

ОЦЕНКА КОГНИТИВНОЙ ЗАГРУЖЕННОСТИ ОПЕРАТОРА НА ОСНОВЕ ВИДЕОИНФОРМАЦИИ С ПОМОЩЬЮ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Б.А. Шишов

Предложен метод оценки когнитивной нагрузки оператора, состоящий в оценивании рабочей нагрузки по информации с видеокамеры с помощью рекуррентных нейронных сетей, обучаемых индивидуально. Для построения модели рабочая нагрузка предварительно оценивается в специальных экспериментальных условиях и записываются лицевые характеристики и направление взгляда испытуемых. Затем когнитивная нагрузка моделируется на основе полученных данных с помощью рекуррентных нейронных сетей с «долгой краткосрочной памятью».

Ключевые слова: когнитивная нагрузка, оценка оператора, рекуррентные нейронные сети, анализ видео.

ВВЕДЕНИЕ

Большинство производственных систем управления можно представить, как комбинацию двух взаимодействующих друг с другом составляющих — человека и машины. Производительность и надежность таких систем сильно зависят от обеих составляющих.

Если моделирование производительности машины может быть сравнительно простым, то построение обобщенной модели производительности человека не представляется возможным несмотря на многочисленные исследования. Многие исследователи сосредоточены на оценке влияния отдельных человеческих факторов и разрабатывают модели для конкретных задач, ориентированных на практическое применение конкретной системы управления.

Проблема оценки функционирования человеко-машинных систем управления актуальна для многих производств, особенно при управлении непрерывными процессами. В нефтегазовом комплексе, в химических производствах, в атомной энергетике большое распространение получили автоматизированные системы диспетчерского управления (АСДУ) [1, 2]. Исследование действий диспетчера или оператора позволит получить новые возможности для оценки качества функционирования АСДУ в целом.

Рабочая когнитивная нагрузка — один из важнейших индикаторов когнитивной активности. Ее

можно охарактеризовать как ментальную характеристику, которая отражает психическое напряжение, возникающее в результате выполнения задачи в конкретных условиях окружающей среды и эксплуатации, в сочетании с возможностью оператора реагировать на эти условия. Основная причина потребности в измерениях когнитивной нагрузки состоит в необходимости качественной оценки умственных затрат на выполнение задач в целях прогнозирования работы оператора и системы. Это особенно важно в случаях, когда высокие требования к решению задач могут привести к неприемлемой производительности системы.

Различают три основных подхода к измерению когнитивной нагрузки:

- с помощью субъективных рейтинговых шкал (вопросников);
- на основе анализа эффективности выполнения задач;
- путем измерения психофизиологических характеристик.

Различные подходы чувствительны к различным аспектам когнитивной нагрузки. В результате разные подходы оценивают несколько разные характеристики. Основная причина — отсутствие строгого определения рабочей когнитивной нагрузки.

Модели на основе субъективных рейтинговых шкал опираются на оценку, которую дает сам оператор. Одна из таких моделей — NASA-TLX (Task



Load Index) — индекс, отражающий загруженность оператора в ходе выполнения задачи, разработанный в NASA для оценки загруженности пилотов [3]. Анализ когнитивной нагрузки с ее помощью предполагает единовременную субъективную оценку самим оператором, что делает такие модели не подходящими для долгосрочного мониторинга в режиме реального времени.

Существует большое число моделей, основанных на анализе эффективности выполнения задач оператором, которые часто представлены как генеративные или предсказательные модели, но они в значительной степени зависят от подробного описания поведения оператора относительно целей и задач. Известные примеры таких моделей:

- семейство моделей Goals Operators Methods and Selection rules (GOMS), в которых для анализа требуется последовательно описать выполняемую задачу с помощью специальных языковых операторов [4];

- Model Human Processor (МНР), позволяющая оценить время, требуемое для решения задачи путем ее декомпозиции на элементарные действия, для которых имеется временная оценка [5];

- когнитивная архитектура ACT-R для разработки и симуляции моделей когнитивных процессов [6];

- сетевые модели массового обслуживания, такие как QN-MHP (Queueing Network MHP), в которой информация представлена в виде заявок, обработкой которых занимаются параллельно функционирующие модули, моделирующие отдельные когнитивные функции [7].

Однако для всех этих моделей требуется подробное разбиение задач на простые операторы, что возможно только тогда, когда задача уже хорошо изучена; т. е. речь идет об объектах, структура которых допускает декомпозицию. В сложных, частично наблюдаемых или неизвестных средах такие модели могут быть непригодными.

Подход, опирающийся на измерение психофизиологических характеристик, позволяет измерять определенные характеристики в динамическом режиме без необходимости описания задачи. Эти методы часто требуют надежных алгоритмов преобразования входных сигналов и зависят от чувствительности данных сигналов к когнитивной нагрузке. В рамках данного метода для измерения когнитивной нагрузки наиболее часто применяется электроэнцефалография [8]. Известны исследования, в которых используются результаты измерений сердечного ритма и других его показателей, частоты дыхания, частоты моргания, движения глаз, температуры лица и расширения глазного зрачка. Однако для этого часто требуются дополнительные аппаратные средства, такие как электроэнцефалографы или высокоскоростные инфра-

красные камеры, поэтому их практическое применение вызывает затруднения.

