

История разработки и перспективы системной инженерии нейросетевых архитектур

С.А.К. Диане ^{1,2}, А.Д. Воронков ², Д.А. Акимов ², Е.О. Гурьянова ²

¹Институт проблем управления им. В. А. Трапезникова РАН, Москва, Россия

²МИРЭА – Российский технологический университет, Москва, Россия

Аннотация – В статье дана классификация типовых архитектур искусственных нейронных сетей, а также рассмотрены ключевые исторические этапы их развития. Выявлены базовые функциональные свойства и представлен метод многокритериальной оценки потенциальной эффективности нейросетевых архитектур. Предлагаются несколько подходов к системной инженерии нейронных сетей. Показано, что наряду с хорошо зарекомендовавшими себя принципами многослойности, иерархичности, рекуррентности на первый план выходят вопросы обеспечения модульности, структурной адаптивности, а также вычислительной эффективности нейросетевых моделей.

Ключевые слова – нейросетевые архитектуры, история нейронных сетей, системная инженерия, машинное обучение, многокритериальная оценка.

I. ВВЕДЕНИЕ

Многолетние исследования в области разработки и применения искусственных нейронных сетей (ИНС) показали, что вычислительные структуры подобного типа способны эффективно решать задачи обработки информации и управления в широком спектре прикладных задач [1-50].

Общая методология применения нейросетевых технологий в составе технических систем того или иного типа включает следующие этапы: (1) выбор или синтез архитектуры ИНС под решаемую прикладную задачу; (2) разведочный анализ данных и подготовка обучающей выборки для ИНС; (3) обучение, валидация и оптимизация ИНС по целевым критериям качества; (4) реализация ИНС с применением вычислительно эффективных аппаратно-программных средств.

Всю совокупность вариантов использования ИНС в технических системах можно независимым образом классифицировать (1) по области применения (автономные роботы, киберфизические системы, программные комплексы); (2) по способу применения (управление, обработка информации, генерация образов); (3) по формату обрабатываемых данных (сигнальные, векторные, символьные, растровые, разреженно точечные, графовые).

Существующие архитектуры нейронных сетей также сильно варьируются и могут быть разделены, по меньшей мере, на 15 отдельных классов (Рис. 1).



Рис. 1. Классификация архитектур нейронных сетей

Разнообразие сфер применения и архитектур ИНС обуславливает необходимость в системной инженерии нейросетевых моделей на основе тщательного анализа потребностей отдельно взятых прикладных областей и обоснованного выбора функциональных возможностей, предоставляемых современным уровнем техники.

II. ИСТОРИЯ РАЗРАБОТКИ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР

Анализируя историческую последовательность появления нейронных сетей различного типа, можно прийти к выводу о том, что данная область науки планомерно развивается, охватывая все новые аспекты искусственного интеллекта и расширяя границы применимости ИНС, как с алгоритмической, так и с аппаратной точки зрения.

Исторические вехи разработки базовых архитектур и алгоритмов ИНС отражены в Табл. I. В привязке к событиям указаны оригинальные теоретические публикации, а также более поздние публикации с примерами практического применения ИНС.

