

На правах рукописи



Дуденков Владимир Михайлович

**РАЗРАБОТКА НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ
ЧЕЛОВЕКОМАШИННОГО ОБЩЕНИЯ**

Специальность

05.13.17 – Теоретические основы информатики

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата физико-математических наук

Воронеж – 2016

Работа выполнена в ФГБОУ ВО «Воронежский государственный университет»

Научный руководитель: доктор технических наук,
профессор Новикова Нелля Михайловна

Официальные оппоненты: Тулупьев Александр Львович, доктор физико-математических наук, доцент, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук, лаборатория теоретических и междисциплинарных проблем информатики, заведующий;

Душкин Александр Викторович, доктор технических наук, доцент, ФГБОУ ВПО «Воронежский институт Федеральной службы исполнения наказаний», кафедра управления и информационно-технического обеспечения, заведующий.

Ведущая организация: ФГАОУ ВО «Российский университет дружбы народов»


Защита состоится «25» января 2017 г. в 15:10 на заседании диссертационного совета Д 212.038.20 ФГБОУ ВО «Воронежский государственный университет» по адресу: 394018, Воронеж, Университетская пл.1, ауд. 335.

С диссертацией и авторефератом можно ознакомиться в Зональной научной библиотеке ФГБОУ ВО «Воронежский государственный университет», а также на сайте

http://www.science.vsu.ru/dissertations/3715/Диссертация_Дуденков_В.М..pdf

Автореферат разослан «_10_»_ноября_2016г.

Ученый секретарь
диссертационного совета Д 212.038.20
кандидат физико-математических наук,
доцент



С.А. Шабров

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Сближение технологий, повышающих возможности человека в освоении биотехнологий, информационных технологий и когнитивных наук, является одним из современных направлений в развитии науки и техники. О важности и актуальности этого направления можно судить потому, что эта проблема обсуждалась в докладе министерства обороны Великобритании «Стратегический контекст будущего» и на конференции «Конвергенция технологий, повышающих возможности человека: нанотехнологии, биотехнологии, информационные технологии и когнитивные науки». Эту конференцию провела группа по изучению проблем национальной безопасности США. Реальное сближение фундаментальных и прикладных исследований осуществляется Агентством по прорывным исследовательским проектам в области обороны США (DARPA) и японской программой ERATO (Exploratory Research for Advanced Technology).

Результаты таких исследований приведут в кибернетике к тому, что максимально сближаются возможности естественного и искусственного интеллектов за счет освоения нейроподобных вычислительных технологий. В военной технике – к развитию интерфейсов для системы «человек-машина», разработке систем, имитирующих поведение водителей и пилотов бронетанковой и авиационной техники, подводных лодок, кораблей, развитию более совершенных методов обучения и тренировки военнослужащих. Идет реальное финансирование исследований, направленных на сближение био- и информационных технологий, в результате чего наступает качественно новый этап развития и исследования человеко-машинных систем, в которых перераспределяются функции между искусственным и естественным интеллектами.

Степень разработанности темы исследования. Исследованию информационных человеко-машинных систем посвящено много работ как психологов (Ломов Б.Ф., Венда В.Ф., Зинченко В.П. и др.), так и инженеров (Губинский А.И., Павлов В.В., Поспелов Д.А., Новикова Н.М. и др.). Среди зарубежных исследователей следует отметить работы Шеридана Т., Феррел У., Левис А. и др. В этих исследованиях рассматриваются различные модели работы человека-оператора в информационной системе, но ни в одном из исследований не рассмотрены нейросетевые модели и работа человека с дисплеем. Следовательно, построение и анализ математических моделей работы человека-оператора на основе нейронных сетей при работе с дисплеем является весьма актуальной задачей. Решению этой актуальной задачи посвящена данная диссертация.

Диссертационная работа выполнена в рамках одного из основных научных направлений Воронежского государственного университета «Математическое моделирование, программное и информационное обеспечение, методы вычислительной и прикладной математики и их применение к фундаментальным исследованиям в естественных науках».

Цель и задачи. Целью диссертационной работы является разработка модели работы человека-оператора при решении задачи распознавания и классификации изображений, а также нейросетевого распознавательного комплекса, основанного на данной модели.

Для достижения поставленной цели необходимо решить следующие задачи.

1. Разработать структурную модель работы человека-оператора в системе «человек-дисплей» на основании исследований в области инженерной и когнитивной психологии, психофизики и теории статистических решений.
2. На основе структурной модели разработать методику натурального эксперимента и провести эксперимент с группой операторов.
3. Исследовать возможность и особенности применения нейронных сетей для имитационного моделирования работы человека-оператора, распознающего изображения, предъявляемые на экране дисплея компьютера.
4. Разработать и реализовать нейросетевой распознавательный комплекс для моделирования работы человека-оператора.
5. Создать программное обеспечение компьютерных моделей и оценить их работоспособность на основе вычислительных экспериментов.