Для оценки когнитивной нагрузки оператора целесообразно совместить подход, основанный на анализе эффективности, с подходом, предусматривающим измерение психофизиологических характеристик. Предлагается строить модель оценки когнитивной нагрузки по видеоизображению лица оператора, поскольку положение головы, выражение лица [9—11] и направление взгляда [12] дают много информации о физическом и психическом состоянии оператора. Такое совмещение стало возможным благодаря недавним достижениям в машинном обучении [13].

К числу хорошо известных моделей для имитации выражений лица относится, в частности, система кодирования лицевых движений FACS (Facial Action Coding System) [14], которая дает количественную оценку в терминах двигательных единиц — Action Units (AU). Широкие возможности современных методов компьютерного зрения [9—12] позволяют воспользоваться камерой потребительского уровня (например, веб-камеру) в качестве единственного дополнительного оборудования для захвата изображения лица.

В настоящей работе предлагается модель, описывающая динамику когнитивной нагрузки в зависимости от характеристик лица на основе модели FACS и информации о направлении взгляда [12]. Она может рассматриваться как модель прогнозирования временного ряда по входным временным рядам (чертам лица). Поэтому целесообразно применить рекуррентные нейронные сети (Recurrent Neural Networks, RNN) с долгой краткосрочной памятью (Long short-term memory, LSTM) [15], поскольку они эффективны [13] при прогнозировании временных рядов. Для оценки когнитивной нагрузки необходимо сначала обучить LSTM-модель, предоставив ей некоторую целевую переменную (когнитивную нагрузку). Чтобы получить значения моделируемой переменной (когнитивной нагрузки), проводился специально организованный эксперимент, в котором когнитивная рабочая нагрузка моделировалась методом подробного анализа задач.

1. ПРЕДЛАГАЕМЫЙ МЕТОД

Основная идея предлагаемого метода состоит в моделировании когнитивной рабочей нагрузки на основе рекуррентных нейронных сетей RNN с «долгой краткосрочной памятью» (LSTM) [15], используя в качестве входных данных характеристики лица (Action Units) по модели FACS и информацию о направлении взгляда, которые извлекаются из видео (рис. 1). Во время проведения эксперимента фиксируются данные для обучения

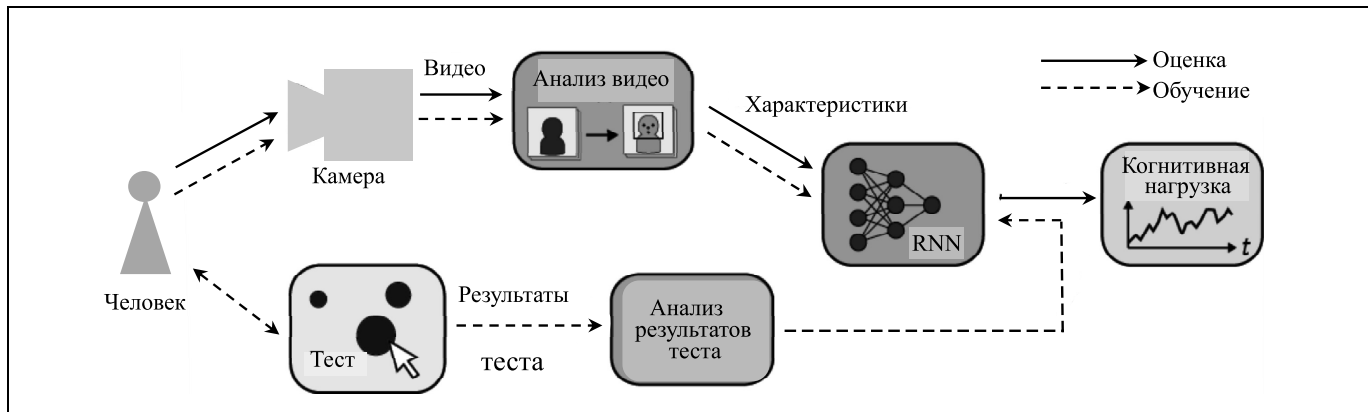


Рис. 1. Схема предложенного метода анализа когнитивной нагрузки

(черты лица, направление взгляда и целевая когнитивная нагрузка). Для получения значения целевой переменной (когнитивной нагрузки) применяется специально разработанная модель на основе МНР [5], в основе которой лежит детальный анализ активности при выполнении задач. На рис. 1 показано, что предлагаемый метод содержит два отдельных процесса: обучение модели на основе результатов специального теста и непосредственно применение модели (оценки) когнитивной нагрузки.

