Islamic Azad University , Shiraz Branch
نشریه تحلیل مدارها، داده ها و سامانه ها
Journal of Circuits, Data and Systems Analysis
sanad.iau.ir/journal/jcdsa

Modeling the speech recognition system using the deep learning technique of spiking neural networks

Melika Hamian*[1], Karim Faez[2], Sohila Nazari[3], Maliheh Sabeti[4]

[1] Department of Engineering, Payame Noor University (PNU), Hamedan, Iran.
hamian.melika@gmail.com
[2] Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran.
karim.faez@gmail.com
[3] Department of Electrical Engineering, Shahid Beheshti University, Tehran, Iran.
msoheilanazari21@yahoo.com
[4] Department of Computer Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran.
malihe.sabeti@gmail.com

Abstract: The architecture of spiking neural network (SNN) is introduced inspired by dynamic spiking neurons. SNNs have great potential to understand time-dependent entanglement pattern by dynamic spiking neurons and can process coded data according to time event. However, training deep SNNs is not straightforward. In this paper, we propose a new layered SNN learning framework for fast and efficient pattern recognition, which uses optimization algorithms to learn deep SNNs. In the mentioned method in the deep learning problem of our deep SNN layers, with the help of different algorithms of gradient-based optimization (GBO) and wild horse optimization (WHO), the two main parameters of spike neurons are searched and calculated for different layers. We use SNN to model the digital speech recognition system and compare and evaluate their performance in different scenarios with other deep learning methods. The results of SNN training for data extracted from different datasets show an increase in identification and estimation accuracy compared to the performed tasks. Comparing the results, the proposed SNN-WHO method was able to achieve accuracies of 95.47% and 92.3% among its counterparts, and they show an increase in the accuracy of identification and estimation compared to the performed works.

Keywords: spiking Neural Networks (SNN); Gradient Based Optimization (GBO) ; Wild Horse Optimization (HWO).

JCDSA, Vol. 2, No. 3, Autumn 2024	Online ISSN: 2981-1295	Journal Homepage: https://sanad.iau.ir/en/Journal/jcdsa
Received: 2023-12-06	Accepted: 2024-09-23	Published: 2024-12-20

CITATION

Hamian, M., et. al., " *Modeling the speech recognition system using the deep learning technique of spiking neural networks*", Journal of Circuits, Data and Systems Analysis (JCDSA), Vol. 2, No. 3, pp. 43-52, 2024.
DOI: 00.00000/0000

COPYRIGHTS

©2024 by the authors. Published by the Islamic Azad University Shiraz Branch. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0)
https://creativecommons.org/licenses/by/4.0

* Corresponding author

Extended Abstract

1- Introduction

The architecture of spiking neural network is introduced inspired by dynamic spiking neurons. SNNs have great potential to understand time-dependent entanglement pattern by dynamic spiking neurons and can process coded data according to time event. However, training deep SNNs is not straightforward. In this paper, we propose a new layered SNN learning framework for fast and efficient pattern recognition, which uses optimization algorithms to learn deep SNNs. In the mentioned method in the deep learning problem of our deep SNN layers, with the help of different algorithms of gradient based optimization (GBO) and wild horse optimization (WHO), two main parameters of spike neurons are searched and calculated for different layers. We use SNN to model the digital speech recognition system and compare and evaluate their performance in different scenarios with other deep learning methods. Comparing the results, the proposed SNN-WHO method was able to achieve accuracies of 95.47 and 92.3 among its counterparts, and they show an increase in the accuracy of identification and estimation compared to the performed works.

2- Methodology

Considering the problem of SNN training according to systematic data, providing an optimization approach with the help of optimization algorithms is the main solution to adapt the trained systems to the expected values. Therefore, in this article, an effective error measure is used to calculate the objective function. First, we prepare the training data for the desired system. In this step, with the help of a genetic algorithm, the data sets or features that have the closest match with the system outputs are normalized to train the SNN network in the interval [0,1]. Other input features are removed from the training set. After preparing the training data for the SNN network and defining the network structure for the number of layers and neurons in each layer, we determine the learning parameters of the DSNN deep spiking neural network by algorithms.

In this method, by determining the number of layers of the deep SNN network and the number of neurons in each layer, a general definition of the network is obtained, which is used to match the results of the desired system to estimate the output of the network. The system will use the optimization algorithms proposed in this article (GBO and WHO).

To match the simulation results, we need to calculate the generated network for each test data and calculate the effective error or deviation for their equivalent outputs. In the next step, the objective function of aggregating all errors is determined based on the following relationship. The important point for calculating and valuing the output of the network in this work is to calculate the RMS value of the turn spike in the output of the output neuron, which is calculated for the interval 0 to 0.21.

$$F(x) = (\exp(\mu * \sum error(i)^2) - 1) \quad \text{for} \quad i = 1, \ldots, \text{number of training set}$$

3- Results and discussion

In this article, we have used two data sets to analyze the performance of our proposed design for different samples. Three machine learning methods have been applied, which include feedforward ANN and ANFIS, adaptive neural network, and the proposed SNN method with two GBO and WHO algorithm approaches. For this case study, the defined network of a three-layer network with the number of neurons [5 3 1] has been used for all machine learning networks. Comparing the results, the proposed SNN-WHO method was able to achieve accuracies of 95.47 and 92.3 among its counterparts, and they show an increase in the accuracy of identification and estimation compared to the performed works.

4- Conclusion

Preliminary results have shown that the detection performance of SNNs is either comparable or slightly worse than that of ANNs with the same network architecture. A possible reason for this performance reduction is the reduced representation power of the discrete neural representation (i.e. spike count) compared to the continuous floating point representation of artificial neural networks. In addition, the identification performance of ANFIS and ANN and SNN models in a scenario with few resources is also investigated. In this scenario, SNN noise models outperform conventional artificial neural networks, which can be attributed to the noisy training of the burst learning framework. The neural encoding scheme adopted in this work allows the input features to be encoded in a short encoding time window for fast processing by SNNs. Recurrent neural networks by exploring the long temporal context information in the input signals have excellent modeling capability for the signals. They have shown once. As future work, we will investigate recurrent networks of spiking neurons for speech recognition applications for digit classification to improve recognition performance.

Islamic Azad University , Shiraz Branch
نشریه تحلیل مدارها، داده ها و سامانه ها
Journal of Circuits, Data and Systems Analysis
sanad.iau.ir/journal/jcdsa

مدل‌سازی سیستم تشخیص گفتار با استفاده از تکنیک یادگیری عمیق شبکه‌های عصبی اسپایکینگ

ملیکا حامیان*[1] ، کریم فایز[2]، سهیلا نظری[3]، ملیحه ثابتی[4]

۱- عضو هیات علمی دانشگاه پیام نور همدان، ایران(hamian.melika@ pnuhp.ac.ir)
۲- گروه مهندسی برق، دانشگاه صنعتی امیرکبیر، تهران، ایران(karim.faeez@gmail.com)
۳- دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران، ایران(msoheilanazari21@yahoo.com)
۴- مهندسی کامپیوتر، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران(malihe.sabeti@gmail.com)

