

УДК 625.7/8.004

Обзор современного оборудования

Дустмуродов Э. Э., Махмудова Д. Х., Дустмуродова Х. Э.
Ташкентский государственный транспортный университет
Ташкент, Республика Узбекистан

Растущая сложность интеллектуальных транспортных систем (ИТС), которые включают в себя широкий спектр приложений и услуг, вызвала необходимость в высокопроизводительных современных аппаратных устройствах (МНД). Проблема производительности стала более заметной после интеграции методов машинного обучения (ML), развернутых в крупномасштабных условиях. Машинное обучение эффективно поддержало сферу ИТС, предоставив эффективные и оптимизированные решения проблем, которые в противном случае решались с использованием традиционных статистических и аналитических подходов. Потребность в развертывании аппаратного обеспечения ИТС в эпоху машинного обучения – сложная проблема, которая включает в себя временные, пространственные, экологические и экономические факторы. В этом обзоре рассматривается новейшая литература ИТС на основе машинного обучения, в которых использовались МНД, с упором на показатели эффективности.

За последние несколько десятилетий ИТС вызвали растущий исследовательский интерес как многообещающая дисциплина, способная произвести революцию в транспортном секторе и решить общие проблемы, связанные с дорожным движением и транспортными средствами. ИТС включает в себя множество взаимосвязанных инженерных решений, которые функционируют как единое целое для оптимизации путешествий в масштабе сети с технической, социальной, экономической и экологической точек зрения. Такая оптимизация требует развития информационных и коммуникационных технологий, электронных датчиков, систем управления и компьютеров, что подчеркивает управляемую данными природу современных ИТС. Собирая и анализируя соответствующие данные, ИТС может эффективно управлять ресурсами вычислительного оборудования с помощью нескольких алгоритмов управления и координации, что приводит к повышению удобства путешественников, снижению расхода топлива и увеличению транспортных потоков. ИТС – это широкая тема, охватыва-

ющая множество предметов, каждая из которых представляет собой отдельную исследовательскую задачу. Этим темам посвящено множество исследований, но их можно разделить на несколько областей, включая передовые системы управления дорожным движением, передовые информационные системы для путешественников, передовые системы управления транспортными средствами, передовые системы общественного транспорта, эксплуатацию коммерческих транспортных средств и системы сельского транспорта [1].

Сложность приложений ИТС требует наличия мощных в вычислительном отношении алгоритмов, способных анализировать огромные объемы данных, особенно в эпоху больших данных, а также компетентных устройств реализации для получения точных результатов в реальном времени. Удобно, что ML включает в себя широкий спектр таких алгоритмов, предназначенных для классификации, регрессии, ранжирования, кластеризации и уменьшения размерности в различных сценариях обучения, таких как контролируемое, неконтролируемое, полуконтролируемое, онлайн-обучение, подкрепление или активное обучение, а также федеративное обучение. Таким образом, машинное обучение, включая глубокое обучение (DL), имеет решающее значение для различных приложений, включая, помимо прочего, классификацию текста или документов, обработку естественного языка, обработку речи, компьютерное зрение и другие.

МНД для ИТС-приложений на основе машинного обучения



Рис. 1. Набор приложений ИТС для контроля транспортных средств и управления дорожным движением с акцентом на разнообразное внедрение методов МНД и ML

(рис. 1). Тщательный обзор недавней литературы обращает внимание на то, что обширные исследования направлены на внедрение методов машинного обучения на МНД для решения общих проблем ИТС, связанных с контролем транспортных средств и управлением дорожным движением. Эти проблемы связаны с получением данных для дальнейшей обработки и анализа при одновременном выполнении требований реального времени. Таким образом, выполняется обучение и тестирование подходящих алгоритмов ML, которые сравниваются с аналогичной работой посредством оценки производительности. Авторы выбрали различные аппаратные устройства,

которые подпадают под три основные категории: многоцелевые устройства (MD), специализированные устройства (SD) и гибридные устройства (HD). MD определяются как аппаратные устройства, способные выполнять

разнообразные вычислительные задачи благодаря их перепрограммируемости. Напротив, SD – это устройства, специально разработанные и оптимизированные для желаемого применения. HD, однако, объединяют два вышеупомянутых типа для достижения уникальных результатов за счет совместного проектирования аппаратного и программного обеспечения, предварительной обработки и совместной обработки (рис. 1).

