

УДК 004.93

КОНЦЕПТУАЛЬНЫЙ ПОДХОД К РАСПОЗНАВАНИЮ ОБРАЗОВ И АНАЛИЗУ СИТУАЦИЙ НА ОСНОВЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Филин Н.Н.*ФГБОУ ВО «Ростовский государственный экономический университет», Ростов-на-Дону,
e-mail: filin_nn@mail.ru*

За последние несколько десятков лет благодаря разработкам и исследованиям в области искусственных нейронных сетей достигнуты значительные успехи в решении задач распознавания и классификации образов. Вместе с тем проблемы, связанные с распознаванием сцен и ситуаций, их анализом и пониманием, остаются актуальными. Рассмотренные в статье особенности восприятия у человека позволили сформулировать ряд необходимых принципов построения нейронных сетей в виде концептуального подхода к распознаванию сложных образов и анализу ситуаций: иерархичность; двунаправленность; принцип последовательно-параллельной обработки. Что позволяет реализовать на основе нейросетей ряд новых функций: осуществлять предобработку входных данных с использованием механизма сканирования; рассматривать пространственные и временные паттерны как частный случай пространственно-временных, применяя для их представления в сети один и тот же механизм обработки и кодирования; реализовывать на нейросети быструю параллельную обработку информации одного типа и более медленную – другого (требующего последовательных методов). Предложена оригинальная архитектура сети с сопряженными афферентными и эфферентными каналами обработки информации, в которой воспроизведение входной информации на выходе на каждом шаге сканирования, может рассматриваться как объяснение принимаемого решения.

Ключевые слова: биологический прототип, нейронная сеть, распознавание образов, интериоризация, последовательно-параллельная обработка информации

CONCEPTUAL APPROACH TO PATTERN RECOGNITION AND SITUATION ANALYSIS BASED ON ARTIFICIAL NEURAL NETWORKS

Filin N.N.*Rostov State economic university, Rostov-on-Don, e-mail: filin_nn@mail.ru*

Over the past few decades, through the development and research in the field of artificial neural networks, significant progress has been made in solving problems of image recognition and classification. However, problems related to the recognition of scenes and situations, their analysis and understanding remain relevant. The features of human perception discussed in the article made it possible to formulate a number of necessary principles of neural network construction in the form of a conceptual approach to recognition of complex images and analysis of situations: hierarchy; bidirectionality; Principle of serial-parallel processing. This allows a number of new functions to be implemented on the basis of neural networks: to pre-process the input data using the scanning mechanism; Consider spatial and temporal patterns as a particular case of spatial-temporal, applying the same processing and coding mechanism to represent them in the network; Implement fast parallel processing of information of one type and slower – of another (requiring sequential methods) on neural networks. The invention proposes an original network architecture with conjugated afferent and efferent information processing channels, in which reproduction of input information at the output at each scanning step can be considered as an explanation of the decision.

Keywords: biological prototype, neural network, image recognition, interiorization, serial parallel information processing

За последние несколько десятков лет благодаря разработкам и исследованиям в области искусственных нейронных сетей достигнуты значительные успехи в решении задач распознавания и классификации изображений, звуковых сигналов разной природы, распознавании речи, прогнозировании временных рядов и др. Существует точка зрения, что нейронные сети уже превосходят возможности человека в ряде случаев (например, скорости распознавания лиц). Наверное, это так. Но является ли это показателем интеллектуальных «достижений» нейросетей? В работе [1] отмечается, что «...вряд ли задача понимания изображения под силу какому-либо классификатору. Любые из них обладают несколькими суще-

ственными ограничениями, которые не позволяют индивидуально применить их для решения задач разбора и понимания изображений». Это утверждение имеет силу применительно практически по всем направлениям использования нейросетевых подходов, идет ли речь о сетях сверточного типа, рекуррентных или же их модификациях.

Как известно, в основе нейросетевой парадигмы лежит биологический прототип. И такие характеристики, как многослойность (перцептроноподобные сети), ассоциативная память (сеть Хопфилда), обучение и правила обучения (например, правило Хэбба), однообразие связей и элементов сети (упрощенная модель реального нейрона) и др., есть тому подтверждение.

Однако, как известно, системы восприятия человека и животных (имеющих существенное сходство), хотя и демонстрируют упомянутые характеристики, но качественно иначе: комплексно, объединяя процессы передачи информации от слоя к слою по прямым и обратным путям (вертикально, двунаправленно), а также внутри и между слоями (горизонтально).