چکیده: ساختار شبکه عصبی اسپایکینگ با الهام از نورون‌های اسپایکینگ پویا معرفی شده است. شبکه‌های عصبی اسپایکینگ پتانسیل فوق‌العاده‌ای برای درک الگوی درهم وابسته به زمان توسط نورون‌های اسپایکینگ پویا دارند و می‌توانند داده‌های رمزگذاری شده را مطابق با رویداد زمان پردازش کنند. با این حال، آموزش شبکه‌های عصبی اسپایکینگ عمیق ساده نیست. در این مقاله، یک چارچوب جدید یادگیری لایه‌ای شبکه عصبی اسپایکینگ برای تشخیص الگوی سریع و کارآمد پیشنهاد می‌شود که از الگوریتم‌های بهینه‌سازی برای یادگیری شبکه‌های عصبی اسپایکینگ عمیق استفاده می‌کند. در روش اشاره شده در مساله یادگیری عمیق، به کمک الگوریتم‌های مختلف بهینه‌سازی مبتنی بر گرادیان و بهینه‌سازی اسب وحشی، دو پارامتر اصلی نرون‌های اسپایک برای لایه‌های مختلف جستجو و محاسبه می‌شود. در این مقاله، از شبکه عصبی اسپایکینگ برای مدل‌سازی سیستم تشخیص گفتار رقمی استفاده و عملکرد آن‌ها در سناریوهای مختلف با سایر روش‌های یادگیری عمیق مقایسه و ارزیابی می‌شود. در مقایسه نتایج، روش پیشنهادی شبکه‌های عصبی اسپایکینگ با بهینه‌سازی اسب وحشی توانسته به دقت‌های ۹۵.۴۷٪ و ۹۲.۳٪ در بین همتایان خود دست پیدا کند؛ که افزایش دقت شناسایی و تخمین را نسبت به کارهای انجام شده نشان می‌دهند.

واژه های کلیدی: شبکه‌های عصبی اسپایکینگ، بهینه‌سازی مبتنی بر گرادیان، بهینه‌سازی اسب وحشی.

نوع مقاله: پژوهشی
DOI: 00.00000/0000
تاریخ ارسال مقاله: ۱۴۰۲/۰۹/۱۵ تاریخ پذیرش مقاله: ۱۴۰۳/۰۷/۰۲ تاریخ چاپ مقاله: ۱۴۰۳/۰۹/۳۰

۱- مقدمه

در سالهای اخیر، شبکه‌های عصبی اسپایکینگ[2] به عنوان نسل جدیدی از شبکه‌های عصبی عمیق کم مصرف به دلیل پردازش پراکنده، ناهمزمان و رویداد محور باینری و قابل اجرا در سخت افزار ظهور کرده‌اند [۱]. شبکه عصبی مصنوعی عمیق به صورت کنترل‌شده از طریق انتشار برگشتی آموزش داده می‌شود [۲] و به‌طور گسترده‌ای به عنوان ابزار محاسباتی برای حل مشکل طبقه‌بندی، رگرسیون، تشخیص الگو و مسئله تخمین تابع و همچنین مسئله بهینه‌سازی پیچیده با ساختار بسیار غیر خطی و ناپیوسته مورد استفاده است. برخی از اشکالات شبکه‌های عصبی مصنوعی به عنوان مجموع چالش‌ها و محدودیت‌های قابل توجه شناخته می‌شوند.

- چند بعدی بودن داده‌ها از جمله محدودیت‌های قابل توجه در یادگیری تحت نظارت است که زمانی اتفاق می‌افتد که میزان ویژگی‌ها و امتیازات آموزشی به طرز وحشتناکی زیاد شود.
- با توجه به حافظه حاصله و نیازهای محاسباتی و حجم غیرقابل مقاومت داده برای آموزش در بعد سخت‌افزاری، یادگیری شبکه‌های عصبی را چالش برانگیزتر می‌کند.
- چالش اضافی در طبقه‌بندی، همپوشانی ویژگی مشخصه بین کلاس‌های مختلف است، زیرا داده‌ها غیرخطی هستند، کار تقسیم کلاس‌ها را چالش‌برانگیز کرده است.
- مدل‌های یادگیری عمیق معمولاً در بسیاری از انواع داده‌ها عملکرد خوبی دارند، اما همواره یک مجموعه داده بزرگ برای آموزش آنها برای تولید نتایج معنی‌دار ضروری است.

* نویسنده مسئول

[2] spiking neural networks (SNN)

با تمرکز بر بهبود این اشکالات با استفاده از روش اکتشافی، مدل‌های شبکه عصبی اسپایکی خوشه‌بندی داده‌ها در این کار بررسی می‌شود. اگرچه شبکه عصبی اسپایکینگ دارای مزایای شباهت مغزی قوی و مصرف کم انرژی به دلیل استفاده از اسپایک‌های گسسته برای نمایش و انتقال اطلاعات است، عملکرد آن هنوز نیاز به بهبود دارد و آموزش مستقیم شبکه عصبی اسپایکینگ چالش‌برانگیز است و در مقایسه با شبکه‌های عصبی از دقت بالایی برخوردار نیست.

- شبکه‌های عصبی اسپایکینگ قابلیت‌هایی را برای پردازش اطلاعات داده‌های منابع مختلف با تعریف این وزن‌ها، نشان داده‌اند.
- علاوه بر این، مدل‌های مبتنی بر شبکه‌های عصبی اسپایکینگ توانایی شبیه‌سازی موفقیت‌آمیز دینامیک عصبی زیربنایی برای کاربردهای تشخیص را نشان داده‌اند.
- با توجه به تعاریف انجام شده برای شبکه عصبی اسپایکینگ و قابلیت اجرای سخت‌افزاری ساده این سیستم در پردازنده‌های کم مصرف، به کاربردهای مهم در این شبکه‌ها برای طبقه‌بندی و پردازش‌های هوش مصنوعی دست یافته‌ایم.

بنابراین تلاش شده است تا مسائل مختلف با کمک شبکه‌های عصبی اسپایکینگ پردازش و حل گردد. اما برای این مدل از هوش مصنوعی با چالش‌هایی برای آموزش و بهینه‌سازی روبرو هستیم. از مهمترین چالش‌ها، برخورد با نحوه آموزش شبکه برای داده‌های مختلف کلاسه‌بندی طبق تعریف وزن‌های مختلف برای نرون‌های لایه‌های مختلف در شبکه عصبی اسپایکینگ عمیق است. ارائه یک مدل آموزش برای شبکه‌های عصبی اسپایکینگ عمیق برای افزایش دقت طبقه‌بندی، از چالش‌های مهم تحقیقاتی اخیر برای کاربردهای کلاسه‌بندی و هوش مصنوعی است. در رویکردهای ارائه شده در مساله آموزش، برای افزایش دقت شبکه نیازمند افزایش نرون‌ها و تعداد لایه‌های شبکه‌های عصبی اسپایکینگ هستیم که این با مشکل پیچیدگی محاسباتی برای آموزش روبرو می‌شود. این مساله ما را بر این داشت تا به دنبال یک روش جدید برای پشتیبانی از شبکه‌های عصبی اسپایکینگ عمیق و پیچیده به منظور افزایش دقت آموزش طبقه‌بندی باشیم.

در [۴]، یک چارچوب جدید تبدیل شبکه‌های عصبی به شبکه‌های عصبی اسپایکینگ و چارچوب یادگیری لایه‌ای را برای تشخیص الگوی سریع و کارآمد پیشنهاد شده است که به عنوان یادگیری پشت سر هم پیشرونده شبکه‌های عصبی اسپایکینگ عمیق شناخته می‌شود. با انگیزه بهره‌وری انرژی بی‌سابقه و قابلیت پردازش سریع اطلاعات، استفاده از شبکه‌های عصبی اسپایکینگ را برای تشخیص گفتار در [۵] بررسی شده است. در این کار، از شبکه‌های عصبی اسپایکینگ برای مدل‌سازی صوتی استفاده شده است و عملکرد آن‌ها را در چندین سناریو تشخیص واژگان بزرگ ارزیابی می‌گردد. نتایج تجربی دقت تشخیص خودکار گفتار واژگان بزرگ رقابتی را به همتایان شبکه‌های عصبی خود نشان می‌دهند.

