разброса точек в кластере от средней точки в относительных единицах; DEL = 0.111.

POR - пифагорово расстояние, настраиваемое для получения оптимального соотношения величин ошибок I и II рода. В данном случае $0 < POR < 300$.

На рис. 12 приведена итоговая диаграмма вероятностей для ошибок 1-го и 2-го рода (E2 и E3 соответственно).

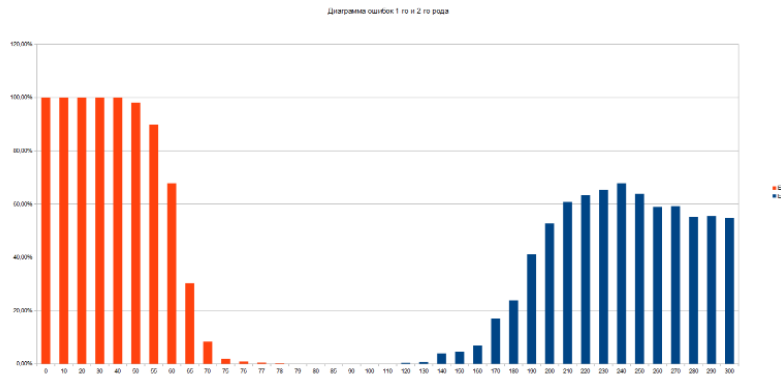


Рис.12. Итоговая диаграмма вероятностей

Стоит отметить, что при одних и тех же начальных данных комитет ГНС выдаёт результаты которые могут отличаться друг от друга незначительно, в отличие от одиночной ГНС.

Из диаграммы на рис. 13 можно сделать вывод, что наилучшие настроечные параметры POR для данного сценария находятся в диапазоне от 80 до 110.

Также стоит отметить, что при увеличении

$POR > 300$ ошибка первого рода не стремится к 100%, а выходит на постоянное значение.

Для получения более точных значений была составлена подробная диаграмма ошибок 1-го и 2-го рода (рис.13). При ее получении были взяты следующие значения: NITER увеличено до 10^6 , POR рассматривается в диапазоне от 83 до 90.

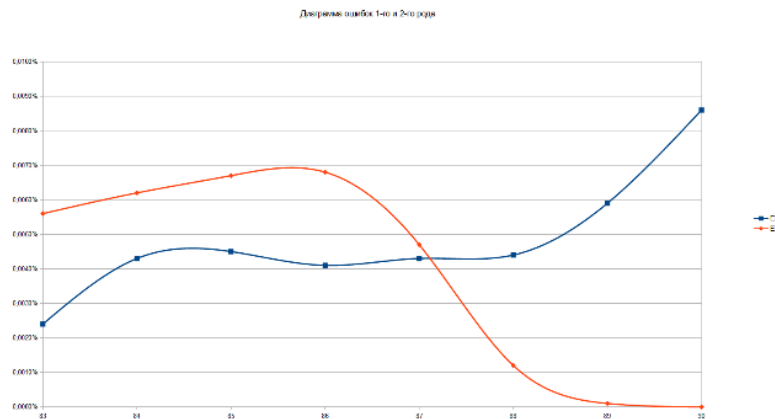


Рис.13. Подробная диаграмма ошибок 1-го и 2-го рода

В данном случае минимальные значения E2 и E3 достигаются при значении $POR = 88 \pm 0.5$. При этом порядок ошибок комитета комитетов ГНС равен 10^{-4} (при условии, что $E \rightarrow 1$).

Однако, представляется возможным уменьшить

порядок ошибки 1-го и 2-го рода можно вплоть до порядка величины $10^{-8} \cdot 10^{-9}$. Но при этом сильно снизится распознавательная способность групп нейронных сетей с требуемого по условиям задачи $E = 0.99+$ до $E = (6,5 \pm 0,2) \cdot 10^{-6}$.

3.4. Сравнение эффективности одиночной ГНС и комитета ГНС на расширенной БД

Для проведения следующей серии экспериментов и сравнения эффективности одиночной ГНС и комитета комитетов ГНС была искусственно создана расширенная база данных. Данная база данных содержит 5000 уникальных «человек». Формат расширенной БД полностью совпадает с форматом, описанным выше.

