

МОДЕЛЬ И МЕТОД ПОСТРОЕНИЯ МНОГОУРОВНЕВОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ**И. В. Шевченко, Н. С. Шкарупа, М. В. Гончар**

Кременчугский национальный университет имени Михаила Остроградского

ул. Первомайская, 20, г. Кременчуг, 39600, Украина. E-mail: ius.shevchenko@gmail.com

Рассмотрены существующие подходы к построению систем распознавания, и предложен метод синтеза многоуровневой архитектуры нейросетевой системы для решения задач диагностики, основанный на вложенности микрокластеров, кластеров и макрокластеров в пространстве признаков. Вторым условием разбиения процесса распознавания на этапы является декомпозиция множества признаков на объединяющие и разделяющие подмножества. При помощи объединяющего подмножества признаков на первом этапе внутри макрокластера формируются кластеры. На втором этапе при помощи разделяющего подмножества признаков внутри кластеров выделяются микрокластеры как группы объектов с определенными свойствами. Двухуровневая система распознавания, полученная в результате синтеза на основе разработанного метода, является логически прозрачной и позволяет повысить точность распознавания по сравнению с классическими перцептронами на том же наборе признаков. В целом разработанная модель и метод синтеза нейросетевой многоуровневой системы является простым и эффективным средством решения практических задач диагностики и распознавания образов.

Ключевые слова: задача распознавания, синтез архитектуры системы, модель, тестирование.

МОДЕЛЬ І МЕТОД ПОБУДОВИ БАГАТОРІВНЕВОЇ СИСТЕМИ РОЗПІЗНАВАННЯ**І. В. Шевченко, Н. С. Шкарупа, М. В. Гончар**

Кременчуцький національний університет імені Михайла Остроградського

вул. Першотравнева, 20, м. Кременчук, 39600, Україна. E-mail: ius.shevchenko@gmail.com

Розглянуто існуючі підходи до побудови систем розпізнавання, і запропоновано метод синтезу багаторівневої архітектури нейромережевої системи для вирішення задач діагностики, що ґрунтується на вкладеності мікрокластерів, кластерів і макрокластерів до простору ознак. Другою умовою розбиття процесу розпізнавання на етапи є декомпозиція множини ознак на підмножини, які об'єднують і розділяють класи. За допомогою підмножини ознак, що об'єднує, на першому етапі всередині макрокластера формуються кластери. На другому етапі за допомогою підмножини ознак, що розділяють, у середині кластерів виділяються мікрокластери як групу об'єктів з певними властивостями. Дворівнева система розпізнавання, отримана в результаті синтезу на основі запропонованого методу, є логічно прозорою і дозволяє підвищити точність розпізнавання порівняно з класичними перцептронами на тому ж наборі ознак. У цілому розроблена модель і метод синтезу нейромережевої багаторівневої системи є простим і ефективним засобом вирішення практичних задач діагностики і розпізнавання образів.

Ключові слова: задача розпізнавання, синтез архітектури системи, модель, тестування.

АКТУАЛЬНОСТЬ РАБОТЫ. Традиционно большинство систем распознавания для решения конкретной задачи используют какой-либо один метод классификации. Однако ни один метод классификации не является универсальным и абсолютно надежным, и разные методы, как правило, основываются на разных концепциях. Вследствие этого, методы, обеспечивающие одинаковую или близкую надежность классификации для всей обучающей и/или контрольной выборки в целом, для отдельных экземпляров, могут давать разные результаты. Это особенно характерно для экземпляров, находящихся на стыках границ классов [1].

Классическим методом распознавания и диагностики в рамках математической статистики является метод на основе байесовской вероятности, которая позволяет вычислить условную вероятность диагностических решений при данном наборе признаков на основании априорной вероятности диагноза и условной вероятности наблюдения признаков [2]. Применение данного подхода бывает затруднительно из-за структурной сложности объектов диагностики и взаимовлияния разных факторов. Кроме того, если реальные функции распределения гипотез в

пространстве входных признаков не согласуются с предположением о нормальности, теоретические оценки не будут соответствовать эмпирическим результатам.

Другой проблемой является неопределенность информации. Это связано со следующими причинами:

- экспертные оценки субъективны;
- степень проявления признаков может быть различной;
- часть признаков может отсутствовать.

Естественно, что системы, построенные на основе одного классификатора, не будут удовлетворять высоким требованиям надежности, предъявляемым во многих областях деятельности, в частности – в медицинской диагностике. Поэтому крайне важно создавать комитетные системы классификации, интегрирующие различные модели и методы [3].

В настоящее время при построении распознающих и диагностических систем широкое распространение получили искусственные нейронные сети (НС). Известно, что нейросетевая реализация экспертной диагностической системы значительно уменьшает затраты времени на разработку и суще-

ственно знижує вимоги до фахівців у вибраній предметній області при побудові системи [4–6]. Так як в штучних нейронних мережах рішення задач виробляється не на основі програм, що реалізують надзвичайно складні обчислювальні алгоритми, а шляхом навчання на наявних прикладах, штучні нейронні мережі інваріантні до функцій густоти розподілу гіпотез, розглядаються як альтернативи.

Однак при побудові нейронних мереж на основі більшості відомих методів виникає ряд проблем, таких, як необхідність задавання користувачем топології та параметрів НС, ініціалізація ваг НС, повільність та ітераційність процесу навчання нейронних мереж, а також надлишковість структури та складність логічної інтерпретації отриманих мереж.

Проблема точності розпізнавання. Для підвищення точності (ступеня розрізнення деталей образу) нейронної мережі розроблені нейронні мережі третього покоління – сверточні мережі глибокого навчання з дуже складною архітектурою [7].

Существует также *проблема залежності ознак*. Якщо ознаки корелюють, то статистичні методи будуть давати сбой. По матрицям взаємозалежностей ознак можна для кожної ознаки виділити набори незалежних одна від одної ознак. В роботі [1] показано, що в формулу Байеса повинні входити тільки взаємозалежні ознаки, т.е. ознаки, що не корелюють між собою. Комплекси таких ознак в роботі [1] названі «ядрами». Такі ядра не завжди можуть бути отримані з допомогою статистичних методів. Однак у даного підходу існує суттєвий недолік. Після визначення найкращого ядра передбачається, що воно буде використовуватися при діагностиці кожного пацієнта, т.е. потрібно, щоб набір ознак хворого відповідав ознакам ядра. На практиці набір реальних ознак пацієнта рідко збігається з ядром.