ازدحام بیش از حد طیف و انسداد شبکه مخابرات بیسیم که باعث قطع تماس می‌شود از نگرانی‌های مهم این شبکه‌ها است. برای مقابله با این مسائل به طور همزمان یک شبکه رله مبتنی بر فناوری رادیویی شناختی در [۶] پیشنهاد شده است. ابتدا یک تکنیک حسگر طیفی مبتنی بر شبکه عصبی اسپایکینگ آموزش دیده با الگوریتم بهینه‌سازی نهنگ اصلاح‌شده جهش یافته برای تشخیص کارآمد حفره‌های طیف پیشنهاد شده است. در اینجا، وزن‌های شبکه‌های عصبی اسپایکینگ با استفاده از این الگوریتم برای پیش‌بینی مؤثر حفره‌های طیف آموزش داده می‌شوند. در[۷]با استفاده از بهینه‌سازی الگوریتم ژنتیک مرتب‌سازی غیرمسلط روی یک مدل شبکه‌های عصبی اسپایکینگ کارآمد خاص، با استفاده از مدل عصبی ایژیکویچ، برای انجام بهینه‌سازی چند هدفه در شبکه‌های عصبی اسپایکینگ، با تمرکز بر جستجوی پارامترهای اتصال شبکه برای دستیابی به نرخ شلیک هدف انواع عصبی تحریک‌کننده و مهارکننده، به کار رفته است. در [۸]، یک معماری عصبی رویداد محور مدولار آموزش‌پذیر انتها به انتها ارائه می‌شود که از قوانین انطباق سیناپسی و آستانه‌ای محلی برای انجام دگرگونی‌های بین الگوهای سنبله مکانی-زمانی دلخواه استفاده می‌کند. این معماری یک مدل بسیار انتزاعی از معماری‌های شبکه عصبی اسپایکی موجود را نشان می‌دهد. معماری شبکه عصبی اسپایکی مبتنی بر رویداد عمیق پیشنهادی (ODESA) می‌تواند به طور همزمان ویژگی‌های مکانی-زمانی سلسله مراتبی را در مقیاس‌های زمانی دلخواه متعدد بیاموزد. این مدل یادگیری آنلاین را بدون استفاده از پس انتشار خطا یا محاسبه گرادیان انجام می دهد. در [۹]، چارچوبی مبتنی بر بهینه‌سازی انرژی آزاد تکراری با شبکه‌های عصبی اسپایکینگ برای مدل‌سازی سیستم فرونتو – مخطط (PFC-BG) برای تولید و یادآوری توالی‌های حافظه صوتی ارائه می‌کند. در راستای مطالعات تصویربرداری عصبی انجام شده در PFC، یک استراتژی کدگذاری واقعی با استفاده از مکانیسم افزایش مدولاسیون پیشنهاد می‌شود تا توالی‌های انتزاعی را تنها بر اساس رتبه و مکان آیتم‌ها در آن‌ها نشان داد. بر اساس این مکانیسم، می‌توان مجموعه‌ای از نورون‌های حساس به ساختار زمانی را در توالی‌هایی ساخت که از آن‌ها بتوان هر دنباله جدیدی را نشان داد. شبکه‌های عصبی اسپایکینگ های پیشرفته برای دستیابی به دقت بالا به حافظه زیادی نیاز دارند، در نتیجه استقرار آنها در سیستم‌های تعبیه شده، به عنوان مثال، در دستگاه‌های تلفن همراه با باتری و گره‌های لبه اینترنت اشیا دشوار می‌شود. در این راستا، مقاله [۱۰]، FSpiNN را پیشنهاد می‌کند، یک چارچوب بهینه‌سازی برای به دست آوردن شبکه‌های عصبی اسپایکینگ‌های کارآمد با حافظه و انرژی کارآمد برای آموزش و پردازش استنتاج، با قابلیت یادگیری بدون نظارت و در عین حال حفظ دقت ارائه شده است.

اکثر تکنیک‌های قبلی شبکه‌های عصبی اسپایکینگ که داده‌ها را مدیریت می‌کنند به شبکه‌های کم عمق محدود می‌شوند و بنابراین، عملکرد پایینی را نشان می‌دهند. به طور کلی، رفتار ادغام و آتش‌سوزی نورون‌های اسپایک، فعالیت اسپایک را در لایه‌های عمیق‌تر کاهش می‌دهد. فعالیت پراکنده اسپایک منجر به یک راه حل بهینه جهانی (به

عنوان مثال، کاهش تابع هدف) در طول آموزش می‌شود. برای پرداختن به این محدودیت، مقاله [۱۱] پیشـــرفت‌های الگوریتمی و معماری جدیدی را برای تسریع آموزش شبکه‌های عصبی اسپایکینگهای بسیار عمیق بر روی داده‌های DVS[1] پیشـــنهاد می‌کند. به طور خاص، با آموزش بالابر فعال‌سازی اسپایک (SALT) که با بهینه‌سازی وزن‌ها و آستانه‌ها در لایه‌های کانولوشن، فعالیت سنبله را در تمام لایه‌ها افزایش می‌دهد. پس از اعمال SALT، وزنه ها را بر اساس افت آنتروپی متقاطع آموزش می یابد.

تاکنون مشخص نیست که این فرآیند خودبهینه‌سازی تا چه حد در مغزهای واقعی نیز مؤثر اســت. در [۱۲] اجرای یک مدل شـــبکه عصـــبی اســـپایکینگ خودبهینه‌ســاز عملی شـــده اســت. در این کار، با استفاده از این شبکه عصبی اسپایک برای شبیه‌سازی شبکه هاپفیلد با یادگیری هبی، سعی می‌شود بین سیستم‌های عصبی مبتنی بر نرخ و مبتنی بر کدگذاری زمانی ارتباط برقرار گردد. اگرچه کارایی فرآیند خودبهینه‌سازی مستقل از فرضیات ساده‌سازی یک شبکه هاپفیلد معمولی است اما، برای واقعی‌تر کردن این مدل در حوزه سخت افزار به کار بیشتری نیاز است.

نوآوری ما در این مقاله اســتفاده از یک چارچوب جدید ســیســتم یادگیری ماشین مبتنی بر شبکه‌های عصبی اسپایکینگ با الگوریتم‌های بهینه‌سازی می‌باشد. در این راستا برای آموزش از یک سیستم تطبیق نتایج طبقه‌بندی با کمک دو الگوریتم بهینه ســازی مبتنی بر گرادیان[2] و بهینه‌ســازی اســب وحشــی[3] برای تعیین پارامترهای شـبکه نرون‌های اســپایکینگ بهره برده اســت. در حقیقت در این تحقیق برای توســعه آموزش‌های نرون‌های اســـپایکینگ از الگوریتم‌های بهینه‌ســـازی بهره برده‌ایم تا برای اطلاعات مختلف پارامترهای شـــبکه‌های عصـــبی اســپایکینگ شــامل وزن و مقدار آســتانه هر نرون محاســبه شــود و با مقدارهای نمونه داده‌های آموزش تطبیق یابد.

۲- مفاهیم اولیه

در این بخش به مفاهیم اولیه مورد اســـتفاده در این مقاله خواهیم پرداخت. در ابتدا شبکه عصبی اسپایکینگ معرفی و مدلسازی می‌شود. ســپس الگوریتم‌های بهینه‌ســازی مورد اســتفاده در این مقاله معرفی می‌شـــود. در آخر، مزایا و محدودیت‌های این شـــبکه‌های عصـــبی اسپایکینگ ها تشریح می‌شود و نیاز به یک تکنیک آموزش جدید برای شبکه‌های عصبی اسپایکینگ بیان می‌شود.