ТАБЛИЦА I
ПЕРЕЧЕНЬ ЗНАЧИМЫХ СОБЫТИЙ В РАЗРАБОТКЕ
НЕЙРОННЫХ СЕТЕЙ

№	Событие в области разработки ИНС				Литература	
	Год	Авторы	Результат	Страна	Теор.	Прак.
1	1943	Уоррен Мак-Каллок, Уолтер Питтс	Модель искусственного нейрона	США	[1]	[2]
2	1952	Алан Ходжкин, Эндрю Хаксли	Модель нервного импульса	США	[3]	[4]
3	1958	Фрэнк Розенблатт	Многослойный перцептрон	США	[5]	[6]
4	1970	Сеппо Линнайнена	Алгоритм обратного распространения ошибки	Финляндия	[7]	[8]
5	1970	Виктор Гладун	Растущие пирамидальные сети	СССР	[9]	[10]
6	1980	Кунихико Фукусима	Сверточные нейронные сети	Япония	[11]	[12]
7	1982	Джон Хопфилд	Рекуррентные нейронные сети	США	[13]	[14]
8	1983	Скотт Фалман и др.	Машины Больцмана	США	[15]	[16]
9	1988	Леон Онг Чуа	Ячеистые нейронные сети	США	[17]	[18]
10	1989	Джордж Цыбенко	Универсальная теорема аппроксимации	США	[19]	[20]
11	1990	Ларс Кай Хансен, Петер Саламон	Методология нейросетевых ансамблей	Дания	[21]	[22]
12	1991	Томас Мартинац, Клаус Шультен	Алгоритм растущего нейронного газа	Германия	[23]	[24]
13	1997	Зепп Хохрайтер, Юрген Шмидхубер	Долгосрочно-краткосрочные сети	Германия	[25]	[26]
14	1997	Вольфганг Маасс	Спайковые нейронные сети	Германия	[27]	[28]
15	2005	Марко Гори и др.	Графовые нейронные сети	Италия	[29]	[30]
16	2013	Дидерик Кингма, Макс Веллинг	Вариационные автоэнкодеры	Дания	[31]	[32]
17	2015	Яша Золь-Дикштейн и др.	Диффузионные модели	США	[33]	[34]
18	2016	Чарльз Руйчжунтай Ци и др.	Нейронные сети для облаков точек	США	[35]	[36]
19	2017	Ашиш Васвани и др.	Нейросетевые трансформеры	США	[37]	[38]
20	2017	Сара Сабур и др.	Капсульные нейронные сети	США	[39]	[40]
21	2017	Мазнар Ранси и др.	Физико-информированные сети	США	[41]	[42]
22	2020	Рамин Хасани и др.	Жидкие нейронные сети	США	[43]	[44]

А. Многослойные перцептроны

За точку отсчета в теории искусственных нейронных сетей традиционно берут разработку модели искусственного нейрона У. Мак-Каллоком и У. Питтсом [1] в 1943 г. Дальнейший импульс развитие нейросетевых технологий получило в 1958 г. после выхода публикации Ф. Розенблатта, посвященной многослойному перцептрону (МП) [5] – простейшей форме ИНС, состоящей из одного или нескольких слоев нейронов, способных принимать входные значения, вычислять их суммарное воздействие и активировать выходной сигнал при превышении порога [2, 6]. Крайне важным при настройке МП является определение совокупности весовых коэффициентов, которая позволяет с требуемой точностью аппроксимировать многомерное отображение, задаваемое набором обучающих примеров. Поэтому последующее развитие данной технологии было связано с разработкой алгоритма обратного распространения ошибки для автоматической настройки весов, основанного на методе стохастического градиентного спуска и подробно описанного в работе С. Линнайнена [7].

В. Растущие пирамидальные сети

На стыке нейросетевого и логического подходов к построению интеллектуальных систем возникли растущие пирамидальные сети (РПС), применяемые для классификации объектов, с несколькими независимыми признаками [9] (В. Гладун, 1970), [10]. Вычислительная модель такого типа представляет собой иерархическую структуру нейронов, растущую в зависимости от получаемых данных. РПС обычно состоит из нескольких уровней, каждый из которых представляет собой набор нейронов, отвечающих за различные пространственные, временные или логические масштабы. На основе РПС могут моделироваться процессы ассоциативного мышления и обобщения семиотической информации – за счет соответствующих правил выбора активируемых или соединяемых контрольных элементов (нейронов) в составе сети. Создание новых нейронов также происходит в соответствии с правилами логик первого и второго порядка и нацелено на различение всевозможных комбинаций входных признаков.

С. Сверточные нейронные сети

Важной вехой в развитии нейросетевых технологий стало появление сверточных нейронных сетей (СНС), эффективно распознающих многомерные растровые образы. СНС являются развитием архитектуры МП. Сети такого типа обладают меньшим числом соединений, благодаря учету пространственной локальности, характерной для изменений яркостей точек изображения. СНС используют слои свертки для извлечения значимой информации из входных данных

и слои подвыборки для уменьшения размерности формируемых карт признаков. К ранним архитектурам СНС можно отнести *Neocognitron* К. Фукусимы [11] 1980 г. и *LeNet-1* Я. Лекуна [8] 1989 г. Среди современных архитектур СНС [45, 46, 47] выделяются архитектуры базовых классификаторов *VGG*, *ResNet*, *GoogLeNet*, детекторов *Faster R-CNN*, *YOLO*, сегментирующих сетей *FastFCN*, *U-Net* и др.

