

Сарманбетов С.А., Хохлов А.А., Кожажулов Е.Т., Ибраимов М.К.

Казахский национальный университет имени аль-Фараби,

Алматы, 050040, Казахстан

*E-mail: sarmanbetov.sanzhar@mail.ru

КЛАССИФИКАЦИЯ СПЕКТРОВ НОРМАЛЬНЫХ ЗВЕЗД ГЛАВНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ НЕЙРОННОЙ СЕТЬЮ ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация. В данной статье рассмотрена задача классификации спектральных классов F и A нормальных звезд главной последовательности и дано краткое описание о методе исследования. Convolutional Neural Network (CNN) были протестированы в классификации на двух общеизвестных наборах данных и достигнуты положительные результаты даже при очень небольшом количестве обучающих образцов в классе. На основе полученных результатов можно утверждать, что CNN модель глубокого обучения является эффективным инструментом в задачах классификации спектров звезд главной последовательности. Использование большого количества обучающих выборок повышают вычислительную точность вне зависимости от сложности формы сигнала. Использование искусственных нейронных сетей показывает преимущества в задачах обработки, классификаций, идентификаций большого массива данных сложных сигналов, таких как спектров звезд главной последовательности и других типов, радиоизлучения Солнца, а также сигналов различных астрофизических объектов. Эффективная обработка большого массива данных, использования меньшей ресурсной базы за короткое время обработки – это все является актуальной задачей современных исследований обработки сложных сигналов астрофизических объектов.

Ключевые слова: Сверточные нейронные сети, глубокое обучение, спектр, звезды главной последовательности

Введение

С момента разработки прибора астрономического наблюдения, можем получить огромное количество спектров. Это явно требует извлечения астрофизической информации из спектров автоматически и точно. Различные методы были введены в спектральной обработке данных [1-4], которая включает в себя извлечение физических параметров и спектральную классификацию.

Нейронные сети (НС) [5] как инструмент классификации данных возникли как альтернатива известным условным методам классификации. Преимущества НС определяются тем, что они могут адаптироваться к данным без какой-либо их спецификации, а также они могут аппроксимировать любую функцию с любой точностью. НС преобразуют данные нелинейно, что делает их гибким инструментом моделирования сложных реальных данных и их преобразований. Кроме того, НС могут оценивать апостериорные вероятности и могут использоваться для создания правил статистической классификации медицинских данных для меди-

цинской диагностики. Нейронные сети могут состоять из произвольного количества нейронов, сгруппированных в один или несколько слоев. Выбор количества нейронов и конфигурации сети зависит от конкретной задачи и может варьироваться от десятков до десятков тысяч в задачах классификации спектров. Изначально, наиболее популярными были нейронные сети на основе самообучающихся карт Кохонена и НС обученные по правилу обратного распространения ошибки. В настоящее время, огромный интерес вызывают сверточные нейронные сети, но их применение ограничивается областью распознавания образов [6]. Основным недостатком нейронных сетей является отсутствие гарантированного положительного результата при обучении, то есть, нет явного указания, как нужно выбирать конфигурационные параметры нейронной сети, чтобы получить хороший классификатор. Однако сложности при обучении компенсируются эффективностью классификации.

Эта работа подходит к проблеме автоматической звездной классификации, соче-

тающая в себе традиционные и нечеткие методы рассуждения.

Цель данной работы — построить модель (классификатор) с помощью сверточных нейронных сетей [7] для классификации спектров нормальных звезд главной последовательности.

Экспериментальные результаты исследования

Для исследования были выбраны два спектральных класса F и A нормальных звезд главной последовательности. При этом в качестве данных были выбраны спектры выбранных объектов, которые были получены на телескопе диаметром 1.93 м обсерватории Observatoire de Haute Provence со спектрографом ($R \sim 42,000$). Проекционная скорость выбранных звезд не превышала 100 км/с. Отношение сигнала/шум спектров составлял ~ 100 . Спектральный диапазон проводимого исследования составляет 4000-7000 Å.

На рисунке 1, 2 представлен пример одного из спектров нормальной звезды с ELODIE. Данные спектры приводились к континууму, т.е. были нормализованы, чтобы в первую очередь избавиться от влияния межзвездного поглощения. Наряду с этим из спектров звезд были убраны особенности, не принадлежащие самому объекту (теллурические линии, космики, межзвездные особенности). Данные процедуры были проделаны с использованием стандартных функций программы IRAF.