۲-۱- مفهوم عصب‌های مصنوعی

ساده‌ترین مدل نورون که معمولاً برای کاربرد عملی مورد بررسی قرار گرفته می‌شود، مدل مک‌کالوخ-پیتس است که در سال ۱۹۴۳ توصیف شد [۱۳] و توسط روزنبلت در سال ۱۹۵۸ [۱۴] پیاده‌سازی شد. در این مدل، یک نورون به عنوان یک ماژول جمع وزنی معرفی می‌شود که تمام مقادیر ورودی را جمع می‌کند و آن را در مقداری به نام وزن ضرب می‌کند. سپس این مقدار از یک تابع انتقال عبور می‌کند و به عنوان یک خروجی محاسبه می‌شود. مدل را می‌توان به صورت (۱) نشان داد [۱۵]:

$$y = f(u), \mathrm{u} = \sum_{k=0}^{n} wixi + w0x0 \qquad (۱)$$

در (۱)، xi نشان‌دهنده سیگنال‌های ورودی، wi نشان‌دهنده وزن‌های ورودی و $\mathrm{f(u)}$ یک تابع انتقال است. پارامترهای $x0$ و $w0$ مقدار اولیه یک نورون را معرفی می‌کنند. تابع انتقال تعیین می‌کند که سیگنال خروجی چگونه به مجموع وزنی سیگنال‌های ورودی وابسته باشد. معمولاً نیاز به ارضای چند شرط مانند محدود بودن ارزش، افزایش یکنواختی، تعیین بر روی همه آرگومان‌های واقعی و قابل تمایز بودن برای تسهیل الگوریتم های یادگیری تعریف می‌شود.

۲-۲- مدل نورون های اسپایکینگ

شبکه‌های عصبی اسپایکینگ، شبکه های عصبی مصنوعی هستند که از نزدیک، شـبکه‌های عصـبی طبیعی را تقلید می‌کنند. علاوه بر حالت عصبی و سیناپسی، شبکه‌های عصبی اسپایکینگ مفهوم زمان را در مدل عملیاتی خود قرار می‌دهند. ایده این اســت که نورون‌ها در شـبکه‌های عصـبی اسـپایکینگ اطلاعات را در هر چرخه انتشـار منتقل نمی‌کنند، بلکه اطلاعات را فقط هنگامی انتقال می‌دهند که پتانســیل غشای - کیفیت ذاتی نورون مربوط به بار الکتریکی غشای آن - به یک مقدار خاص می‌رسد که آستانه نامیده می‌شود. وقتی پتانسیل غشا به آســتانه می‌رســد، نورون آتش می‌کند و ســیگنالی تولید می‌کند که به نورون های دیگر می‌رود که به نوبه خود، در پاســخ به این ســیگنال، پتانسیل‌های خود را افزایش یا کاهش می‌دهند. به یک مدل نورونی که در لحظه عبور از آســتانه آتش‌ســوزی می‌کند، مدل نورون ســنبله گفته می‌شـود. برجسته‌ترین مدل نورون سـنبله‌ای، مدل ادغام و آتش است. در مدل ادغام و آتش، ســطح فعال‌ســازی لحظه‌ای (به عنوان معادله دیفرانسیل مدل‌سازی شده) به طور معمول حالت نورون در نظر گرفته می‌شـود. با افزایش ســنســورهای ورودی، این مقدار را بالاتر یا پایین می‌کند تا اینکه در نهایت حالت پوســیدگی یا در صــورت رســیدن به آستانه شلیک - نورون شلیک می‌کند. پس از شلیک ، متغیر حالت به مقدار کمتری تنظیم می‌شود.

روش‌های مختلف رمزگشـایی برای تفسـیر سـنبله یا اسـپایک به عنوان یک عدد با ارزش واقعی وجود دارد، با تکیه بر فرکانس سـنبله‌ها (کد نرخ)، زمان ســنبله برای اولین بار پس از تحریک، یا فاصــله بین خوشـــه در این مقاله برای بخش مقداردهی نهایی نرون آخر از مفهوم میانگین پالس‌های نرون خروجی برای ارزش‌گذاری سیستم شبکه‌های عصـبی اسـپایکینگ پیشـنهاد و اسـتفاده شـده اسـت. برای درک نحوه عملکرد مغز، باید مطالعات تجربی سیستم عصبی حیوان و انسان را با شـــبیه‌ســـازی عددی مدل‌های مغز در مقیاس بزرگ ترکیب کنیم.

[1] Dynamic Vision Sensor
[2] Gradient Based Optimization (GBO)
[3] Wild Horse Optimization (HWO)

همانطور که ما چنین مدل‌های مغزی در مقیاس بزرگی متشـــکل از نورون‌های پراکنده را توسـعه می‌دهیم، باید مصـالحه‌ای بین دو الزام به ظاهر منحصـر به فرد پیدا کنیم: مدل یک نورون منفرد باید: ۱) از نظر محاسـباتی سـاده، در عین حال ۲) قادر به تولید الگوهای شـلیک غنی باشد. نورون‌های بیولوژیکی واقعی استفاده از مدل‌های دقیق بیوفیزیکی نوع هوچکین-هاکسـلی از نظر محاسـباتی بسـیار منع اسـت، زیرا ما می‌توانیم تنها تعداد انگشـــت شـــماری از نورون‌ها را در زمان واقعی شـبیه‌سـازی کنیم. در مقابل، اسـتفاده از یک مدل ادغام و آتش از نظر محاسباتی مؤثر است، اما این مدل به‌طور غیرواقعی ساده است و قادر به تولید پویایی‌های پرشتاب و انفجاری است که توسط نورون‌های قشر مغز به نمایش گذاشته می‌شود.

در این مقاله، یک مدل اسپایکی ساده استفاده شده است که از نظر بیولوژیکی به اندازه مدل هوچکین-هاکسلی قابل قبول است، اما از نظر محاســباتی به اندازه مدل ادغام و آتش کارآمد اســت. بســته به چهار پارامتر، این مدل رفتار جهش و انفجار انواع شـــناخته شـــده نورون‌های قشر مغز را بازتولید می کند. تجزیه و تحلیل ریاضی مدل در مونوگراف ایژیکویچ [۱۶] منتشر شده است.

۲-۳- نورون‌های ادغام و آتش

این مدل، پرکاربردترین مدل است که عموماً به عنوان شبکه عصبی اسپایکینگ استفاده می‌شود. این مدل بر اصول الکترونیک تکیه داشت. یک اسپایک از آکسون پایین می‌رود و توسط یک کانال پایین گذر تغییر می‌کند، که در طول پالس کوتاه به یک پالس جریان $I(t-t_j(f))$ تبدیل می‌شود که مدار مختصات و آتش را شارژ می‌کند. یک پتانسیل پس سیناپسی $\varepsilon(t-t_i{}^f)$ را می‌توان با افزایش ولتاژ حاصل افزایش داد. با این حال، نورون یک پالس ارسال می‌کند که ولتاژ از مقدار آستانه بیشتر شود [۱۸، ۱۹].

$$\tau_m \frac{\partial u}{\partial m} = -u(t) + RI(t) \qquad (۲)$$

برای به تصویر کشیدن پیامدهای یک پتانسیل غشایی u بعد از مدتی، با τ_m زمان لایه‌ای است که در آن ولتاژ "نشت" می‌کند. به طور مشابه با مدل واکنش اسپایک، هنگامی که u از حد عبور کرد و یک پالس کوتاه σ تحریک شد، نورون شلیک می‌شود.

$$I_i(t) = \sum_{j\in\Gamma_i} c_{ij} \sum_{t_j} \delta(t - t_j^{(0)}) \qquad (۳)$$

جریان ورودی I برای نورون I به طور منظم صفر خواهد بود، زیرا پالس‌های نزدیک به طول کوتاه محدودی دارند. هنگامی که یک اسپایک ظاهر می شود، توسط عامل زنده ماندن سیناپسی c_{ij} افزایش می‌یابد که پتانسیل پس سیناپسی را که خازن را شارژ می‌کند، شکل می‌دهد. این مدل از نظر محاسباتی ساده است و بدون شک می‌توان آن را برای سخت افزارهای چندبعدی اعمال کرد [۱۸].