Научная новизна. В работе получены следующие результаты, характеризующиеся научной новизной:

1. Разработана структурная модель работы человека-оператора решающего задачу распознавания изображений, отличающаяся тем, что использует как статистические, так и психологические законы восприятия, и позволяющая оценить эффективность систем «человек-дисплей».
2. Структурная модель представлена в виде двух подсистем – когнитивной и решающей, что позволило применить нейросетевые модели для их компьютерной реализации.
3. Разработана архитектура построения нейросетевого комплекса, основанного на самоорганизующихся картах Кохонена и гибридной нейронной сети, позволяющая моделировать и оценивать работу человека-оператора.
4. Создано специальное программное обеспечение для проведения вычислительных экспериментов, позволяющее провести анализ работы нейросетевых моделей и человека-оператора.

Содержание диссертации соответствует специальности 05.13.17. – «Теоретические основы информатики» по следующим областям исследований: исследование информационных структур, разработка и анализ моделей информационных процессов и структур (п. 2 паспорта специальности); разработка методов, языков и моделей человеко-машинного общения; разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения данных из текстов на естественном языке (п. 6 паспорта специальности).

Теоретическая и практическая значимость работы. Работа имеет теоретический и практический характер. Теоретические обоснования разработки структурной модели работы человека-оператора по распознаванию изображений позволяют проектировать и использовать нейросетевые системы классификации в системе «человек-дисплей». Теоретические результаты диссертации используются в учебном процессе Воронежского государственного университета при чтении спецкурсов и выполнении курсовых, дипломных работ и магистерских диссертаций.

Практические результаты работы используются и тестируются на предприятия ООО «Тэга» в г. Воронеж. По результатам диссертационной работы получены свидетельства о

государственной регистрации программ для ЭВМ № 2015618275 «DirectAndConvolutionNet 1.0» от 4 августа 2015г. и № 2015618276 «FuzzyAndCohonenNet 1.0» от 4 августа 2015г.

Методология и методы исследования. Для решения поставленных в диссертационной работе задач были использованы методы теории статистических решений, психофизики, инженерной и когнитивной психологии; методы теории распознавания образов и обработки изображений; технологии программирования, а также аппарат искусственных нейронных сетей.

Положения, выносимые на защиту. На защиту выносятся следующие результаты, впервые достаточно полно развитые и полученные в диссертации.

1. Модель работы человека-оператора в задаче распознавания изображений, основанная на применении теории статистических решений и инженерной психологии, психофизики и когнитивной психологии.
2. Алгоритм работы нейросетевого распознавательного комплекса на основе сегментации изображений, карт Кохонена, нечеткого гибридного классификатора, позволяющий распознавать изображения в условиях малого объема обучающей выборки.
3. Методика проведения вычислительных экспериментов по распознаванию зашумленных изображений с операторами и созданными нейросетевыми моделями, анализ результатов экспериментов.
4. Программный комплекс для проведения вычислительных экспериментов по разработанным алгоритмам.

Достоверность результатов работы. Результаты исследований, сформулированные в диссертации, получены на основе корректного использования взаимно дополняющих друг друга теоретических и экспериментальных (имитационное моделирование) методов исследований.

Личный вклад автора. Основные результаты по теме диссертации были получены лично автором и опубликованы в соавторстве с научным руководителем. Научный руководитель определил основные направления исследований.

Степень достоверности и апробация результатов. Основные положения диссертации докладывались и обсуждались на следующих конференциях: XIII, XIV, XVI Международной научно-технической конференции «Кибернетика и высокие технологии XXI века» (Воронеж, 2012, 2013, 2015), седьмой Международной научно-технической конференции «Информатизация процессов формирования открытых систем на основе САПР, АСНИ, СУБД и систем искусственного интеллекта» (Вологда, 2013), Воронежской весенней математической школе «Понtryгинские чтения - XXIV» (2013), XIV Международной научно-технической конференции «Информатика: проблемы, методология, технологии» (Воронеж, 2014), VII Международной научной конференции «Современные методы математики, теории управления и компьютерных технологий» (Воронеж, 2014), научных сессиях Воронежского государственного университета (2012-2016).

Публикации. По теме диссертации (лично и в соавторстве) опубликовано 12 работ, из них 2 работы – в изданиях, рекомендованных ВАК РФ. Получены 2 свидетельства о государственной регистрации программ для ЭВМ.

Объем и структура работы. Диссертация состоит из введения, четырех глав, заключения и списка литературы из 80 наименований. Объем диссертации составляет 131 страницу, включая 114 страниц основного текста, содержащего 27 рисунков и 13 таблиц.

Основное содержание работы

Во введении к диссертации обоснована актуальность темы, сформулированы цель и задачи работы, методы исследования, научная новизна, практическая значимость полученных результатов, результаты, выносимые на защиту.

Первая глава посвящена анализу систем «человек-машина» и «человек-компьютер». Рассмотрены основные понятия инженерной психологии показано, что инженерная психология является научной дисциплиной, находящейся на стыке технических и психологических наук. Представлена классификация систем «человек-машина» по основным группам признаков: целевое назначение системы, характеристики человеческого звена, тип и структура машинного звена, тип взаимодействия компонентов системы. Рассмотрен отдельный класс информационных систем «человек-компьютер», различные подходы к проектированию этих систем, представлен деятельностный подход, использующий принципы теории действий. Показано, что наибольший интерес представляют информационные системы, которые используются для получения информации, необходимой для распознавания и принятия решений. В таких системах человек-оператор воспринимает информацию со средств отображения (индикаторы, дисплеи), преобразует ее в концептуальную модель, на основании которой принимает решение об изменении состояния воспринимаемого объекта. Исследование таких систем представляется достаточно трудной и актуальной задачей и особое внимание при этом следует уделять человеку-оператору.

