

УДК 004.932.2

РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ СИСТЕМ НЕЙРО-НЕЧЕТКОЙ КЛАССИФИКАЦИИ К ЗАДАЧЕ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ АНАЛИЗОВ МОКРОТЫ, ОКРАШЕННЫХ ПО МЕТОДУ ЦИЛЯ – НИЛЬСЕНА

^{1,2}Шеломенцева И.Г., ²Ченцов С.В., ¹Наркевич А.Н.

¹Красноярский государственный медицинский университет
им. проф. В.Ф. Войно-Ясенецкого, Красноярск;

²Сибирский федеральный университет, Красноярск, e-mail: inga.shell@yandex.ru

Поиск оптимальной модели искусственного интеллекта для решения задач медицинской диагностики является, как правило, длительным процессом. В данной статье приведены результаты исследования применимости систем нейро-нечеткого вывода ANFis и NEFClass в качестве классификаторов для задачи распознавания изображений анализов мокроты, окрашенных по методу Циля – Нильсена. Система ANFis представлена тремя функциями и рассматривалась на примере своей реализации в программе Matlab. Система NEFClass рассматривалась на примере реализации в программе NefClass-PC. Был проведен сравнительный анализ алгоритмов обучения рассматриваемых нейро-нечетких систем по критериям величин среднеквадратичной ошибки, точности классификации и частоты ошибок. Признаковое описание экспериментальной выборки составлено из характеристик цвета и формы. В работе приведены результаты работы классификаторов при различных наборах параметров вектора признаков. Полученные результаты находятся в диапазоне достаточных требований к своему качеству и свидетельствуют о применимости нейро-нечетких моделей классификации к задаче диагностики туберкулеза. Сделано заключение о том, что использование моделей нейро-нечеткой логики может стать основой для построения оптимального классификатора для задачи распознавания изображений анализов мокроты, окрашенных по методу Циля – Нильсена.

Ключевые слова: метод Циля – Нильсена, распознавание образов, нейро-нечеткие сети, ANFis, NEFClass, среднеквадратичная ошибка, частота ошибок

THE RESULTS OF THE APPLICATION OF NEURO-FUZZY SYSTEM FOR DETECTION AND CLASSIFICATION OF ZIEHL-NIELSEN STAINED IMAGES OF SPUTUM SMEAR SAMPLES

^{1,2}Shelomentseva I.G., ²Chentsov S.V., ¹Narkevich A.N.

¹Krasnoyarsk State Medical University named after professor V.F. Voino-Yasenetsky, Krasnoyarsk;

²Siberian Federal University, Krasnoyarsk, e-mail: inga.shell@yandex.ru

Finding an optimal model of artificial intelligence for valid diagnosis is mostly a prolonged process. This article presents the results of experiment for using neuro-fuzzy system as classifiers for image recognition of sputum smear samples obtained using a light microscope. The ANFis system represent by three models. We use the Matlab implementation of the ANFis system and the NefClass-PC implementation of the NEFClass system. The comparison of these algorithms has been performed using values of the mean-square error, classification accuracy and error rate as criteria of the quality of classification. The extracted features are described by color and shape characteristics. The paper presents the results of classification for different sets of parameters of the feature vector. The results indicate the effectiveness of using the neuro-fuzzy classification models for diagnosing tuberculosis. The proposed system can be used for building an optimal intelligent system for detection and classification of Ziehl-Nielsen stained images.

Keywords: Ziehl-Nielsen, image processing, neuro-fuzzy system, ANFis, NEFClass, mean-square error, error rate

В настоящее время искусственные нейронные сети успешно применяются в решении биомедицинских проблем – они используются в таких прикладных областях медицины, как функционирование диагностических систем, биохимический анализ, анализ изображений и разработка лекарств.

Среди моделей искусственных нейронных сетей отдельно выделяются нейро-нечеткие сети (ННС), характеризующиеся гибридизацией методологий нечеткой логики и классических нейронных сетей. Диапазон использования нейро-нечетких моделей достаточно широк – прогнозирование, нечеткое управление и классификация. Как

правило, системы на базе нечеткого вывода дают лучшие результаты в случае неопределенности и неполноты входных данных или качественных критериев. Согласно крупнейшей библиографической базе статей по медицинским наукам Medline, ННС используются в качестве классификаторов автоматизированных медицинских систем для диагностики опухоли головного мозга, рака молочной железы, остеопороза, глаукомы и т.д.