Обучение модели требуется для «настройки» нейронной сети на конкретную задачу и конкретного человека. Выбор именно рекуррентных нейронных сетей LSTM обусловлен тем, что эти сети позволяют учесть динамику характеристик во времени. Отметим, что обучение осуществляется индивидуально для каждого испытуемого с целью учета индивидуальных особенностей; более подробно вопрос индивидуальности различий в зависимости от лицевых характеристик и когнитивной нагрузки рассмотрен в работе [1].

Как только модель будет обучена, ее можно будет использовать для оценки и прогнозирования когнитивной нагрузки только по видео с обычной веб-камеры. Для построения модели необходимо выполнить несколько шагов (см. далее п. 1.1—1.3).

1.1. Разработка тестового эксперимента для оценки когнитивной нагрузки

Подборка данных для обучения нейронных сетей во многом определяет качество моделирования. Учитывая особенности поставленной задачи, предлагается для получения данных в целях обучения нейронной сети провести эксперимент, моделирующий когнитивную нагрузку.

Поскольку, с одной стороны, необходимо учитывать индивидуальность испытуемых и что каждая модель — нейронная сеть имитирует его действия, а, с другой стороны, этап сбора и подготовки

данных представляет собой лишь фрагмент задачи, то на основании анализа существующих моделей [3—7] было решено создать достаточно простой тест, чтобы уменьшить общую сложность системы моделирования. Участникам предлагается выполнять примитивные задачи на компьютере — щелкать «мышью» по кругам на экране (рис. 2). Эти круги появляются в случайных местах и исчезают после определенного временного интервала (1300 мс). Круги увеличиваются в радиусе на протяжении их существования (от 30 до 40 пикселей). Каждый новый круг появляется с задержкой (временем между появлением разных кругов), которая меняется во время эксперимента. Начиная с задержки в 1200 мс, в течение 3 мин эта задержка линейно уменьшается до 300 мс, что приводит к тому, что появление новых задач ускоряется. Точная формула для расчета задержки имеет вид:

$$\text{delay}(t) = 1200 - (1200 - 300) \frac{t}{180\,000},$$

$$0 \leq t \leq 180\,000,$$

где t — время эксперимента, мс. Во время сеанса эксперимента (всего 3 мин) записывается производительность выполнения задач. Она состоит из

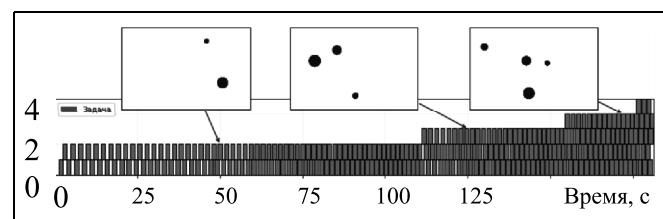


Рис. 2. График возникающих задач во время эксперимента и снимков экрана в разные моменты: каждый серый прямоугольник представляет собой время существования отдельного кружка (1300 мс); в пике одновременно может отображаться на экране до пяти кругов

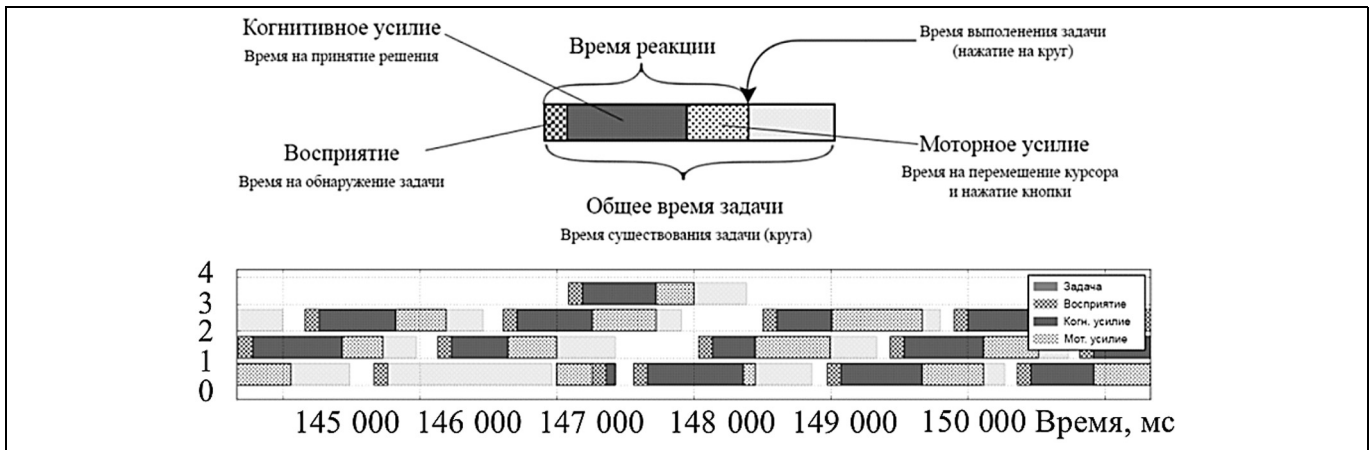


Рис. 3. Разбиение каждой задачи на этапы: задачи состоят из восприятия, когнитивного усилия и моторной реакции; моторные реакции не перекрываются; на 147-й с событие щелчка мышью мимо круга, для которого требуется некоторое время для проверки; в момент 145,8 с — пропущенный кружок

времени реакции (время от появления круга до щелчка мышью, мс), расстояния между позициями щелчков мышью (требуется для дальнейшей оценки времени моторной реакции), числа пропущенных кругов и числа щелчков мимо кругов. Кроме того, в процессе эксперимента записывается видео лица.