۲-۴- الگوریتم بهینه‌ساز اسب وحشی

الگوریتم‌های بهینـه‌ســـازی معمولاً از رفتار طبیعی یک عامل الهام می‌گیرند که می‌تواند انســـان، حیوان، گیاه یا یک عامل فیزیکی یا شـیمیایی باشـد. بسـیاری از الگوریتم‌های ارائه شـده در دهه گذشـته از رفتار حیوانات الهام گرفته شـده‌اند. در این مقاله از یک الگوریتم جدید به نام بهینه‌سـاز اسـب وحشـی اسـتفاده می‌کنیم که از رفتار اجتماعی اسب‌های وحشی الهام گرفته شده است. اسب‌ها معمولاً در گروه‌هایی متشـکل از یک اسـب نر و چند مادیان و کره اسـب زندگی می‌کنند. اســـب‌ها رفتارهای زیادی از خود نشـــان می‌دهند، مانند چرا، تعقیب، تســـلط، رهبری و جفت گیری. رفتار جذابی که اســـب‌ها را از ســـایر حیوانات متمایز می‌کند، ادب آنهاسـت. رفتار پرورش اسـب به گونه‌ای است که کره اسب‌ها قبل از رسیدن به سن بلوغ گروه را ترک کرده و به گروه‌های دیگر می‌پیوندند. این خروج برای جلوگیری از جفت شـــدن پدر با دختر یا خواهر و برادر است. الهام‌بخش اصلی الگوریتم پیشنهادی رفتار مؤدبانه اسب است[۲۰].

۲-۵- الگوریتم مبتنی بر گرادیان

در الگوریتم مبتنی بر گرادیان پیشـــنهادی کـه ترکیبی از روش‌های گرادیان و جمعیت است، جهت جستجو با روش نیوتن مشخص می‌شود تا دامنه جســـتجو را با اســـتفاده از مجموعه ای از بردارها و دو عملگر اصلی (یعنی قانون جستجوی گرادیان و عملگرهای محلی فرار) جستجو کند. به حداقل رساندن تابع هدف در مسائل بهینه‌سازی در نظر گرفته می‌شود. [۲۱].

۲-۶- مزایا و محدودیت‌های شبکه‌های عصبی اسپایکینگ‌های عمیق

یک انگیزه برای مطالعه شـبکه‌های عصـبی اسـپایکینگ‌ها این اسـت که مغزهای طبیعی عملکرد خوبی را برای شناخت در کارهای دنیای واقعی از خود نشـــان می‌دهنـد. با تلاش‌های مـداوم برای بهبود درک مـا از محاسبات مغزی، انتظارات وجود دارد که مدل‌هایی که به زیست‌شناسی نزدیک‌تر هســـتند، نســـبت به مدل‌های انتزاعی‌تر به هوش طبیعی نزدیک‌تر می‌شـــوند و توانایی بیشـــتری برای پیشـــبینی و مدلســـازی سیسـتم‌ها و داده های مختلف را دارند. شـبکه‌های عصـبی اسـپایکینگ برای پردازش اطلاعات مبتنی بر رویداد مکانی-زمانی از حســـگرهای نورومورفیک، که خود کارآمد هســـتند، مناســـب هســـتند. حســـگرها اطلاعات دقیق زمانی را از محیط ضـبط می‌کنند و شـبکه‌های عصـبی اســـپایکینگ می‌توانند از کدهای زمانی کارآمد در محاســـبات خود نیز اســـتفاده کننـد و پردازش اطلاعات را انجام دهنـد [۲۲]. این پردازش اطلاعات نیز رویداد محور اســـت به این معنی که هر زمان که اطلاعات کمی ثبت شـــود یا هیچ اطلاعاتی ثبت نشـــده باشـــد، محاســـبات در شبکه‌های عصبی اسپایکینگ به مقدار زیادی کاهش می‌یابد، اما زمانی که رویدادها یا اسـپایک‌های ناگهانی فعالیت ثبت می‌شـود، شـبکه‌های عصبی اسپایکینگ اسپایک های بیشتری ایجاد می‌کند. با این فرض که

معمولاً اطلاعات از دنیای خارج شامل رویدادهای پراکنده است، این منجر به یک ایجاد روش محاسباتی بسیار کارآمد می‌شود. علاوه بر این، استفاده از ورودی حوزه زمان در مقایسه با رویکردهای فریم محور، که در آن یک مرحله زمانی مصنوعی تحمیل شده توسط سنسور معرفی می‌شود، اطلاعات ارزشمند بیشتری است.

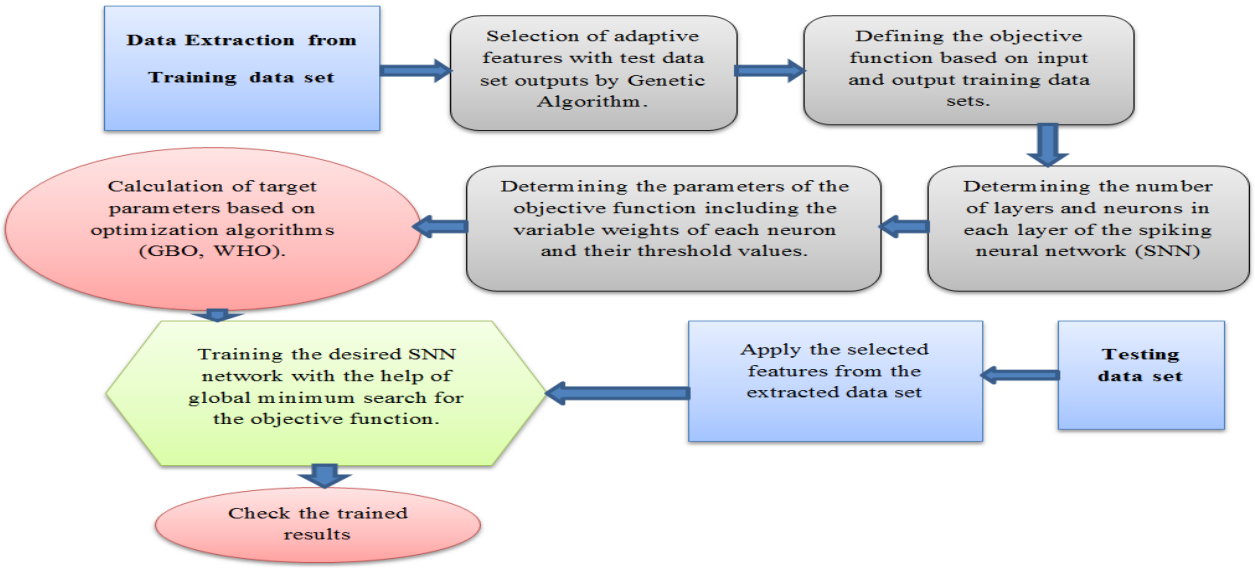
یکی از بزرگترین اشکالات شبکه‌های عصبی اسپایکینگ عمیق این است که علیرغم پیشرفت اخیر [۲۳، ۲۴] دقت آنها در معیارهای معمولی مانند MNIST [۲۵]، CIFAR [۲۶] یا ImageNet [۲۷] به سطوح مشابهی با همتایان یادگیری ماشین خود نمی‌رسند. تا حدودی می‌توان این را به ماهیت این بنچمارک‌ها نسبت داد که بر روی داده‌های ورودی معمولی اعمال می‌شوند. بنابراین، نوعی تبدیل از سیگنال‌های اطلاعاتی به قطارهای سنبله مورد نیاز است که معمولاً دارای تلفات و ناکارآمدی است. عامل محدودکننده دیگر فقدان الگوریتم‌های آموزشی است که از قابلیت‌های نورون‌های اسپایکی، به عنوان مثال، کدهای زمانی کارآمد استفاده می‌کنند. در بحث آموزش شبکه‌های عصبی اسپایکینگ، بیشتر رویکردها از تقریب‌های مبتنی بر نرخ شبکه عصبی عمیق معمولی استفاده می‌کنند، که به این معنی است که هیچ افزایش دقتی در این روش‌های آموزش نسبت به شبکه عصبی عمیق، قابل انتظار نیست. شبکه‌های عصبی اسپایکینگ عمیق ممکن است در چنین سناریوهایی مفید باشند، زیرا نتایج تقریبی ممکن است سریعتر و کارآمدتر از سیستم های معمولی به دست آید، به خصوص اگر شبکه‌های عصبی اسپایکینگ بر روی سخت افزار نورومورفیک اجرا شود. اما به‌عنوان یک روش آموزش مستقل و مجزا قابل تعمیم نمی‌باشد. از سوی دیگر، طراحی و تجزیه و تحلیل الگوریتم‌های آموزشی برای شبکه‌های عصبی اسپایکینگ نیز دشوارتر است، زیرا روش محاسباتی ناهمزمان و ناپیوسته است، که استفاده مستقیم از تکنیک‌های پس‌انتشار موفقیت‌آمیز را که برای شبکه عصبی عمیق استفاده می‌شود، دشوار می‌کند.