В результате рассмотрения многочисленных математических моделей работы человека-оператора, как линейных, так и нелинейных, детерминированных и статистических, получено, что наиболее перспективным является имитационное моделирование с использованием нейронных сетей.

Вторая глава посвящена рассмотрению структурной модели работы человека-оператора при восприятии информации с экрана дисплея и методике проведения экспериментов по распознаванию изображений с человеком-оператором.

Пусть на входе системы имеется пространство сигналов \vec{S} , пространство наблюдаемых данных \vec{X} , а также пространство шумов N и пространство решений.

Рассмотрены основные положения теории статистических решений, на основании которых строятся алгоритмы обнаружения, распознавания образов и принятия решений. Применение фундаментальной леммы Неймана–Пирсона показывает, при каких условиях эти алгоритмы являются байесовскими оптимальными алгоритмами. Пусть \vec{S} и \vec{X} представляют соответственно n -мерные векторы сигнала и наблюдаемых данных в пространствах Ω и Γ , Ω - пространство сигналов, Γ – пространство наблюдений. Рассматривается двухальтернативная задача: гипотеза H_0 утверждает, что имеется только шум, ее альтернатива H_1 утверждает, что имеются сигнал и шум, так что в пространстве решений есть только две точки (γ_0, γ_1) соответственно. Пусть \vec{S} сигнал на входе, который мог присутствовать на протяжении времени наблюдения, тогда кратко обе гипотезы можно

записать так: $H_0 | \vec{S} \in \Omega_0$ и $H_1 | \vec{S} \in \Omega_1$, где Ω_0 и Ω_1 - соответствующие неперекрывающиеся друг друга классы сигналов или гипотез.

Байесовская система подчинена байесовскому правилу выбора решения δ , которому соответствует наименьшая средняя оценка потерь при априорной функции распределения $W(\vec{S})$. Применение статистической теории проверки гипотез позволяет ввести отношение правдоподобия, при помощи которого можно сформулировать байесовское правило выбора решения в более сжатом виде:

выбирать решение γ_1 , если $\Lambda_n(\vec{X}) \geq \Lambda_0$

выбирать решение γ_0 , если $\Lambda_n(\vec{X}) < \Lambda_0$

Отношение правдоподобия:

$$\Lambda_n(\vec{X}) = \frac{p \cdot W_n(\vec{X} / \vec{S})}{q \cdot W_n(\vec{X} / O)}, \quad (1)$$

где n - объем обучающей выборки, q и $p = 1 - q$ - соответственно априорные вероятности того, что появится некоторый сигнал из классов Ω_0 и Ω_1 . Величина Λ_0 называется порогом и зависит от выбранных заранее стоимостей и априорных вероятностей появления сигнала и шума.

$$\Lambda_0 = \frac{q \cdot (P_{01} - P_{00})}{p \cdot (P_{10} - P_{11})} > 0 \quad (2)$$

где P_{ij} - элементы платежной матрицы - стоимости, прописываемые каждой из возможных комбинаций сигнала на входе и решения:

$$P_{ij} = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix} \quad P_{01} > P_{00} > 0; P_{10} > P_{11} \geq 0. \quad (3)$$

По главной диагонали расположены выигрыши за правильное решение, а по побочной - платы (потери) за ошибочные решения.

Байесовское правило по существу представляет разделение пространства наблюдений Γ на две области, причем граница проходит по значениям \vec{X} , которые удовлетворяют уравнению $\Lambda_n(\vec{X}) = \Lambda_0$.

Применение теории статистических решений к задачам распознавания образов позволяет построить, как и в случае обнаружения сигналов, оптимальный байесовский классификатор. Если величина потери при правильном выборе класса принимается нулевой, а при ошибочном решении составляет некоторую ненулевую величину, например 1 (симметричная функция потерь), то байесовский классификатор вычисляет дискриминантные функции вида:

$$d_j(\vec{x}) = W(\vec{x} / \omega_j) P(\omega_j), \quad j = 1, 2, \dots, M \quad (4)$$

Классификатор относит образ \vec{x} к тому классу ω_j , у которого значение дискриминантной функции $d_j(\vec{x})$ оказывается наибольшим. В формуле (4) $P(\omega_j)$ - вероятность появления образа из класса ω_j , $W(\vec{x} / \omega_j)$ - плотность распределения образов

каждого класса. Как и в случае обнаружения сигналов байесовская дискриминантная функция представляет собой разделяющую поверхность в виде линий для двумерных классов, если выполняется следующие условия:

- 1) классы имеют нормальное распределение;
- 2) все ковариационные матрицы единичные.