В табл. представлена классификация обследованных статей по категориям аппаратного обеспечения в хронологическом порядке, где были представлены используемые аппаратные устройства, алгоритм ML и применение ИТС.

Таблица

Классификация категорий аппаратного обеспечения в хронологическом порядке (*A* – обнаружение трафика; *B* – мониторинг трафика; *C* – анализ автомобиля; *D* – распознавание; *E* – оптимизация)

Группа	Год	Аппаратное устройство	Аппаратное устройство. Выбор				Алгоритм машинного обучения	ИТС-приложение
			Проводимость	Глобость	Мощность Кон.	Ресурсы		
1	2	3	4	5	6	7	8	9
Многоцелевые устройства	2014	ARM Cortex A9	+				Модель гауссовой смеси	<i>A</i>
		Xilinx Virtex-5	+				Адабуст	<i>B</i>
		Intel Altera Stratix II	+			+	Машина опорных векторов	<i>A</i>
	2015	ARM Cortex M4F					SVM	<i>A</i>
		Xilinx Spartan-6	+				SVM	<i>C</i>
	2016	Intel Altera Cyclone IV			+		Модель гауссовой смеси	<i>A</i>
		Xilinx Spartan-6	+		+		SVM	<i>B</i>
	2017	Xilinx Virtex-5	+	+			Адабуст	<i>A</i>
		Intel Altera Cyclone IV	+			+	SVM	<i>A</i>
	2018	Xilinx Virtex-6	+	+			ANN	<i>D</i>
		Xilinx UltraScale	+				SVM & NN	<i>C</i>
	2019	Xilinx Spartan-3	+		+		EZ-Net	<i>A</i>
		Raspberry Pi 4	+				CNN	<i>B</i>
	2020	Raspberry Pi 3B+			+		M2-tMTL	<i>D</i>

Продолжение таблицы

1	2	3	4	5	6	7	8	9	
Специализированные устройства	2014	NVIDIA Tesla C2075	+			+	SVM	A	
		NVIDIA Tesla K20	+			+	CNN	D	
	2017	NVIDIA Tesla K40	+					LSM	A
		NVIDIA GeForce GTX 970	+					FM	A
		NVIDIA Tesla C2075	+				+	FCSNet	B
		NVIDIA GeForce GTX	+					CNN	D
	2018	NVIDIA GeForce GTX 1080	+				+	AggNet	D
		TSMC 90nm	+			+	+	SVM	A
		Snapdragon TM 820A (GPU)	+					CNN	D
		NVIDIA Titan Black	+				+	CNN	A
	2019	NVIDIA Tesla K20c	+				+	CNN	B
		NVIDIA GeForce GTX 1080Ti	+					CNN	A
		NVIDIA Titan Xp					+	CNN	D
		NVIDIA GeForce GTX 1080	+					Составной автоэнкодер	D
	2020	NVIDIA Tesla K80					+	MSRA-NN	B
		NVIDIA Tesla P100-PICE					+	L-CNN	E
NVIDIA Titan Xp		+					GCGA	E	
NVIDIA GeForce GTX 1060						+	SE_ResNet	C	
Гибридные устройства	2015	FriendlyARM Tiny4412	+			+	CNN	B	
	2016	Xilinx Zynq-7000	+			+	ELM	D	
		Xilinx Zynq-7000 SoC	+				+	SE-ResNet	A
	2017	NVIDIA Jetson TX	+			+	CNN & RLMS	D	
	2018	Xilinx Zynq-7000 SoC	+				RLMS	D	
	2019	NVIDIA Jetson TK1	+				+	CNN	B
Xilinx Zynq-7000 SoC					+		SVM	B	
NVIDIA Tegra TX1						+	CNN	B	

Выявленные аппаратные устройства являются основными аппаратными устройствами, использованными в расследованиях, поскольку в не-

скольких расследованиях использовались вторичные аппаратные устройства для таких целей, как обучение и/или предварительная обработка.

