

Нейросетевые решения интеллектуального анализа разнородной информации

В. Д. Петухов¹, А. Мирко²

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

¹vova.96@mail.ru, ²antonmirko.ukr@gmail.com

Аннотация. Данная статья является анализом существующих нейросетевых решений для анализа разнородной информации на примере обработки данных со встроенных датчиков планшетного компьютера. Отобрано два метода на базе построения последовательностей нейронных сетей, а также технология гибридной нейронной сети, именуемая как Hybrid deep neural network (HDNN).

Ключевые слова: нейронные сети, разнородные входные данные, гибридные глубокие нейронные сети, анализ разнородной информации

I. ВВЕДЕНИЕ

Мир наполнен информацией, представленной в самых разных формах. Например, в визуальной форме, в форме звука, в форме сигналов. Наиболее полным, представление о любой ситуации, становится при использовании всех доступных источников информации сразу. При этом один из самых непредсказуемых и всесторонних потоков информации формируется при перемещении наблюдателя, или иными словами, сенсоров, в физическом пространстве.

Обширный набор сенсоров, готовых к такой мобильности, присутствует в планшетных компьютерах и смартфонах. Поскольку речь идёт не только о сборе, но и об обработке информации, а фактически, о продолжительных вычислениях в условиях мобильности, разумно отдать предпочтение устройствам с более ёмкой батареей — планшетным компьютерам. С программной точки зрения, в условиях большого потока данных самой разной формы, хорошо себя показал нейросетевой подход.

В настоящее время нейронные сети нашли огромное распространение в различных сферах деятельности человека. Нейронные сети хорошо справляются с задачами распознавания образов, используя источники видео сигнала, в частности, автопилот в автомобилях Tesla, использует для своей работы информацию с видео [1], происходит обработка аудиосигналов, сетями, которые перед этим изучили большое количество аудиосигналов, однако с ростом возможностей современных вычислительных машин, также и повышается сложность задачи, решение которой ожидают от нейронных сетей.

Благодаря быстрому развитию Big Data и появлению высокопроизводительных вычислительных устройств, появилась возможность извлекать гораздо больше информации, и получать новые результаты из этих обширных наборов данных. Для обработки такого объема данных требуется использование глубоких

нейронных сетей, которые отличаются от стандартного машинного обучения с точки зрения эффективности по мере увеличения объема данных. Однако в нашем случае требуются еще более нестандартные решения для обработки разнородных данных [2] и такие решения существуют, и будут рассмотрены далее в данной статье, а также будет проведен анализ.

II. ТРЕБОВАНИЯ К ЗАДАЧЕ

Интеллектуальный анализ информации в условиях мобильности, в практическом смысле, применим к широкому спектру отраслей и процессов, включая логистику, продажи, услуги наземного, воздушного транспорта, судоходство, но не ограничиваясь ими. Спектр применения настолько широк, что задача начинает носить универсальный характер. Интеллектуальный анализ подразумевает выделение из доступной информации ранее неизвестных знаний. Так как информация поступает непрерывным потоком, существует привязка ко времени. То есть, знание поступает в привязке ко времени, образуя, таким образом, событие.

В рассматриваемом данной статьёй примере, приёмником информации является сам планшетный компьютер: его периферийные устройства и сенсоры. Типовые форматы представления данных собраны в табл. 1.

ТАБЛИЦА I Основные источники информации

Тип	Формат	Объем информации
Камера	Двумерный массив чисел (матрица)	Мегабайты
Микрофон	Одномерный массив чисел (вектор)	Килобайты
Гироскоп, Акселерометр	6 чисел (вектор)	Байты
Глонасс, GPS	2 числа (вектор)	Байты

Формат событий подразумевает связку значений вида: тип события, дата и время. А сам формат будет зависеть от реализации избранного алгоритма. При этом, из-за разнородности данных и сенсоров, а также отсутствия чётко заданных, заранее известных, условий применения конкретного алгоритма, отсутствия гарантий качества данных, невозможно ставить требование о надёжности и устойчивости. Качество работы алгоритма, его скорость, надёжность и прочие характеристики будут определяться в индивидуальном порядке и эмпирически.

В итоге задача формализуется следующим образом: на вход, алгоритм получает непрерывный поток информации. На выходе алгоритм отмечает события. Поскольку задача универсальна, то и данные, и методы

их обработки должны быть универсальны. Как и отмечаемые алгоритмом события. При этом сами пользователи алгоритма решат, что считать событием.

III. СУЩЕСТВУЮЩИЕ РЕШЕНИЯ

Благодаря требованию универсальности, и взяв во внимание неопределённость входных данных, а также отсутствие заранее определённых событий-результатов, рассмотрены только нейросетевые алгоритмы, как самые универсальные.