D. Рекуррентные нейронные сети

Отдельную нишу в области нейросетевых технологий занимают рекуррентные нейронные сети (РНС), наиболее ранней разновидностью которых принято считать полносвязную рекуррентную сеть Хопфилда [13] (Дж. Хопфилд, 1982 г.), [14]. За счет наличия обратных связей рекуррентные нейронные сети способны сохранять информацию о предыдущих своих состояниях для более эффективной обработки последующих входных значений. РНС позволяют моделировать последовательности данных, такие как временные ряды или текст, учитывая контекст и порядок следования отдельных элементов. Улучшенная версия РНС – сети долгосрочно-краткосрочной памяти позволяют выявлять закономерности в анализируемых данных на более протяженных временных отрезках за счет внедрения в структуру сети вентилях (однослойных перцептронов) для записи, считывания и стирания значимой информации.

E. Стохастические нейронные сети

Стохастические нейронные сети (СтНС) представляют собой достаточно широкий класс нейронных сетей, в которых некоторые элементы модели (связи, функции активации, функции потерь) имеют вероятностный характер, что позволяет им более точно обрабатывать и моделировать неопределенности в данных, генерировать статистически правдоподобные векторы значений и временные последовательности. Одним из ранних примеров стохастических нейронных сетей с рекуррентными связями являются машины Больцмана [15] (С. Фалман и др., 1983 г.), [16]. ИНС данного типа функционально близки к вероятностным конечным автоматам, так как способны динамически менять состояния и выходы в соответствии с решаемой прикладной задачей. Более поздние архитектуры стохастических ИНС включают вариационные автоэнкодеры [31] и диффузионные модели [33] для генерации изображений. Функционирование данных архитектур основано на теории гармонии для динамических когнитивных систем [48].

F. Ячеистые нейронные сети

Одна из сложностей разработки и реализации нейронных сетей заключается в большом количестве настраиваемых параметров. В этой связи уже в конце

прошлого века рассматривались варианты высокоскоростной аппаратной реализации ИНС. В качестве одного из решений данной задачи были предложены ячеистые нейронные сети (ЯНС) [17] (Л. О. Чуа, 1988 г.), [18]. ЯНС состоят из ячеек, каждая из которых содержит один или несколько нейронов и способна обрабатывать информацию в своем окружении. Эти сети, удачно вписывающиеся в парадигму параллельных вычислений, используются для обработки изображений, прогнозирования, управления, моделирования и анализа различных систем. Интересным свойством таких сетей является способность формировать устойчивые паттерны активации в ответ на входные стимулы, изменяющиеся во времени. По этой причине такие сети хорошо подходят для обработки видео данных.

G. Нейросетевые ансамбли

Вариативность данных, обрабатываемых ИНС, и структурная сложность нейронных сетей неизбежно приводят к неточностям их функционирования. В качестве одного из способов минимизации таких неточностей, а также в целях обеспечения лучшей их диагностируемости применяется ансамблирование нейронных сетей [21] (Л. К. Хансен, П. Саламон, 1990 г.), [22]. При этом выделяют три основных стратегии их объединения в ансамбль: пакетирование (модели малой вычислительной сложности параллельно изучают независимые друг от друга данные, после чего на этапе непосредственного применения ансамбля выходы сетей агрегируются); усиление (нейросетевые модели малой вычислительной сложности поочередно изучают входные данные, но с расширением вектора входных данных выходами предыдущей модели); надстройка (адаптивное усреднение параллельных вычислений слабых моделей при помощи обученной на их выходах супервизорной сети). Углубленный подход к построению ансамблевых моделей и мультиагентных нейронных сетей связан с рассмотрением ИНС в составе проектируемой системы в качестве взаимозависимых агентов, принимающих на себя определенные роли в процессе обучения или выполнения прикладных вычислений.

H. Растущий нейронный газ

Многослойные нейронные сети, как правило, строятся с некоторым запасом информационной емкости, что нерационально с вычислительной точки зрения. Поэтому весьма перспективна идея роста структуры ИНС по мере возникновения в этом потребности. Наряду с РПС одной из технологий, алгоритмически реализующей биологические принципы роста и перестроения естественных нейронных сетей, является растущий нейронный газ (РНГ) [23] (Т. Мартинец, К. Шультен, 1991 г.) – алгоритм машинного обучения без учителя, который

применяется для кластеризации данных [24]. В рамках алгоритма растущего нейронного газа создается и итеративно обновляется сеть нейронов с учетом структуры входной информации. Нейроны становятся центрами кластеров, постепенно смещаясь в сторону наблюдаемых многомерных точек данных. В отличие от самоорганизующихся карт Кохонена, имеющих фиксированную размерность [49], число кластеров РНГ подбирается автоматически за счет своевременного создания новых нейронов в местах фиксации наибольшей ошибки рассогласования между положением ближайшего нейрона и регистрируемыми в его окрестности целевыми точками.