В качестве алгоритма классификации был использован CNN (Convolutional neural network). Основной идеей метода является распознавание и классификации изображений [8-15]. Перед реализацией классификации проводился этап обучения CNN. Для этого каждая группа спектров случайным образом была поделена на две равные выборки, 18 из которых использовалась для обучения метода (обучающая выборка), а вторая (тестовая выборка) – для классификации. После этого проводилась классификация тестовой выборки [16-21]. Классификация проведена попарно, и в каждом случае найден класс звезд. Полученные результаты

представлены в Таблице 1.

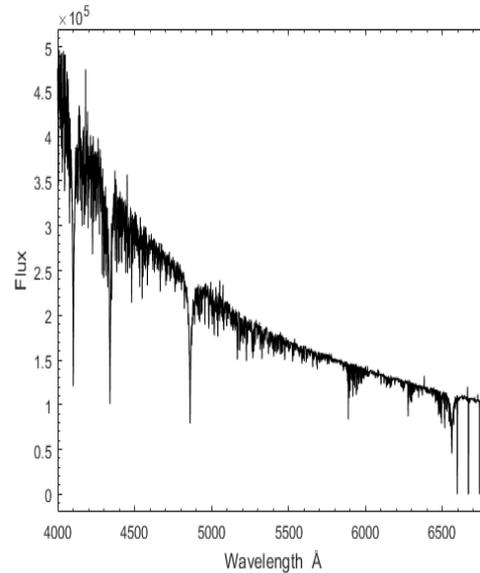


Рисунок 1 – Спектр звезды HD 210210 со спектрографа ELODIE. Ось Y – интенсивность, Ось X – длина волны в Ангстремах.

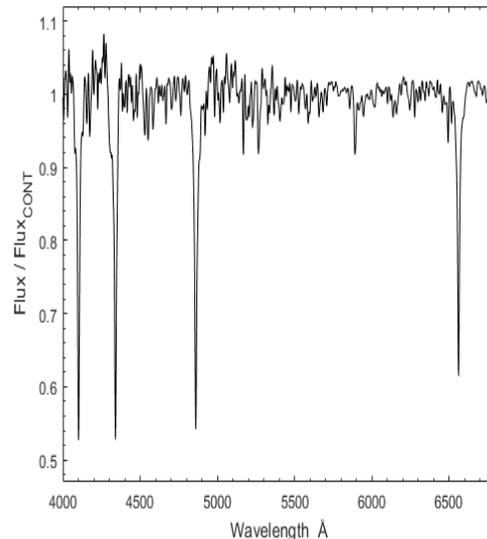


Рисунок 2 – Спектр звезды HD 210210 со спектрографа ELODIE приведенный к континууму.

В данной работе, аналогично исследовались 2 (Рис.1, Рис.2) группы спектров звезд общим количеством 20, количество спектров в каждой группе одинаково. В первую группу (группа 1) вошли спектры класса O. Во вторую группу (группа 2) вошли спектры класса B.

Таблица 1. Полученные результаты классификации исследуемых групп

True test1: 10 False test1: 10 True test2: 10 False test2: 10 4.675081253051758	True test1: 10 False test1: 0 True test2: 10 False test2: 10 4.6085426807403564	True test1: 10 False test1: 10 True test2: 0 False test2: 10 4.563679218292236
Пример 1	Пример 2	Пример 3
True test1: 10 False test1: 0 True test2: 0 False test2: 10 4.47825813293457	True test1: 0 False test1: 10 True test2: 10 False test2: 0 4.497290849685669	True test1: 10 False test1: 0 True test2: 0 False test2: 0 4.425152063369751
Пример 4	Пример 5	Пример 6

Основным показателем качества обучения служит первичное тестирование. Именно, первоначальное тестирования отражает то, насколько хорошо обученная нейронная сеть обобщает входные данные и не совершает ошибочных классификаций.

Для реализации данной задачи было выбрано следующее программное обеспечение:

1. язык программирования Python, на котором создано множество библиотек для работы с данными и нейронными сетями;
2. библиотека «Keras» [22], которая позволяет имплементировать CNN с помощью «TensorFlow» [23] на более высоком программном уровне.