Задача выделения событий представляет собой задачу классификации. Алгоритм определит, на какой класс событий указывает поступившая в указанный момент времени совокупность информации с сенсоров. Разметка данных будет осуществляться пользователем алгоритма: разметка пар вида событие-время его наступления сформирует обучающую выборку. В таком случае, уместно «обучение с учителем».

Существуют доказавшие свою эффективность нейросетевые алгоритмы: многослойные перцептроны, свёрточные нейросети, сети с остаточными связями, рекуррентные сети. Но все они доказали эффективность в рамках конкретных задач в рамках заданного типа данных. При переходе от изображений к звукам, от звуков к сигналам, архитектура нейронных сетей значительно меняется. Должно пройти время, прежде чем специализированные модели, готовые к эффективной обработке разнородных данных станут пригодны к широкому продуктивному использованию на предприятиях. Так, несмотря на наличие единичных исследований и публикаций, подобные архитектуры, по состоянию на март 2022 года, всё ещё не включены в список доступных пред обученных моделей популярных программных библиотек. К примеру, они отсутствуют в рамках библиотеки машинного обучения Keras [3], несмотря на продолжающееся по сей день её развитие [4].

При разработке любого программного обеспечения, необходимо учитывать долгосрочные перспективы эксплуатации и поддержки, не говоря уже о предсказуемости и качества результата. Использование новаторских методов — риск для предприятия. Появляется вопрос: возможно ли объединить существующие эффективные нейросетевые модели для одновременной обработки информации так, чтобы не потерять главные преимущества нейросетевых методов — гибкость и обучаемость? И ответ: да, существует проверенный метод — ансамбли нейронных сетей.

Причина, почему ансамбль (усреднение моделей) работает, и работают лучше, чем отдельно взятая модель в том, что обычно разные модели ошибаются на разных примерах [5]. Где одна модель ошибётся, другая даст правильный ответ. Такой подход уменьшает дисперсию [5]. Обучение таких моделей менее чувствительно к характеристикам набора обучающих данных [5]. Победители соревнований с помощью ансамблей показывают лучшие результаты, в сравнении с отдельными нейросетевыми моделями [6]. Построение ансамбля может быть представлено двумя шагами: выбор сетей-компонентов, и поиск оптимального способа объединить их результаты [7].

Задача каждой нейросети-участника — выделение признаков из данных (feature-extraction). Для выделения признаков, во многих программных модулях, реализующих нейросетевые модели, предусмотрена настройка исключения последнего слоя для получения высокоуровневых признаков из нейросети. В модуле Keras, это реализуется параметром **include_top=False**.

Запуская алгоритмы на планшетном компьютере, придётся работать в условиях мобильности и ограниченности ресурсов: важно минимизировать вычисления и размеры моделей. Обучение типовой сети с миллионами параметров в таких условиях слишком затратно и длительно [5]. Поэтому, сопоставив требования предсказуемости, энергоэффективности, объёмов модели и её сложности, на практике, используют предварительно обученные модели, поставляемые в комплекте с программными модулями, реализующими нейросети (Keras, TensorFlow и другие). Для таких модулей поставляются сводные таблицы, сочетающие информацию о количестве параметров и точности на формальном наборе данных (ImageNet, CIFAR и другие).

Чем больше параметров, тем больше ресурсов потребуются: батарея, память, вычислительные ресурсы. Руководствуясь ограниченностью ресурсов, выбирать модель приходится по соотношению ресурсоёмкости к точности даваемых результатов. Что отображается в показатель «количество операций над параметрами сети за полученную единицу точности».

Для изображений анализ таблицы из официальной документации Keras показал, что наиболее выгодными в своих классах, являются алгоритмы MobileNetV2, EfficientNetB0 и NASNetMobile с показателями количества параметров за единицу точности 4,91E+06, 6,87E+06 и 7,12E+06 соответственно. Набор данных ImageNet.

Для звуков, специализированных предобученных моделей, в комплекте к популярным программным модулям не поставляется. Однако свою эффективность показали методы, работающие не с необработанными данными о звуковой волне (обычный .wav формат), а с графическим их представлением, спектрограммами [8]. Так, обученные для работы с графическими данными модели показывают свою эффективность и при работе со звуками. Например, в одном из решений модель XResnet18 успешно использовалась в качестве экстрактора признаков, после дополнительного обучения (transfer learning) [8]. В стандартной поставке Keras, присутствует модель со сравнимыми характеристиками — ResNet50V2 и показателем количества параметров за единицу точности 33,7E+06, что входит в 10 лучших таблицы.