В обзорной статье, опубликованной в широко известном журнале «Tuberculosis» издательства Elsevier и посвященной использованию разнообразных моделей нейронных

сетей для диагностики туберкулеза, затронуты проблемы использования и других методов интеллектуального анализа – нейро-нечетких сетей, генетических алгоритмов и средств имитационного моделирования как «перспективных современных технологических инструментов для борьбы с глобальной угрозой туберкулеза» [1].

Проведенное исследование было направлено на оценку эффективности использования нейро-нечетких моделей для задачи диагностики туберкулеза методом классификации изображений микроскопии препаратов, подготовленных по методу Циля – Нильсена.

Материалы и методы исследования

Медицинские данные были получены у пациентов туберкулезного диспансера и представляют собой цветные изображения, полученные при помощи тринокулярного микроскопа Микромед 1 вар. 3–20 при увеличении 10x60.

В основе функционирования нейро-нечетких систем лежат базовые алгоритмы и методы нейронных сетей, только используются они для поиска параметров нечеткой системы, а именно базы правил, параметров фаззификации и дефаззификации. Сама нейро-нечеткая система ориентирована на аппроксимацию неизвестной n -мерной функции, которая частично определяется данными обучения. Нечеткие правила, закодированные внутри системы, представляют собой нечеткие образцы и нечеткие прототипы данных обучения [2].

В качестве базовых нейро-нечетких систем были рассмотрены системы ANFis и NEFClass, базирующиеся на диаметрально противоположных принципах построения и используемых алгоритмов.

NEFClass

Одна из наиболее популярных систем нейро-нечеткой классификации – это систе-

ма NEFClass, основанная на обобщенной архитектуре нечеткого перцептрона. Система NEFClass используется для построения нечетких правил из набора данных, обладающих свойством неполноты измерений, которые необходимо разделить на некоторое количество четких непересекающихся классов [3].

Система NEFClass имеет трехслойную архитектуру, состоящую из входного слоя, скрытого слоя с нечеткими правилами и выходного слоя и построена на основе треугольных функций принадлежности. Активация нейронов происходит во втором и третьем слоях по формулам (1) и (2) [4].

$$a_R^{(p)} = \min_{x \in U_1} \{W(x, R)(a_x^{(p)})\}, \quad (1)$$

$$a_c^{(p)} = \min_{R \in U_2} \{W(c, R)(a_R^{(p)})\}, \quad (2)$$

где $W(x, R)$ – нечеткий вес соединения входного нейрона x с нейроном правила R , а $W(R, c)$ – нечеткий вес соединения нейрона правила R с нейроном выходного слоя.

Вычислительный эксперимент строился на основе программы NEFClass-PC, которая является свободно распространяемым программным обеспечением. Первичные параметры работы этой системы представлены в табл. 1. Подбор параметров был осуществлен экспериментальным путем.

Система ANFis

Система ANFis (adaptive neuro-fuzzy inference system, адаптивная нейро-нечеткая система) предназначена для моделирования различных типов систем управления, экспертных систем и других сложных систем экономики, сельского хозяйства, медицины и социальных наук. Она является наиболее популярной нейро-нечеткой системой вследствие гибкости, простоты и легкости в понимании основ ее функционирования [5].

Таблица 1

Значение параметров для работы программы NEFClass-PC

Параметр	Значение
Алгоритм генерации правил	Лучший для класса
Алгоритм обучения	Классический
Количество термов для каждого признака	Индивидуально
Функция агрегации	Взвешенная сумма
Максимальное количество эпох	500
Значение параметра сигма	0,01
Критерий останова	Количество эпох, при которых ошибка обучения не уменьшается = 100

Система ANFis построена на базе продукционных правил If-Then (Если – То), которые формируют базу нечетких правил выбранной модели. Данная база может быть создана вручную или автоматически сгенерирована системой. В качестве моделей системы выступают три функции – Genfis 1, Genfis 2 и Genfis 3. Все три модели основаны на структуре типа Сугено [6].

$$Z = \frac{(\mu_{A_1}(x)\mu_{B_1}(y)(a_1x + b_1y) + (\mu_{A_2}(x)\mu_{B_2}(y)(a_2x + b_2y))}{\mu_{A_1}(x)\mu_{B_1}(y) + \mu_{A_2}(x)\mu_{B_2}(y)}, \tag{3}$$