И. Спайковые нейронные сети

Углубленные исследования физиологии передачи информации в нервной системе животных и человека показали, что модель искусственного нейрона может быть существенно усложнена. Это породило направление исследований, связанное со спайковыми нейронными сетями (СпНС) [27] (В. Маасс, 1997 г.), [28]. В отличие от традиционных ИНС, где данные обрабатываются непрерывными значениями, спайковые сети используют стробоскопические сигналы, которые передают информацию в виде кратковременных импульсов (спайков). Это позволяет не только более правдоподобно симулировать работу участков мозга, но и в технических задачах достичь высокой энергоэффективности и параллельности вычислений. В то же время учет динамики передачи сигналов в ИНС данного типа открывает возможности для оперативного формирования в составе нейронной сети таких информационных элементов, как линии задержки, осцилляторы, фильтры помех. Упрощается и задача моделирования аттракторов состояний, реализующих механизмы внимания и динамической памяти у биологических прототипов ИНС.

Л. Графовые нейронные сети

Другим активно развивающимся направлением разработки нейросетевых архитектур, являются графовые нейронные сети (ГНС) [29] (М. Гори и др., 2005 г.), [30] – класс ИНС, которые были разработаны для обработки данных, представленных в виде графов, где узлы представляют объекты, а ребра – отношения между ними. По существу, графовые нейронные сети, являющиеся развитием рекурсивных ИНС [50]. В процессе функционирования графовой ИНС нейроны обрабатывают информацию от смежных узлов (векторы признаков узлов и информацию о связях между узлами), в результате чего обновляют собственные векторы признаков. При этом взамен многослойности в графовых сетях применяется итеративный расчет каждого из нейронов, до момента стабилизации выходных значений сети. Графовые нейронные сети могут использоваться для различных

задач, таких как анализ социальных сетей, сегментация облаков точек, принятие групповых управляющих решений в распределенных системах и т.д.

К. Нейронные сети для облаков точек

Нейронные сети для облаков точек (НСОТ), например, *PointNet* [35] (Ч. Р. Ци и др., 2016 г.) и ее усовершенствованная версия *PointNet++* подходят для обработки данных, не имеющих явно определенной матричной или графовой структуры, таких как неупорядоченные облака точек в трехмерном пространстве. Подобно сверточным нейронным сетям НСОТ используют слои абстракции для иерархического извлечения признаков из подмножеств точек и слои распространения признаков с механизмами интерполяции для восстановления размерности обработанных данных. НСОТ способны учитывать глобальный вектор признаков трехмерной сцены для каждой точки данных, что обеспечивает высокую точность и эффективность визуального анализа трехмерных объектов. Еще одна архитектура НСОТ – *FFHNet* [36] – работает на схожих принципах, но более производительна за счет эффективного кодирования входных данных в формате значений удаленности точек облака от опорных вершин.

Л. Нейросетевые трансформеры

В качестве альтернативы рекуррентным и стохастическим сетям могут быть рассмотрены нейросетевые трансформеры (НТ) [37] (А. Васвани и др., 2017 г.), [38]. В общем случае сеть-трансформер состоит из кодирующей и декодирующей частей. Перед подачей на вход кодирующей части, символьные представления (токены) входных данных НТ преобразуются в векторы признаков (эмбединги) непрерывного пространства с фиксированной размерностью, превышающей исходную. При работе с последовательностями к векторам признаков также добавляется позиционная информация о смещении токенов в пределах входной последовательности. Кодирующая часть НТ, выделяющая признаки более высокого порядка, состоит из МП и модулей многоголового внутреннего внимания (МВВ). Декодирующая часть НТ состоит из модулей МВВ, модулей кросс-внимания (КВ) и МП. Модули МВВ нужны для мультиконтекстной оценки значимости элементов входных данных на каждом этапе обработки информации. Модули КВ сопоставляют элементы входного запроса с релевантными справочными данными. По выходам декодера при помощи перцептрона и softmax-слоя формируются оценочные вероятности предсказанных символов. Общая топология соединений НТ позволяет декодирующей части формировать выходную информацию, как с учетом элементов входного вектора, так и с опорой на уже сгенерированную часть выходных значений.