Опираясь на приведенные данные, полученные из теории статистических решений, психофизики инженерной и когнитивной психологии, была разработана структурная модель работы человека-оператора в системе «человек-машина», которая представлена на рисунке. 1:

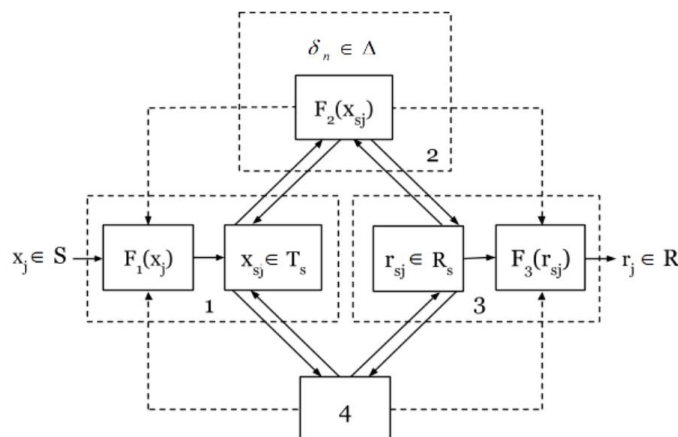


Рисунок 1. Структурная модель работы человека-оператора в задаче распознавания изображений. 1 – когнитивная подсистема, 2 – решающая подсистема, 3 – исполнительная подсистема, 4 – внутренние источники информации и память

В данной модели множество входных сигналов обозначим через \vec{S} . Тогда конкретное значение входного воздействия, поступающее в данный момент на вход наблюдателя, можно обозначить $x_j \in \vec{S}$. Поступившее воздействие x_j преобразуется блоком формирования сенсорного образа с помощью некоторого оператора $F_1(x_j)$ в образ воздействия или субъективный сенсорный результат наблюдения - x_{sj} . Множество $\{x_{sj}\}$ реализуется в сенсорном пространстве наблюдателя T_s , т.е. $x_{sj} \in T_s$.

Утверждение 1. Преобразование $\{x_j\} \rightarrow \{x_{sj}\}$ возможно, когда оператор F_1 является линейным преобразованием поступившего воздействия в сенсорный результат наблюдения.

Для модели работы человека-оператора, решающего задачу обнаружения и распознавания изображений, этот оператор может быть представлен в виде $F_1(x_j) = kx + b$, где k и b – некоторые константы.

При восприятии входного воздействия большое значение имеет порог ощущения или порог восприятия. Порог x^* есть критическое значение стимула, т.е. то минимальное воздействие, которое воспринимает нейронная система человека. Применение теории статистических решений позволило ввести порог Λ_0 , который обобщает понятие порога x^* , введенного Вебером и Фехнером, и используется на практике. В сенсорном пространстве порог как важный фактор включен в процесс принятия решения. На

основании понятия сенсорного пространства можно предложить психофизическую интерпретацию отношения правдоподобия $\Lambda_n(\vec{X})$. Можно считать, что $\Lambda_n(\vec{X})$ задает преобразование пространства стимулов в сенсорное пространство.

Утверждение 2. Оператор $F_2(x_{sj})$ задает однозначное отображение сенсорного пространства в пространство решений с помощью дискриминантных функций.

На основании сенсорных результатов наблюдения в соответствии со сформированным критерием оптимальности выполняется оператор $F_2(x_{sj})$, применяя некоторое правило принятия решения δ из множества возможных правил решения $\delta_n \in \Delta$.

Человек-оператор выбирает один из возможных способов действия r_{si} из множества возможных способов R_s , т. е. $r_{si} \in R_s$. Наконец, с помощью оператора $F_3(r_{sj})$ наблюдатель реализует выбранные действия – дает реакцию r_j из множества R , $r_j \in R$. Блок внутренней информации отражает влияние на другие подсистемы схемы прошлого опыта человека, особенностей процессов памяти и внимания, мотивационно-оценочных структур, а также эмоционального состояния. Для работы модели необходимо осуществить отображение, характер которого определяется работой блока принятия решения. Конечные члены этого отображения фиксирует экспериментатор.

$$S \rightarrow T_s \rightarrow R_s \rightarrow R$$

Данное отображение представляет собой информационный процесс обработки входной сенсорной информации с целью получения конкретных результатов действий человека.

На основании утверждений 1 и 2 работу блока принятия решения можно представить следующим образом. Человек-оператор выбирает такое правило решения $\delta \in \Delta$, которому при известном априорном распределении входного воздействия соответствует наименьшая средняя оценка потерь, т.е. байесовское правило решения.

Таким образом, структурная модель работы человека оператора может быть представлена в виде следующих подсистем: 1 – когнитивная, 2 – решающая, 3 – исполнительная, 4 – внутренние источники информации и память (рисунок 1).

Утверждение 3. Все элементы модели связаны единым информационным процессом и между ними осуществляется согласованное информационное взаимодействие, подразумевающее единство целей, тогда, когда выполняются утверждения 1 и 2.

Рассмотренная модель может служить основой для исследования характеристик работы человека-оператора и для методики проведения эксперимента.

Основная цель эксперимента – сбор данных об эффективности работы человека-оператора в задаче классификации изображений, для дальнейшего анализа и сравнения с результатами функционирования автоматизированного распознавательного комплекса.