Несмотря на то, что задачи примитивны, задача нажатия на кружки по-прежнему может быть подходящей для моделирования рабочей когнитивной нагрузки, поскольку в соответствии с МНР [4] даже эти задачи по-прежнему требуют восприятия (кружок появляется на экране), когнитивной деятельности (решение перемещать курсор) и моторной реакции (перемещение курсора). Таким образом, можно разбить каждую задачу на эти этапы. Сами задачи практически идентичны по содержанию, что приводит к почти одинаковому количеству времени, необходимого для каждой фазы для разных задач. Тем не менее, участнику требовалось перемещать взгляд и курсор на случайное расстояние для каждой задачи (так как круги отображаются в случайном положении на экране), что влияет на время, необходимое для восприятия, и на моторную реакцию. Измерить движение глаз и общее время восприятия с достаточной точностью довольно сложно, поэтому предположим (согласно МНР [5]), что восприятие каждого круга почти равно времени цикла процессора восприятия (100 мс). Тем не менее, требуемое для моторной реакции время, больше времени для восприятия, и, поскольку имеется информация о движении курсора, можно применить закон Фиттса для двумерных задач [16] для оценки времени моторной реакции. Данный закон для конкретной задачи представляется в виде:

$$M = a + b \log_2(D/R),$$

где M — среднее время, требуемое для движения мыши; D — расстояние (между координатами нажатий); R — радиус отдельного круга; a и b — константы, определяемые для конкретных условий на основе регрессионного анализа.

После того, как время моторной реакции оценено для каждой задачи, необходимо удалить перекрытия времени, пытаясь держать время моторной реакции как можно ближе к оценке, так как в этом эксперименте человек не может выполнить несколько движений мышью одновременно. Каждая задача не требует усилий памяти, так как щелчок по кругу в основном является просто сенсорно-двигательной операцией, что приводит к упрощенной модели МНР. В результате время реакции (RT) представляет собой сумму времени восприятия (100 мс), времени когнитивной и моторной реакции, последняя моделируется законом Фиттса. Пример такого разбиения показан на рис. 3.

Выполнив данное разбиение, можно оценить динамику когнитивной нагрузки. Для этого «складываем вертикально» когнитивные составляющие из графика, игнорируя перекрываемое время (поскольку перекрытие связано с явлениями психологического рефрактерного периода [17]). Результатом служит диаграмма, где заполненные части означают, что на тот момент участник проявил некоторое когнитивное усилие (рис. 4, сверху). Этот процесс можно формализовать таким образом:

$$\text{effort}(t) = \begin{cases} 1, & \text{если } \exists T_i : t \in [T_i^{\text{start}} + P_i, T_i^{\text{start}} + RT_i - M_i], \\ 0, & \text{иначе,} \end{cases}$$

где t — время, мс, T_i — одна из задач всего эксперимента, T_i^{start} — время начала задачи (время, ког-

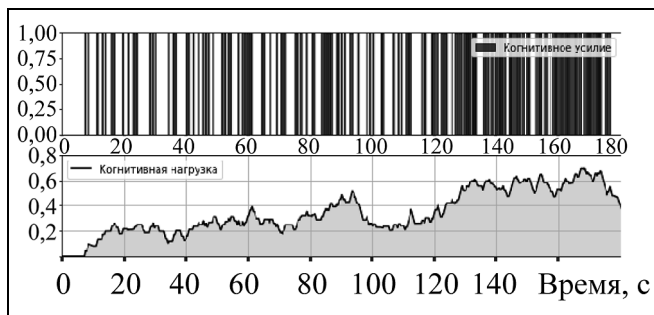


Рис. 4. Результаты обработки производительности выполнения задач, полученной из эксперимента: визуализация когнитивных усилий; каждая заполненная область времени означает, что в это время участник продемонстрировал некоторые когнитивные усилия (вверху); когнитивная рабочая нагрузка, определяемая плотностью когнитивных усилий (внизу)

да на экране появляется круг), RT_i — время общей реакции на эту задачу, P_i — время восприятия, M_i — моторная реакция.