М. Капсульные нейронные сети

Одним из направлений развития сверточных нейронных сетей можно считать капсульные нейронные сети (КНС) [39] (С. Сабур и др., 2017 г.), [40]. Первая ключевая идея капсульных сетей состоит в том, что они используют один или несколько наборов капсул – групп нейронов, работающих совместно для представления геометрических или иных значимых свойств объекта (ориентация, размер, цвет и др.). Причем неявные представления данных свойств в виде подмножеств весовых коэффициентов отдельных капсул формируются автоматически в процессе обучения ИНС градиентным методом. Второе важное свойство капсульных сетей, обеспечивающее специализацию капсул на различных типах входных образов, заключается в применении принципа «победитель забирает все». Этот принцип достигается за счет использования softmax-функций активации для нейронов внутренних слоев ИНС.

Н. Физико-информированные нейронные сети

Физико-информированные нейронные сети (ФНС) [41] (М. Раиси и др., 2017 г.), [42], как правило, основываются на классических архитектурах нейронных сетей, таких как МП, РНС. Основной же особенностью ФНС является учет свойств ассоциированного с ними физического объекта (ФО). В частности, такой учет необходим при обработке измерений, поступающих от объекта или же при формировании управляющих воздействий на него. Интеграция свойств ФО в весовые коэффициенты ФНС достигается в рамках процедуры обучения за счет включения в функцию потерь аналитических выражений или имитационной модели объекта. С одной стороны, подобное разграничение этапов вычисления для нейронной сети и объекта важно с позиций интерпретируемости получаемых результатов. С другой – в повышении точности оценки рассогласования между текущей и желаемой ситуациями, что удобно для коррекции сигналов управления или фильтрации измеряемых значений.

О. Жидкие нейронные сети

Жидкие нейронные сети (ЖНС) представляют собой сравнительно новую разработку в области искусственного интеллекта [43] (Р. Хасани и др., 2020 г.), [44], основанную, тем не менее, на принципах, близких к растущим пирамидальным сетям и растущему нейронному газу. Жидкие ИНС потенциально способны достигнуть высоких результатов в регрессии функций и классификации образов при меньших размерах и потребностях в вычислительных ресурсах. В отличие от основной массы традиционных ИНС, которые требуют огромного объема данных и параметров для достижения работоспособного состояния, жидкие сети

состоят из более компактных групп адаптивных нейронов. Эти нейроны способны динамично изменять свои связи и функции активации в ответ на поступающие входные данные, что позволяет им учиться в режиме реального времени.

III. СИСТЕМНАЯ ИНЖЕНЕРИЯ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР

За основу для анализа, сравнения и системной инженерии различных архитектур ИНС можно взять их 10 ключевых свойств, положительно оцениваемых авторами настоящей статьи: локальность соединений (Л), многослойность сети (М), наличие обратных связей (О), переменность числа нейронов (ПН), переменность числа и направления соединений (ПС), учет временной структуры данных (УВ), учет пространственной структуры данных (УП), импульсность сигналов (И), учет контекста (К), а также диагностируемость (Д), подразумевающая, в том числе лингвистическую интерпретируемость сети.

Для сравнения различных нейронных сетей с возможностью их группировки по выделенным признакам сформируем Табл. II. Значением $q=1$ в ячейках обозначено общепризнанное наличие признака, значением $q=0,5$ – возможность наличия признака в одной из модификаций ИНС для некоторой прикладной задачи. Порядок сортировки ИНС в таблице определяется возрастанием интегральной оценки Q_{Σ} их функциональных возможностей, полученной путем усреднения первичных признаков:

$$Q_{\Sigma} = \frac{1}{N} \sum_{i=1}^N q_i, N=10.$$

Путем визуального поиска подмножеств ненулевых признаков, расположенных на вертикалях (столбцах) Табл. II, можно условно выделить 5 крупных классов нейросетевых моделей: (1) прямopotочные сети №№ 1, 2, 3; (2) пространственно-топологические сети №№ 4, 5, 10; (3) контекстно-зависимые сети №№ 6, 8, 9, 13; (4) структурно-вариативные сети №№ 11, 12, 14; (5) мультиагентные нейронные сети № 7, 15. Отметим, что с увеличением номера класса растет и оценка функциональных возможностей, усредненная по входящим в данный класс типам ИНС.

Системная инженерия новых архитектур ИНС может проводиться (1) с участием эксперта по принципу максимизации степени выраженности некоторых из указанных в Табл. II свойств; (2) с применением нейросетевых генеративных технологий на базе больших языковых моделей; (3) с применением эволюционных методов программирования для автоматического отбора потенциальных решений по априорно заданным критериям. При этом во внимание необходимо удерживать и вопрос обеспечения требуемой вычислительной эффективности ИНС, что может быть достигнуто за счет их инкрементального обучения и дистилляции.