Входными параметрами эксперимента являются искаженные изображения для классификации. Задача классификации изображений ставится следующим образом: имеются изображения двух классов, первый будем условно называть истинными целями (или просто «целями»), а второй ложными («ловушками»).

При распознавании предъявляются искаженные изображения в случайном порядке, необходимо определить принадлежность каждого из объектов к одному из классов. Экземпляры классов представлены на рисунке 2.



Рисунок 2. Образцы изображений классов «цели» (А) и «ловушки» (В, С, D)

Все изображения растровые, приведены к одному размеру и являются бинарными - состоят только из черных и белых пикселей.

В качестве искажающих эффектов используется ряд моделей зашумления изображения: аддитивный гауссов шум, импульсный шум, шум Релея. Входными параметрами в эксперименте выступают искаженные изображения двух описанных выше классов. Степень искажения изображений варьировалась в разных экспериментах. Каждый из операторов участвовал в нескольких экспериментах, с последовательным повышением предельной степени искажения. Эксперименты проводились с контрольными выборками разного объема (40-115 объектов), при этом с увеличением объема контрольной выборки увеличивалась и предельная степень искажения изображений.

Основным оборудованием в эксперименте является дисплей, на котором операторам предъявляются изображения для классификации. Для эксперимента было разработано специальное программное обеспечение на языке программирования C++ в среде Visual Studio 2010.

Представленная программа демонстрирует испытуемому поочередно зашумленные изображения. Интерфейс программы позволяет отнести демонстрируемое изображение к одному из классов («Цели» или «Ловушки»). По окончании эксперимента результаты могут быть сохранены в текстовый файл для дальнейшего анализа.

При проведении эксперимента измерялось зрительное утомление операторов с помощью показателя «устойчивости ясного видения». Расчеты велись с использованием следующей формулы:

$$C_k = \frac{1}{n} \sum_{i=1}^n C_{ik}, \quad (5)$$

где C_{ik} – показатель «устойчивости ясного видения» i -го оператора в k -м эксперименте, n – количество операторов, принимавших в нем участие. Из анализа полученных результатов измерения показателя «устойчивости ясного видения» сделан вывод, что в процессе прохождения эксперимента по классификации изображений операторы не подвергались значительному зрительному утомлению.

В третьей главе рассмотрены основные понятия нейросетевого моделирования, проведен ряд экспериментов по настройке и обучению многослойного персептрона. Многослойный персептрон обучался по алгоритму обратного распространения ошибки. В общем виде алгоритм обратного распространения может быть представлен таким образом.

Коррекция Δw_{ki} , применяемая к синаптическому весу, соединяющему нейроны k и i , определяется следующим правилом

$$\Delta w_{ki} = \eta \delta(n) x_k, \quad (6)$$

где η – параметр скорости обучения, $\delta(n)$ – локальный градиент, x_k – входной сигнал нейрона k , а n – число итераций. Значение локального градиента $\delta(n)$ зависит от положения нейрона в сети. Если нейрон i – выходной, то градиент равен

$$\delta(n) = e_i \cdot f'(v) = (d_i - y_i) \cdot f'(v), \quad (7)$$

где $e_i = d_i - y_i$ – сигнал ошибки для нейрона i , а $f'(v)$ – производная функции активации. Если нейрон i – скрытый, то градиент $\delta(n)$ равен произведению производной $f'(v)$ на взвешенную сумму градиентов, вычисленных для нейронов следующего слоя, которые связаны с данным нейроном

$$\delta(n) = f'(v) \sum_j \delta_j(n) w_{ij}, \quad (8)$$

где w_{ij} – синаптические веса, соединяющие нейроны i и j .

С помощью компьютерного моделирования проведен сравнительный анализ байесовского классификатора, классификатора, основанного на методе потенциальных функций и многослойного персептрона.

Сделан вывод, что помимо архитектуры сети, наибольшее влияние на эффективность работы нейронной сети оказывают обучающие выборки. С помощью нелинейного проекционного метода отображения Семмона был проведен экспериментальный и визуальный анализ влияния обучающих выборок на эффективность работы многослойного персептрона. Получено, что на способность обобщения многослойного персептрона влияет суммарная информативность обучающих примеров, выражающаяся в отличии их друг от друга.

На основании собранных данных сделан вывод о возможности применения нейросетевых алгоритмов для описания работы человека-оператора, решающего задачи распознавания изображений.

Для решения задачи распознавания изображений кораблей из главы II рассмотрен алгоритм работы сверточной нейронной сети. Коррекция синаптических весов сверточной сети рассчитывается по формуле

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta y_i(n) \delta_j(n), \quad (9)$$

где Δw_{ji} – коррекция синаптического веса, связывающего i и j нейрон, y_i – выход i -го нейрона, δ_j – градиент j -го нейрона, α – постоянная момента, η – скорость обучения, n – номер эпохи обучения. Показано, что для упрощения вычислений можно выделить участок изображения размером 32×32 пикселя, являющийся формой лоатора корабля, главным признаком распознавания. Проведен вычислительный эксперимент со сверточной нейронной сетью, распознающей фрагменты изображений и полное изображение. Анализ экспериментов показал, что эффективность работы сверточной нейронной сети уменьшается в 1.2 раза при распознавании полного изображения. Показано, что применение сверточной нейронной сети экономит вычислительные средства по сравнению с многослойным персептроном.