Вычислив временные интервалы когнитивных усилий, необходимо оценить саму когнитивную нагрузку. Для этого когнитивная нагрузка может быть представлена как загруженность некоего ресурса от 0 до 1, который загружен в определенный промежуток времени, если оператор прилагает когнитивные усилия. Общее время когнитивных усилий в течение некоторого периода по отношению ко всему времени эксперимента даст индекс загрузки такого ресурса, который, по сути, является когнитивной нагрузкой. Идея состоит в том, чтобы пройти скользящим окном фиксированной ширины по графику когнитивных усилий (effort), вычисляя нагрузку (workload) таким образом:

$$windowload(t) = \frac{1}{window} \int_{t-window}^t effort(t)dt,$$

где $window$ — интервал окна, мс. Для этого эксперимента использовалось окно, равное 10 000 мс; см. результаты на рис. 4 (внизу).

Затем вычисленная кривая когнитивной нагрузки будет использоваться для обучения модели (рекуррентной нейронной сети). Кроме того, можно выполнить аналогичные шаги для расчета нагрузки восприятия и моторной нагрузки.

1.2. Извлечение лицевых характеристик из видеосигнала

Для получения данных обучения, помимо кривой рабочей нагрузки, также необходимо извлечь черты лица из видео, снятого во время эксперимента. Само выражение лица содержит много информации о человеческом состоянии. Основной подход к описанию выражения лица заключается в применении системы кодирования лицевых дейст-

вий (FACS) [15]. Он дает подробное объяснение выражения лица с помощью параметров, называемых «Action Units» (AU), которые могут быть численно оценены, что дает возможность использовать их в качестве признаков для нейронной сети. Последние достижения в области компьютерного зрения [10] предоставляют надежные модели для извлечения AU из видео. Они базируются на модели оценки маркеров лица [9]. Полезным фактором для описания рабочей нагрузки также может служить движение глаз. Доказано [18], что движение глаз может быть надежной функцией для оценки когнитивной нагрузки. В настоящей работе была применена существующая модель [12] для извлечения подобной информации.

Цель этого процесса — извлечение характеристик лица и движения глаз из записанного видео. Поскольку эта проблема уже хорошо изучена, разработано программное обеспечение для извлечения данных на основе инструментария OpenFace [11]. Характеристики, извлекаемые из видео, включают в себя: информацию о положении головы (положение и вращение в трехмерных координатах), направление взгляда (две трехмерные точки, по одному для каждого глаза) и интенсивность 16-ти параметров AU, что дает 28 признаков для каждого кадра. Средняя частота кадров камеры составляет 30 с^{-1} , а эксперимент длится 3 мин, что дает 5400 кадров. Это означает, что для каждого участника в качестве набора для обучения получено 28 временных рядов длиной в 5400 значений.

1.3. Модель на основе LSTM RNN

После получения всех данных для обучения, т. е. кривой когнитивной нагрузки и извлеченных характеристик из видео, необходимо подготовить модель для оценки рабочей нагрузки из черт лица. Поскольку когнитивная нагрузка и исходные лицевые характеристики меняются последовательно, то необходимо в модели учесть динамику их изменений, а не просто получить соответствие между чертами лица и когнитивной нагрузкой; т. е. конкретизировать задачу как задачу прогнозирования временного ряда (когнитивной нагрузки) по множеству временных рядов (извлеченных характеристик).

Одним из мощнейших инструментов для решения таких задач служат рекуррентные нейронные сети (RNN) с «долгой краткосрочной памятью» (LSTM) [13, 15]. Сети RNN отличаются от классических перцептронов несколько иной архитектурой — они содержат в себе элемент памяти. Это делает RNN-сети эффективными для анализа последовательности, поскольку они могут использовать информацию о предыдущих входных данных, тогда как классические нейронные сети работают только с текущим образцом. А применение

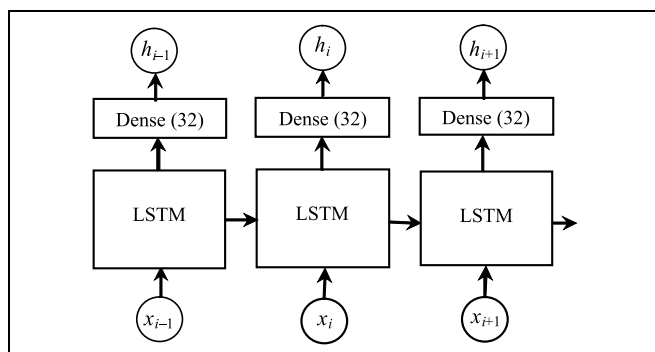


Рис. 5. Сетевая архитектура: сеть состоит из двух слоев: LSTM и Dense layer (32 входа, один выход); X — значения 28-ми входных признаков; h_i — оценка выходной когнитивной нагрузки для каждого шага i

LSTM-модулей позволяет сети запоминать значения как на короткие, так и на длинные промежутки времени. Применение именно этой архитектуры сети позволит учесть в модели динамику изменения когнитивной нагрузки.

