

ИСПОЛЬЗОВАНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ПРИ РАСПОЗНАВАНИИ И АНАЛИЗЕ СОСТОЯНИЙ ДИНАМИЧЕСКИХ ОБЪЕКТОВ

Тюшкевич М. В.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: markt93@mail.ru

В данной статье рассмотрены подходы к решению задач распознавания динамических объектов и анализ их состояний. Дополнительной сложностью в распознавания динамических объектов является то, что помимо статических признаков объекта существуют и динамические характеристики. При решении задачи следует так же следить за изменением пространственных данных во времени. В качестве решение было предложено использование рекуррентных нейронных сетей. Для обучения сети применяется алгоритм обратного распространения ошибки.

ВВЕДЕНИЕ

В настоящее время широко развиваются системы распознавание динамических объектов, которые нашли применение в различных отраслях. Для построения современных систем определения динамических объектов характерны следующие основные проблемы и задачи: обнаружение объекта интереса на сложном изменяющемся фоне; измерение пространственных параметров и оценка динамических параметров объектов по наборам последовательных кадров; сопровождение объектов; классификация и распознавание динамических объектов. Для решения такого рода задач широкое применение получили искусственные нейронные сети.

В задаче распознавания динамических объектов входной сигнал, подлежащий распознаванию, рассматривается не в статике, а в динамике и основная информация об образе содержится в траектории изменения входного сигнала во времени. В результате образ получает новое измерение - время, что значительно усложняет решение задачи распознавания. Анализ пространственно-временных данных позволяет выявлять не только статические, но и динамические признаки объектов наблюдения. В этом случае задачу распознавания можно определить как классификацию совокупностей состояний или как классификацию траекторий, решение которой не может быть найдено классическими методами распознавания, т.к. временные переходы могут порождать преобразования изображений, не описываемые известными аналитическими зависимостями[1].

Также наряду с задачей распознавания динамических объектов возникают задачи распознавания активных действий объектов и событий сцены, например, для выявления несанкционированных действий в местах скопления людей или определении жанра сцены для индексации в мультимедийных базах данных. Если рассматривать задачу распознавания объектов и собы-

тий по последовательностям изображений в виде единого процесса, то наиболее целесообразным является системный иерархический подход с элементами параллельной обработки на каждом уровне[2].

1. РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ КАК ДИНАМИЧЕСКАЯ СИСТЕМА

Решение задачи идентификации динамического объекта является одной из ключевых задач при построении нейросетевой системы управления. В качестве идентификатора используются методы теории нейронных сетей. Применительно к динамическим объектам сеть должна обладать архитектурой, позволяющей аппроксимировать временные последовательности. Из теории нейронных сетей известна теорема об универсальной аппроксимации многослойной нейронной сети, согласно которой достаточно применения сети с одним скрытым слоем для построения равномерной аппроксимации с точностью для любого обучающего множества, которое может быть задано входами и желаемыми откликами системы[3]. Дальнейшая работа по созданию модели идентификации динамического объекта опирается на вышеприведенную теорему.

Выбор архитектуры нейронной сети сводится к выбору среди рекуррентных многослойных сетей. Подобные сети можно получить из классической многослойной нейронной сети путем введения обратных связей (см. рис. 1), благодаря которым они имеют собственную динамику, что и отличает их от статических НС. Структура рекуррентной нейронной сети является её основной характеристикой, от которой напрямую зависят все её свойства. Обратные связи могут быть от выхода сети ко входу и от внутреннего слоя сети ко входу. При этом объект может быть рассмотрен в пространстве состояний либо, при отсутствии информации о внутренней структуре, как модель «вход-выход».