Мощним інструментом розпізнавання є теорія розмитого логіки та розмитих множин, що дозволяє описувати розмиті поняття та знання, оперувати цими знаннями та здійснювати розмитий або чіткий висновок. Використання методів розмитого логіки має наступні переваги [8]:

- дозволяє сформулювати експертні знання об'єктах розглядаємої предметної області в лінгвістичній формі, т.е. знання, описувані не тільки кількісними, але й якісними характеристиками;
- забезпечує можливість обчислення значення ступеня належності елементів до розмитій множині з використанням операцій заперечення, кон'юнкції, диз'юнкції та інших, що враховують концепцію невизначеності;
- забезпечує рішення задач розпізнавання з

використанням розмитих відстаней між об'єктами.

Бази знань класичних розмитих продукційних систем заповнюються експертами. Якщо є можливість генерувати приклади на основі чисельних значень ознак, то, використовуючи розмиті нейронні мережі з алгоритмом Сугено, можна синтезувати систему розпізнавання з мінімальним залученням експертів [9]. Однак архітектура такої системи достатньо складна, навіть якщо використовувати ієрархічну базу знань. Що стосується точності розпізнавання розмитих систем, то для її підвищення необхідно збільшувати кількість термів лінгвістичних змінних, що призводить до прояву «прокляття розмірності» при побудові бази знань.

З урахуванням вищесказаного можна сформулювати проблему – прагнення збільшити точність розпізнавання призводить до збільшення обчислювальної складності статистичних, нейронних та розмитих алгоритмів розпізнавання образів в багатовимірному просторі ознак.

Ціллю роботи є створення вдосконалих мереж та методів синтезу систем розпізнавання, що є актуальною задачею.

МАТЕРІАЛ І РЕЗУЛЬТАТИ ІСЛЕДУВАНЬ. При побудові систем розпізнавання образів, до яких належать і діагностичні системи, необхідно вирішити три основні задачі:

- вибір інформативних ознак (традиційно – оффлайн-методи);
- розробка архітектури системи розпізнавання (структури системи та способу взаємодії її складових частин);
- навчання та оптимізація налаштувань системи (мінімізації повної помилки на репрезентативній тестуючій вибірці).

В даній роботі пропонується модель та метод синтезу багатовимірної нейронної мережі для рішення задач діагностики, архітектура якої ґрунтується на двох гіпотезах:

- вложенности мікрокластерів, кластерів та макрокластерів в просторі ознак;
- декомпозиції множини ознак на об'єднуючі та розділювальні підмножини.

Ідея розпізнавання образів в будь-якому випадку ґрунтується на гіпотезі компактності образів в просторі ознак в залежності від ступеня визначеності входних класів, які потрібно розпізнавати. Задачі розпізнавання та класифікації образів є типовими для штучних нейронних мереж (ІНС). Розвиток теорії ІНС пов'язаний з появою нових архітектур мереж. Архітектура мережі, призначеної для розпізнавання образів, залежить від наявності та потужності навчального множини. При цьому розглядаються дві парадигми: навчання з вчителем та навчання без вчителя. При наявності навчальних пар «вхід/вихід» місце займає навчання з вчителем. В межах цієї па-

радиґмы можно выделить две наиболее популярны архитектуры: многослойная сеть прямого распространения и двухслойная сеть, использующая радиально-базисные нейроны (RBF-сеть).

В *сети прямого распространения* часто применяется обучение с использованием алгоритма обратного распространения ошибки, в котором подстройка всех весовых коэффициентов производится одновременно при каждом предъявлении обучающего примера [6]. Процедура обучения такой сети требует большого числа эпох и заканчивается нахождением локального минимума, тогда как целью обучения является поиск глобального минимума для критерия распознавания. Поэтому часто для отыскания глобального минимума используют и методы случайного поиска, такие, как генетические алгоритмы [9]. При большом числе классов, большом числе входов и скрытых слоев обучение сети прямого распространения занимает много времени и ресурсов.

RBF-сеть – это *двухслойная сеть*, содержащая распознающий слой радиально-базисных нейронов и интерпретирующий слой, в котором производится формирование решения. Каждый нейрон первого слоя принимает всю информацию от входов сети и в процессе обучения запоминает ее в синапсах одного из нейронов, который выбран для обучения конкретному образу. Отклик обученного нейрона на пример своего класса близок к 1, а на примеры из других классов – к 0. При этом при предъявлении любого входного сигнала отклики разных уровней будут получены во всех нейронах первого слоя. Второй слой является, по сути, однослойным персептроном, который можно обучать любым известным методом. Выходы нейронов второго слоя можно интерпретировать как степень уверенности или вероятности в распознавании определенного класса объектов.

Классическая последовательность обработки данных для распознавания образов выглядит следующим образом [2]:

1. Формирование массивов исходных данных для оценки информативности признаков.
 2. Кластеризация данных.
 3. Оценка информативности признаков.
 4. Ранжирование признаков по информативности.
 5. Отбор признаков для распознавания сигналов.
- Малоинформативные признаки отбрасываются по очереди общего ранжирования, независимо от того, к какому классу сигналов они принадлежат.

Анализ различных предметных областей показывает, что целесообразный состав признаков для разных классов может быть разным. В задачах распознавания нередки случаи, когда в пространстве признаков нужно вначале выделить макрокластеры (например, классы заболеваний – заболевания органов дыхания, заболевания органов пищеварения и т.п.). Затем внутри макрокластера (класса) необходимо выделить кластер как подкласс (например, острые респираторные инфекции верхних дыхательных путей), а затем распознать конкретное заболевание и его стадию и, с некоторой долей вероятности, поставить диагноз.

На рис. 1 показано пространство признаков, в котором подмножества объектов А и Б, В и Г перекрываются.