Исследования в направлении комплексного подхода при построении нейронных были и во времена второй волны повышенного интереса к НС (если считать, что первая завершилась проблемой «сложение по модулю 2» в конце 1960-х гг.). Одним из таких результатов было построение многослойной искусственной нейронной сети, частично интегрирующий свойства систем восприятия (с акцентом прежде всего на зрительную систему) стала нейронная сеть сверточного типа неокогнитрон (в развитие когнитрона) К. Фукушимы (1980 г.).

Сегодня ведутся исследования и разработки в направлении интеграции классических методов распознавания и методов концептуального анализа, построения гибридных моделей, комбинации уже упомянутых сверточных и рекуррентных сетей, в которые явно или не явно включаются компоненты структур сетей и связей с новыми характеристиками, способствующими повышению качества распознавания, классификации или прогноза.

Цель исследования: рассмотрение уже известных принципов функционирования систем восприятия биологического прототипа с позиций их комплексного использования при создании искусственных нейронных сетей. Определение необходимого минимума требований к архитектуре и принципам функционирования нейронных сетей, способных если не работать со смыслом, то по крайней мере способных к более глубокому анализу и пониманию сложных объектов, сцен и ситуаций. Построение концептуального подхода распознавания образов и анализа ситуаций на основе искусственных нейронных сетей.

Материалы и методы исследования

В качестве информационной базы исследования были использованы публикации, касающиеся специфики функционирования систем восприятия биологического прототипа. Методологической основой исследования заявленного подхода к распознаванию и анализу ситуаций послужили концепции и парадигмы искусственных нейронных сетей.

Результаты исследования и их обсуждение

Изучение функционирования системы восприятия в живых системах позволило

сформулировать ряд особенностей и принципов: многоуровневость организации и функционирования сенсорных и двигательной систем; сопряженность в работе систем обработки афферентной и эфферентной информации; сложное сканирование воспринимаемого образа.

Многоуровневость организации и функционирования сенсорных и двигательной систем

Известно, что специфическая часть сенсорной системы состоит из нескольких уровней обработки входной информации: рецепторной системы, центральных нервных путей (стволовые ядра, таламические сенсорные ядра) и корковой сенсорной проекции [2].

Структурным уровням сенсорных систем ставят в соответствие этапы обработки входной информации, которая заключается в выделении все более сложных компонент внешних воздействий по мере продвижения вверх по афферентным каналам к проекционным зонам коры и далее в I, II, III ассоциативные области.

Сходная схема организации характерна и для двигательной системы.

В [3] отмечается, что соединения между эфферентными органами и нервными волокнами, которые их иннервируют, в принципе аналогичны синапсам между нейронами, а также между рецепторами и нейронами.

Если анализаторные системы в качестве обязательного компонента имеют свою собственную подчиненную двигательную подсистему сканирования или реализуют эту функцию, используя любые другие двигательные органы, то двигательная система включает анализаторные системы или свои специфические сенсорные входы для реализации нужной двигательной реакции.

Сопряженность в работе систем обработки афферентной и эфферентной информации

Согласно теории моторного изоморфизма при восприятии образа какой-либо модальности происходит постепенный переход к его моторному образу. Ряд работ [3] посвящен развитию этих идей. В процессе восприятия речи осуществляется переход от слухового образа речевого сигнала к его моторному образу – набору управляющих команд, требуемых для произнесения сигнала, подобного услышанному.

Таким образом, многоуровневость строения и функционирования систем обработки афферентной и эфферентной информации и их взаимодополняемость является принципиальным свойством организации

анализаторных и двигательных систем. Между сенсорной и моторной системами существуют как внутренние, так и внешние взаимосвязи. Каждой структуре возбуждения высших центров, ответственной за инициацию движения, соответствует определенная структура возбуждения сенсорных центров и каждое осуществленное движение вызывает изменение структуры сенсорного входа.

*Сложное сканирование
воспринимаемого образа*

Изучение процессов опознавания зрительных объектов взрослыми здоровыми людьми показывает, что в реальных условиях человек пользуется наряду с локальными (первичными) признаками и набором других признаков.