где $\mu_{A_i}(x)$ и $\mu_{B_i}(y)$ – функции принадлежности, которые могут быть гауссианами (формулы (4) и (5)). Основная задача нейронной сети – подобрать параметры γ (формула (6)).

$$\mu_{A_i}(x) = \exp\left[-\left(\frac{x - \bar{x}_i^A}{\delta_i^A}\right)^2\right], \tag{4}$$

$$\mu_{B_i}(y) = \exp\left[-\left(\frac{y - \bar{y}_i^B}{\delta_i^B}\right)^2\right], i = \overline{1, 2}, \tag{5}$$

$$\gamma = \{\bar{x}_1^A, \bar{x}_2^A, \bar{y}_1^B, \bar{y}_2^B, \delta_1^A, \delta_2^A, \delta_1^B, \delta_2^B, a_1, a_2, b_1, b_2\}. \tag{6}$$

Genfis 1 генерирует нейро-нечеткую структуру (FIS) классического типа – с использованием функций принадлежности, где количество правил зависит от количества входов и числа термов. Genfis 2 генерирует FIS, основываясь на методе субтрактивной кластеризации, а Genfis 3 – генерирует нечеткую систему, основываясь на классическом алгоритме кластеризации метода К-средних. Параметры работы моделей Genfis 1 и Genfis 3 приведены в табл. 2 и 3. Подбор параметров также был осуществлен экспериментальным путем.

Вычислительный эксперимент

Перед процедурой классификации исходные изображения были обработаны при помощи операций фильтрации и сегментации. Оптимальными алгоритмами и методами фильтрации и сегментации при этом являются фильтрация при помощи свертки с ядром Гонсалеса 3x3 и сегментация методами FAST и Mexican hat [7, 8]. Результаты предобработки изображений анализов мокроты пациентов специализированного учреждения были представлены в виде текстовых файлов с расширениями .dat и .csv.

Дальнейшие исследования проводились в области построения оптимального вектора признаков по параметрам размерности и информативности в рамках решаемой задачи. В результате были отобраны в качестве классификаторов признаки цвета и формы участков интереса (ROI) – числовые значения параметров цвета RGB, HSV, значения площади и компактности [9].

Для оценивания результатов вычислительного эксперимента применения различных моделей и систем ННС к решаемой задаче были выбраны критерий RMSE (квадратный корень из среднеквадратичного отклонения) и критерий точности и частоты ошибок из-за ограничений программы NEFClass-PC [10]. Под частотой ошибки подразумевается величина, равная отношению суммы критериев FP (количество случаев ложноположительной классификации) и FN (количество случаев ложноотрицательной классификации) к общему размеру

Таблица 2

Значение параметров для работы функции Genfis 1

Параметр	Значение
Количество термов входных переменных	2–3
Тип функции принадлежности	gaussmf – гауссиана
Тип выходной функции	constant – константа

Таблица 3

Значение параметров для работы функции Genfis 3

Параметр	Значение
Количество кластеров	2 – 3
Степень перекрытия кластеров	2
Максимальное количество операций	200
Минимальное приращение функции	1e-5

исследуемой выборки. Под точностью подразумевается доля правильных ответов алгоритма классификации.

Для проверки качества классификаторов из первичного набора, в связи с ограничениями ННС на количество обрабатываемых данных, генерировалась выборка случайным образом. На каждый классификатор было проведено не менее 10 испытаний. Среднее значение результатов проведенных исследований представлено на рисунке и табл. 4 и 5, где отражена зависимость выбранного параметра сравнения от набора признаков.

Результаты исследования и их обсуждение

Все три модели системы ANFis классифицируют исходный набор данных, сформированный на основе признаков цвета и признаков формы, при этом критерий ошибочной классификации, вне зависимости от размерности вектора признаков и количества записей, не выходит за пределы диапазона значений 9,9–17,8% (рисунок).

Минимальное значение RMSE, равное 0,27578, система ANFis показала на

наборе данных из средних значений цветовых характеристик RGB и HSV, компактности и площадь с использованием модели genfis1 (табл. 4). Максимальное значение точности, равное 85,94%, система NefClass показала на том же наборе данных, как и модель genfis1 системы ANFis – 90,1% (табл. 5).