Для каждой ссылки также выделены различные факторы, влияющие на выбор аппаратного устройства и алгоритма ML для каждой статьи. Эти факторы выбора были идентифицированы в статьях либо путем выводов из текста, либо путем обнаружения этих факторов явно упомянутых.

Для ML алгоритмов факторы выбора следующие: производительность отражает время обучения и вывода, точность отражает способность модели давать правильные результаты, надежность отражает способность алгоритма обрабатывать изменения входных данных, а сложность отражает то, является ли модель легкой для вычислений.

В табл. L-CNN обозначает сверточную нейронную сеть на основе поиска, SE-ResNet – сеть сжатия и возбуждения-остатка, а GCGA – сверточный генеративный автоэнкодер графа.

Для ML алгоритмов факторы выбора следующие: производительность отражает время обучения и вывода, точность отражает способность модели давать правильные результаты, надежность отражает способность алгоритма обрабатывать изменения входных данных, а сложность отражает то, является ли модель легкой для вычислений. В табл. L-CNN обозначает сверточную нейронную сеть на основе поиска, SE-ResNet – сеть сжатия и возбуждения-остатка, а GCGA – сверточный генеративный автоэнкодер графа.

Верхний индекс в столбце «Аппаратное устройство» соответствует цели машинного обучения, для которой использовалось устройство, где «train» означает обучение, «inf» – вывод, «on» – онлайн-обучение, «pre» – предварительная обработка и «данные» для сбора данных.

Аналогично, верхний индекс в столбце приложения ИТС соответствует области применения ИТС, где «sim» означает «Моделирование», «pho» – «Физический прототип ИТС», а «der» – физическое развертывание ИТС.

Таксономия ИТС и модель производительности. Тщательный обзор приложений ИТС в широком спектре задач контроля транспортных средств и управления дорожным движением выявил набор отдельных подкатегорий [2], представленных на рис. 2, которые представляют собой общие проблемы, решаемые в литературе, сгруппированные по целям.

В разделе «Управление транспортным средством» анализ транспортного средства относится к сбору и анализу данных для водителя или транспортного средства, распознавание инструкций относится к распознаванию светофора и знаков, а помощь водителю включает в себя задачи транспортного средства, которые не подпадают под другие подкатегории. Для управления дорожным движением обнаружение трафика подразуме-

вает определение типа или номерного знака транспортного средства, а оптимизация включает, среди прочего, изучение состояния дороги, погоды, качества воздуха. Остальные подкатегории говорят сами за себя.

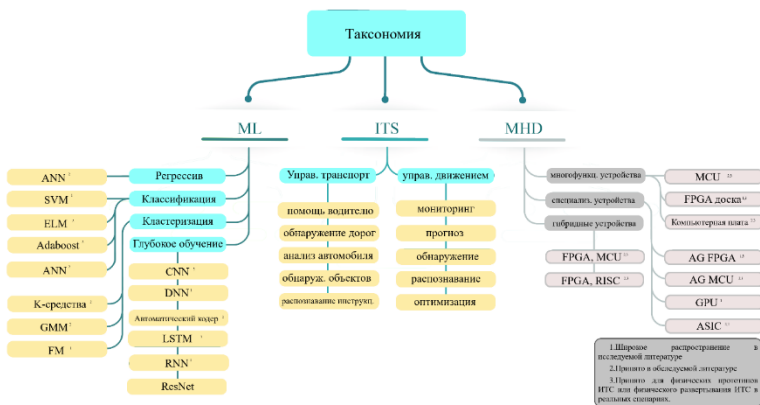


Рис. 2: Таксономия методов MHD и ML, используемых в рассмотренных приложениях ИТС, где ELM означает «Машина с экстремальным обучением», GMM для модели гауссовой смеси, FM для нечетких минимумов, DNN для глубокой нейронной сети, LSTM для долгосрочной краткосрочной памяти, и RNN для рекуррентной нейронной сети