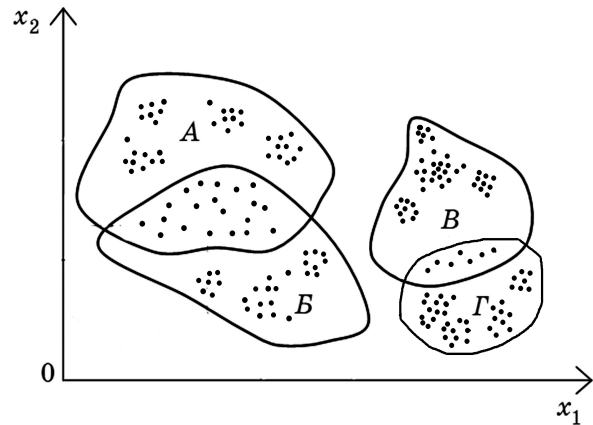


Рисунок 1 – Иллюстрация к разделению макрокластеров, кластеров и микрокластеров

Для распознавания классов (подмножеств) А и Б (В и Г) признаки x_1 и x_2 согласно классической парадигме считаются малоинформативными. Однако, если подмножества объектов $A \cup B \cup G$ образуют макрокластер, а подмножества $A \cup B$ и $B \cup G$ образуют кластеры, то признаки x_1 и x_2 могут быть использованы для формирования «ядра признаков», указывающего на кластер. По другим признакам (не показаны на рис. 1) можно распознавать объекты на уровне микрокластеров, показанных на рис. 1 в виде сгустков точек.

В связи со сказанным выше можно разделить признаки по двум аспектам.

Первый аспект – способность признака выделить на множестве образов макрокластер, включающий ряд пересекающихся кластеров. Каждый кластер соответствует подклассу образов. Внутри каждого кластера имеются сходные образы, различающиеся, однако по некоторым признакам – микрокластеры. Микрокластеры содержат объекты, сходные по всем существенным для данного микрокластера признакам.

Второй аспект – способность к тонкому различению образов внутри отдельного кластера. Если такое разделение возможно, то концепция распознавания предусматривает, как минимум, три этапа.

На первом этапе требуется распознать макрокластер, на втором – кластер, на третьем – объект внутри кластера. В реальных задачах часто на уровне макрокластера задачу выбора решает эксперт, формирующий онтологию предметной области [10, 11]. Ограничим задачу построения системы распознавания двумя уровнями, учитывая, что операции, описанные для этих уровней, могут итеративно повторяться на более высоких и более низких уровнях.

Общий подход к анализу разделительных свойств признаков заключается в следующем. Множество признаков $S_S \subset S$ является *разделителем*, если подмножества объектов, состоящие из значений признаков из S_S , попарно различны для всех объектов, принад-

лежащих разным классам [12]. При оценке разделяющей способности признака s_k на множестве гипотетических объектов H учитывается количество признаков, которыми различаются те же пары объектов, что и признаком s_k . Чем меньше признаков различают те же пары объектов, что и признак s_k , тем выше разделяющая способность (ценность) признака s_k . Однако если признаки зависимы, то при выборе наиболее информативной подсистемы признаков оценками их индивидуальной информативности руководствоваться нецелесообразно [12].

Метод построения системы распознавания. Сформулируем этапы метода формирования многоуровневой системы распознавания, предполагая, что для анализа признаков на разных уровнях можно использовать либо перцептрон, либо RBF-сеть.

Этап 1. Обозначим множество распознаваемых внутри макрокластера образов $H = \{h_i, i = \overline{1, I}\}$. Обозначим множество кластеров $C = \{c_z, z = \overline{1, Z}\}$ внутри макрокластера MC . Представим множество признаков как три подмножества:

$$S = \langle SM, SC, SO \rangle, \quad (1)$$

где SM – множество признаков, указывающих на определенный макрокластер, который, по сути, определяет узкую предметную область; SC – подмножество признаков, разделяющих макрокластер на кластеры; $SC = \{s_{c_g}, g = \overline{1, G}\}$; SO – подмножество признаков, разделяющих образы внутри кластера на микрокластеры: $SO = \{s_{o_k}, k = \overline{1, K}\}$. Каждый микрокластер есть подмножество объектов, относящихся к одному образу или группе подобразов (например, различные стадии одного заболевания).

Очевидно, что в пределах кластера у признаков из подмножества SC должна быть малая дисперсия D_{SC} (признаковое ядро кластера), т.е. объекты внутри кластера по этим признакам малоразличимы. Данное подмножество признаков должно быть использовано для идентификации кластера на первом уровне распознавания. У признаков из подмножества SO внутри кластера должна быть значительная дисперсия: $D_{SO} > D_{SC}$.

Этап 2. Анализ признаков на разделяющую способность. Для анализа сформируем таблицу-отношение RHS , в которой каждая строка – объект, а каждый столбец – признак (симптом, свидетельство). Тем самым получим отношение, которое ставит в соответствие образы и значения признаков:

$$R \subseteq H \times S. \quad (2)$$

Сформируем верхнетреугольные матрицы DM_k учета различимости образов по каждому признаку s_k . Таких матриц размером $I \times I$ должно быть K – по числу признаков.

Поставим данному признаку s_k в соответствие совокупность различимых k -м признаком пар объек-

тов. Для этого каждый признак (столбец таблицы RHS) анализируется на способность к различению объектов. Назначим для каждого признака s_k величину порога различимости объектов DT_k (discrimination threshold). При помощи порогового преобразования можно выделить те строки таблицы RHS , для которых признак s_k имеет заметно разные значения. То есть, если $\Delta s_{kij} \geq DT_k$, считаем, что признак различает объекты H_i и H_j . Соответственно, если $\Delta s_{kij} < DT_k$, то признак на первом уровне анализа может быть учтен в признаковом ядре кластера SC : $s_{kij} \in SC$. Значения порогов DT_k можно настраивать во время тестирования и отладки системы.

Для формирования подмножества разделяющих признаков осуществляем вложенный цикл обработки по таблице RHS , сравнивая значения s_{ki} и s_{kj} для разных объектов и, если выполняется условие $\Delta s_{kij} > DT_k$, заносим результат в виде 1 в k -ю матрицу различимости DM_k .