۳- روش پیشنهادی

با توجه به مساله آموزش شبکه‌های عصبی اسپایکینگ منطبق با داده‌های سیستماتیک ارائه یک رویکرد بهینه‌سازی با کمک الگوریتم‌های بهینه‌سازی راهکار اصلی برای تطبیق سیستم‌های آموزش یافته با مقادیر مورد پیش بینی می‌باشد. بنابراین در این مقاله از یک معیار خطای موثر برای محاسبه تابع هدف استفاده شده است. در شکل (۱) نمای کلی طرح پیشنهادی معرفی شده است. هدف این بخش ارائه یک رویکرد آموزشی برای شبکه‌های عصبی اسپایکینگ عمیق می‌باشد. بنابراین ابتدا داده‌های آموزش برای سیستم مورد نظر را آماده‌سازی می‌کنیم. در این مرحله با کمک یک الگوریتم ژنتیک مجموعه داده‌ها یا ویژگی‌هایی که نزدیکترین انطباق با خروجی‌های سیستم را دارند برای آموزش شبکه شبکه‌های عصبی اسپایکینگ در بازه [0,1] انتخاب و نرمالسازی می‌شوند. سایر ویژگی‌های ورودی هم از مجموعه آموزش حذف می‌شوند. بعد از آماده‌سازی داده‌های آموزش برای شبکه‌های عصبی اسپایکینگ و تعریف ساختار شبکه برای تعداد لایه‌ها و نرون‌های هر لایه به تعیین پارامترهای یادگیری شبکه عصبی اسپایکینگ عمیق توسط الگوریتم‌ها می‌رسیم که در ادامه اشاره می‌شود.

۳-۱- مدل یادگیری پیشنهادی

در این روش با تعیین تعداد لایه‌های شبکه‌های عصبی اسپایکینگ عمیق و تعداد نرون‌های هر لایه به یک تعریف کلی از شبکه دست می‌یابیم که برای انطباق نتایج سیستم مورد نظر جهت تخمین خروجی سیستم از الگوریتم‌های بهینه‌سازی پیشنهاد شده در این مقاله (الگوریتم بهینه‌ساز اسب وحشی و الگوریتم مبتنی بر گرادیان) بهره خواهیم برد.



شکل (۱): نمای کلی روش پیشنهادی برای آموزش شبکه‌های عصبی اسپایکینگ.

برای انطباق نتایج شبیه‌سازی می‌بایست شبکه ایجاد شده را برای تک تک داده‌های تست محاسبه کرده و برای خروجی‌های معادل انها خطا یا انحراف موثر را محاسبه کنیم. در مرحله بعد تابع هدف از تجمیع تمام خطاها بر اساس رابطه زیر تعیین می‌شود. نکته مهم برای محاسبه و ارزش‌گذاری خروجی شبکه در این کار محاسبه مقدار جذر میانگین مربعات قطار اسپایک در خروجی نرون خروجی می‌باشد که برای محدوده ۰ تا ۰/۲۱ محاسبه می‌شود. برای تغییر محدوده عملیاتی شبکه‌های عصبی اسپایکینگ از تغییر محدوده زمانی محاسباتی یا فرکانس عملیاتی اسپایک های ورودی می‌توان بهره برد. در این کار ما فرکانس عملیاتی برای تعریف اسپایک‌های ورودی برابر ۱۰ کیلوهرتز در نظر گرفته‌ایم. این فرکانس انتخابی برای سخت‌افزارهای مختلف با توجه تکنولوژی مورد استفاده قابل تغییر است. محدوده زمانی محاسبه برای انجام عملیات پاسخ‌دهی شبکه برابر ۰/۵ ثانیه محاسبه شده است.

۳-۲- استخراج خروجی داده‌های آزمایشی از شبکه عصبی اسپایکینگ آموزش دیده

در این بخش بعد از آموزش شبکه با تکنیک پیشنهادی نوبت به بررسی نتایج عملکرد سیستم برای داده‌های آزمایشی می‌شودکه به اثبات نتایج برای تکنیک پیشنهادی می‌پردازد. در بحث یادگیری عمیق داده‌های ورودی و خروجی به دو گروه داده‌های اموزشی و داده‌های آزمایشی تقسیم می‌شوند. در این بخش تابع هدف با توجه به پاسخ‌های سیستم اسپایکینگ به ازای تک‌تک داده‌های آموزش بررسی و خروجی آنها با مقدار واقعی مقایسه می‌شود. خروجی تابع هدف با مقدار مجموع انحراف پاسخ‌های داده‌های آموزش به توان دوم و طبق رابطه تابع نمایی زیر تعریف می‌شوند. این مقادیر برای مجموعه داده‌های آموزش محاسبه و پارامترهای هدف به عنوان پاسخ‌های بهینه نهایی الگوریتم‌های بهینه‌سازی محاسبه میشود.

$$F(x) = (\exp(\mu * \sum error(i)2) - 1)$$
$$for \;\; i = 1, \ldots\ldots, number\ of\ training\ set$$

که برای سیستم‌های مختلف مقدار μ متغیر و مثبت است.

۴- نتایج و بحث

در این مقاله دو مساله یادگیری ماشین شامل مجموعه داده‌های IRIS در کتابخانه نرم افزار متلب و همچنین پردازش سیگنال صوتی برای تشخیص ارقام ۰-۹ از مجموعه داده های ۲۰۰ نمونه صوت از افراد مختلف TIDIGITS برای [۳۰] ارائه شده است. نمونه‌ها از مجموعه‌های مختلف زن و مرد استفاده شده است. هدف از این کار، یافتن نقطه شروع و پایان هر رقم گفتاری در ضبط رشته‌های رقم‌های ۰ تا ۹ و ساختن سیستمی برای تشخیص این رقم‌های گفتاری است. بر اساس شناخت این سخنان گفتار، می‌توان کاربرد تلفن صوتی را پیاده‌سازی نمود. همچنین از روش‌های پردازش سیگنال دیجیتال برای استخراج ویژگی‌های سیگنال و ایجاد یک سیستم تشخیص وابسته به بلندگو استفاده می‌گردد.