Для обучения модели были выбраны следующие параметры:

1. Метрика для ошибки: «categorical_crossentropy». Используется, когда желательна вероятностная интерпретация оценок. Данная метрика измеряет сходство между истинным классом и прогнозируемым.
2. Оптимизатор: «adam». Алгоритм сочетает в себе идею накопления движения градиента и идею более слабого обновления весов для типичных признаков.
3. Метрика для оценки точности классификатора: «F-мера». Для многоклассовой классификации выделяют два подхода расчета F-меры: micro-усреднение и macro-усреднение. При micro-усреднении значения

матрицы несоответствий усредняются по всем классам, а затем вычисляется F-мера. При macro-усреднении сначала вычисляется F-мера для каждого класса, а затем результаты усредняются по всем классам. Все исходные данные были разделены на обучающую (80 %) и тестовую выборку (100 %). Обучение прекращалось, когда ошибка на тестовой выборке начинала расти, а на обучающей продолжала падать (метод ранней остановки). Обучение модели проходило на процессоре core i7 с видеокартой NVIDIA 1080. Графический процессор этой видеокарты поддерживает параллельные вычисления (CUDA) [24], что сокращает время обучения в несколько раз по сравнению с вычислениями на CPU. 10 эпох длилась 4,6 секунды (рисунок 3).

```
import os
import numpy as np
from keras.preprocessing.image import ImageDataGenerator, load_img, img_to_array
from keras.models import Sequential, load_model
import time
start_time = time.time()

img_width, img_height = 250, 250
model_path = './models/model.h5'
model_weights_path = './models/weights.h5'
model = load_model(model_path)
model.load_weights(model_weights_path)

def predict(file):
    x = load_img(file, target_size=(img_width, img_height))
    x = img_to_array(x)
    x = x / 255.
    x = x.reshape((1, img_height, img_width, 3))
    prediction = model.predict(x)
    return prediction

# Example usage
file = 'test1.jpg'
prediction = predict(file)
print(prediction)
```

Рисунок 3 – Обучение сети

Модель CNN была проверена на контрольных примерах, неиспользованных при обучении.

В Таблице 1 приведены 6 примеров, в примере 1 были выбраны в обеих папках спектры двух типов, сеть разделила на четыре: True test1, False test1, True test2, False test2. True уточняет что спектр принадлежит классу, а False отрицает. В остальных примерах сделаны эксперименты как видно в Таблице 1. Поскольку вход - это не один объект, показатель успеха для одного класса

приблизительно 99% и 0% для оставшегося класса.

Заключение

В большинстве приложений глубокого обучения CNN [25] используются для огромного количества обучающих примеров для изучения мощных классификаторов. В этой статье обрабатывались ограниченные обучающие образцы. CNN были протестированы в классификации на двух популярных наборах данных и достигнуты положительные результаты даже при очень небольшом количестве обучающих образцов в классе. Наша работа показывает, что CNN модели глубокого обучения могут применены к проблемам классификации с помощью нескольких обучающих выборок, не теряя своего обобщения, если сеть хорошо спроектирована.