Результат – данные в матрицах DM_k , на основании которых можно определить, сколько признаков нужно для уверенного распознавания каждого объекта (разделяющая способность каждого признака относительно каждого объекта).

Этап 3. Формирование признакового ядра кластера. В матрицах DM_k выделяем ячейки, помеченные нулями, и заполняем массив CF (core features) учета признаков с низкой разделяющей способностью. Для отбора признаков s_k в массив CF установим порог ND количества пар объектов, которые не разделяются данными признаком. Будем считать, что признак s_k имеет низкую разделяющую способность, если для $ND\%$ пар объектов выполняется условие $\Delta s_{kij} < DT_k$.

Этап 4. Определение подмножеств признаков для гарантированного распознавания всех образов.

Шаг 4.1. Проверка условия различимости. Эта задача решается поэлементной дизъюнкцией матриц DM_k . Результат – матрица R , показывающая «суммарную» различимость объектов на заданном множестве признаков. Нули в ячейках этой матрицы показывают, какие объекты нельзя различить при заданных значениях порогов различимости DT_k . При необходимости возможна коррекция значений DT_k с целью устранить неразличимость объектов.

Шаг 4.2. Определение подмножеств признаков для гарантированного распознавания каждого образа. Эта задача решается прямым перебором покрытий признаками подмножеств объектов и отдельных объектов внутри кластера. Размер и количество подмножеств объектов не регламентированы в пределах I . Результат – таблица RR , в которой сопоставлены объекты (строки) и признаки (столбцы), и единицами отмечены в каждой строке необходимые для распознавания объекта признаки.

Для распознавания образов (микрокластеров) внутри выбранного кластера можно использовать как метод опорных векторов (линейный перцептрон

с выборочным соединением рецепторов и распознающих нейронов), так и нейроны с ядерными функциями.

Этап 5. Учет влияния дополнительных факторов (additional factors). Обозначим множество дополнительных факторов $AF = \{af_l, l = \overline{1, L}\}$. Для учета влияния сопутствующих факторов (в медицине это возраст, пол, длительность наблюдения симптомов, профессия, алкоголь, курение, наркотики и пр.), влияющих на вероятность того или иного заболевания, необходимо сформировать отдельный матричный классификатор, выходы которого будут давать дополнительные признаки для более точного распознавания образов.

Шаг 5.1. Построим отображение в виде матрицы, в которой столбцы соответствуют дополнительным факторам AF_l , а строки – суммарному влиянию факторов на степень достоверности результата распознавания образа H_i :

$$R_{AFH} \subseteq AF \times H \quad (3)$$

Шаг 5.2. Элементы матрицы R_{AFH} представим как коэффициенты влияния факторов на достоверность распознавания образа H_i . Определим диапазон изменения степени влияния $a_{ij} \in [0, 1]$. Аналогично определим диапазон изменения степени уверенности в наличии фактора $af_l \in [0, 1]$.

Шаг 5.3. Экспертная оценка степени влияния каждого фактора на достоверность распознавания образов. Используется группа экспертов, каждый из которых заполняет таблицу «фактор-образ» и десятибальную шкалу оценки влияния. Далее выставленные оценки усредняются и нормируются к диапазону $[0, 1]$.

Шаг 5.4. Алгоритм оценки влияния сводится к уточнению степеней уверенности наличия факторов и вычислению скалярных произведений входного вектора степеней уверенности AF и строк матрицы R_{AFH} . Полученные результаты подаются на входы персептронов второго слоя НС для взвешенного суммирования с другими признаками.

Этап 6. Построение нейросетевой модели.

Шаг 6.1. Получение и кластеризация данных. Определение наличия метакластеров, кластеров, микрокластеров, их количества и классификация (присвоение имен метакластерам, кластерам и микрокластерам).

Шаг 6.2. Формирование первого слоя НС. Используя таблицу RR , формируем для каждого кластера множество SC входов нейрона с ядерной функцией активации вида

$$f(x) = \exp(-\alpha \|x - \bar{c}\|^2), \quad (4)$$

где \bar{c} – вектор центров множества одномерных радиально-симметричных функций; $\|x - \bar{c}\|$ – норма

вектора отклонений входной переменной от центров одномерных функций, расположенных на координатах пространства признаков. Параметр α связан с радиусом рассеяния входных переменных r соотношением: $\alpha = \frac{1}{2r^2}$. Норма вектора отклонений рассчитывается как евклидово расстояние:

$$\|x - \bar{c}\| = \sqrt{(x_1 - c_1)^2 + (x_2 - c_2)^2 + \dots + (x_m - c_m)^2}.$$

Выход каждого нейрона первого слоя должен активизировать отдельный блок второго слоя НС, распознающего образы на уровне микрокластера.

Шаг 6.3. Формирование второго слоя НС.

При использовании метода опорных векторов требуется I распознающих нейронов второго слоя, каждый из которых соединен синапсами со своим подмножеством признаков, которое указано в таблице RR (см. шаг 4.2).

Выходы нейронов первого слоя являются сигналами, играющими роль смещения (модуляторов) для нейронов второго слоя.

Шаг 6.5. Формирование матричного классификатора для учета влияния сопутствующих факторов, выходы которого будут дополнительными входами нейронов второго слоя НС.

Модель многоуровневой системы распознавания. Как видно из предыдущих материалов, система распознавания имеет три функциональных блока. Первый блок должен распознавать кластер внутри макрокластера, второй блок распознает микрокластеры внутри кластера. Третий блок корректирует выходные значения нейронов второго блока с учетом влияния сопутствующих факторов. Связь выхода d_i , указывающего на степень вероятности правильного распознавания образа h_i с признаками so, sc, caf , показана в выражении (5).

$$d_i(so, sc, caf) = \psi \left(\sum_{k=1}^{K_i} w_k so_k + \exp(-\alpha \|sc - c\|^2) + caf_i \right), \quad (5)$$

где w_k – веса нейронов второго слоя; so_k – степени уверенности или значения признаков, разделяющих микрокластеры; sc – степени уверенности или значения признаков, разделяющих кластеры; caf_i – смещение, учитывающее степень влияния дополнительных факторов; K_i – количество признаков, отделяющих образ h_i ; $\psi(s) = s$ – активационная характеристика нейрона выходного слоя.