Для сигналов возможно ограничиться кодированием в «удобный» для восприятия нейросетью формат: векторное представление текущих значений или скользящего окна, затрагивающего некоторое количество предыдущих значений с сенсоров. Финальный классификатор, при необходимости примет их во внимание. Возможно и графическое представление информации с предобученным графическим классификатором, либо текстовое представление в совокупности с предобученной языковой моделью (NLP).

Правильное сочетание таких моделей в единый ансамбль, даст возможность нивелировать ошибки отдельных моделей, давая более точные результаты. Следующий шаг – комбинирование выделенных признаков в единый ответ, формируя, таким образом, ансамбль. Далее, представлены популярные методы формирования ансамблей. Согласно высказыванию одного из Data Scientist из компании Майкрософт, есть следующие подходы, представленные ниже:

- Объединение нескольких нейронных сетей в единую сеть с обновлением слоев, и их весов.
- Использование дифференцирующей операции, для результатов, которые требуется объединить, для дальнейшего принятия решений или классификации.
- Использование технологии Hybrid Deep Neural Network (HDNN).

IV. АНАЛИЗ РЕШЕНИЙ

Ансамбль нейронных сетей – технология, описанная в обзорной статье «Объединение нескольких классификаторов: уроки и следующие шаги» [9], показывает, что использование нескольких классификаторов в большинстве случаев является более точным методом, нежели использование индивидуальных классификаторов. Исходный вариант такого метода предполагает, что ансамбль принимает идентичные данные, и обрабатывает их на различных классификаторах, каждый из которых отвечает за свой класс, и далее объединяется в конечный результат.

Для использования данного метода для обработки разнородной информации, требуется разобраться с двумя основными проблемами в данном решении, а именно:

- вопрос разнородных входных данных;
- вопрос объединителя.

Для обработки разнородной информации, предлагается, что каждый классификатор будет обрабатывать свои исходные данные, как если бы мы объединили различные нейронные сети, однако важным вопросом такого решения будет синхронизация данных так, чтобы различные входные данные обрабатывали одно событие в точный момент времени, как это представлено на рис. 1.

Второй вопрос объединения является тем, что разработчику требуется определить вариант, каким образом будет осуществляться объединение результатов классификаторов.



Рис. 1. Вариант различных входных данных

Среди возможных методов объединения можно выделить три основных:

- Необучаемый – одним из таких методов может являться алгоритм «большинства голосов».
- Обучаемый – является одним из методов, который сочетает в «простоту» и анализ результатов. Благодаря данному методу возможно провести анализ и взвесить результаты, получаемые от классификаторов.
- Мета-классификатор – добавляется дополнительная нейронная сеть, которая является объединителем результатов от исходных классификаторов. Является более сложным среди вариантов объединения результатов.

Опираясь на озвученные во вступительной части статьи требования о работе в условиях мобильности, несмотря на большую сложность, мета-классификатор является единственным применимым методом из представленных, поскольку выделенные предобученными классификаторами признаки не предназначены для дальнейшего использования человеком, или даже алгоритмами: они трудно интерпретируемы человеком, а так же имеют большую размерность представления признаков: для MobileNetV2 это вектор из 1024 элементов на предпоследнем слое [10]. Таким образом, результат работы одной нейросети, может эффективно интерпретироваться только другой нейросетью. Обучение такого объединителя концептуально не будет отличаться от концепции адаптации или переноса знаний (transfer learning).

Следующая технология, которая специально создана для работы с разнородными данными, это технология гибридной нейронной сети (Hybrid Deep Neural Network – HDNN) [11]. Это перспективная технология, которая ещё не поставляется с популярными программными модулями в готовом виде, и приводится здесь для сравнения.

Архитектура Гибридной глубокой нейронной сети представлена на рис. 3, и представляет собой усовершенствованный подход использования ансамбля нейронных сетей и целевой нейронной сети для выполнения обучения.

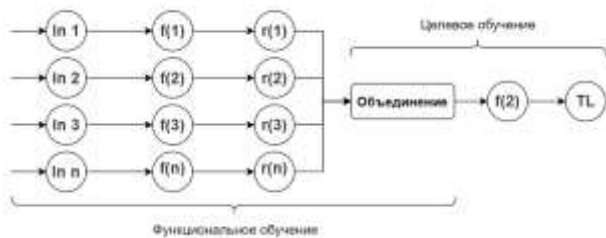


Рис. 2. Архитектура HDNN

На рис. 3 представлены In разнородные входные данные от 1 до N , f нелинейные уравнения, определяемые нейронными сетями для каждого входных данных, r результат каждой отдельной нейронной сети функционального обучения, g нелинейное целевое уравнение для целевых данных, TL целевой результат обработки данных.