Архитектура сети (рис. 5), используемая для оценки когнитивной нагрузки, содержит слой LSTM с 32 выводами, с «выпрямляющей» функцией активации Rectified Linear Unit (ReLU) и с параметром отсева 0,2, необходимым для предотвращения переобучения. За ним следует полносвязный слой с 32 нейронами и одним выходом (когнитивная нагрузка) и с функцией активации ReLU. Входные данные на каждом этапе обучения имеют размерность в 28 параметров (скаляров), а размерность выходных значений — одна скалярная величина в диапазоне значений от 0 до 1.

Показано [1], что выражение лица при когнитивной нагрузке зачастую носит индивидуальный характер. Поэтому для учета индивидуальных особенностей LSTM-модель обучается отдельно для каждого человека. Обучаясь на данных конкретного человека, сеть может «настроиться» на инди-

видуальные особенности мимики, что делает сеть точнее по отношению к обобщенной модели. Кроме того, обучение модели индивидуально позволяет увеличить общую скорость сходимости при обучении.

Модель была реализована на языке Python с помощью системы глубокого обучения Keras [19]. Обучение проводилось в пять эпох методом Adam с размером партии в 64 образца. В качестве функции потерь была выбрана средняя абсолютная ошибка Mean Average Error (MAE). Для обучения модели по данным одного участника требуется около 2 мин. Оценка модели может выполняться более чем с 60 образцами в секунду, что позволяет применять эту систему для оценки когнитивной нагрузки в режиме реального времени.

2. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА

Эксперимент проводился с участием 20-ти испытуемых (15 мужчин и 5 женщин со средним возрастом 20 лет) — студенты 2—4 курсов технических вузов. Все участники имели схожий и достаточный для выполнения задачи уровень компьютерной грамотности. Каждому участнику было предложено сидя за компьютером выполнить всю последовательность задач (щелкая мышью круги в течение 3 мин) перед камерой. Во время теста участникам не разрешалось разговаривать. Сам тест реализован с помощью HTML и JavaScript для удобного запуска в браузере.

В результате эксперимента получен видеофайл и файл-журнал с событиями мыши для каждого участника, на основе которых описанными ранее методами были получены лицевые характеристики, характеристики взгляда, а также кривая когнитивной нагрузки. Когнитивная нагрузка каждого участника показала схожую динамику с общими (усредненными) результатами (рис. 6) ввиду того, что тест сам по себе не предполагал большой ва-

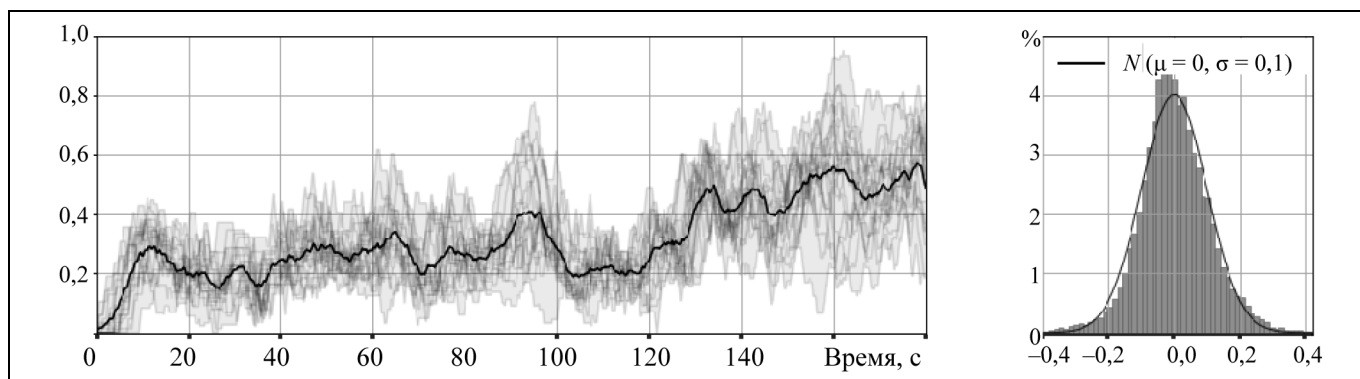


Рис. 6. График когнитивной нагрузки участников в ходе эксперимента (слева) и распределение отклонений от средней нагрузки (справа): отклонения от средней когнитивной нагрузки по участникам описываются нормальным распределением $N(\mu = 0, \sigma = 0,1)$

риативности и был одинаков для всех. При этом отклонения от среднего значения описываются нормальным распределением. Отметим, что на интервале со 100 по 120 с у всех испытуемых наблюдается снижение когнитивной нагрузки, хотя частота задач на данном интервале продолжает линейно расти (см. рис. 2), это связано с проявлением своего рода усталости. Кроме того, в этом интервале совершалось наибольшее число ошибок. Однако подробное рассмотрение данного эффекта лежит за рамками настоящей статьи.