ТАБЛИЦА II
МНОГОПРИЗНАКОВЫЙ АНАЛИЗ НЕЙРОННЫХ СЕТЕЙ

№	Тип ИС	Признак классификации										Q _Σ
		Л	М	О	ПН	ПС	УВ	УП	И	К	Д	
1	МП	-	1	-	-	-	0,5	0,5	-	-	0,5	0,25
2	СНС	1	1	-	-	-	0,5	1	-	-	0,5	0,4
3	КНС	1	1	-	-	-	0,5	1	-	-	0,5	0,4
4	ГНС	1	0,5	-	-	0,5	0,5	1	-	-	0,5	0,4
5	НСОТ	1	0,5	-	-	0,5	0,5	1	-	-	0,5	0,4
6	РНС	0,5	0,5	1	-	-	1	0,5	-	1	0,5	0,5
7	ФНС	0,5	1	0,5	-	-	0,5	0,5	-	1	1	0,5
8	СтНС	0,5	0,5	1	-	-	1	0,5	-	1	0,5	0,5
9	НТ	0,5	1	0,5	-	-	1	0,5	-	1	0,5	0,5
10	ЯНС	1	-	1	-	1	0,5	1	-	-	1	0,55
11	РНГ	1	-	-	1	1	0,5	1	-	-	1	0,55
12	РПС	1	1	-	1	1	0,5	0,5	-	-	1	0,6
13	СпНС	1	0,5	0,5	-	-	1	0,5	1	1	0,5	0,6
14	ЖНС	1	0,5	0,5	1	1	0,5	0,5	0,5	0,5	0,5	0,65
15	НА	0,5	1	0,5	0,5	0,5	0,5	0,5	0,5	1	1	0,65

IV. ВЫВОДЫ И ЗАКЛЮЧЕНИЕ

В рамках представленного обзора дано перечисление основных задач, решаемых с помощью ИНС, и формализована методология их использования. Проанализированы 15 базовых нейросетевых архитектур, которые были выделены по результатам рассмотрения 22 исторических событий, относящихся к предметной области.

Дополнительно определены 10 ключевых свойств нейросетевых моделей для проведения их многопризнакового анализа. С опорой на введенную шкалу, было выделено 5 укрупненных классов ИНС и предложено 3 ключевых подхода к системной инженерии нейронных сетей – экспертный, генеративный и эволюционный.

По результатам сравнения интегральных оценок функциональных возможностей различных ИНС отметим, что в настоящее время особый приоритет приобретают вопросы обеспечения их модульности, структурно-функциональной адаптивности, точности и вычислительной эффективности.

