ТЕХНОЛОГИЯ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ ЭЛЕМЕНТОВ ТЕКСТА НА ОСНОВЕ НЕЙРОННОЙ СЕТИ ДЛЯ СИСТЕМ ОБУЧЕНИЯ ЕСТЕСТВЕННЫМ ЯЗЫКАМ

Жабборов Канкел Гулмуродович ,Эшонкулов Шерзод Умматович.
Джизакский политехтический институт, Республика Узбекистан

Современный уровень теоретического понимания и практического использования нейронных сетей (НС) в мировой информационной индустрии все явственнее проникает в область педагогических технологий. Учитывая тот факт, что опыт по организации и проведению исследований, связанных с лингвистическими проблемами, решаемыми на основе НС очень скромен, на наш взгляд, для преподавателей иностранных языков будет весьма полезно построить новые нейросетевые технологии классификации элементов теста, способствующие повышению качества адаптивного обучения.

В настоящей работе изложена постановка проблемы визуализации элементов текста (букв, символов, слов, предложений) для автоматического распознавания и классификации в процессах обработки текстовой информации, которые являются ключевым моментом при построении систем познания и обучения иностранным и вообще, естественным языкам. Определены модели распознавания и классификации и основные этапы моделирования процессов визуализации изображений.

Анализ литературных данных показал, что работы в выбранном нами направлении пока не ведутся в большинстве стран СНГ и, в том числе, в нашей республике.

Разработанная нами система позволяет быстро и точно определить разновидности объекта и принадлежность к какому-либо классу, используя их геометрические формы и другие специфические характеристики.

Следует отметить, что существует много способов, которые можно использовать при создании автоматизированной системы классификации и распознавания элементов текста и каждый из них может найти самостоятельное применение. Причем, основным принципом классификации и распознавания изображений является непосредственное сравнение неизвестных объектов с объектами, расположенными в ссылающейся базе данных. В этом случае может быть использован стандартный метод в распознавании изображения, заключающийся в расчете инвариантных особенностей (характеристик) неизвестного объекта для сравнения и классификации.

Традиционный способ извлечения характеристики заключается в вычислении инвариантных особенностей и состоит в проведении сложной предварительной обработки подобно бинаризации или обнаружению кривых. Нами вычислены особенности из полученного двоичного изображения и кривых.

Для предварительной обработки изображение деструктируется на части, которые могут быть использованы при исследовании внутренних особенностей объектов. Для обнаружения кривых используются вычисленные характеристики, которые отображаются в данных серой шкалы.

Извлеченные характеристики объекта позволяют в свою очередь сравнивать неизвестные объекты с объектами в ссылающейся базе данных по паре номеров и исследовать полученные множества различных особенностей для каждого объекта.

Следующей задачей является разработка процедуры сравнения, которая составляет важную часть системы и требует особого внимания при распознавании изображения на основе базы данных и знаний.

Уместно отметить, что между объектами одного и того же класса могут существовать некоторые различия. Это означает, что номера характеристик различных объектов никогда не будут соответствовать точно. В связи с этим степень соответствия номеров проверяется по определенным критериям, например, по критерию "лучшее соответствие".

Теперь рассмотрим модели построения системы классификации и распознавания изображений элементов текста. Для распознавания изображений могут применяться два метода.

Первый - приближенный, основанный на использовании модулей поиска «двойных границ». Для установления наличия таких двойных границ, используется модель «Snake».

Второй метод использует нейронную сеть, т.е. распознающую модель «Paradise», которая позволяет классифицировать объекты по их схожести с заданными особенностями, хранящимися как эталон. Такими особенностями могут являться границы, текстура и др.

Исследования показывают, что НС дает положительные результаты, сравнимые с человеческими показателями, подтверждение чего требует детального описания каждого метода и тестирования с использованием большого набора тестовых изображений.

Теперь опишем основные принципы использования модели, для применения которых проводится тестирование большого набора картинок, полученных из текста, с образцами известных и неизвестных символов и букв.