۴-۱- استخراج ویژگی‌ها

در کارهای مختلف [۵و۹] عمدتاً از ضرایب کپسترال فرکانسی مل [1] برای تمایز ارقام مختلف گفتاری استفاده شده است. کپسترال فرکانسی مل طیف قدرت کوتاه مدت یک گفتار است که بر اساس تبدیل کنوانسیون خطی از یک طیف قدرت ورود به سیستم در مقیاس مل غیرخطی از فرکانس است. ضرایب کپسترال فرکانسی مل ضرایبی هستند که کپسترال فرکانسی مل را تشکیل می دهند [۳۱]. برای ضرایب محاسبه شده از پنجره سیگنال صوت در این کار از یک مدل ویژگی های آماری استفاده شده است که در کنار ویژگی های مختلف ZCR [2] و توان سیگنال‌های صوتی جمعاً ۷۷ ویژگی برای هر سیگنال استخراج می‌شود. به دلیل بالا بودن حجم پردازش محاسباتی در این مطالعه موردی به کمک الگوریتم ژنتیک تنها ۷ ویژگی اصلی شناخته و انتخاب می‌شود. معیار انتخاب ویژگی‌ها هم بر اساس بالاترین تطبیق نتایج برای بسط تیلور چندجمله‌ای مرتبه ۵ برای تک تک ویژگی‌ها می‌باشد. بنابراین هر ویژگی‌ای که کمترین انحراف از مقدار خروجی داده‌های تست داشتند، به‌عنوان ویژگی برتر انتخاب می‌گردد.

برای سیستم دیتابیس IRIS برای آموزش از کل مجموعه ورودی‌های که شامل ۴ ویژگی است استفاده شده است. در این مقاله از دو مجموعه داده استفاده کرده‌ایم تا عملکرد طرح پیشنهادی خود را برای نمونه‌های مختلف بررسی و تحلیل گردد. در زمینه داده های IRIS که از دیتابیس کتابخانه متلب استفاده شده است، سه روش یادگیری ماشین اعمال شده که شامل روش شبکه‌های عصبی پیشخور و شبکه فازی عصبی تطبیقی [3] و روش پیشنهادی شبکه‌های عصبی اسپایکینگ با دو رویکرد الگوریتم بهینه‌ساز اسب وحشی و الگوریتم مبتنی بر گرادیان آورده شده است. برای این مطالعه موردی شبکه تعریف شده یک شبکه سه لایه با تعداد نرون‌های [5 3 1] برای تمام شبکه‌های یادگیری ماشین استفاده شده است. جدول (۱) نتایج مقایسه را برای دو مطالعه مورد نظر نشان می‌دهد. همانطور که نشان داده شده است شبکه عصبی اسپایکینگ با الگوریتم مبتنی بر گرادیان [4] توانسته بالاترین دقت را در مقایسه سایر روش‌های یادگیری ماشین ایجاد کند.

در این میان روش شبکه عصبی اسپایکینگ با الگوریتم بهینه‌ساز اسب وحشی [5] توانسته است بالاترین دقت را ایجاد کند. در نتایج محدوده برچسب گذاری برای شبکه شبکه‌های عصبی اسپایکینگ مطابق با محدوده ۱ تا ۱۱ برای مشخصه اعداد به ترتیب۱ تا ۹ و O و صفر برچسب گذاری شده است که به محدوده عملیاتی شبکه در بازه [0 0.2] نرمالسازی شده است.

[1] Mel-frequency cepstral coefficients (MFCC)
[2] zero-crossing rate (ZCR)
[3] adaptive network-based fuzzy inference system (ANFIS)
[4] SNN-GBO
[5] SNN-WHO

جدول ۱- مقایسه نتایج دقت روش‌های مختلف یادگیری ماشین.

ML methods (%)	ANFIS	ANN	SNN-GBO	SNN-WHO
Digit recognize	81.2	84.38	88.3	**92.3**
IRIS	95.85	96.27	97.13	**95.47**

۵- نتیجه‌گیری

تقاضاهای رو به رشد سریع خدمات تشخیص گفتار نگرانی‌هایی را در مورد کارایی محاسباتی، عملکرد بلادرنگ، و امنیت داده‌ها و غیره ایجاد کرده است. همانطور که از محاسبات مبتنی بر رویداد که در سیستم‌های عصبی بیولوژیکی مشاهده می‌شود، این مقاله با استفاده از شبکه‌های عصبی اسپایکی الهام‌گرفته از مغز و الگوریتم‌ها بهنیه‌سازی الهام گرفته از طبیعت برای وظایف مختلف در حوزه شناسایی ارقام از سیگنال‌های صوتی و سایر موارد بررسی کرده است. برای این منظور، ما یک چارچوب جدید سیستم یادگیری ماشین مبتنی بر شبکه‌های عصبی اسپایکینگ با الگوریتم بهینه‌ساز اسب وحشی و الگوریتم مبتنی بر گرادیان را پیشنهاد کردیم، که در آن شبکه‌های عصبی اسپایکینگ برای طبقه بندی سیگنال‌های صوتی و دیتابیس ISIR استفاده می‌شود و ویژگی‌های مختلف را در مجموعه‌ای از واحدهای صوتی استخراج می‌کند. این خروجی‌ها از استخراج ویژگی‌های مختلف، اطلاعات سطح رقم برچسب گذاری شده را از مدل زبان مربوطه با الگوریتم ژنتیک انتخاب می‌کنند تا نزدیکترین ویژگی‌های مربوط به سیگنال گفتار ورودی را پیدا کنند. در مجموعه دیگر که از دیتابیس ISIR کتابخانه متلب استفاده شده است، سیستم یادگیری ماشین پیشنهادی را مدل کرده و برای اعتبارسنجی طرح پیشنهادی در کنار سایر روشهای یادگیری ماشین استفاده می شود.

نتایج اولیه نشان داده است که عملکرد تشخیص شبکه‌های عصبی اسپایکینگ یا قابل مقایسه یا کمی بدتر از شبکه‌های عصبی با همان معماری شبکه است. یک دلیل احتمالی برای این کاهش عملکرد، کاهش قدرت نمایش بازنمایی عصبی گسسته (یعنی تعداد سنبله) در مقایسه با نمایش ممیز شناور پیوسته شبکه‌های عصبی مصنوعی [۳۲] است. این شکاف عملکرد به طور بالقوه می‌تواند با گسترش پنجره کدگذاری شبکه‌های عصبی اسپایکینگ بستگی داشته باشد. علاوه بر این، عملکرد شناسایی مدل‌های شبکه فازی عصبی تطبیقی و شبکه‌های عصبی و شبکه‌های عصبی اسپایکینگ در سناریویی با منابع کم نیز بررسی می‌شود. در این سناریو، مدل‌های صوتی شبکه‌های عصبی اسپایکینگ از شبکه‌های عصبی مصنوعی معمولی که می‌تواند به آموزش پر نویز چارچوب یادگیری پشت سر هم نسبت داده شود، بهتر عمل کند. طرح رمزگذاری عصبی اتخاذ شده در این کار به ویژگی‌های ورودی اجازه می‌دهد تا در یک پنجره زمانی رمزگذاری کوتاه برای پردازش سریع توسط شبکه‌های عصبی اسپایکینگ ها کدگذاری شوند. برای کارهای کلاسه‌بندی سیگنال گفتار، همگام زمانی که نیاز به عملکرد زمان واقعی دارند جذاب است. شبکه‌های عصبی مکرر با کاوش اطلاعات بافت زمانی طولانی در سیگنال‌های ورودی، قابلیت مدل‌سازی بسیار خوبی برای سیگنال‌های زمانی نشان داده‌اند [۳۳]. به عنوان کار آینده، ما شبکه‌های مکرر نورون‌های اسپایکینگ را برای کاربرد تشخیص گفتار برای کلاسه‌بندی ارقام را بررسی خواهیم کرد تا عملکرد تشخیص را بهبود ببخشیم.