В качестве основы для такого подхода, количество входных данных и нейронных сетей не ограничено, в зависимости от задачи разработчика могут быть использованы различные входные данные, такие как аудио, видео, текст или различные изображения. Используемые сети для такого подхода могут быть любыми, начиная от традиционных алгоритмов машинного обучения, такими как метод опорных векторов или другие. Исходя из вышеописанных методов, можно составить сравнительную таблицу (табл. 2), и провести параллели между технологиями, а также вывести их плюсы и минусы.

ТАБЛИЦА II СРАВНЕНИЕ ТЕХНОЛОГИЙ

Критерий	Ансамбль	HDNN
Входные данные	Любые	Любые
Синхронизация событий	Да	Не заложена в изначальной архитектуре
Объединение данных	Да	Да
Обработка после объединения	Не заложена в изначальной архитектуре	Да
Использование различных нейронных сетей	Да	Да
Сложность разработки	Средняя/ Высокая	Высокая
Сложность обучения	Низкая	Высокая
Длительность обучения	Низкая	Высокая
Готовые обученные модели	Да	Нет

В условиях мобильности и ограниченности ресурсов, представляется более разумным использование проверенных и хорошо оптимизированных моделей, идущих в комплекте с популярными нейросетевыми программными модулями, совместно с метаклассификатором в роли обучаемого объединителя,

нежели самостоятельная проработка и обучение HDNN-модели.

V. ВЫВОДЫ

Исходя из результатов, описанных выше, метод чистого ансамбля, на базе предобученных нейросетевых моделей, является предпочтительным при разработке универсальной системы, выделяющей (регистрающей) события, исполняемой в условиях мобильности на планшетном компьютере. Такой ансамбль может эффективно использоваться для выделения событий, появляющихся в конкретный момент времени, так как предполагает в себе предварительную обработку и синхронизацию данных.

Перспективный метод HDNN хорошо подходит для задач, когда данные уже обработаны, и содержат информацию о конкретном событии (задачи), где не требуется синхронизация, но требуется более глубокое обучение для получения результатов обработки разнородных данных.

Также, дальнейшим шагом разработки является изучение и экспериментальный перебор различных архитектур «Объединителей», для отбора лучших при обработке реальных данных.

СПИСОК ЛИТЕРАТУРЫ

- [1] PyTorch at Tesla - Andrej Karpathy, Tesla (2019). Available at: <https://www.youtube.com/watch?v=oBkltKXtDE> (accessed 6 March 2022).
- [2] Iqbal H. Sarker. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. Nature Public Health Emergency Collection, SN Comput Sci. 2021; 2(6): 420 Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8372231/> (Accessed 1 March 2022). DOI: 10.1007/s42979-021-00815-1
- [3] Keras Applications, (2022). Available at: <https://keras.io/api/applications/> (Accessed 15 March 2022).
- [4] Keras Release 2.8.0, (2022). Available at: <https://github.com/keras-team/keras/releases> (Accessed 15 March 2022).
- [5] Jason Brownlee. Ensemble Learning Methods for Deep Learning Neural Networks. August 6, 2019. Available at: <https://machinelearningmastery.com/ensemble-methods-for-deep-learning-neural-networks/> (Accessed 14 March 2022).
- [6] Francois Chollet. Deep Learning With Python, USA, Manning Publications Co., 2017, 264p.
- [7] Chang Shu, Donald H. Burn. Artificial neural network ensembles and their application in pooled flood frequency analysis, Water Resources Research, Volume 40, Issue 9 September 2004. Available at: <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2003WR002816> (Accessed 21 March 2022). DOI: 10.1029/2003WR002816
- [8] David Puiggròs, Comparison between simple audio classification methods, Sep 19, 2021. Available at: <https://medium.com/@puiggròs/comparison-between-simple-audio-classification-methods-f2989e9647a8> (Accessed 22 March 2022).
- [9] T. K. Ho. Multiple Classifier Combination: Lessons and Next Steps. World Scientific Publishing Company; 1st edition (June 15, 2002) - pp.171-198.
- [10] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv:1704.04861v1 [cs.CV] 17 Apr 2017. Available at: <https://arxiv.org/abs/1704.04861> (Accessed 10 March 2022).
- [11] Zhenyu Yuan, Yuxin Jiang, Jingjing Li, Handong Huang. Hybrid-DNNs: Hybrid Deep Neural Networks for Mixed Inputs. (18 May 2020) Available at: <https://arxiv.org/abs/2005.08419> (Accessed 22 March).