Таким образом, выражения (1)–(5) определяют математическую модель многоуровневой системы распознавания. Количество уровней распознавания может быть большим, в зависимости от решаемой задачи. При этом порядок и содержание этапов синтеза каждого уровня системы остаются такими, как они описаны выше. Пример схемы НС показан на рис. 2.

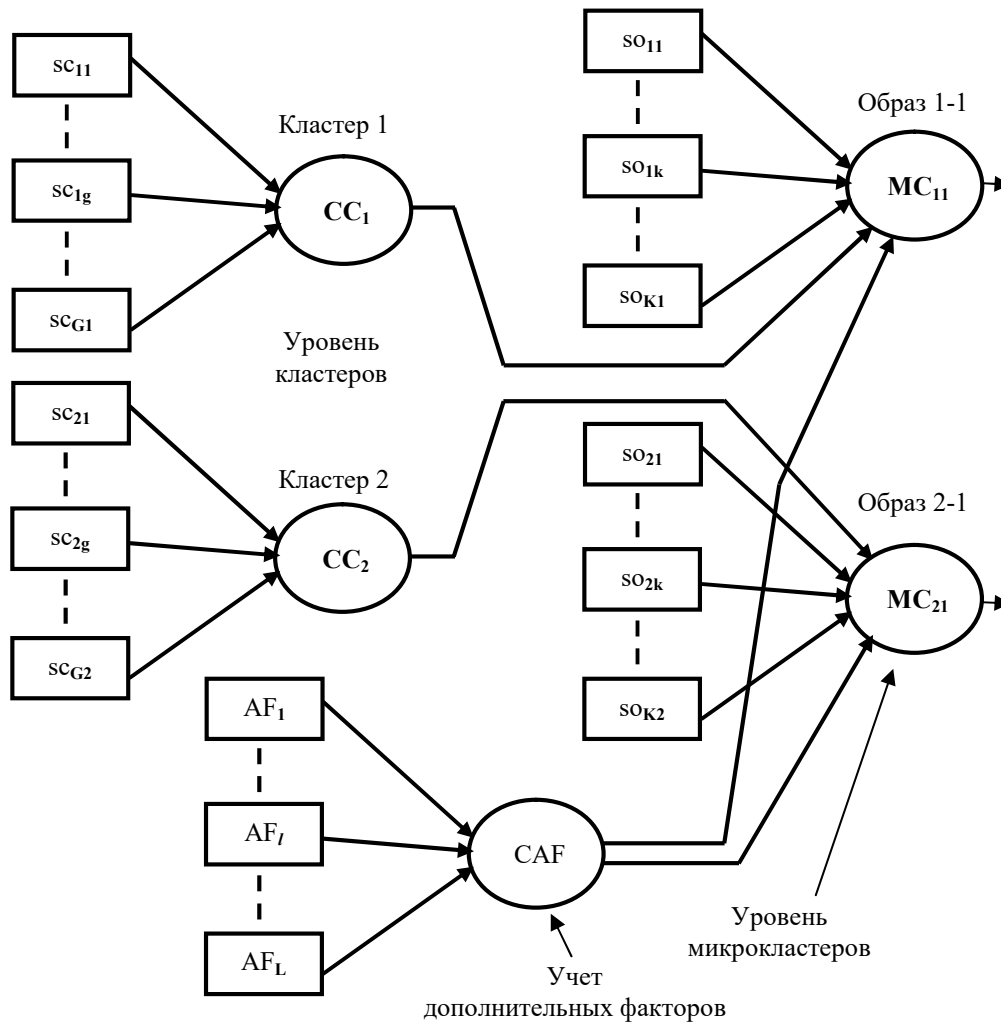


Рисунок 2 – Пример структуры многоуровневой НС

Для простоты показаны два ядра для кластеров и два персептрона второго слоя, распознающие образы на уровне микрокластеров.

Обучение системы распознавания. Обучение проводится в несколько этапов. На первом этапе формируется подмножество SC признаков ядра каждого кластера. Координаты объектов обучающей выборки по данному подмножеству признаков используются для формирования ядерной функции (4) каждого кластера.

На втором этапе формируется матрица влияний дополнительных факторов R_{AFH} , поскольку её выходные сигналы должны войти в примеры обучающих выборок для нейронов второго слоя.

На третьем этапе формируются примеры для обучения и тестирования нейронов второго слоя. Каждый пример состоит из трех массивов данных:

- массив значений степени уверенности в наличии признаков из подмножества SO . Этот массив дополняется значением смещения, полученным от ядерного нейрона первого слоя, распознающего соответствующий кластер;
- массив значений степени уверенности в наличии дополнительных факторов AF .
- массив шаблонных значений выходов нейронов

второго слоя.

На четвертом этапе происходит обучение нейронов второго слоя с использованием правила Уидроу-Хоффа [6].

Пример практической реализации предложенного метода построения системы распознавания.

Для проверки работоспособности и точности многоуровневой системы распознавания была выбрана область медицинской диагностики.

Хорошо известно, что многие диагностические программы используют для определения диагноза метод вычисления условной вероятности гипотезы о том или ином заболевании. Вычисление осуществляется по формуле Байеса [1]. Степень выраженности симптома врач фиксирует отдельно, и эта величина учитывается при вычислении вероятности наличия конкретного заболевания.

В качестве исходных данных для расчета условной вероятности гипотезы используются известные значения условных вероятностей симптомов, сопутствующих данному заболеванию. Часто это субъективные оценки, отражающие мнение эксперта, но, тем не менее, они достаточно точно отражают его опыт.