С помощью мыслительных операций категоризации [3, 4], свойственных структурам памяти, первичный код объекта преобразуется этими операциями и из него выделяется группа вторичных признаков. Вторичные признаки характеризуют свойства контура в целом.

Формирование в памяти вторичных признаков приводит к изменению характера обследования объектов при последующем их восприятии, т.е. к изменению стратегии сканирования.

Выделение в памяти вторичных признаков видоизменяет восприятие таким образом, что обследование объекта происходит по-прежнему одноканально, но не последовательно по линии контура, а скачкообразно, по характерным точкам объекта. Траектория сканирования определяется на этом этапе восприятия не только контуром самого объекта, но и хранящейся в памяти обобщенной информацией о классе, к которому он принадлежит.

Таким образом, происходит переход от пассивного осмотра всех точек контура к активной проверке гипотез, связанной с хранящимися в памяти вторичными признаками. Первичный признак характеризует чувственно воспринимаемые свойства объекта, вторичный – отношения между свойствами, а третичный – отношение между этими отношениями. Третичные признаки, так же как вторичные, характеризуют объект в целом, но они еще слабее, чем вторичные, связаны с деталями конфигурации конкретной фигуры. Как видно из изложенного, оперативное поле для выявления первичных признаков определяется размером двух смежных частей фигуры, для вторичных – размером группы частей и становится соизмеримым с «размером» фигуры, для третичных – это поле увеличивается еще

больше: оно включает не только весь объект, но и часть внешнего поля.

В процессе формирования признаков любого уровня важнейшее место занимает природа операций по обработке воспринятых свойств объекта [3, 4]. Число и состав этих операций и определяют в основном каждый уровень иерархии и его отличие от предыдущего и последующего. Чем выше уровень, тем большая часть операций (в том числе и сканирование) производится мысленно, в памяти, с отрывом во времени от восприятия объекта, предъявленного для узнавания. Движение от низшего уровня к все более высоким соответствует постепенному переходу при узнавании от использования информации, полученной только при взаимодействии с объектом, через включение в процесс, кроме того, мыслительных операций, (вычленивающих различную информацию о взаимодействии элементов объекта) к полному исчезновению внешних действий на мыслительные операции обобщения и абстрагирования. Поскольку наиболее медленные операции – внешние сканирующие действия – занимают все меньшее место при формировании обобщенных признаков, то по мере овладения человеком признаками более высоких уровней повышается скорость узнавания и возрастает скорость реакции на изменение внешних событий.

Рассмотренные нейрофизиологические аспекты восприятия у человека и животных позволяют сформулировать ряд необходимых принципов построения искусственных нейронных сетей на основе биологического прототипа.

Три принципа или механизма распознавания в живых системах:

1. Иерархичность. Существует, по меньшей мере, три уровня системы и процесса распознавания: на первом уровне определяются элементарные признаки среды, свойства, атрибуты; на втором уровне – объекты, как композиции или конструкции из элементарных признаков; на третьем уровне – сцены, объекты + контекст.

2. Двуправленность. В процессах восприятия, распознавания объектов и понимания ситуаций окружающей среды учувствуют афферентные и эфферентные каналы в равной степени. Проходя по афферентным каналам от уровня к уровню, информация об окружении определяется и записывается на языках внутреннего представления. По эфферентным каналам каждого уровня проверяется на соответствие внутреннего представления об объектах и процессах среды.

3. Принцип последовательно-параллельной обработки. При распознавании задействованы два механизма снятия, обработки и воспроизведения информации: параллельный анализ и синтез элементарных стимулов и простейших конструкций из них (т.е. их восприятие и воспроизведение), на нижних уровнях; последовательный перебор и сборка сложных объектов и сцен (восприятие, анализ, синтез, восприятие), на верхних уровнях.

На рис. 1 схематично отображены уровни и каналы потока афферентной и эфферентной информации, соответствующие нейрофизиологическим особенностям строения и функционирования сенсорных систем человека. Схема включает упомянутые выше особенности биологического прототипа: многоуровневость организации и функционирования сенсорных и двигательной систем; сопряженность в работе систем обработки афферентной и эфферентной информации.