3. МОДЕЛИРОВАНИЕ

Перед обучением моделей входные данные были предварительно обработаны. Интерполированы значения признаков в кадрах, в которых система анализа видео не смогла дать предсказания с большой степенью уверенности. Такие признаки, как поворот и положение головы заданы относительно камеры, но нас лишь интересуют отклонения от начальных положений, поэтому значения для этих признаков были переведены в «относительные» путем вычитания из каждой кривой среднего значения за первые 5 с эксперимента. Также все значения признаков по всем участникам были отмасштабированы, чтобы соответствовать диапазону от 0 до 1. Кроме того, из набора входных данных для обучения также были исключены данные двух испытуемых, для которых система анализа видео предсказывала слишком аномальные значения. Это связано с тем, что в ходе эксперимента эти испытуемые закрывали лицо рукой или же существенно смещались по отношению к камере. Итоговая выборка насчитывала 18 испытуемых, для каждого из которых извлечены значения 28-ми признаков на протяжении 5400 кадров входного видео.

Для сравнения, помимо предложенной модели, моделирование проводилось с помощью многофакторной линейной регрессионной модели. Обучение и валидация моделей проводились тремя способами:

1) обучение и валидация по каждому испытуемому отдельно, где данные отдельного испытуемого случайным образом разбивались на обучающую и проверочную выборку в соотношении 75 к 25%, а метрики для модели усреднялись по всем испытуемым;

2) обучение и валидация по всем испытуемым сразу, где все данные случайным образом разбива-

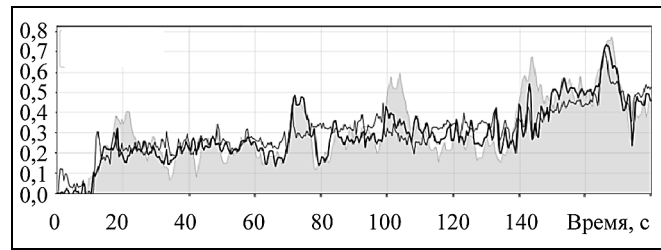


Рис. 7. Предсказания моделей, обученных индивидуально на отдельно взятом испытуемом: — линейная регрессия; — LSTM; — целевая нагрузка

лись на обучающую и проверочную выборку в соотношении 75 к 25%;

3) обучение на всем наборе, где для обучения использовались данные 75% случайно выбранных испытуемых, а валидация проводилась на оставшихся 25% испытуемых (не включенных в обучающую выборку), причем такое разбиение проводилось несколько раз, а значения метрик усреднялись (метод кросс-валидации).

Для оценки качества моделей были выбраны две метрики: средняя абсолютная ошибка (MAE), так как имеет наиболее интуитивную интерпретацию, а также коэффициент детерминации R^2 , показывающий долю дисперсии зависимой переменной, объясняемой рассматриваемой моделью.

Как видно из результатов моделирования, представленных в таблице (жирным выделены лучшие результаты метрик среди моделей), предложенная модель на основе LSTM имеет более высокую точность (рис. 7) по сравнению с обычной регрессионной моделью при обучении на отдельных испытуемых, при этом модель на основе LSTM объясняет большую часть дисперсии зависимой переменной. Однако обе модели имеют плохие (отрицательный коэффициент детерминации) результаты при предсказании когнитивной нагрузки испытуемых, которые совсем не включены в обучающую выборку, т. е. модели не в состоянии должным образом обобщить зависимость между лицевыми параметрами и когнитивной нагрузкой на столь малой выборке, однако в случае отдельного тестирования по испытуемым модели хорошо «настраиваются» на определенного человека и определенную задачу. Связано это, в первую очередь, с тем что каждый участник ведет себя по-разному (т. е. индивидуально) в ходе теста, говоря о мимике

Результаты моделирования

Способ Модель	1) индивидуально		2) по всем испытуемым		3) по части испытуемых	
	MAE	R^2	MAE	R^2	MAE	R^2
Линейная регрессия	0,091 ± 0,013	0,605 ± 0,153	0,118	0,191	0,130 ± 0,013	-0,06 ± 0,13
LSTM	0,056 ± 0,011	0,717 ± 0,166	0,065	0,637	0,148 ± 0,011	-0,34 ± 0,11



и положении тела. Но в случае, когда все испытываемые включены в общую выборку, обе модели в состоянии обобщить известную им информацию, демонстрируя неплохую, пусть и более низкую, точность предсказаний по сравнению с обучением индивидуально.

В дальнейшем для создания обобщенной модели необходимо провести более длительные эксперименты с большим числом участников. Они позволят также снизить влияние факторов переобучения моделей и в полной мере раскрыть возможности моделей глубокого обучения.

ЗАКЛЮЧЕНИЕ

Предложен подход для оценки когнитивной нагрузки оператора, выполняющего интенсивные задачи, требующие когнитивных усилий. Нагрузка оценивается по видеосигналу (например, с веб-камеры), что позволяет легко применять его на практике. Метод оценки разработан в целях учета индивидуальных особенностей путем обучения предложенной модели отдельно для каждого пользователя. В подходе применяются новейшие методы машинного обучения, такие как рекуррентные нейронные сети и методы анализа лицевых характеристик по видеосигналу.