Различные методы обучения, используемые в литературе для приложений ИТС с реализациями MHD, сгруппированы в четыре категории, как показано на рис. 2: регрессия, классификация, кластеризация и DL. Эта группировка основана на отношениях сопоставления ввода-вывода, где регрессия прогнозирует действительные метки невидимых данных, классификационные метки невидимых данных с учетом дискретного набора возможных классов, а кластеризация идентифицирует закономерности или области в наборе данных. Эти категории включают в себя классические алгоритмы машинного обучения в дополнение к простым архитектурам ИТС. Предлагается, чтобы четвертая категория DL включала более сложные и сложные архитектуры нейронных сетей, которые можно использовать для любой из трех других категорий. В исследованных документах использовались различные показатели эффективности применяемых методов ML.

Различные MHD, представленные в литературе, делятся на три класса: MD, SD и HD. Эти классы, показанные на рис. 2, включают устройства, которые являются неотъемлемой частью любой системы ИТС, каждое из

которых обладает особым набором функций, определяющих компетентность устройства. В MD микроконтроллеры, которые считаются устройствами с низким энергопотреблением, предлагают ограниченные вычислительные возможности, в отличие от более сложных компьютерных плат, которые могут решать сложные вычислительные задачи. MD также включают в себя FPGA, которые поддерживают программируемые логические блоки, обеспечивая реконфигурацию оборудования после производства. В то время как MD выполняют задачи общего назначения, SD занимаются конкретными задачами для целевых приложений. В SD, платы AG FPGA и микроконтроллеры включают в себя специализированные вычислительные блоки, которые облегчают работу приложений ИТС, уменьшая сложность их проблем. В этот класс также входят графические процессоры, которые используются в приложениях компьютерного зрения, и ASIC, которые обладают специально разработанными функциональными блоками для конкретного использования. Третий класс MHD предлагается для решения HD, которые связывают комбинации MD и SD.

Концепция модели выбора. Модель оценки перекрестного сопоставления, представленная выше, направлена на ранжирование методов ML и MHD, которые обещают определенный уровень производительности в зависимости от сложности ИТС. Концептуальная модель выбора представлена на рис. 3 с целью выбора (или рекомендации) метода ML и MHD, подходящих для применения ИТС.

Предлагаемая концепция модели выбора подчеркивает необходимость дополнительных показателей, выходящих за рамки диаграммы перекрестного соответствия и аспекта производительности процесса оценки, которые включают сферу применения (реальное развертывание или тестирование и проверка), бюджет и выбор технологии, который соответствует типу проблемы.

Подмодель выбора MHD учитывает показатели, которые не были включены в структуру перекрестного сопоставления, такие как процессор (общего и специального назначения) и технологии ИС (программируемые логические устройства или ASIC), а также подходящие ранги MHD полученные из таблицы перекрестных сопоставлений. Примечательно, что основным показателем выбора MHD, оказывающим наибольшее влияние на процесс выбора, является стоимость целевого устройства. Что касается подмодели выбора ML, она учитывает тип проблемы (обучение с учителем, обучение без учителя или обучение с подкреплением) и сценарий ее работы, а также ранги подходящего метода ML, выведенные из диаграммы перекрестного сопоставления.

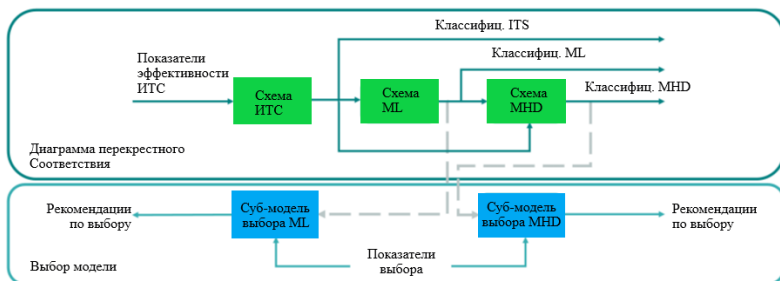


Рис. 3. Концептуальная модель отбора, которую можно интегрировать с предлагаемой моделью перекрестной оценки

Предлагаемые подмодели могут сочетать в себе разнородные, качественные и количественные показатели [5]. Кроме того, подмодели могут быть сформулированы с использованием традиционных методов принятия решений, основанных на множестве критериев, объединенных статистически [3–5] или с использованием машинного обучения [5]. Стоит отметить, что модель перекрестной оценки и модель отбора были разделены, чтобы обеспечить разделение задач; поскольку первый демонстрирует ориентацию на производительность, а второй направлен на отражение концепции помощи в рекомендациях по выбору.