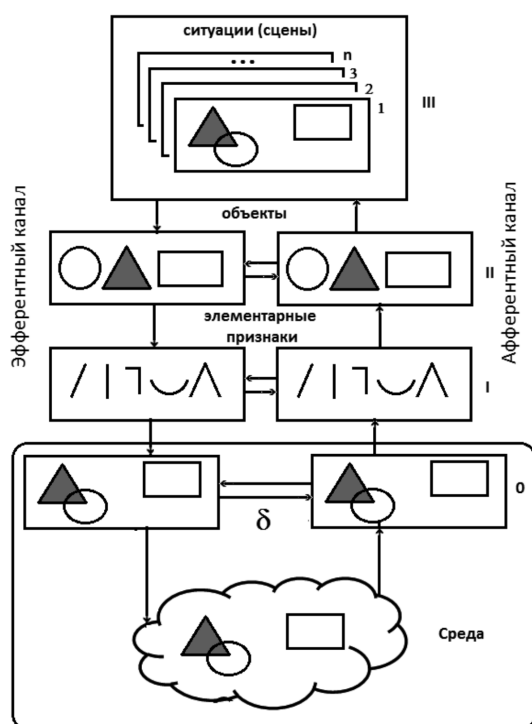


Рис. 1. Концептуальная схема распознавания образов и анализа ситуаций. Где δ – степень расхождения результатов распознавания и воспроизведения

На первом уровне выполняется детекция по афферентному и отображение по эфферентному каналу элементарных признаков (предварительно обученной сети). Замыкание осуществляется по контуру:

среда – афферентный блок 1-го уровня – эфферентный блок 1-го уровня – среда. Параметр δ указывает на степень расхождения результатов распознавания и воспроизведения и может быть использован как критерий обученности.

На втором уровне выполняется распознавание и воспроизведение пространственно-временных объектов (состоящих из элементарных признаков). В результате интериоризации замыкание осуществляется по контуру: афферентный блок 1-го уровня – афферентный блок 2-го уровня – эфферентный блок 2-го уровня – эфферентный блок 1-го уровня.

На третьем уровне формируются сцены (ситуации) в виде семантической сети или онтологии предметной области. Описание отношений между объектами или процессами осуществляется на базе пространственно-временной логики (включающей отношения расстояния и взаимного положения и причинно-следственные связи в условиях времени). В результате интериоризации замыкание осуществляется по контуру: афферентный блок 2-го уровня – афферентный блок 3-го уровня – эфферентный блок 3-го уровня – эфферентный блок 2-го уровня.

Таким образом, анализ нейрофизиологических оснований восприятия и распознавания образов позволяет сформулировать не бесспорное, но при решении сформулированных задач в исследовании достаточно обоснованное утверждение: с нашей точки зрения, распознать – значит суметь воспроизвести.

Система распознавания и анализа ситуаций, построенная на основе сформулированных принципов, включает четыре уровня: РП-среду и три нейросетевых модуля. Каждый нейросетевой модуль соответствует определенному уровню обработки образов и устроен однотипно: состоит из двух подсетей афферентной (классифицирующей, работающей со входной информацией) и эфферентной (воспроизводящей, работающей с выходной информацией).

Рассмотрим на конкретном примере функционирование нейронной сети, с реализацией некоторых принципов распознавания и воспроизведения биологического прототипа.

Как уже отмечалось, предлагаемая архитектура сети имеет два канала афферентный и эфферентный. Эфферентный канал для задач распознавания служит как вспомогательный или подчиненный. Для решения задач моделирования двигательной активности каналы меняются ролями. В данном исследовании мы рассматриваем только проблемы касающиеся распознава-

ния образов и их комбинаций, находящихся в среде внешней по отношению к системе распознавания.

На вход первого афферентного уровня поступает информация непосредственно из среды анализа как результат сканирования; для второго – входом являются выходы элементов первого уровня; для третьего – входом являются выходы элементов второго уровня, активность которых соответствует распознаванию объектов (структур связанных элементарных признаков).

Таким образом, уровни системы распознавания отличаются по объему и качеству обрабатываемой информации: элементарные признаки, объекты и сцены. Из этого следует, что скорость обработки, поступающей на вход разных уровней информации, имеет и разную длительность, зависящую от количества шагов сканирования. Этот факт отражается в постоянной времени τ уравнений динамики состояний элементов. Так для первого уровня постоянная $\tau = \tau_1$ близка к нулю, поскольку детекция элементарных признаков (по крайней мере в данной модели) происходит практически за один шаг. Для обнаружения объекта (структуры связанных признаков) и определения сцены (композиции объектов) необходимо больше шагов сканирования, что определяет соотношение $\tau_1 \ll \tau_2 \ll \tau_3$, где τ_2 и τ_3 – постоянные времени для элементов сети второго и третьего уровня соответственно.