Литература

1. S. Kumar, J. Ghosh, M.M. Crawford, Best-bases feature extraction algorithms for classification of hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 39 (7) (2001) 1368–1379.
2. D. Manolakis, D. Mardon, G.A. Shaw, Hyperspectral image processing for automatic target detection applications, *Lincoln Lab. J.* 14 (1) (2003) 79–115.
3. J.M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N.M. Nasrabadi, J. Chanussot, Hyperspectral remote sensing data analysis and future challenges, *IEEE Geosci. Remote Sens. Mag.* 1 (2) (2013) 6–36.
4. Q. Du, L. Zhang, B. Zhang, X. Tong, P. Du, J. Chanussot, Foreword to the special issue on hyperspectral remote sensing: theory, methods, and applications, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 6 (2) (2013) 459–465.
5. X. Li, R. Rakkiyappan, G. Velmurugan, Dissipativity analysis of memristor-based complex-valued neural networks with time-varying delays, *Inform. Sci.* 294 (2015) 645–665.
6. N. Younan, S. Aksoy, R. King, Foreword to the special issue on pattern recognition in remote sensing, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 5 (5) (2012) 1331–1334.
7. ABritz D. Implementing a CNN for Text Classification in TensorFlow, 2015
8. G.F. Hughes, On the mean accuracy of statistical pattern recognizers, *IEEE Trans. Inf. Theory* 14 (1) (1968) 55–63.
9. S. Kumar, J. Ghosh, M.M. Crawford, Best-bases feature extraction algorithms for classification of hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 39 (7) (2001) 1368–1379.
10. L.O. Jimenez-Rodriguez, E. Arzuaga-Cruz, M. Velez-Reyes, Unsupervised linear feature-extraction methods and their effects in the classification of high-dimensional data, *IEEE Trans. Geosci. Remote Sens.* 45 (2) (2007) 469–483.
11. A. Agarwal, T. El-Ghazawi, H. El-Askary, J. Le-Moigne, Efficient hierarchical PCA dimension reduction for hyperspectral imagery, in: *IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 353–356.
12. M. Fauvel, J. Chanussot, J.A. Benediktsson, Kernel principal component analysis for the classification of hyperspectral remote sensing data over urban areas, *EURASIP J. Adv. Signal Process.* 2009 (2009) 1–14.
13. J. Wang, C.-I. Chang, Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis, *IEEE Trans. Geosci. Remote Sens.* 44 (6) (2006) 1586–1600, <http://dx.doi.org/10.1109/TGRS.2005.863297>.
14. C. Xu, C. Lu, J. Gao, W. Zheng, T. Wang, S. Yan, Discriminative analysis for symmetric positive definite matrices on lie groups, *IEEE Trans. Circuits Syst. Video Technol.* 25 (10) (2015) 1576–1585.
15. B.S. Serpico, L. Bruzzone, A new search algorithm for feature selection in hyperspectral remote sensing images, *IEEE Trans. Geosci. Remote Sens.* 39 (7) (2001) 1360–1367.
16. I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
17. C.-I. Chang, S. Wang, Constrained band selection for hyperspectral imagery, *IEEE Trans. Geosci. Remote Sens.* 44 (6) (2006) 1575–1585.
18. JS. Jia, G. Tang, J. Zhu, Q. Li, A novel ranking-based clustering approach for hyperspectral band selection, *IEEE Trans. Geosci. Remote Sens.* 54(1) (Jan. 2016), <http://dx.doi.org/10.1109/TGRS.2015.2450759>.

19. Y. Qian, F. Yao, S. Jia, Band selection for hyperspectral imagery using affinity propagation, *IET Comput. Vis.* 3 (4) (2009) 213–222, <http://dx.doi.org/10.1049/iet-cvi.2009.0034>.
20. S. Jia, Z. Ji, Y. Qian, L. Shen, Unsupervised band selection for hyperspectral imagery classification without manual band removal, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 5 (2) (2012) 531–543.
21. M. Fauvel, Y. Tarabalka, J.A. Benediktsson, J. Chanussot, J.C. Tilton, Advances in spectral-spatial classification of hyperspectral images, *Proc. IEEE* 101 (3) (2013) 652–675
22. Keras: The Python Deep Learning library [Электронный ресурс] / Официальный сайт библиотеки «Keras». — Режим доступа : <https://www.keras.io/> (дата обращения : 23.04.17).
23. About TensorFlow [Электронный ресурс] / Официальный сайт фреймворка «Tensorflow». — Режим доступа : <https://www.tensorflow.org/> (дата обращения : 23.04.17).
24. Параллельные вычисления от NVIDIA [Электронный ресурс] / NVIDIA. — Режим доступа : <https://www.nvidia.ru/object/cuda-parallel-computing-ru.html> (дата обращения : 23.04.17).
25. Corinna Cortes and Vladimir N. Vapnik. Support vector networks. *Machine Learning*, 20, 1995, pp. 1

Принято в печать 04.04.2019

Сарманбетов С.А., Хохлов А.А., Кожугулов Е.Т., Ибраимов М.К.