Применение нейросетевого подхода к распознаванию образов дало возможность представить имитационную структурную модель работы человека-оператора в виде двух подсистем моделируемых сетью извлечения признаков (когнитивная подсистема) и сетью, отвечающей за принятие решений (решающая и исполнительная подсистема).

Помимо непосредственно нейросетевых алгоритмов в когнитивной системе используется такой метод работы с изображениями, как сегментация. Значительная часть изображений для классификации повторяется, поэтому имеет смысл работать только с различающимися областями, которые разделяются на сегменты с помощью агломеративного метода иерархической кластеризации (рис. 4).



Рисунок 4. Выделение областей распознавания с помощью сегментации

При агломеративной кластеризации для определения меры близости между кластерами использовалось взвешенное евклидово расстояние

$$\rho_{BE}(x_i, x_j) = \sqrt{\sum_{l=1}^k \omega_l (x_{il} - x_{jl})^2}, \quad (10)$$

где x_{il}, x_{jl} - значения l -го признака у i -го (j -го) объекта ($l = 1, 2, \dots, k, i, j = 1, 2, \dots, n$), ω_l - вес, пропорциональный степени важности признака в задаче классификации. Обычно принимают $0 \leq \omega_l \leq 1$, где $l = 1, 2, \dots, k$. Рассмотрена работа гибридной нейронной сети нечеткой логики и оценена возможность ее применения для моделирования решающей подсистемы. Гибридная сеть должна осуществить следующее отображение:

$$y^k = f(x^k) = f(x_1^k, x_2^k, \dots, x_n^k), k = 1, 2, \dots, N, \quad (11)$$

при обучающем множестве

$$\{(x^1, y^1), \dots, (x^N, y^N)\}. \quad (12)$$

Для моделирования отображения f используется алгоритм нечеткого вывода с предикатными правилами вида:

Π_i : если x_1 есть A_{i1} , x_2 есть A_{i2} , ..., x_n есть A_{in} , то $y = z_i, i=1, 2, \dots, m$,

где A_{ij} - нечеткие числа треугольной формы, z_i - вещественные числа. Степень истинности i -го правила определяется с помощью операции умножения:

$$\alpha_i = \prod_{j=1}^n A_{ij}(x_j^k). \quad (13)$$

При этом выход нечеткой системы определяется дискретным аналогом центроидного метода:

$$o^k = \frac{\sum_{i=1}^m \alpha_i z_i}{\sum_{i=1}^m \alpha_i}, \quad (14)$$

Введение функции ошибки для k -го предъявленного вида

$$E_k = \frac{1}{2} (o^k - y^k)^2, \quad (15)$$

позволяет использовать градиентный метод для настройки параметров предикатных правил. Величины z_i корректируются соотношением:

$$z_i = z_i - \eta \frac{\partial E_k}{\partial z_i} = z_i - \eta(o^k - y^k) \frac{\alpha_i}{\alpha_1 + \alpha_2 + \dots + \alpha_m}, i = 1, 2, \dots, m, \quad (16)$$

где η – константа скорости обучения.

Сигмоидные функции принадлежности нечетких понятий A_1, A_2, \dots, A_N и B_1, B_2, \dots, B_N имеют следующий вид:

$$A_k(x) = \frac{1}{1 + e^{-b_k(x-a_k)}}, \quad (17)$$

$$B_k(x) = \frac{1}{1 + e^{b_k(x-a_k)}}$$

где b_k и a_k – настраиваемые параметры, $k = 1 \dots N$.

На основе нейросетевых моделей, гибридных нейронных сетей и сегментации разработан алгоритм классификации изображений, согласующийся с исследованиями в области инженерной и когнитивной психологии, психофизики и теории статистических решений, используемый для моделирования работы человека-оператора в системе «человек-дисплей».

Четвертая глава посвящена рассмотрению экспериментов по моделированию распознавательного комплекса, представленного в главе III. В результате вычислительных экспериментов и математического анализа сделан вывод о том, что в качестве ядра когнитивной подсистемы следует использовать самоорганизующиеся карты Кохонена (SOM): нейросетевой алгоритм, обучающийся без учителя, со сравнительно простой архитектурой. Каждый нейрон характеризуется n -мерным вектором весовых коэффициентов $w = [w_1, w_2, \dots, w_n]^T$, где n определяется размерностью входных векторов. На каждом шаге обучения из исходного набора данных выбирается один из образов, а затем производится поиск наиболее похожего на него вектора коэффициентов нейронов. Найденный нейрон признается нейроном-победителем. Под похожестью понимается расстояние между векторами, вычисляемое в евклидовом пространстве. Таким образом,

$$\|x - w_c\| = \min_i \{\|x - w_i\|\}, \quad (18)$$

где w – вектор весовых коэффициентов, i – индекс узла, c – нейрон-победитель. После определения нейрона-победителя веса сети корректируются. Для модификации весовых коэффициентов используется формула:

$$w_i(t+1) = w_i(t) + h_{ci}(t) \cdot [x(t) - w_i(t)], \quad (19)$$

где i – индекс узла, t – номер эпохи, вектор $x(t)$ выбирается случайно из обучающей выборки на итерации t , а $h(t)$ – функция соседства нейронов. Примером функции соседства является функция, определяющая окрестность узла-победителя:

$$h_{ci}(t) = \begin{cases} \alpha(t), i \in N_c(t) \\ 0, i \notin N_c(t) \end{cases}, \quad (20)$$

где $N_c(t)$ – множество индексов точек, соседствующих с узлом-победителем, $\alpha(t)$ – коэффициент скорости обучения. $N_c(t)$ и $\alpha(t)$ монотонно уменьшаются с течением времени.