Для проверки метода разработан и проведен специальный эксперимент с участием 20-ти человек. Для обучения модели для каждого участника был проведен детальный анализ задач, выполняемых в ходе эксперимента, на основе модифицированной модели МНР. Данный анализ позволил получить оценку когнитивной нагрузки, которая впоследствии использовалась для обучения рекуррентной нейронной сети.

Предложенная модель на основе рекуррентных сетей с «долгой краткосрочной памятью» LSTM позволяет получить большую точность при индивидуальном анализе когнитивной нагрузки для конкретной задачи по сравнению с регрессионной моделью, так как позволяет учитывать динамику показателей во времени. Однако построение обобщенной модели пока остается еще довольно сложной задачей.

ПРИЛОЖЕНИЕ

Исходный код программ, используемых для изложенного исследования, доступен в Интернете по адресам <https://github.com/bshishov/MentalWorkload> и <https://github.com/bshishov/Emotions>.

ЛИТЕРАТУРА

1. Шишов Б.А., Колесникова А.С., Пулин М.А. Разработка модели психоэмоционального состояния диспетчера АСДУ // Автоматизация, телемеханизация и связь в нефтяной промышленности. — 2017. — № 1. — С. 15–23.

2. Григорьев Л.И., Чернышева О.Н., Кучерявый В.В. Оценка профессионально важных психологических характеристик диспетчера в человеко-машинных системах АСУТП нефтегазового производства // Проблемы управления. — 2017. — № 1. — С. 57–64.
3. Hart S.G., Staveland L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research // Advances in psychology. — 1988. — Vol. 52. — P. 139–183.
4. John B.E., Kieras D.E. The GOMS family of user interface analysis techniques: Comparison and contrast // ACM Transactions on Computer-Human Interaction. — 1996. — Vol. 3, N 4. — P. 320–351.
5. Card S.K., Moran T.P., and Newell A. The model human processor: an engineering model of human performance: In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), Handbook of Perception and Human Performance. Vol. 2: Cognitive Processes and Performance, 1986. — P. 1–35.
6. Anderson J.R. ACT: A simple theory of complex cognition // American Psychologist. — 1996. — Vol. 51, N 4. — P. 355–365.
7. Liu Y., Feyen R., Tsimhoni O. Queueing Network-Model Human Processor (QN-MHP): A computational architecture for multitask performance in human-machine systems // ACM Trans. on Computer-Human Interaction. — 2006. — Vol. 13, N 1. — P. 37–70.
8. Roy R.N., Charbonnier S., Campagne A. and Bonnet S. Efficient mental workload estimation using task-independent EEG features // Journal of Neural Engineering. — 2016. — Vol. 13, N 2. 026019.
9. Baltrušaitis T., Robinson P., Morency L.-P. Constrained local neural fields for robust facial landmark detection in the wild // Proc. of the IEEE Intern. Conf. on Computer Vision Workshops. — 2013. — P. 354–361.
10. Baltrušaitis T., Mahmoud M., Robinson P. Cross-dataset learning and person-specific normalisation for automatic action unit detection. In: 2015 11th IEEE Intern. Conf. and Workshops on Automatic Face and Gesture Recognition (FG). — IEEE, 2015. — Vol. 6. — P. 1–6.
11. Baltrušaitis T., Robinson P., Morency L.P. Openface: an open source facial behavior analysis toolkit. In: 2016 IEEE Winter Conf. on Applications of Computer Vision (WACV). — IEEE, 2016. — P. 1–10.
12. Wood E., Baltrušaitis T., Zhang X., et al. Rendering of eyes for eye-shape registration and gaze estimation // Proc. of the IEEE Intern. Conf. on Computer Vision. — 2015. — P. 3756–3764.
13. Greff K., Srivastava R.K., Koutnik J., et al. LSTM: A search space odyssey // IEEE Trans. on Neural Networks and Learning Systems. — 2017. — Vol. 28, iss. 10. — P. 2222–2232.
14. Ekman P., Rosenberg E.L. (ed.) What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). — Oxford University Press, USA, 1997.
15. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. — 1997. — Vol. 9, N 8. — P. 1735–1780.
16. MacKenzie I.S., Buxton W. Extending Fitts' law to two-dimensional tasks // Proc. of the SIGCHI Conf. on Human Factors in Computing Systems. — ACM, N.-Y., 1992. — P. 219–226.
17. Meyer D.E., Kieras D.E. A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena // Psychological Review. — 1997. — Vol. 104. — N 4. — P. 749–791.
18. May J.G., Kennedy, R.S., Williams, M.C., et al. Eye movement indices of mental workload // Acta Psychologica. — 1990. — Vol. 75, N 1. — P. 75–89.
19. Chollet F. et al. Keras // Github: [сайт]. URL: <https://github.com/keras-team/keras>. — 2015.

Статья представлена к публикации членом редколлегии Н.Н. Бахтадзе.

Шишов Борис Александрович — аспирант, Российский государственный университет нефти и газа (национальный исследовательский университет) им. И.М. Губкина, ✉ shishov.b@gubkin.ru.