Казахский национальный университет имени аль-Фараби,

Алматы, 050040, Казахстан

**E-mail: sarmanbetov.sanzhar@mail.ru*

КЛАССИФИКАЦИЯ СПЕКТРОВ НОРМАЛЬНЫХ ЗВЕЗД ГЛАВНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ НЕЙРОННОЙ СЕТЬЮ ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация. В данной статье рассмотрена задача классификации спектральных классов F и A нормальных звезд главной последовательности и дано краткое описание о методе исследования. Convolutional Neural Network (CNN) были протестированы в классификации на двух общеизвестных наборах данных и достигнуты положительные результаты даже при очень небольшом количестве обучающих образцов в классе. На основе полученных результатов можно утверждать, что CNN модель глубокого обучения является эффективным инструментом в задачах классификации спектров звезд главной последовательности. Использование большого количества обучающих выборок повышают вычислительную точность вне зависимости от сложности формы сигнала. Использование искусственных нейронных сетей показывает преимущества в задачах обработки, классификаций, идентификаций большого массива данных сложных сигналов, таких как спектров звезд главной последовательности и других типов, радиоизлучения Солнца, а также сигналов различных астрофизических объектов. Эффективная обработка большого массива данных, использования меньшей ресурсной базы за короткое время обработки – это все является актуальной задачей современных исследований обработки сложных сигналов астрофизических объектов.

Ключевые слова: Сверточные нейронные сети, глубокое обучение, спектр, звезды главной последовательности

Сарманбетов С.А., Хохлов А.А., Кожажулов Е.Т., Ибраимов М.К.

Казахский национальный университет имени аль-Фараби,

Алматы, 050040, Казахстан

**E-mail: sarmanbetov.sanzhar@gmail.com*

БАС ТІЗБЕКТЕГІ ҚАЛЫПТЫ ЖҰЛДЫЗДАРДЫҢ СПЕКТРЛЕРІНІҢ КЛАССИФИКАЦИЯСЫН НЕЙРОНДЫ ЖҮЙЕЛЕР АРҚЫЛЫ ТЕРЕҢ ЗЕРТТЕУ

Аннотация. Бұл мақалада қалыпты негізгі жүйелі жұлдыздардың спектрлік класстарын F, A классификациялау мәселесі қарастырылады және зерттеу әдісінің қысқаша сипаттамасы келтірілген. Оралған нейрондық желі (CNN) екі танымал деректер жиынтығымен сынақтан өткізілді және класстағы оқу үлгілерінің өте аз санымен оң нәтижелерге қол жеткізді. Алынған нәтижелер негізінде CNN терең оқу үлгісі негізгі жүйеленген жұлдыздардың классификациялық мәселелеріндегі тиімді құрал болып табылады деп айтуға болады. Оқу үлгілерінің көп мөлшерін пайдалану толқындардың күрделілігіне қарамастан есептік дәлдігін арттырады. Жасанды нейрондық желілерді пайдалану, негізгі тапсырмаларды өңдеуде, классификациялауда, негізгі жүйелі жұлдыздардың спектрі сияқты күрделі сигналдардың үлкен деректер жиынтығын анықтайтын артықшылықтарын көрсетеді. Шағын дерек қорларын тиімді өңдеу, қысқа өңдеу уақытында ресурстық базаны қолдану, заманауи зерттеулердің, оның ішінде астрофизикалық сигналдарды өңдеудің өзекті міндеті болып табылады.

Түйін сөздер: Оралған нейрондық желілер, терең оқыту, спектр, бас тізбек жұлдыздары

S.A. Sarmanbetov, A.A. Khokhlov, E.T. Kozhagulov, M.K. Ibraimov

al-Farabi Kazakh National university,

Almaty, 050040, Kazakhstan

**E-mail: sarmanbetov.sanzhar@gmail.com*

CLASSIFICATION OF SPECTRA OF NORMAL STARS OF THE MAIN SEQUENCE OF THE NEURAL NETWORK, DEEP LEARNING

Abstract: This article discusses the problem of classifying spectral classes F and A of normal main sequence stars and gives a brief description of the research method. The convolutional Neural Network (CNN) was tested in a classification on two well-known data sets and achieved positive results even with a very small number of teaching samples in the classroom. On the basis of the obtained results, it can be argued that the CNN deep learning model is an effective tool in the classification problems of the main sequence stars. The use of a large number of training samples increases computational accuracy regardless of the complexity of the waveform. The use of artificial neural networks shows its advantages precisely in processing tasks, classifications, identifying large data sets of complex signals, such as spectra of main sequence stars, etc.. Effective processing of large data sets, using a smaller resource base in a short processing time, is an urgent task of modern research, including the processing of astrophysical signals.

Key words: Convolutional neural networks, deep learning, spectrum, main sequence stars