При реализации данного метода в базе знаний экспертной системы хранятся для каждого заболевания списки симптомов. Для каждого симптома в БЗ фиксируются две величины – $P(S:H)$ – условная вероятность наличия данного симптома при наличии данного заболевания и $P(S:\bar{H})$ – условная вероятность наличия данного симптома при отсутствии данного заболевания. Эти величины можно использовать для вычисления отношения правдоподобия [1]:

$$ОП(H/E) = \frac{P(E/H)}{P(E/\bar{H})}, \quad (6)$$

которое, в свою очередь, можно интерпретировать как значимость симптомов для диагностики конкретного заболевания, то есть, использовать, как значения признаков.

Для проверки возможности использования отношений правдоподобия и уровней уверенности наличия симптомов как сигналов на входе нейронной сети был построен и испытан два перцептрона,

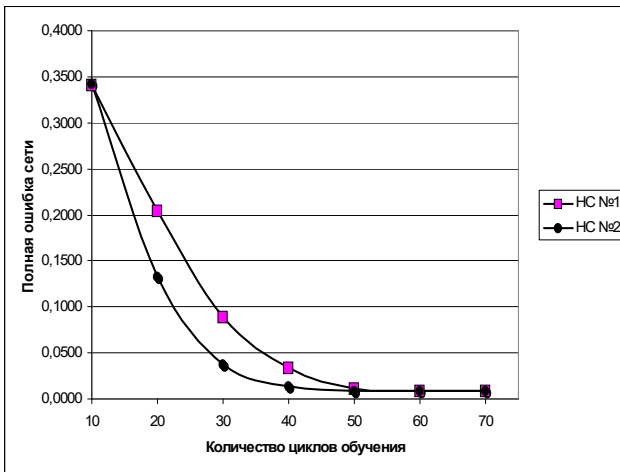


Рисунок 3 – Динамика обучения перцептронів

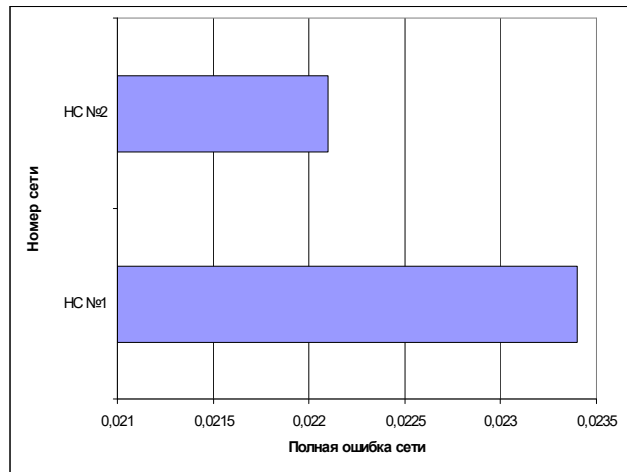


Рисунок 4 – Повна помилка перцептронів

Для обучения и тестирования генерировались выборки, содержащие по 100 примеров. Полная ошибка сети вычислялась как сумма средних квадратических ошибок сети по каждому предъявленному примеру, отнесенная к числу предъявленных примеров. Видно, что сеть № 2 несколько превосходит классический перцептрон по скорости обучения и точности, что дало уверенность в том, что использование отношений правдоподобия в качестве признаков вполне возможно.

После вычисления отношения правдоподобия (ОП) по каждому симптому для каждого заболевания, и проведения процедуры кластеризации отдельно по каждому симптому, были выделены симптомы, имеющие близкие значения ОП для двух групп заболеваний. Эти симптомы служат признаками для выделения двух кластеров. Соответствующие ячейки табл. 1 заполнены цифрой 1 для первого

стандартной структуры с одним скрытым слоем.

На входы первого перцептрона подавались степени выраженности симптомов в диапазоне от 0 до 1. Перцептрон должен распознавать диагноз заболевания. Обучению подвергались выходной и скрытый слои. Второй перцептрон был построен по той же схеме, однако весовыми коэффициентами скрытого слоя были отношения правдоподобия, и скрытый слой не подвергался обучению. Для экспериментов была выбрана группа заболеваний дыхательных путей и характерные для этой группы симптомы: 1. Насморк, 2. Боль в горле. 3. Кашель. 4. Головная боль. 5. Повышенная или высокая температура. 6. Выраженное общее недомогание. 7. Гиперемия и покраснение миндалин. 8. Одышка.

Процедуры обучения перцептронів описаны ниже. Результаты обучения и тестирования всех сетей приведены на графиках рис. 3 и 4.

кластера и цифрой 2 для второго кластера. Символ d обозначает разделимость заболеваний по данному признаку и, соответственно, возможность использования данного признака на втором уровне системы распознавания для дифференциальной диагностики (табл. 1).

В соответствии с полученным разделением на кластеры первый уровень системы имел два ядра признаков, так, как это показано на рис. 2.

На втором уровне каждое заболевание, относящееся к одному из кластеров, распознается нейроном с линейной функцией активации. К каждому нейрону подключены признаки, согласно таблице RR разделяющие образы внутри кластера. К этим нейронам подключены также выходы соответствующего ядра первого слоя и выходы матрицы учета влияния дополнительных факторов.

Таблица 1 – Разделение заболеваний и симптомов на кластеры

Заболевания	Симптомы							
	1	2	3	4	5	6	7	8
ОРВИ	1	<i>d</i>	1	<i>d</i>	<i>d</i>	1	<i>d</i>	<i>d</i>
Фарингит	<i>d</i>	1	<i>d</i>	<i>d</i>	1	<i>d</i>	<i>d</i>	<i>d</i>
Острый тонзилит	<i>d</i>	1	<i>d</i>	1	1	1	1	<i>d</i>
Грипп	<i>d</i>	<i>d</i>	<i>d</i>	1	1	1	<i>d</i>	<i>d</i>
Ларингит	<i>d</i>	1	1	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>
Острый бронхит	<i>d</i>	<i>d</i>	1	<i>d</i>	1	1	<i>d</i>	<i>d</i>
Хронический бронхит	<i>d</i>	<i>d</i>	2	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>	2
Ангина	<i>d</i>	1	<i>d</i>	1	1	1	1	<i>d</i>
Пневмония	<i>d</i>	<i>d</i>	2	<i>d</i>	2	2	<i>d</i>	<i>d</i>
Туберкулез	<i>d</i>	<i>d</i>	2	<i>d</i>	<i>d</i>	2	<i>d</i>	2

Для обучения и тестирования нейронов второго слоя были сгенерированы обучающие и тестирующие выборки. Генерация примеров осуществлялась по «принципу варьирования данных» – методике, идея которой почерпнута из работы [13]. Применение варьирования имеющихся данных к решению прикладных задач идентификации позволяет получить возможно более полный набор статистических характеристик результатов решения конкретной задачи, включая распределения оценок параметров, моменты соответствующих распределений и т.п.