Расширенная база данных была наполнена синтетическими данными, полученными следующим образом:

1. Числа из БД записываются в массив $T[1..5000, 1..128]$. В массиве $NT[1..5000]$ хранятся идентификаторы лиц.
2. Из этого массива выбираются те записи, у которых есть более двух ракурсов. Первая тройка таких ракурсов усредняется и средний образ записывается в массив T с номера 1202 по номер $KM=1259$ (из 500 мы берем только 58 человек, у которых есть по крайней мере по 3 ракурса).
3. Из этих усредненных ракурсов мы будем «изготавливать» большую БД по формуле: $T[j,i] := (a_1 * T[j,1,i] + a_2 * T[j,2,i] + \dots + a_5 * T[j,5,i]) / (\text{сумма всех } a)$, где $a_1 - a_5$ - случайные числа в диапазоне $[4, 13]$, а $j_1 - j_5$ - случайные номера усредненных компонент векторов, полученных пунктом выше. Числа нормируются к диапазону усредненных чисел, находящихся в 1202-1259 записях и статистически «похожи» на них.

Используемые настроечные параметры для полученной БД следующие: L_V - число компонент вектора, $L_V < 129$. $NITER$ - число своих людей, проходящих через пропускную систему и число чужих, тоже пытающихся пройти через систему. $NITER < 10^9$ - ограничивается временем вычислений. Подбирается исходя из необходимой статистической точности. NB - число людей на предприятии, $NB = 5000$. POR - пифагорово расстояние, настраиваемое для получения оптимального соотношения величин ошибок I и II рода. DEL - величина разброса точек в кластере от средней точки в относительных единицах, $DEL = 0.12$. RAZ - число каналов подаваемых на одиночную ГНС, $RAZ = 64$.

4. Такими записями мы заполняем все 1-5000 записей.

Построим график вероятностей ошибок для ошибок 1-го и 2-го рода соответственно для одиночной ГНС в расширенной БД (рис.14):

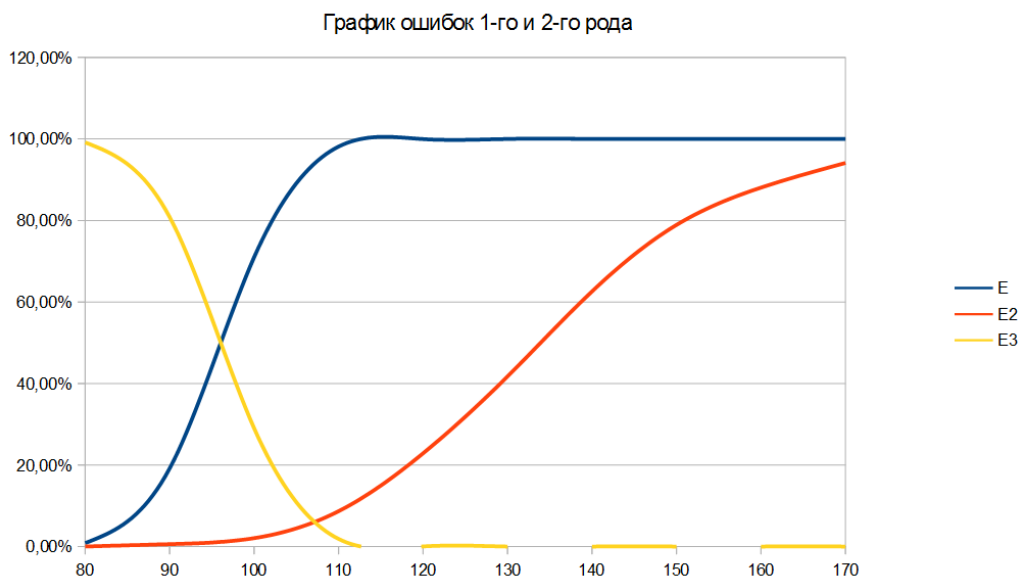


Рис.14. График вероятностей ошибок для ошибок 1-го и 2-го рода

Из графика на рис. 14 можно сделать вывод, что наилучшие настроечные параметры POR для данного сценария находятся в диапазоне от 100 до 110.

Однако в таком случае не соблюдается условие высокой распознавательной способности нейронной сети. Для демонстрации этого факта на график был выведен параметр E - вероятность того, что зарегистрированному человеку будет выдан его пропуск.