Пусть среда (в дальнейшем РП-среда – рецептивная поверхность – среда) на рисунке 2 – это область, подлежащая анализу на предмет наличия объектов (как структур связанных признаков) и/или сцен (т.е. объектов, находящихся в определённых отношениях).

Выделим три области анализа среды.



Рис. 2. Рецептивная поверхность – среда

Рецептивное поле, «рп» – локальный участок, обрабатываемый параллельными процедурами. Область, состоящая из участков связанных с «рп» («рп + ближайшее окружение») обрабатывается с использовани-

ем локальной процедуры сканирования (стохастической или детерминированной). По этой области производится последовательный скользящий перебор параллельно обрабатываемых локальных областей.

«РП-среда» – обрабатывается посредством сложного сканирования (зависящего от результатов распознавания) областями «рп+ближайшее окружение».

Пусть элементарные признаки, которые детектируются на первом уровне, представляют собой линейно разделимое множество (рис. 3): 16 элементарных признаков – полный набор возможных состояний активности «рп» размером 2×2 . Соответственно первый слой состоит из 16 элементов настроенных (или обученных) на детекцию элементарных признаков.

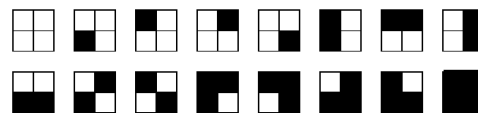


Рис. 3. Элементарные признаки

Сканирование «РП-среды» осуществляется областями размером 5×5 («рп + ближайшее окружение»), сканирование «рп + ближайшее окружение» осуществляется областями 2×2 («рп») одним из возможных способов или их комбинаций в зависимости от результатов распознавания или воспроизведения (например, линейным, стохастическим, по развертке Гильберта). В демонстрационной модели ограничимся линейным скользящим сканированием. На сканирование среды анализа демонстрационной модели потребуется не менее 250 шагов $((4 \times 4) \times 16)$.

Известно, что использование сканирования со сложной разверткой в системах распознавания в качестве процедуры предобработки может обеспечить их информацией инвариантной к группе движения твердого тела [5].

Рассмотрим процессы, происходящие вдоль афферентного и эфферентного каналов второго уровня. Каждый уровень состоит из линейки взаимосвязанных тормозными связями нейроподобных элементов (рис. 4).

Обозначения афферентного канала второго уровня: Y_{1ai} – входы к афферентной сети (линейки связанных элементов), Y_{2ai} – выходы афферентной сети, C_{ij} – веса связей (тормозных) линейки элементов, W_{1aij} – веса связей входов к афферентной сети второго уровня от первого уровня, A_{2aji} – коэффициент связи афферентных с соответствующими эфферентными элементами.

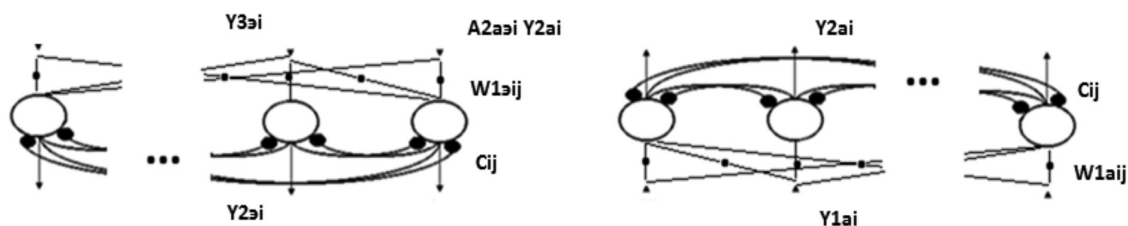


Рис. 4. Архитектура второго уровня демонстрационной модели

Обозначения эфферентного канала второго уровня: Y_{3ai} – входы к афферентной сети (линейки связанных элементов), Y_{2ai} – выходы эфферентной сети, C_{ij} – веса связей (тормозных) линейки элементов, W_{1aij} – веса связей выходов эфферентной сети второго уровня к первому уровню.