Применительно к данной задаче суть метода заключается в том, что эксперт сообщает возможный диапазон значений коэффициентов уверенности для каждого симптома по каждому заболеванию, а также среднее значение этого параметра. Для квалифицированного специалиста это сделать нетрудно. Далее специальный программный модуль генерирует псев-

доданные по каждому симптому, используя нормальное или равномерное распределение плотности вероятностей в указанном диапазоне.

Все сети обучались на одной и той же обучающей выборке, в которую входило 400 примеров, по 40 примеров на каждое заболевание. В каждом цикле обучения сети использовалось 100 примеров, выбираемых таким образом, чтобы по каждому заболеванию было предъявлено 10 примеров. Внутри подмножества примеров по одному заболеванию примеры выбирались случайным образом.

На рис. 5 показана доля правильно диагностированных заболеваний системой, построенной на вычислении условной вероятности по формуле Байеса (А), обученным персептроном без разделения признаков (В) и двухуровневой системой, построенной по предложенному методу (С).

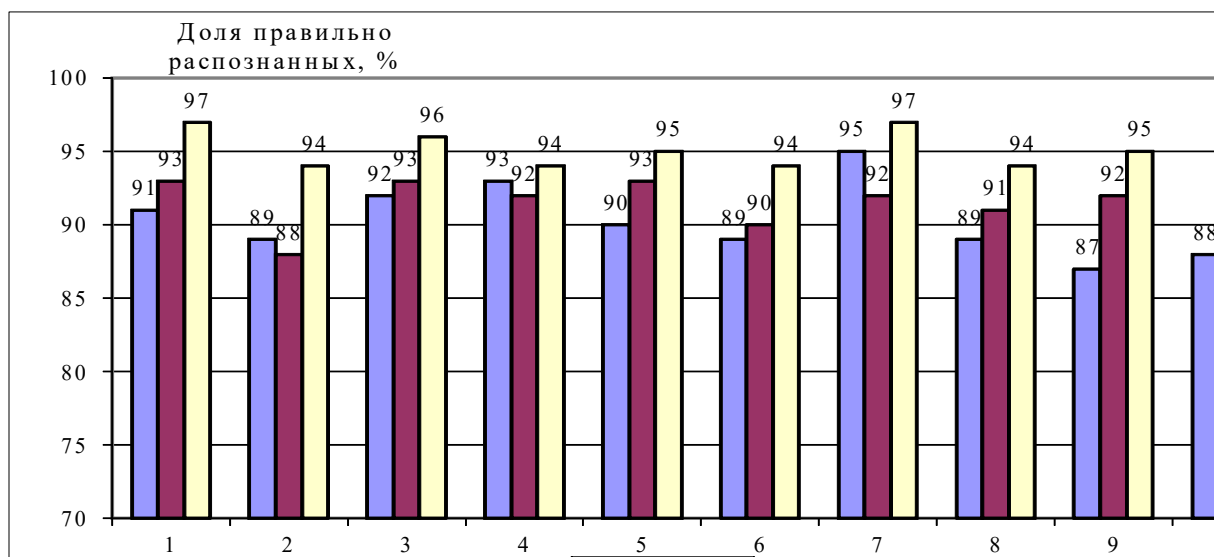


Рисунок 5 – Сравнение точности распознавания: А – вычисление условных вероятностей по Байесу; В – обученный персептрон; С – двухуровневая система

Эксперименты показали работоспособность и заметное преимущество предлагаемой системы. Следует отметить, что система распознавания, использующая в качестве признаков отношения правдоподобия, логических прозрачна и может быть ис-

пользована как основа для построения дерева решений.

ВЫВОДЫ. В работе рассмотрены подходы к построению систем распознавания и предлагается метод синтеза многоуровневой архитектуры нейросе-

тевой системы для решения задач диагностики, отличающийся тем, что в его основе лежит иерархическая вложенная структура пространства признаков, содержащая микрокластеры, кластеры и макрокластеры, а также декомпозиция множества признаков на объединяющие и разделяющие подмножества. Объединяющее подмножество признаков служит для выделения кластеров, а разделяющее подмножество – для выделения микрокластеров. Двухуровневая система распознавания, полученная в результате синтеза и обучения на основе разработанного метода, является логически прозрачной и позволяет повысить точность распознавания по сравнению с классическими перцептронами на том же наборе признаков.

В целом разработанные модель и метод синтеза нейросетевой многоуровневой системы являются простым и эффективным средством решения практических задач диагностики и распознавания образов. Предложенный подход может быть рекомендован для использования в разработке автоматизированных систем распознавания.

ЛІТЕРАТУРА

1. Диагностика заболеваний методами теории вероятностей / М.Л. Жмудяк, А.Н. Повалихин, А.В. Стребуков и др. – Барнаул: Изд-во АлтГТУ, 2006. – 168 с.
2. Лорьер Ж.Л. Системы искусственного интеллекта. – М.: Мир, 1991. – 342 с.
3. Pranke J., Mandler E.A. Comparison of Two Approaches for Combining the Votes of Cooperating Classifiers // Proceedings 11-th IAPR International Conference on Pattern Recognition. – 1992. – Vol. 2. – PP. 611–614.
4. Байдык Т.Н. Нейронные сети и задачи искусственного интеллекта. – К.: Наукова думка, 2001. – 263 с.
5. Уоссермен Ф. Нейрокомпьютерная техника:

Теория и практика. – М.: Мир. 1992. – 423 с.