По результатам эксперимента минимальные значения ошибок 1-го и 2-го рода для предприятия в 5000 человек достигаются при $POR = 108 \pm 0.5$. При этом порядок ошибок одиночной ГНС = $10^{-1} - 10^{-2}$ (при

условии, что $E \rightarrow 1$). Уменьшить данный порядок ошибок не представляется возможным ввиду разрешающей способности одиночной нейронной сети.

Ни при каких других значениях POR ошибка 1-го рода и 2-го рода не становится меньше 10^{-6} , как и в случае предприятия из 100 человек.

Теперь рассмотрим результаты, полученные от комитета комитетов ГНС на расширенной базе из 5000 человек.

Для предприятия в 5000 человек были выбраны следующие параметры: L_V - число компонент вектора, $L_V = 128$. RAZ - число каналов, подаваемых на

одиночную ГНС, $RAZ = 64$. NITER - число своих людей, проходящих через пропускную систему и число чужих, тоже пытающихся пройти через систему. Данный параметр будет изменяться для получения более точных результатов и сравнения вероятностей ошибок 1-го и 2-го рода с аналогичными входными данными у одиночной ГНС. $NITER = 10^4 - 10^8$. NB - число людей на предприятии,

$NB = 5000$. DEL - величина разброса точек в кластере от средней точки в относительных единицах, $DEL = 0.114$. POR - пифагорово расстояние, настраиваемое для получения оптимального соотношения величин ошибок I и II рода. В данном случае $0 < POR < 300$.

Построим график вероятностей ошибок для E2 и E3 для комитета комитетов ГНС в расширенной БД (рис.15):

График ошибок 1-го и 2-го рода

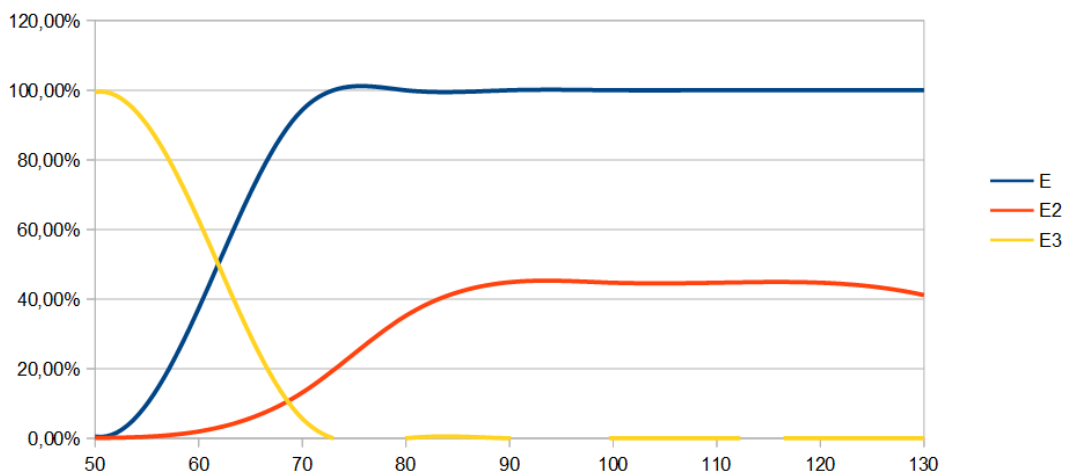


Рис.15. График вероятностей ошибок для E2 и E3

Из графика можно сделать вывод, что наилучшие настроечные параметры POR для данного сценария находятся в диапазоне от 65 до 75. Однако в таком случае, как и при одиночной ГНС, не соблюдается условие высокой распознавательной способности сети.

Минимальные значения ошибок 1-го и 2-го рода (E2 и E3 соответственно) для предприятия в 5000 человек достигаются при $POR = 71 \pm 0.5$. При этом порядок ошибок комитета ГНС равен 10^{-1} .

Еще одной особенностью графика становится выход ошибки 1-го рода на «плато». Таким образом, уменьшить порядок до наперед заданных значений не представляется возможным.

Исходя из полученных графиков, результаты текущей пары исследований говорят о проигрышном положении группы ГНС по сравнению с одиночной ГНС. Однако в следующем исследовании мы изменим число каналов, подаваемых на каждую ГНС из комитета, и сравним полученные данные.