Динамика состояний элементов второго уровня описывается следующим уравнением:

$$\tau_2 \frac{dy_{2ai}}{dt} = -y_{2ai} + \sum_{j \neq i} C_{ij} Y_{2aj} + \sum_{j=1} W_{1aij} Y_{1aj}, \quad (1)$$

где y_{ai} – уровень состояния элемента детектора i , C_{ij} – матрица тормозных связей внутри слоя, Y_{aj} – выход элемента конкурента j , такой что

$$Y_{2aj} = \begin{cases} 0, & \text{если } y_j \leq h \\ y_{2aj}, & \text{если } h < y_{2aj} \leq h+1, \\ 1, & \text{если } y_{2aj} > h+1 \end{cases} \quad (2)$$

W_{1aij} – веса связей между выходами Y_{1aj} элементов предшествующего и элементами текущего уровней, которые формируются в процессе обучения по правилу подобному правилу Хэбба, с той разницей, что связь рассматривается как односторонняя (от элементов предшествующего к элементам текущего уровня):

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W, \quad (3)$$

где ΔW – приращение, величина и знак которого зависит от активности элементов, между которыми устанавливается связь.

Теперь рассмотрим строение и назначение сетей эфферентного канала второго уровня.

Каждая сеть эфферентного канала повторяет архитектуру афферентного канала соответствующего уровня. Отличия в направлении связей и уравнениях описывающих состояние элементов. То есть связи в эфферентных каналах направлены от верхних уровней к нижним, при этом значения весов сохраняются. Кроме того,

каждый элемент афферентной сети имеет одностороннюю положительную связь со своим двойником в эфферентной сети.

Динамика состояний элементов описывается следующим уравнением:

$$\tau \frac{dy_{2ai}}{dt} = -y_{2ai} + \sum_{j \neq i} C_{ij} Y_{2aj} + A_{2ai} Y_{2ai}, \quad (4)$$

где y_{2ai} – уровень состояния элемента сети воспроизведения i , Y_{1aj} – выход элемента сети распознавания, A_{2ai} – технический коэффициент, предназначенный для выбора алгоритма функционирования сети, C_{ij} – матрица тормозных связей внутри слоя, Y_{2aj} – выход элемента конкурента j , такой, что

$$Y_{2aj} = \begin{cases} 0, & \text{если } y_j \leq h \\ y_{2aj}, & \text{если } h < y_{2aj} \leq h+1. \\ 1, & \text{если } y_{2aj} > h+1 \end{cases} \quad (5)$$

На рис. 5 графически представлена активность элементов афферентного и эфферентного каналов второго уровня при тестировании после обучения сети распознаванию объектов «С», «Т» и «Х» (моделирование демонстрационных процедур проведено в VBA MS Excel). Стрелками на графиках отмечены моменты распознавания и воспроизведения объектов.

На рис. 6 представлена анализируемая среда на входе и результат воспроизведения на выходе. Цифрами 1, 4 и 16 обозначены шаги сканирования и соответствующие им результаты распознавания и воспроизведения объектов анализируемой среды.

Реальные объекты и сцены значительно сложнее, поэтому при сохранении общей логики работы сети необходимо усложнение сетевых структур уровней системы. Например, в качестве сетей распознавания линейно неразделимых объектов могут потребоваться многослойные перцептроноподобные структуры или ассоциативные нейронные сети, восстанавливающие по искаженному и/или зашумленному образу ближайший к нему эталонный.