ЛИТЕРАТУРА

- [1] W.S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biology*, vol. 5 (4), pp. 115-133, 1943.
- [2] M.G. Quiles and R.A.F. Romero, "A computer vision system based on multi-layer perceptrons for controlling mobile robots," *Proc. of 18th International Congress of Mechanical Engineering*, pp. 1-8, 2005.
- [3] A.L. Hodgkin and A.F. Huxley, "Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*," *The Journal of Physiology*, vol. 116 (4), pp. 449-472, 1952.
- [4] G.H. Rutherford, Z.D. Mobile, J. Brandt-Trainer, R. Follmann and E. Rosa, "Analog implementation of a Hodgkin-Huxley model neuron," *American Journal of Physics*, pp. 918-923, 2020.
- [5] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65 (6), pp. 386-408, 1958.
- [6] M.K. Bugeja and S.G. Fabri, "Multilayer perceptron adaptive dynamic control of mobile robots: experimental validation," *Proc. of the 2nd European Robotics Symposium EUROS 2008*, pp. 165-174, 2008.
- [7] S. Linnainmaa, "The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors," *Master's Thesis (in Finnish)*. Univ. Helsinki, 1970.
- [8] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1 (4), pp. 541-551, 1989.
- [9] V.P. Gladun, "Formation of concepts by growing learning nets," *Kibernetika*, vol. 2, 1970, pp. 99-104 (In Russ.).
- [10] Y.V. Selyavskiy, "Dynamic growing pyramidal networks for assessing innovative projects feasibility in metallurgy," *Journal of Legal and Economic Studies*, vol. 2, pp. 246-250, 2017 (In Russ.).
- [11] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36 (4), pp. 193-202, 1980.
- [12] C. Lin, J. Jhang, "Intelligent traffic-monitoring system based on YOLO and convolutional fuzzy neural networks," *Proc. of IEEE Access*, vol. 10, pp. 14120-14133, 2022.
- [13] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc Natl Acad Sci USA*, vol. 79 (8), pp. 2554-2558, 1982.
- [14] Z. Li and B. Deng, "A networked smart home system based on recurrent neural networks and reinforcement learning," *Systems Science & Control Engineering*, vol. 9 (1), pp. 775-783, 2021.
- [15] S.E. Fahlman, G.E. Hinton and T.J. Sejnowski, "Massively parallel architectures for AI: NETL, Thistle, and Boltzmann machines," *Proc. of AAAI-83* 109, vol. 113, pp. 1-10, 1983.
- [16] O. Alemi, W. Li and P. Pasquier, "Affect-expressive movement generation with factored conditional Restricted Boltzmann Machines," *Proc. of 2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 442-448, 2015.
- [17] L. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Transactions on Circuits and Systems*, vol. 35, pp. 1257-1272, 1988.
- [18] Z. Szlavik, R. Tetzlaff, A. Blug and H. Hoefler, "Visual inspection of metal objects using cellular neural networks," *Proc. of Int'l Workshop on Cellular Neural Networks and Their Applications*, pp. 1-5, 2006.
- [19] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, vol. 2 (4), pp. 303-314, 1989.
- [20] Z. Dlugosz and R. Dlugosz, "Nonlinear activation functions for artificial neural networks realized in hardware," *Proc. of 2018 25th International Conference "Mixed Design of Integrated Circuits and System" (MIXDES)*, pp. 381-384, 2018.
- [21] L.K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 993-1001, 1990.
- [22] M. Aljasim and R. Kashef, "E2DR: A deep learning ensemble-based driver distraction detection with recommendations model," *Sensors*, vol. 22 (5), pp. 1858, 2022.
- [23] T. Martinetz and K. Schulten, "A "Neural Gas" network learns topologies," *Artificial Neural Networks*, T. Kohonen, K. Makisara, O. Simula, J. Kangas (Eds.). Elsevier, pp. 397-402, 1991.
- [24] H. Sasaki, T. Fukuda, M. Satomi, N. Kubota and N. Kubota, "Growing neural gas for intelligent robot vision with range imaging camera," *Proc. of 2009 International Conference on Mechatronics and Automation*, pp. 3269-3274, 2009.
- [25] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9 (8), pp. 1735-1780, 1997.
- [26] F. Nicola, Y. Fujimoto and R. Oboe, "A LSTM neural network applied to mobile robots path planning," *Proc. of 2018 IEEE 16th*