3.5. Сравнение эффективности комитета ГНС на расширенной БД в зависимости от числа каналов

Для проведения последней серии экспериментов и

сравнения эффективности комитета комитетов ГНС при разном количестве подаваемых на каждую сеть каналов была взята та же расширенная база данных.

Для повышения точности распознавания на каждую ГНС будем подавать не 64, а 128 каналов.

Для предприятия в 5000 человек были выбраны следующие параметры: LV - число компонент вектора, $LV=128$. RAZ - число каналов подаваемых на одиночную ГНС, $RAZ = 128$. NITER - число своих людей, проходящих через пропускную систему и число чужих, тоже пытающихся пройти через систему. Данный параметр будет изменяться для получения более точных результатов и сравнения вероятностей ошибок 1-го и 2-го рода с аналогичными входными данными у одиночной ГНС.

$NITER = 10^4 - 10^8$. NB - число людей на предприятии, $NB = 5000$. DEL - величина разброса точек в кластере от средней точки в относительных единицах, $DEL = 0.111$. POR - пифагорово расстояние, настраиваемое для получения оптимального соотношения величин ошибок I и II рода. В данном случае $0 < POR < 200$.

Построим график вероятностей для E2 и E3 для описанного выше случая (рис.16):

График ошибок 1-го и 2-го рода

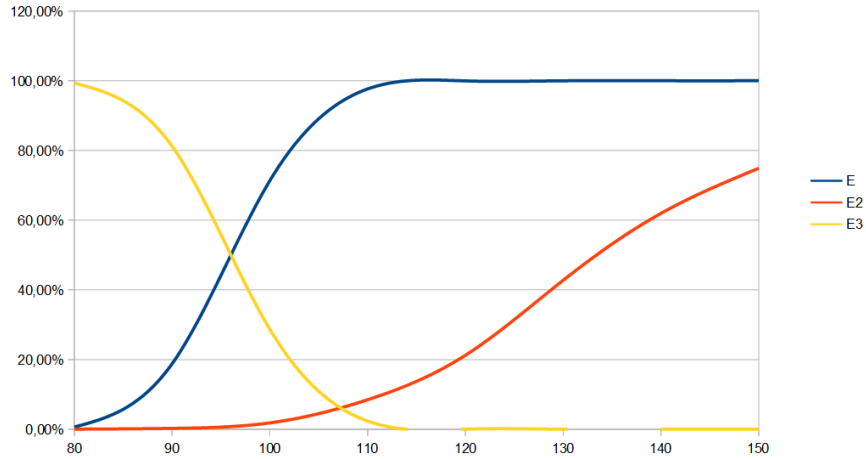


Рис.16. График вероятностей для E2 и E3

Из графика на рис. 16 можно сделать вывод, что наилучшие настроечные параметры POR для данного сценария находятся в диапазоне от 100 до 110. Однако в таком случае не соблюдается условие высокой распознавательной способности сети.

Минимальные значения ошибок 1-го и 2-го рода для предприятия в 5000 человек достигаются при $POR = 107 \pm 0.5$. При этом порядок ошибок комитета ГНС составляет 10^{-1} .

Однако, уменьшить порядок ошибки 1-го и 2-го рода представляется возможным вплоть до порядка величины равной $10^{-5} - 10^{-6}$. Но при этом сильно падает распознавательная способность групп нейронных сетей с требуемого по условиям задачи $E = 0.99+$ до $E = 10^{-4} - 10^{-5}$.

Теперь рассмотрим случай уменьшения числа каналов, подаваемых на каждую ГНС. Опытным путём было выяснено, что оптимальным значением является $RAZ = 24$. При меньшем значении RAZ распознавательная способность ГНС сильно ухудшается.

Для предприятия в 5000 человек были выбраны следующие параметры: LV - число компонент вектора, $LV=128$. RAZ - число каналов подаваемых на одиночную ГНС, $RAZ = 24$. NITER - число своих людей, проходящих через пропускную систему и число чужих, тоже пытающихся пройти через систему. Данный параметр будет изменяться для получения более точных результатов и сравнения вероятностей ошибок 1-го и 2-го рода с аналогичными входными данными у одиночной ГНС. $NITER = 10^4 - 10^8$. NB - число людей на предприятии, $NB = 5000$. DEL - величина разброса точек в кластере от средней точки в относительных единицах, $DEL = 0.111$. POR - пифагорово расстояние, настраиваемое для получения оптимального соотношения величин ошибок I и II рода. В данном случае $0 < POR < 200$.