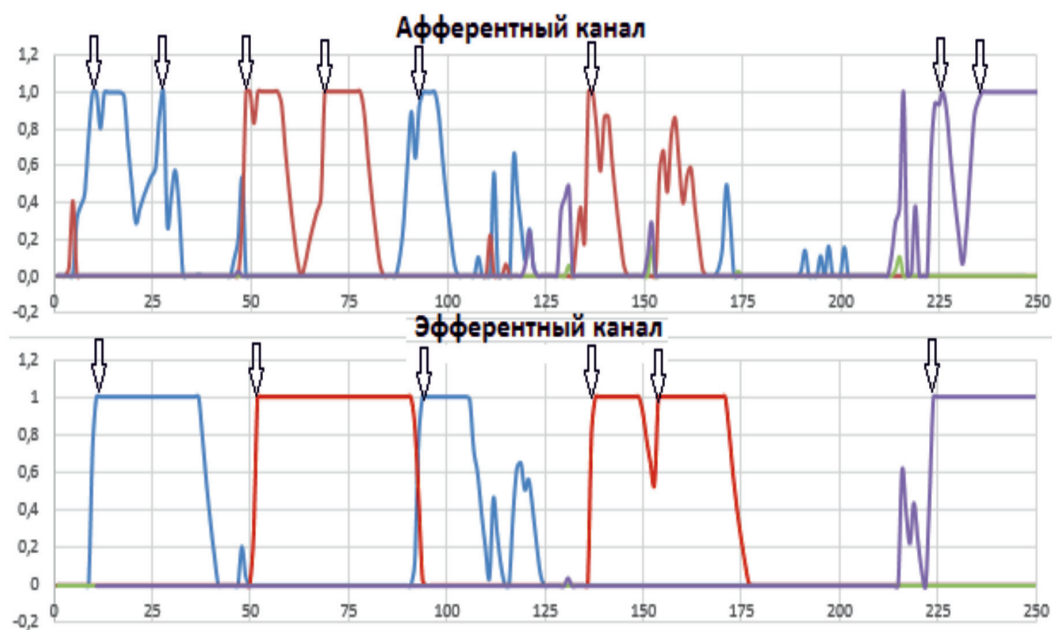


Рис. 5. Активность элементов афферентного и эфферентного каналов второго уровня

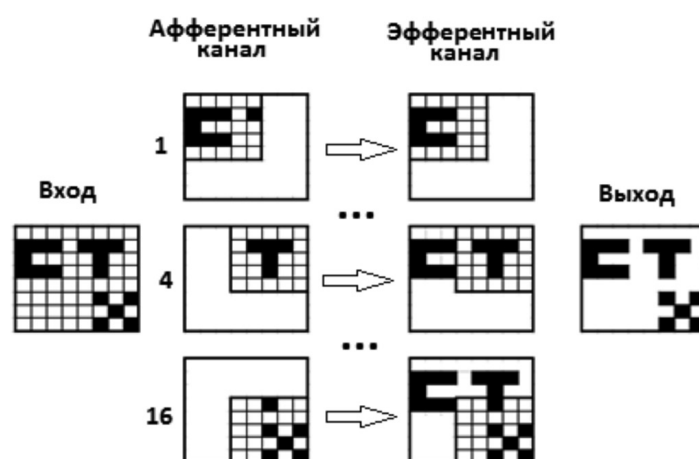


Рис. 6. Шаги сканирования 1, 4 и 16 входной информации и соответствующие им результаты распознавания и воспроизведения объектов анализируемой среды

Выводы

В рассмотренном подходе построения нейросетей распознавания образов и анализа ситуаций предлагается использовать взаимодополняющие принципы присущие системам восприятия биологического прототипа: иерархичность; двунаправленность; механизм последовательно-параллельной обработки.

Эти механизмы позволяют реализовать на нейросети ряд комплексных функций:

- осуществлять предобработку входных данных с использованием механизма сканирования;

- рассматривать пространственные и временные паттерны как частный случай пространственно-временных, применяя для их представления в сети один и тот же механизм обработки и кодирования;

- реализовывать на нейросети быструю параллельную обработку информации одного типа и более медленную – другого (требующего последовательных методов).

Предложена оригинальная архитектура сети с сопряженными афферентными и эфферентными каналами обработки информации, в которой воспроизведение входной информации на выходе на каждом шаге сканирования может рассматриваться как объяснение принимаемого решения.

Список литературы

1. Хахалин Г.К., Лобзин А.П. Комплекс для гетерархической системы разбора и понимания изображений // Пятнадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2016 (г. Смоленск, 3–7 октября 2016 г.). Труды конференции. В 3-х т. Т. 1. Смоленск: Универсум, 2016. С. 121–131.
2. Ковалева А.В. Нейрофизиология, физиология высшей нервной деятельности и сенсорных систем: учебник для академического бакалавриата. М.: Издательство Юрайт, 2018. 365 с.
3. Моисеева Л.А. Основы нейрофизиологии: учебное пособие для студентов психологических и медицинских учебных заведений. М.: Московский институт психоанализа, 2014. 200 с.
4. Шевелев ИЛ. Нейроны-детекторы зрительной коры: ревизия свойств и механизмов. М.: Наука, 2010. 182 с.
5. Федотов Н.Г. Методы стохастической геометрии в распознавании образов. М.: Радио и связь, 1990. 144 с.