- International Conference on Industrial Informatics (INDIN), pp. 349-354, 2018.
- [27] W. Maass, "Networks of spiking neurons: The third generation of neural network models," *Neural Networks*, vol. 10, pp. 1659-1671, 1997.
- [28] G. Tang and K.P. Michmizos, "Gridbot: an autonomous robot controlled by a Spiking Neural Network mimicking the brain's navigational system," *ICONS '18: Proceedings of the International Conference on Neuromorphic Systems*, pp. 1-8, 2018.
- [29] M. Gori, G. Monfardini and F. Scarselli, "A new model for learning in graph domains," *Proc. of the 2005 IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 729-734, 2005.
- [30] B. Platten, M. Macfarlane, D. Graus and S. Mesbah, "Automated personnel scheduling with reinforcement learning and graph neural networks," *Proc. of RecSys in HR'22: The 2nd Workshop on Recommender Systems for Human Resources*, pp. 1-10, 2022.
- [31] D.P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv:1312.6114*, pp. 1-14, 2013.
- [32] X. Chen, J. Xu, R. Zhou, W. Chen, J. Fang and C. Liu, "TrajVAE: A Variational AutoEncoder model for trajectory generation," *Neurocomputing*, vol. 428, pp. 332-339, 2021.
- [33] J. Sohl-Dickstein, E.A. Weiss, N. Maheswaranathan and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," *Proc. of the 32nd International Conference on Machine Learning*, vol. 37, pp. 1-10, 2015.
- [34] I. Reutov, "Generating of synthetic datasets using diffusion models for solving computer vision tasks in urban applications," *Procedia Computer Science*, vol. 229, pp. 335-344, 2023.
- [35] C.R. Qi, H. Su, K. Mo and L. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," *arXiv:1612.00593*, pp. 1-19, 2016.
- [36] V. Mayer, Q. Feng, J. Deng, Y. Shi, Z. Chen and A. Knoll, "FFHNet: Generating multi-fingered robotic grasps for unknown objects in real-time," *Proc. of 2022 International Conference on Robotics and Automation (ICRA)*, pp. 762-769, 2022.
- [37] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998-6008, 2017.
- [38] H. Kim, Y. Ohmura and Y. Kuniyoshi, "Transformer-based deep imitation learning for dual-arm robot manipulation," *Proc. of 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 8965-8972, 2021.
- [39] S. Sabour, N. Frosst and G.E. Hinton, "Dynamic routing between capsules," *arXiv:1710.09829*, pp. 1-11, 2017.
- [40] X. Zhang, L. Zheng, Z. Tan and S. Li, "Loop closure detection based on residual network and capsule network for mobile robot," *Sensors*, vol. 22 (19), pp. 7137, 2022.
- [41] M. Raissi, P. Perdikaris and G.E. Karniadakis, "Physics informed deep learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations," *arXiv:1711.10561*, pp. 1-22, 2017.
- [42] A.K. Sahoo and I. Klein, "MoRPI-PINN: A Physics-informed framework for mobile robot pure inertial navigation," *arXiv:2507.18206*, pp. 1-9, 2025.
- [43] R.M. Hasani, M. Lechner, A. Amini, D. Rus and R. Grosu, "Liquid time-constant networks," *Proc. of AAAI Conference on Artificial Intelligence*, vol. 35, pp. 7657-7666, 2020.
- [44] M. Chahine, R. Hasani, P. Kao, A. Ray, R. Shubert, M. Lechner, A. Amini and D. Rus, "Robust flight navigation out of distribution with liquid neural networks," *Science Robotics*, vol. 8 (77), pp. 1-15, 2023.
- [45] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *arXiv:1506.01497*, pp. 1-14, 2015.
- [46] R. Sapkota and M. Karkee, "Comparing YOLO11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment," *arXiv:2410.19869*, pp. 1-17, 2024.
- [47] L. Fu and S. Li, "A new semantic segmentation framework based on UNet," *Sensors*, vol. 23 (19), pp. 8123, 2023.
- [48] P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory," *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations*, pp. 194-281, 1986.
- [49] T. Kohonen, "Self-organization and associative memory," *Springer Series in Information Sciences*. 3d ed, 1989.
- [50] V. Di Massa, G. Monfardini, L. Sarti, F. Scarselli, M. Maggini and M. Gori, "A comparison between recursive neural networks and graph neural networks," *Proc. of the 2006 IEEE International Joint Conference on Neural Networks*, pp. 778-785, 2006.

Информация об авторах

Диане Секу Абдель Кадер, канд. техн. наук, старший научный сотрудник ИПУ РАН, доцент кафедры проблем РТУ МИРЭА, Москва, Россия, e-mail: diane1990@yandex.ru, ORCID: 0000-0002-8690-6422

Воронков Андрей Дадашевич, аспирант 4 курса, кафедры проблем управления РТУ МИРЭА, Москва, Россия, e-mail: a.voronkov.rtu@yandex.ru, ORCID: 0000-0003-4688-9346

Акимов Дмитрий Александрович, доцент кафедры автоматических систем, РТУ МИРЭА, Москва, Россия, e-mail: akimovdmritri@gmail.com, ORCID: 0000-0001-6889-618X

Гурьянова Екатерина Олеговна, старший преподаватель кафедры автоматических систем, РТУ МИРЭА, Москва, Россия, e-mail: guryanova@mirea.ru, ORCID: 0000-0002-8809-8801

History of development and prospects for system engineering of neural architectures

Sekou Abdel Kader Diane^{1,2}, Andrey Dadashevich²
Voronkov, Dmitry Alexandrovich Akimov², Ekaterina
Olegovna Guryanova²

¹*V.A. Trapeznikov Institute of Control Sciences, Moscow, Russia*

²*MIREA – Russian Technological University, Moscow, Russia*

Abstract – In this article typical neural network architectures are classified and key historical stages in their development are examined. Fundamental functional properties of neural architectures are identified and method for multi-criteria evaluation of their potential effectiveness is presented. Several approaches to system engineering of neural networks are also proposed. It is shown that, along with well-established principles of multilayering, hierarchy and recurrence, new issues are becoming increasingly important such as modularity, structural adaptability, and computational efficiency of neural networks.

Keywords – neural network architectures, history of neural networks, systems engineering, machine learning, multicriteria assessment.