На основе представленного нейросетевого классификатора был реализован программный комплекс, общая схема которого представлена на рисунке 5.

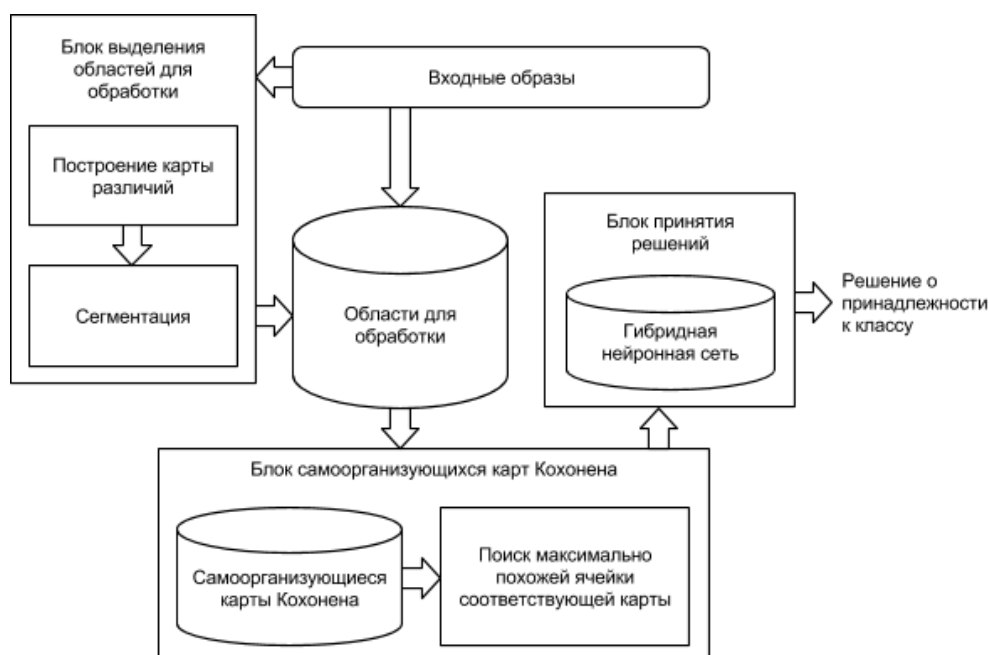


Рисунок 5. Структура программного комплекса по классификации изображений

Программное обеспечение было разработано на языке программирования C++ в среде Visual Studio 2010. Оно представляет собой библиотеки классов, реализующие ряд нейросетевых алгоритмов: многослойные сети прямого распространения, сверточные нейронные сети, самоорганизующиеся карты Кохонена, гибридные нейронные сети.

С помощью указанного программного обеспечения были проведены эксперименты с нейросетевым классификатором на основе самоорганизующихся карт Кохонена и гибридной сети нечеткой логики.

Эксперименты проводились с контрольными выборками разного объема, при этом с увеличением объема контрольной выборки увеличивалась и предельная степень искажения изображений. Результаты экспериментов представлены в таблице.

Таблица 2. Результаты экспериментов с человеком-оператором и нейронной сетью

Объем контрольной выборки		40 примеров, 15% макс. процент зашумления	55 примеров, 25% макс. процент зашумления	70 примеров, 35% макс. процент зашумления	85 примеров, 45% макс. процент зашумления	100 примеров, 55% макс. процент зашумления	115 примеров, 65% макс. процент зашумления
Человек-оператор	Р _{прав.}	0.99	0.98	0.96	0.91	0.85	0.80
Человек-оператор	Р _{ош.}	0.01	0.02	0.04	0,09	0.15	0.20
Нейронная сеть	Р _{прав.}	0.99	0.99	0.96	0.92	0.92	0.87
Нейронная сеть	Р _{ош.}	0.01	0.001	0.04	0.08	0.08	0.13

Статистический анализ полученных результатов, с помощью непараметрического рангового критерия Вилкоксона показал, что работу человека-оператора по распознаванию изображений можно описать предложенной автором моделью с вероятностью 0,95.

В заключении сформулированы основные результаты диссертационной работы.

