

ДОСЛІДЖЕННЯ КОМІТЕТУ НЕЙРОМЕРЕЖ У АВТОМАТИ- ЗОВАНІЙ СИСТЕМІ РОЗПІЗНАВАННЯ МОВЦІВ КРИТИЧНО- ГО ЗАСТОСУВАННЯ

Вінницький національний технічний університет;

Анотація

У тезах наведено результати синтезу комітету нейромережових класифікаторів для оптимізації процесу прийняття рішень автоматизованою системою розпізнавання мовців критичного застосування, який адаптовано до специфіки застосування критичних систем і дозволяє об'єднувати у комітет класифікатори різних типів.

Ключові слова: автоматизована система розпізнавання мовців критичного застосування, інформативні ознаки, комітет штучних нейронних мереж, розпізнавання мовців.

Abstract

The results of synthesis of neural network classifiers committee for decision-making process optimization by automated speaker recognition system of critical use are presented in the theses, which is customized to the specific of critical systems use and allows to combine in a committee of different types of classifiers.

Keywords: automated speaker recognition system of critical use, information features, neural networks committee, speaker recognition.

Вступ

Структура автоматизованої системи розпізнавання мовця критичного застосування у першому наближенні містить блок аналого-цифрового перетворення мовного сигналу, блок попереднього його оброблення, блок виділення інформативних для розпізнавання мовця ознак і блок класифікації. Кожен із цих блоків має значну кількість параметрів, вплив яких визначає якість розпізнавання мовця автоматизованою системою, яку традиційно оцінюють імовірністю правильного розпізнавання та імовірністю виникнення помилок першого та другого роду тестуючи новостворену систему на загальновідомих базах даних мовних сигналів.[1] Втім, остаточне рішення при функціонуванні автоматизованої системи приймає класифікатор, вид, архітектура, параметри та методи навчання якого потребують дослідження і оптимізації. Огляд існуючих методів побудови комплексних систем класифікації у задачі розпізнавання образів показав, що жоден з них не може бути безпосередньо використаним для оптимізації процедури класифікації у автоматизованій системі розпізнавання мовців критичного застосування, т.я. інформативні ознаки, що виділяються із мовного сигналу для подальшого розпізнавання мовця мають різне представлення – значення, множина, матриця і т.д., що вимагає використання класифікаторів різного типу, наприклад, загортальні нейромережі глибокого навчання і перцептрони, які за класичними методами у комітети не об'єднуються.[3] Також специфіка критичних систем полягає у максимізації достовірності роботи системи на протязі життєвого циклу її функціонування, що також слід урахувати при генеруванні комітета класифікаторів

Результати дослідження

В результаті проведених досліджень автори запропонували алгоритм синтезу оптимального нейромережевого комітета для автоматизованої системи розпізнавання мовців критичного застосування, тестування якого провели за даними із бази еталонних записів NOIZEUS. Для навчання нейромережових класифікаторів використовувалися 10% бази еталонних записів рівномірно розподілених за класами мовців та за співвідношенням шум/сигнал. Решта 90% еталонних мовних записів розділялися навпіл між контрольною та тестуючою вибірками.

При виборі множини інформативних для розпізнавання мовця ознак автори врахували специфіку процесу мовотворення (частота/період основного тону на вокалізованих фрагментах мовного

сигналу, коефіцієнти лінійного прогнозу, спектр Фур'є), особливості сприйняття мови людиною (кепстральні коефіцієнти) та динамічними характеристиками мови. Кожна із цих ознак має як переваги так і недоліки при її використанні для розпізнавання мовця, залежно від способу проведення процедури запису мовних сигналів, емоційного стану мовця, параметрів навколишнього середовища, наявними обчислювальними ресурсами і т.ін., але в комплексі вони утворюють цілісне представлення особи мовця у факторному просторі.[4] Загалом кількість елементів факторного простору, що описували особу мовця, дорівнювала 72 коефіцієнтам, які, відповідно до створеного алгоритму, авторами було рівномірно розподілено по чотирьом альтернативним факторним просторам (АФП) АФП1 - АФП4 відповідно.

Для створення як окремих нейромереж так і їх комітетів, автори застосовували можливості крос-платформної бібліотеки чисельного аналізу ALGLIB, яка підтримує основні мови програмування (C++, C#, Pascal, VBA) і кілька операційних систем (Windows, Linux, Solaris). Для навчання нейромереж у ALGLIB можна використати дві підпрограми - MLPEBaggingLM і MLPEBaggingLBFGS. Перша навчає індивідуальні нейромережі із допомогою алгоритму із сімейства Лівенберга-Марквардта, друга - із допомогою L-BFGS алгоритму, суттєвою перевагою останнього у випадку навчання комітетів нейромереж є те, що в результаті його роботи окрім безпосередньо навченого комітету також відбувається і внутрішнє оцінювання похибки узагальнення.[2]

Відповідно до авторського алгоритму було синтезовано базову нейромережу (трьохшаровий перцептрон), яка по чергово формувалася під класифікацію у кожному із чотирьох сформульованих вище факторних просторах.

Висновки

Проведені експериментальні дослідження із розпізнавання мовців доводять, що комітет нейромереж, сформований за авторським алгоритмом, показує меншу помилку узагальнення, ніж самостійна нейромережа оптимальної архітектури у виконанні задачі розпізнавання мовців для автоматизованої системи критичного застосування; Комітет нейромереж, навчений за допомогою L-BFGS алгоритму, показує меншу схильний до перенавчання, ніж традиційна нейронна мережа.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Фурман М.С. Дослідження нейромережевого класифікатора цифрових зображень / М.С. Фурман, В.В. Ковтун ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2016. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/10974>.
2. Биков М.М. Метод оптимізації процесу навчання нейромережі в задачі розпізнавання мовців / М.М. Биков, В.В. Ковтун, А. Раїмі ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2015. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/3584> .
3. Биков М.М. Метод нормалізації тривалості звучання паролних фраз для системи розпізнавання мовців / М.М. Биков, В.В. Ковтун, А. Раїмі ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2015. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/3585?locale-attribute=en> .
4. Биков М.М. Метод підвищення ефективності роботи пам'яті в системах пошуку ключових слів у мовному сигналі / М.М. Биков, В.В. Ковтун, К. Конате ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2015. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://visnyk.vntu.edu.ua/index.php/visnyk/article/view/1221> .

Гафурова Анжеліка Дамірівна — студентка групи 2АВ-136, Факультет автоматизації та комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця, e-mail.: fkca.2av13.gad@gmail.com

Ковтун В'ячеслав Васильович — доцент кафедри комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця

Hafurova Anzhelika D. — Department of Automation and computer control systems, Vinnytsia National Technical University, Vinnytsia, e-mail: fkca.2av13.gad@gmail.com

Kovtun Vyacheslav V. — docent of the Department of Computer control systems, Vinnytsia National Technical University, Vinnytsia