Модель «Paradise» сеть - нейронная сеть. Нейронная сеть «Paradise» была первоначально разработана для распознавания статических отпечатков пальцев. Но позже модель продемонстрировала свою более общую применимость для классификации объектовразличного характера.

Разработчики нейронной сети FRN (Feature Recognition Network) Хунейн и Кабуна, доказали ее преимущества для распознавания особенностей объекта. Тем не менее, Paradise сеть в отличии от FRN сети может применяться к серым изображением и обучаться как во включенном состоянии так и в выключенном.

Данная модель использует метод, основанный на создании маленьких шаблонов (Pattern Detection Modules), которые предназначены для распознавания важных особенностей объектов.

Распознавание происходит на основе связывания нескольких из этих шаблонов вместе с классификационной средой.

Сеть имеет трехуровневую архитектуру: распознавание «Особенностей» - уровень (FE); определение «Шаблона» - уровень (PD). уровень классификации (C)

Уровень распознавания особенностей (FE). Этот уровень состоит из единичных уровней FE- панелей. Каждая панель определяет конкретный тип и особенность входного изображения.

Для определения символа или буквы используется четыре FE панели. Одна - для определения горизонтальных линий, другая - вертикальных линий и две - для двух частот.

Такие частоты выбираются в результате исследования линий, определенных из изображения. Причем, выбираются те, которые несут наибольшее количество информации для распознавания.

Сегодня современные компьютерные технологии активно внедряются во все сферы жизнедеятельности человека, в том числе в образовании. [4]

Эти шаблоны имеют довольно таки маленький размер. Они также могут использоваться для представления частей многих распознаваемых объектов. Шаблоны используются для обучения Paradise сети. В период обучения шаблоны генерируются автоматически, а структура Paradise сети делает их доступными к малым преобразованиям и деформациям.

Уровень классификации (C). При получении информации об объекте используется все шаблоны, определенные на уровне PD.

Каждый класс объектов представляется классификационной ячейке. Информация об объекте кодируется, используя связи между ячейками классификации. Порезанные объекты представляются в сеть. Существующие классы исследуются на наличие хорошего представления. Если его нет, то создается новый класс.

Параметры Paradise сети. Это число параметров, определяемое сетью, которые устанавливаются для определения типа распознавания.

Они устанавливаются, как правило, на основе эмпирических данных и сильно зависят от приложения.

Это параметры - «классификационные пороги», которые используются для определения оценки схожести между входным объектом и внутренней шаблонной моделью.

«Классификационный порог» рефлектирует переменчивость картинки изображения. Поведение изменения объекта в сети градиентно с изменениями порога.

Литература

1. Электрон журнал. **International Journal of Academic and Applied Research (IJAAR).** Published in Vol. 4 Issue: 10 October 2020 г. http://ijeais.org/wp-content/uploads/2020/10/IJAAR201028.pdf

2. Монахов В.М. Аксиоматический подход к проектированию педагогических технологий. Журн. «Педагогика», 1997.№6.

3. Савурбаев, А., Дангалов, Н. А., Шертойлоков, Г. М., & Эшонкулов, Ш. У. (2014). Алгоритм расчета переходного процесса при ударе цилиндрического кольца о жесткое полупространство. *Молодой ученый*, (8), 246-250.

4. Савурбаев, А., Мухаммадиев, М. Т., Эшанкулов, Ш. У., & Гулиев, А. А. (2015). Косой удар цилиндрического кольца о жесткое полупространство. *Молодой ученый*, (1), 97-102.

5. Жабборов, К. Г., Эгамов, А., Насриддинов, Б. Б., & Жабборова, Б. К. (2016). обоснования параметров грейдера планировщика методом математические планирования. *ученый xxi века*, 14.

6. Жабборов, К. Г. (2020). экономическое-математическое моделирование в экономических исследований. *академическая публицистика*, (8), 49-52.

7. Norqo‘ziyev, Q., Aynakulov, T., Mahkamov, S., & Sodiqov, D. (2024). IMPLEMENTING MOBILE ROBOT WALKING PATH CONTROL AND TRACKING IN C# WINDOWS FORM APPLICATION. International Journal of scientific and Applied Research, 1(3), 95-99.