6. Каллан Р. Основные концепции нейронных сетей / Пер. с англ. – М.: Изд. дом «Вильямс», 2001. – 291 с.
7. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение – Deep Learning. – М.: ДМК Пресс, 2017. – 652 с.
8. Zadeh L.A. Computing With Words. Principal Concepts and Ideas. – Berlin: Springer, 2012. – PP. 52–70.
9. Митюшкин Ю.И., Мокин Б.И., Ротштейн А.П. Soft Computing: идентификация закономерностей нечеткими базами знаний: монография. – Винница: УНИВЕРСУМ-Винница, 2002. – 145 с.
10. Тертышный В.А., Тертышный Н.Ю., Шкарупа Н.Л. Логическое представление документов и запросов поискового модуля медицинской ERP-системы // Вісник Кременчуцького національного університету імені Михайла Остроградського. – 2016. – Вип. 5/2016 (100). – С. 33–43.
11. Шевченко И.В., Грицаков С.А., Дымченко Н.Н. Диагностика неблагоприятных ситуаций в процессе выращивания монокристаллов полупроводников // Вісник Кременчуцького національного університету імені Михайла Остроградського. – 2014. – Вип. 1/2014 (84). – С. 34–43.
12. Немирко А.П., Манило Л.А., Калиниченко А.Н. Математические методы анализа медицинских данных. – СПб: Изд-во СПбЭТУ «ЛЭТИ», 2013. – 175 с.
13. Dubrovin V., Subbotin S. Choice of neuron transfer functions and research of their influence for learning quality of neural networks // Proceedings of International Conference on Modern Problem of Telecommunications, Computer Science and Engineers Training TCSET'2000, February 14–19, 2000, Lviv Slavsko, pp. 114–115.

MODEL AND METHOD OF A MULTI-LEVEL RECOGNITION SYSTEM CONSTRUCTION

I. Shevchenko, N. Shkarupa, M. Gonchar

Kremenchuk Mykhailo Ostrohradskyi National University

vul. Pervomayskaya, 20, Kremenchuk, 39600, Ukraine. E-mail: ius.shevchenko@gmail.com

Purpose. To consider existing approaches to the construction of recognition systems. The task of pattern recognition can be solved step by step, if in the analysis of data, hierarchical structures are distinguished in the feature space. A method has been proposed for synthesizing a multilevel architecture of a neural network system for solving diagnostic problems. **Results.** The method is based on the nesting of microclusters, clusters and macroclusters in the feature space. The second condition for splitting the recognition process into stages is the decomposition of characteristics set into unifying and separating subsets. With the help of a unifying subset of features, clusters are formed within the macrocluster at the first stage. In the second stage, using a separating subset of features within clusters, microclusters are distinguished, which are groups of objects with certain properties. **Practical value.** A two-level recognition system, obtained as a result of synthesis based on the developed method, is logically transparent and allows increasing recognition accuracy in comparison with classical perceptrons on the same set of features. The developed model and method of synthesis of a neural network multilevel system is a simple and effective means of solving practical problems of diagnostics and image recognition. References 13, figure 5, table 1.

Key words: recognition problem, system architecture synthesis, model, testing.

REFERENCES

1. Zhmudiyak, M.L., Povalihin, A.N., Strebukov, A.V. etc. (2006), *Diagnostika zabolevaniy metodami teorii veroyatnostey* [Diagnostics of diseases the methods of theory of chances], Izd-vo AltGTU, Barnaul, Russia.
2. Lorer, Zh.L. (1991), *Sistemyi iskusstvennogo intellekta* [Intelligence systems], Mir, Moscow, Russia.
3. Pranke, J., Mandler, E. (1992), "A Comparison of Two Approaches for Combining the Votes of Cooperating Classifiers", *Proceedings 11-th IAPR International Conference on Pattern Recognition*, vol. 2, pp. 611–614.
4. Baydyk, T.N. (2001), *Neyronnyie seti i zadachi iskusstvennogo intellekta* [Neural networks and tasks of artificial intelligence], Naukova dumka, Kiev, Ukraine.
5. Uossermen, F. (1992), *Neyrokompyuternaya tehnika: Teoriya i praktika* [Neurocomputer technique: Theory and practice], Mir, Moscow, Russia.
6. Kallan, R. (2001), *Osnovnyie kontseptsii neyronnyih setey* [Basic conceptions of neural networks], Izd. dom «Vilyams», Moscow, Russia.
7. Gudfellou, Ya., Bendzhio, I., Kurvill, A. (2017), *Glubokoe obuchenie* [Deep Learning], DMK Press, Moscow, Russia.
8. Zadeh, L.A. (2012), [Computing With Words. Principal Concepts and Ideas], Springer, Berlin, Germany.
9. Mityushkin, Yu.I., Mokin, B.I., Rotshteyn, A.P. (2002), *Soft Computing: identifikatsiya zakonomernostey nechetkimi bazami znaniy: monografiya* [Soft Computing: authentication of conformities to law the unclear bases of knowledge: monograph], UNIVERSUM-Vinnitsa, Vinnitsa, Ukraine.
10. Tertyishnyiy, V.A., Tertyishnyiy, N.Yu., Shkarupa, N.L. (2016), "Logical presentation of documents and queries of thesearching module of the medical ERP-system", *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, iss. 5, no. 100, pp. 33–43.
11. Shevchenko, I.V., Gritsakov, S.A., Dymchenko, N.N. (2014), "Diagnostics of unfavorable situations in the process of growing of single-crystals of semiconductors", *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, iss. 1, no. 84, pp. 34–43.
12. Nemirk, A.P., Manilo, L.A., Kalinichenko, A.N. (2013), *Matematicheskie metody analiza meditsinskih dannyih* [Mathematical methods of analysis of medical data], Izd-vo SPGETU «LETI», SP, Russia.
13. Dubrovin, V., Subbotin, S. (2000), "Choice of neuron transfer functions and research of their influence for learning quality of neural networks", *Proceedings of International Conference on Modern Problem of Telecommunications, Computer Science and Engineers Training TCSET'2000, February 14–19*, Lviv Slavsko, pp. 114–115.

Стаття надійшла 13.12.2017.