Построим график вероятностей для E2 и E3 для описанного выше случая (рис.17).

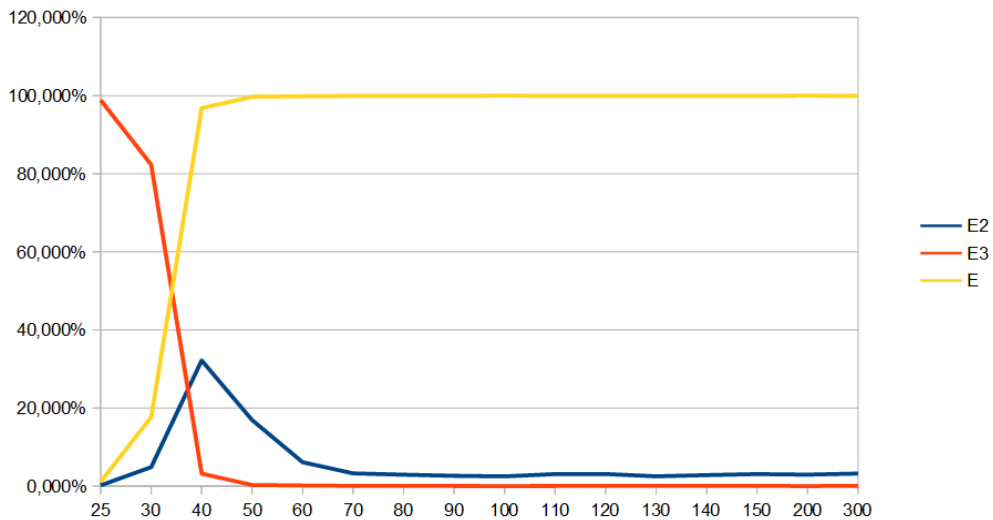


Рис.17. График вероятностей для E2 и E3

На основе графика сделать вывод, что наилучшие настроечные параметры POR для данного сценария находятся в диапазоне $POR > 90$. При этом значении POR достигаются наименьшие значения ошибок 1-го рода в 2-3% ($E_2 \sim 10^{-2}$) и 2-го рода в 0.05-0.1% ($E_2 \sim 10^{-3}$) при распознавательной способности комитета ГНС стремящейся к единице ($E \rightarrow 1$).

Однако, в таком случае уменьшить ошибку 1-го и 2-го рода до наперёд заданного значения невозможно. Как видно из графика, с увеличением POR значения ошибок остаются неизменными.

4. Заключение

По результатам проведенных исследований можно сделать вывод, что одиночная ГНС проявляет себя лучше, чем комитет ГНС при малом значении числа сотрудников (на малой базе данных).

Однако на расширенной БД становится возможным настройка комитета ГНС таким образом, чтобы получить меньшие значения ошибок, чем на одиночной ГНС.

По условию задачи необходимо добиться высокой распознавательной способности нейронной сети ($E \rightarrow 0.99+$), однако, если ставить задачу минимизации ошибок 1-го и 2-го на первый план, то метод эволюционного согласования ГНС позволяет значительно понизить порядок этих ошибок.

Стоит отметить, что работа комитета ГНС занимает в разы больше времени, чем работа одиночной ГНС (в работе численно этот параметр не фиксировался). Потому в задачах, где время работы сети играет решающую роль либо ограничено, применение комитета ГНС не рекомендуется.

Благодарность

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов 18-07-00225, 18-07-0909 и 18-07-01111.

Библиографический список

- [1] Протасов В.И. Конструирование метасистемных переходов. - М.: Изд-во института физико-технической информатики, 2009. - 186 с.
- [2] Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests /Expanded Edition, with Foreword and Afterword by B.D. Wright. Chicago: University of Chicago Press, 1980.
- [3] Condorcet, marquis Marie-Jean-Antoine-Nicolas de Caritat. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Imprimerie Royale, Paris, 1785.
- [4] В.И. Протасов. Метод эволюционного согласования решений. Компьютерная и математическая модель. 2011г. 360-379 с.