مراجع

[1] J. L. Lobo, J. Del Ser, A. Bifet, and N. Kasabov, "Spiking Neural Networks and online learning: An overview and perspectives," *Neural Networks*, vol. 121, pp. 88–100, Jan. 2020, doi: https://doi.org/10.1016/j.neunet.2019.09.004.

[2] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986, doi: https://doi.org/10.1038/323533a0.

[3]Stork, "Is backpropagation biologically plausible?," *International Joint Conference on Neural Networks*, 1989, doi: https://doi.org/10.1109/ijcnn.1989.118705.

[4] J. Wu, C. Xu, D. Zhou, H. Li, and K. C. Tan, "Progressive Tandem Learning for Pattern Recognition with Deep Spiking Neural Networks," *arXiv.org*, 2020. https://arxiv.org/abs/2007.01204

[5]J. Wu, E. Yilmaz, M. Zhang, H. Li, and K. C. Tan, "Deep Spiking Neural Networks for Large Vocabulary Automatic Speech Recognition," *arXiv.org*, 2019. https://arxiv.org/abs/1911.08373

[6] G. Eappen, S. T, and R. Nilavalan, "Cooperative relay spectrum sensing for cognitive radio network: Mutated MWOA-SNN approach," *Applied Soft Computing*, vol. 114, p. 108072, Jan. 2022, doi: https://doi.org/10.1016/j.asoc.2021.108072.

[7] J. Fitzgerald and KongFatt Wong-Lin, "Multi-Objective Optimisation of Cortical Spiking Neural Networks With Genetic Algorithms," *Ulster University Research Portal (Ulster University)*, vol. 71, pp. 1–6, Jun. 2021, doi: https://doi.org/10.1109/issc52156.2021.9467860.

[8] Yeshwanth Bethi, Y. Xu, G. Cohen, A. V. Schaik, and S. Afshar, "An Optimized Deep Spiking Neural Network Architecture Without Gradients," *IEEE Access*, vol. 10, pp. 97912–97929, Jan. 2022, doi: https://doi.org/10.1109/access.2022.3200699.

[10]A. Pitti, Mathias Quoy, C. Lavandier, and Sofiane Boucenna, "Gated spiking neural network using Iterative Free-Energy Optimization and rank-order coding for structure learning in memory sequences (INFERNO GATE)," *Neural Networks*, vol. 121, pp. 242–258, Jan. 2020, doi: https://doi.org/10.1016/j.neunet.2019.09.023.

[11]Y. Kim and P. Panda, "Optimizing Deeper Spiking Neural Networks for Dynamic Vision Sensing," *Neural Networks*, vol. 144, pp. 686–698, Dec. 2021, doi: https://doi.org/10.1016/j.neunet.2021.09.022.

[12]A. Woodward, T. Froese, and T. Ikegami, "Neural coordination can be enhanced by occasional interruption of normal firing patterns: A self-optimizing spiking neural network model," *Neural Networks*, vol. 62, pp. 39–46, Feb. 2015, doi: https://doi.org/10.1016/j.neunet.2014.08.011.

[13]W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, Dec. 1943, doi: https://doi.org/10.1007/bf02478259.

[14]"APA PsycNet," *psycnet.apa.org*. https://psycnet.apa.org/record/1959-09865-001

[15]K. S. Sayarkin, A. V. Popov, and A. A. Zhilenkov, "Spiking neural network model MATLAB implementation based on

Izhikevich mathematical model for control systems," vol. 8, pp. 979–982, Jan. 2018, doi: https://doi.org/10.1109/eiconrus.2018.8317253.
[16]"Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting (Computational Neuroscience) by Izhikevich, Eugene M. (2010) Paperback (Computational Neuroscience Series): Izhikevich, Eugene M. M: 9780262514200: Amazon.com: Books," *Amazon.com*, 2024. https://www.amazon.com/Dynamical-Systems-Neuroscience-Excitability-Computational/dp/0262514206 (accessed Nov. 08, 2024).
[17] www.izhikevich.com.
[18] J. Vreeken, "Spiking neural networks, an introduction." Accessed: Nov. 08, 2024. [Online]. Available: https://webdoc.sub.gwdg.de/ebook/serien/ah/UU-CS/2003-008.pdf
[19] H. Paugam-Moisy and S. Bohte, "Computing with Spiking Neuron Networks," *Handbook of Natural Computing*, pp. 335–376, 2012, doi: https://doi.org/10.1007/978-3-540-92910-9_10.
[20] I. Naruei and F. Keynia, "Wild horse optimizer: a new meta-heuristic algorithm for solving engineering optimization problems," *Engineering with Computers*, Jun. 2021, doi: https://doi.org/10.1007/s00366-021-01438-z.
[21]I. Ahmadianfar, O. Bozorg-Haddad, and X. Chu, "Gradient-based optimizer: A new metaheuristic optimization algorithm," *Information Sciences*, vol. 540, pp. 131–159, Nov. 2020, doi: https://doi.org/10.1016/j.ins.2020.06.037.
[22] H. Mostafa, "Supervised learning based on temporal coding in spiking neural networks," *arXiv.org*, 2016. https://arxiv.org/abs/1606.08165 (accessed Nov. 08, 2024).
[23] B. Rueckauer, I.-A. Lungu, Y. Hu, M. Pfeiffer, and S.-C. Liu, "Conversion of Continuous-Valued Deep Networks to Efficient Event-Driven Networks for Image Classification," *Frontiers in Neuroscience*, vol. 11, Dec. 2017, doi: https://doi.org/10.3389/fnins.2017.00682.
[24] A. Sengupta, Y. Ye, R. Wang, C. Liu, and K. Roy, "Going Deeper in Spiking Neural Networks: VGG and Residual Architectures," *arXiv.org*, 2018. https://arxiv.org/abs/1802.02627 (accessed Nov. 08, 2024).
[25]"Gradient-based learning applied to document recognition - IEEE Journals & Magazine," *Ieee.org*, 2019. https://ieeexplore.ieee.org/document/726791
[26] A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images," Apr. 2009. Available: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf
[27] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *arXiv.org*, 2014. https://arxiv.org/abs/1409.0575
[28] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. Maida, "Deep learning in spiking neural networks," *Neural Networks*, vol. 111, pp. 47–63, Mar. 2019, doi: https://doi.org/10.1016/j.neunet.2018.12.002.
[29] "Spiking Neural Networks: Learning, Applications, and Analysis: 9783845405155: Computer Science Books @ Amazon.com," *Amazon.com*, 2024. https://www.amazon.com/Spiking-Neural-Networks-Learning-Applications/dp/3845405155 (accessed Nov. 08, 2024).
[30] D. Ellis. Clean Digits and Digit Strings (Sound Examples), http://www.ee.columbia.edu/~dpwe/sounds/tidigits/
[31] M. Xu, L.-Y. Duan, J. Cai, L.-T. Chia, C. Xu, and Q. Tian, "HMM-Based Audio Keyword Generation," *Advances in Multimedia Information Processing - PCM 2004*, pp. 566–574, 2004, doi: https://doi.org/10.1007/978-3-540-30543-9_71.
[32] M. Xu, L.-Y. Duan, J. Cai, L.-T. Chia, C. Xu, and Q. Tian, "HMM-Based Audio Keyword Generation," *Advances in Multimedia Information Processing - PCM 2004*, pp. 566–574, 2004, doi: https://doi.org/10.1007/978-3-540-30543-9_71.
[33] A. Graves and N. Jaitly, "Towards End-To-End Speech Recognition with Recurrent Neural Networks," *proceedings.mlr.press*, Jun. 18, 2014. https://proceedings.mlr.press/v32/graves14.html