Исследования проводились на персональном компьютере со следующими характеристиками: процессор Intel® Core™ i7, 2,70 GHz, оперативная память 10,00 ГБ, 64-разрядная операционная система Microsoft Windows 7. При этом часть вычислений не давала результатов по причине переполнения памяти из-за проблемы проклятия размерности (табл. 4–5).

Выводы

Применение нейро-нечетких моделей классификации к задаче диагностики туберкулеза при помощи микроскопии по методу Циля – Нильсена может обеспечить результаты, которые могут соответствовать качеству тех, которые получены с помощью других методов, таких как дискриминационный анализ или нейронные сети.

Таблица 4

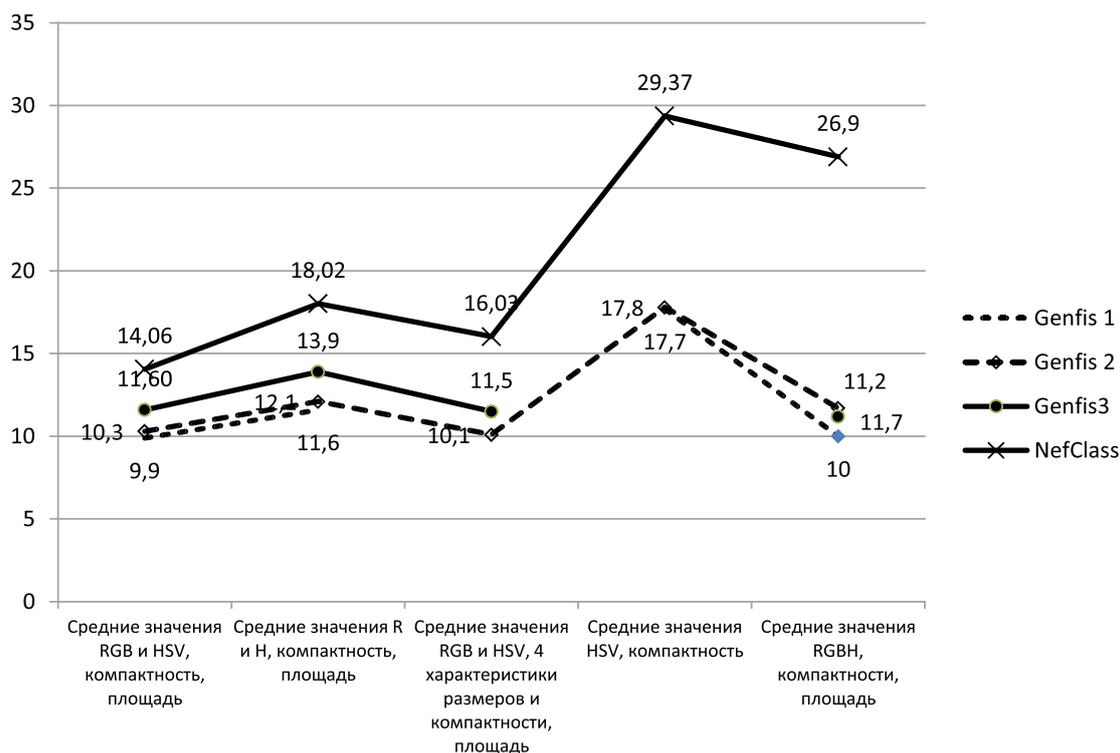
Сравнение моделей ННС по параметру RMSE

	Genfis 1	Genfis 2	Genfis3	NefClass
Средние значения HSV, максимальное значение HSV, минимальные значения HSV, компактность, площадь	Out of memory	0,29	Out of memory	0,4569
Средние значения HSV, максимальное значение H, минимальные значения SV, компактность, площадь	Out of memory	0,28586	Out of memory	0,379962
Средние значения RGB и HSV, компактность, площадь	0,27578	0,277	0,29741	0,2344
Средние значения R и H, компактность, площадь	0,30067	0,30338	0,334	0,3579
Средние значения RGB и HSV, 4 характеристики размеров и компактности, площадь	Out of memory	0,276	0,29599	0,25
Средние значения HSV, компактность	0,35063	0,35352	Out of memory	0,4029
Средние значения RGBH, компактности, площадь	0,27147	0,29208	0,29782	0,4088