Ожидаемый объем данных, которые будут производиться ИТС в эпоху больших данных, стимулировал заметный исследовательский интерес к применению методов машинного обучения для работы определенных приложений. Благодаря своей способности обрабатывать большие объемы данных и находить нелинейные связи в сложных системах, машинное обучение вызвало волну, прокатившуюся по ИТС, предоставляя жизнеспособные решения сложных проблем. Хотя методы ML играют решающую роль в качестве услуг, предлагаемых ИТС, аналогичная роль возлагается на устройства обработки, используемые для запуска этих методов, основанных на данных, особенно с учетом недавних разработок в области автономного вождения и подключенных транспортных средств. В этой статье мы предоставляем всесторонний обзор недавней литературы по приложениям ИТС, где методы ML применялись вместе с МНД для запуска этих методов. Вдохновленные обзорной литературой, мы предлагаем таксономию на рис. 2, которая классифицирует подполя ИТС, ML и МНД таким образом, чтобы это ограничивалось контекстом исследования. Для решения нетривиальной задачи выбора подходящего метода ML и

MHD для приложения ИТС определенной сложности мы разрабатываем структуру оценки производительности, диаграмма оценки перекрестного сопоставления которой показана на табл. Предлагаемая структура обеспечивает рекомендация о том, какую компетенцию техники ML и MHD выбрать для приложения ИТС, сложность которой можно определить по предлагаемому уровням производительности. После проверки наша система оказалась достаточно эффективной, имея лишь несколько небольших различий в классификации из пула ссылок и только два существенных различия между рекомендациями системы.

Несмотря на обширные исследования в литературе, мы выявили несколько пробелов и ограничений, которые могут определять направления будущих исследований. С точки зрения машинного обучения дополнительные усилия должны быть направлены на повышение надежности моделей, которые будут использоваться для развертывания в реальном мире. Большинство опрошенных приложений ИТС, которые полагаются на визуальные данные, выиграют от этого, поскольку они в основном обучаются на наборах данных, которые не в полной мере отражают реальные проблемы. Следовательно, исследование таких методов, как адаптация предметной области, может быть полезно для адаптации моделей к реальной среде. Аналогичным образом, в области аппаратного обеспечения MHD, такие как TPU, AG FPGA или другие, не использовались в литературе, несмотря на их многообещающие характеристики.

Литература

1. Meneguetto, R. I. Intelligent Transport System in Smart Cities / R. I. Meneguetto, R. E. D. Grande, A. A. F. Loureiro // Springer International Publishing, 2018.
2. Zappone, A. Wireless networks design in the era of deep learning: Model-based, ai-based, or both / A. Zappone, M. Di Renzo, M. Debbah // IEEE-Transactions on Communications. – 2019. – Vol. 67, № 10. – P. 7331–7376.
3. Юлдашев, Б. С. (2020). Понимание бозона Хиггса с помощью LhcC / Б. С. Юлдашев, Э. Э. Дустмуродов, Б. П. Турдиев, Т. Б. Файзиев // Science and Education, 2020. – Vol. 4, № 1. – P. 16–21.
4. Dustmurodov, E. The concept of the FRITIOF (FTF) model / E. Dustmurodov, X. Dustmurodova, N. Valikhanov, D. Nomozova // Science and innovation. – 2023. – № 2(A5). – P. 51–53.
5. Damaj, I. An analytical framework for high-speed hardware particle swarm optimization / I. Damaj, M. Elshafei, M. El-Abd, M. Emin Aydin // Microprocessors and Microsystems. – 2020. – Vol. 72, № 2. – P. 102949.