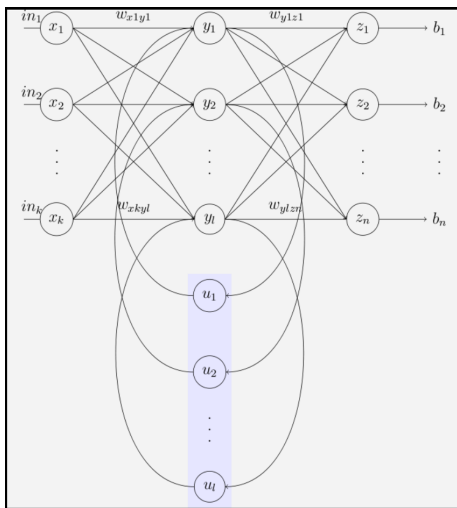


Рис. 1 – Пример простой рекуррентной сети

Необходимым и достаточным условием для устойчивости сети с обратной связью будет наличие единичной задержки в обратной связи. РНС может быть устойчива (не устойчива) в малом и в большом. Устойчивость в большом гарантирует устойчивость во всем пространстве состояний, а устойчивость в малом – лишь в определенных точках. Для устойчивости в большом необходимо подобрать специальную функцию, удовлетворяющую условиям второй теоремы Ляпунова (об асимптотической устойчивости). Для устойчивости в малом необходимо произвести линеаризацию системы в точке равновесия и проверить в ней устойчивость.

II. ОБУЧЕНИЕ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ

Для обучения РНС можно выделить несколько различных по своей сути групп подходов:

- использование известных соотношений между параметрами РНС и её динамикой;
- использование алгоритмов обучения с учителем на базе методов оптимизации по типу алгоритма обратного распространения ошибки;
- использование алгоритма обучения с учителем;
- использование алгоритмов обучения без учителя;
- отсутствие обучения в классическом смысле изменения параметров системы (неявное обучение) – настройка весовых коэффициентов случайными значениями.

Наиболее используемый в настоящее время алгоритм – обучение с учителем на базе методов оп-

тимизации по типу алгоритма обратного распространения ошибки [4]. Он позволяет с достаточной для практического применения точностью оптимизировать веса нейронной сети. В его основе – метод градиентного спуска, когда веса корректируются в соответствии с направлением уменьшения градиента общей для всего обучающего множества ошибки. Он состоит из нескольких основных шагов: инициализации, представления обучающей выборки, прямого прохода расчета сети, обратного прохода и вычисления текущей ошибки.

При инициализации весов используется генерация случайных чисел. После инициализации сети поочередно предъявляются все обучающие пары «вход-выход», предварительно снятые с объекта, подвергаемого идентификации при различных условиях. Для формирования выборки наличие большего количества примеров повышает качество будущей идентификации.

III. ЗАКЛЮЧЕНИЕ

Существует много разновидностей, решений и конструктивных элементов рекуррентных нейронных сетей.

Трудность рекуррентной сети заключается в том, что если учитывать каждый шаг времени, то становится необходимым для каждого шага времени создавать свой слой нейронов, что вызывает серьезные вычислительные сложности. Кроме того многослойные реализации оказываются вычислительно неустойчивыми, так как в них как правило исчезают или сильно перевешивают веса. Если ограничить расчёт фиксированным временным окном, то полученные модели не будут отражать долгосрочных трендов. Различные подходы пытаются усовершенствовать модель исторической памяти и механизм запоминания и забывания.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Дуда, Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт // М.: изд-во «Мир», 1978. – 512 с.
2. Малинецкий, Г. Г. Нелинейная динамика и хаос. Основные понятия: Учеб. пособие / Г. Г. Малинецкий, А. Б. Потапов // М.: Ком-Книга, 2006. – 240 с.
3. Хайкин, С. Нейронные сети: полн. курс / С. Хайкин; Пер. с англ. – 2-е изд., // – М. – СПб. – Киев: ИД «Вильямс», 2006. – 1103 с.
4. Никитин, К. В. Новый подход к применению рекуррентных нейронных сетей для решения задач распознавания образов / К. В. Никитин, Е. Н. Бендерская // Научно-технические ведомости СПбГПУ. – СПб.: Изд-во Политехн. ун-та, – 2007. – № 4-2 (52). – С. 85–92.