Таблица 5

Сравнение моделей ННС по параметру точности классификации

	Genfis 1	Genfis 2	Genfis3	NefClass
Средние значения HSV, максимальное значение HSV, минимальные значения HSV, компактность, площадь	Out of memory	88,9%	Out of memory	79,87%
Средние значения HSV, максимальное значение H, минимальные значения SV, компактность, площадь	89%	89,3	Out of memory	85,76%
Средние значения RGB и HSV, компактность, площадь	90,1%	89,7%	88,4%	85,94%
Средние значения R и H, компактность, площадь	88,4%	87,9%	86,1%	81,98%
Средние значения RGB и HSV, 4 характеристики размеров и компактности, площадь	Out of memory	89,9%	88,5%	83,97%
Средние значения HSV, компактность	82,3%	82,2%	Out of memory	70,63%
Средние значения RGBH, компактности, площадь	90%	88,3%	88,8%	73,1%



Сравнение моделей ННС по параметру частоты ошибочной классификации

Нейро-нечеткие системы чувствительны к проблеме проклятия размерности – модель genfis1 к размерности вектора признаков, модель genfis3 – к объему обрабатываемых данных. Система NEFClass ограничена как с точки зрения объема обрабатываемых данных, так и чувствительна к их набору. Поэтому при использовании нейро-нечетких систем нужно тщательно подходить к выборке исходных и тестовых данных, проводя дополнительные исследования по оценке количества объектов, достаточных для описания исследуемого процесса.

Для дальнейшей работы с системой NEFClass необходима ее оптимизация путем моделирования нечеткого перцептрона средствами современных систем и средств программирования. Для повышения точности результатов можно использовать генетические алгоритмы и оптимизацию базы правил за счет внедрения многослойной структуры.

Список литературы

1. Payal D., Purva S. Acquaintance to Artificial Neural Networks and use of artificial intelligence as a diagnostic tool for tuberculosis: a review. *Tuberculosis*. 2018. № 108. P. 1–9.
2. Omisore M.O., Samuel O.W., Atajeromavwo E.J. A genetic-neuro-fuzzy inferential model for diagnosis of tuberculosis. *Applied Computing and Informatics*. 2017. Vol. 13. № 1. P. 27–37.

3. Keles A., Hasiloglu A.S., Aksoy Y. Neuro-fuzzy classification of prostate cancer using NEFCLASS-J. *Computers in biology and medicine*. 2007. № 11. P. 1617–1628.

4. Nauck D. Detlef Fuzzy data analysis with NEFCLASS. *International Journal of Approximate Reasoning*. 2003. Vol. 32. Issues 2. P. 103–130.

5. Babazadeh K.N. Arabalibeik H., Salehian P., Setayeshi S. Abnormal red blood cells detection using adaptive neuro-fuzzy system. *Studies in health technology and informatics*. 2012. № 173. P. 30–34.

6. Адаптивное нейро-нечеткое моделирование [Электронный ресурс]. URL: <https://www.mathworks.com/help/fuzzy/adaptive-neuro-fuzzy-inference-systems.html> (дата обращения: 03.02.2019).

7. Наркевич А.Н., Шеломенцева И.Г., Виноградов К.А., Сысоев С.А. Сравнение методов сегментации цифровых микроскопических изображений мокроты, окрашенных по методу Циля – Нильсена // *Инженерный вестник Дона*. 2017. № 4. [Электронный ресурс]. URL: <http://ivdon.ru/ru/magazine/archive/n4y2017/4430> (дата обращения: 03.02.2019).

8. Шеломенцева И.Г. Результаты фильтрации и сегментации изображений анализа мокроты, окрашенной по методу Циля – Нильсена // *International journal of advanced studies*. 2017. Т. 7. № 4–2. С. 110–114.

9. Наркевич А.Н., Плотников Д.В., Виноградов К.А., Катаева А.В. Сравнение методов отбора признаков для идентификации объектов на цифровых изображениях микроскопических препаратов // *Инженерный вестник Дона*. 2018. № 2. [Электронный ресурс]. URL: <http://www.ivdon.ru/ru/magazine/archive/N2y2018/4860> (дата обращения: 03.02.2019).

10. Gliwa B., Byrski A. Hybrid neuro-fuzzy classifier based on NEFCLASS model // *Computer Science*. 2011. № 12. P. 115–135.