1. Анализ систем «человек-машина» и «человек-компьютер» показал, что наибольший интерес представляют информационные системы, которые используются для получения информации, необходимой для распознавания изображений и принятия решений человеком-оператором, воспринимающим информацию с экрана дисплея компьютера.
2. На основании теории статистических решений и исследований в области инженерной психологии и психофизики синтезирована структурная модель работы человека-оператора в системе «человек-машина», учитывающая взаимосвязи между памятью, восприятием, генерированием внутренних представлений. Утверждения 1, 2, 3 позволяют рассматривать структурную модель работы человека-оператора как байесовскую оптимальную систему, которая может быть использована для решения задачи классификации изображений.
3. Разработана методика проведения эксперимента по классификации изображений человеком-оператором, отличительной чертой которой является отслеживание зрительного утомления оператора на основании вычисления показателя «устойчивости ясного видения» до и после эксперимента. Для проведения эксперимента разработано специальное программное обеспечение.
4. На основании изучения алгоритмов нейросетового моделирования и в соответствии со структурной моделью работы человека-оператора разработан алгоритм работы нейросетового комплекса распознавания изображений, состоящий из двух блоков, соответствующих когнитивной подсистеме и решающей. Для моделирования когнитивной подсистемы используются самоорганизующиеся карты Кохонена, а гибридная нейронная сеть нечеткой логики моделирует решающую подсистему.
5. Вычислительный эксперимент с нейросетовым распознавательным комплексом и эксперимент с группой операторов показали, что предложенная структурная модель и её программная реализация адекватно описывают работу человека-оператора. На основании статистического анализа можно сделать вывод о том, что работу человека-оператора по распознаванию изображений можно описать предложенной автором моделью с вероятностью 0,95. Проведен эксперимент, в котором показаны преимущества разработанного комплекса над алгоритмом сверточных нейронных сетей.
6. На базе разработанных алгоритмов разработано программное обеспечение для классификации изображений, которое зарегистрировано в федеральной службе по интеллектуальной собственности, патентам и товарным знакам. Разработанные библиотеки для моделирования сетей прямого распространения, самоорганизующихся карт Кохонена и гибридных нейронных сетей написаны с использованием парадигмы ОПП, могут быть с легкостью интегрированы в любые программы на языке C++.

Публикации автора в изданиях, рекомендованных ВАК РФ

1. Дуденков В.М. Распознавание изображений с помощью сверточной нейронной сети и нечеткого гибридного классификатора / Н.М. Новикова, В.М. Дуденков // Нейрокомпьютеры: разработка, применение – 2015. – №2. – С. 43-47.

2. Дуденков В.М. Нейросетевая модель работы человека-оператора в системе «человек-дисплей» / Н.М. Новикова, В.М. Дуденков // Научные ведомости БелГУ – 2015. №7(204)вып.34/1 – С. 153-158.

Свидетельства о регистрации программ

3. Свидетельство о государственной регистрации программы «DirectAndConvolutionNet 1.0» / Дуденков В.М. – М.: ФГУ ФИПС, 2014. Рег. № 2015618275 от 04.08.2015г.
4. Свидетельство о государственной регистрации программы «FuzzyAndCohonenNet 1.0» / Дуденков В.М. – М.: ФГУ ФИПС, 2014. Рег. № 2015618276 от 04.08.2015г.

Статьи и материалы конференций

5. Дуденков В.М. Оптимальный многослойный перцептрон для классификации образов / Н.М. Новикова, В.М. Дуденков // Кибернетика и высокие технологии XXI века (С&Т): материалы XIII Международ. науч.-техн. конф. – Воронеж, 2012. Т.1. – С. 109-119.
6. Дуденков В.М. Нейросетевые и детерминированные алгоритмы распознавания образов / Н.М. Новикова, В.М. Дуденков // Кибернетика и высокие технологии XXI века (С&Т): материалы XIV Международ. науч.-техн. конф. – Воронеж, 2013. Т.1. – С. 298-304.
7. Дуденков В.М. Использование сверточной нейронной сети в задаче распознавания графических изображений / В.М. Дуденков // Информатизация процессов формирования открытых систем на основе САПР, АСНИ, СУБД и систем искусственного интеллекта: материалы седьмой Международ. науч.-техн. конф. – Вологда, 2013. – С. 77-80.
8. Дуденков В.М. Моделирование сверточных нейронных сетей / В.М. Дуденков // Понрягинские чтения: материалы XXIV Воронежской весенней математической школы. – Воронеж, 2013. – С. 71-72.
9. Дуденков В.М. Анализ влияния примеров обучения на эффективность работы многослойного перцептрона / В.М. Дуденков // Информатика: проблемы, методология, технологии: материалы XIV Международ. науч.-метод. конф. – Воронеж, 2014. – С. 51-53.
10. Дуденков В.М. Распознавание изображений с помощью самоорганизующихся карт Кохонена и сетей нечеткой логики / В.М. Дуденков // Современные методы математики, теории управления и компьютерных технологий: материалы VII Международ. науч. конф. – 2014. – С. 131-133.
11. Дуденков В.М. Моделирование нейронной сети для распознавания изображений на основе гибридной сети и самоорганизующихся карт Кохонена / Н.М. Новикова, В.М. Дуденков // Аспирант. – 2015. - №2 – С. 31-34.
12. Дуденков В.М. Универсальный алгоритм классификации образов на основе самоорганизующихся карт Кохонена и гибридных сетей // Кибернетика и высокие технологии XXI века (С&Т): материалы XVI Международ. науч.-техн. конф. – Воронеж, 2015. – С. 1-6.