

## РЕЦЕНЗИЯ

върху дисертационния труд на маг. Десислава Николова Бояджиева на тема „Комбиниран подход за разпознаване на он-лайн подписи”

*Рецензент: проф. д-н Стефан Хаджитодоров*

### *КРАТКИ БИОГРАФИЧНИ ДАННИ*

Десислава Николова Бояджиева завършва висше образование във ФМИ на СУ „Св. Климент Охридски”. От 2000 г. до 2004 г. се обучава в курса за бакалаври по математика. От 2005 следва магистърския курс по био-медицинска информатика, който завършва през 2007 г. От 2008 г. е редовен докторант в Института по информационни технологии (сега И-т по информационни и комуникационни технологии) при БАН. Прекъсва на два пъти поради отпуск по майчинство.

### *АКТУАЛНОСТ НА ПРОБЛЕМА*

В продължение на десетилетия проблемът за идентификация/верификация на подпис се изучава и развива основно за целите на криминалистическите почеркови експертизи и осигуряването на достъп до различни системи, операции и помещения. С развитието на изчислителната техника и методите за обработка на изображения и разпознаване на образи се разработват компютърни системи, целящи обективизиране на експертизите. С появата в последните години на устройства за цифрова регистрация на подписи и използването им в различни сфери на живота: банково дело, бизнес отношения, системи за достъп, биометрична идентификация и др., системите за он-лайн верификация на подпис намират все по-широко приложение. Поради особеностите обаче на електронно поставения и електронно регистриран подпис и подписа, поставен върху документ с обикновено пишешо средство, разработените методи за оф-лайн подписи не могат да се прилагат директно за он-лайн, което налага разработването на нови методи, въвеждането на нови множества от признаци, разработването на бази данни и създаването за удобни графични интерфейси. По тези причини задачата за он-лайн верификация на подписи е с безспорна актуалност.

### *ПОЗНАВАНЕ СЪСТОЯНИЕТО НА ПРОБЛЕМА И ОБЩА ХАРАКТЕРИСТИКА НА ТРУДА*

Използваните 108 литературни източници, основната част от които от последните 10-15 години, са дали възможност на докторантката да се запознае с различните аспекти на изследвания проблем и най-новите резултати в света. Разработката на задачата показва, че тя е усвоила творчески използваните методи и техники за обработка и е в състояние да ги прилага успешно.

Поставеният върху таблет подпис се запаметява като последователност от точки, всяка от които се определя от равнинните си координати и силата на приложения в нея натиск, т.е. съхраняваният подпис има характеристики на двумерно сиво изображение. Това определя и използваната методика, включваща избор на геометрични признаци, признаци на натиск и наклони на писалката, минимизирането на техния брой, избор на метрика за оценка на сходството и правило за вземане на решение. Тези елементи са налице в дисертационния труд, което показва, че избраната методика е обоснована и съответства напълно на формулираната изследователска задача.

Дисертацията обхваща 129 стр., включващи увод, 5 глави, 18 фигури, 19 таблици, заключение, списък на приносите, публикации по темата и библиографска справка.

В първа глава се посочват особеностите на известните варианти за разпознаване на лица по подпис в зависимост от използваните сензори, модалности и методи за класификация. Представени са различията между системите он- и оф-лайн, етапите на изграждането им и съпътстващите този процес трудности. Направен е критически анализ на разработките в тази област от последните години. Подчертано е вниманието, което напоследък се отделя на разработването на комбинирани подходи, паралелни, последователни, хибридни. В резултат на направения анализ в края на главата са посочени целта на дисертационната работа и задачите, които предстои да се решат. Те са коректно формулирани.

Във втора глава се разглежда избора на признаци за описание на он-лайн подписи. Формулирани са основните изисквания, на които трябва да отговарят използваните признаци. Поради различното разположение на подписа върху таблетния екран и различната динамика при изписването се налага предварителна обработка, свързана с отстраняване на повтарящи се точки, ротация и транслация на подписа. Представен е списък на използваните първични геометрични, динамични и свързани с натиска признаци. Поради това, че някои от признаците са свързани функционално или статистически, т.е не носят допълнителна информация, се налага тяхното отстраняване. За тази цел се прилагат последователно метода на корелационните плеяди и регресионния анализ със Ср – статистиката.

В трета глава се разглежда въпроса за построяване на класификационни правила и оценка на качеството им. Основно внимание е обърнато на апарата на невронните мрежи. Дискутирани са въпросите за архитектурата и обучението. Анализирани са предимствата и недостатъците им и приложението им за разпознаване на подписи. Вторият основен метод за вземането на решение е К-най-близките съседи (K-NN). Следва да се отбележи, че K-NN метода е свързан до голяма степен със статистическите характеристики на класовете, тъй като фактически представлява емпирична оценка на условните плътности на разпределение на вероятности на класовете в дадената точка. Посочени са основните параметри за оценка на точността на класификаторите.

За подобряване на качеството на класификаторите при малки извадки се прилага метода на N- кратната крос-валидация. Този метод е подходящ за намиране на „добра“ стойност на K при K-най-близки съседи.

Главата завършва с представяне на модел на процеса на разпознаване на подпис в постановката на теорията на обобщените мрежи. Моделът представлява принос към приложението на апарата на обобщените мрежи в теорията на разпознаване на образи.

В четвърта глава са описани резултатите от проведените експерименти. Използвани са две бази данни – собствена, състояща се от по 10 истински подписа на 8 потребители и по 9 умели фалшификати и интернет достъпната база SUsig, състояща се от 8-10 истински и 10 умело фалшифицирани подписа на 89 участници. За намаляване броя на признаците и в двата случая се прилагат последователно методът на корелационните плеяди и Ср статистиката. Получено е значително редуциране на броя на признаците: за първата база при първоначален брой 26, общият брой признаци се свежда до 17, а необходимият индивидуален брой варира от 5 до 13 признака, а за втората от 24 начални общият брой се свежда до 13, а индивидуалният варира от 3 до 9 признака. Това показва, че само малка част от първоначално избраните признаци е достатъчна за задоволително

решение на задачата. Сравнението на средната точност между невронните мрежи (НМ) и  $k$ -NN по общи и индивидуални признаци показва превъзходство на НМ.

В пета глава е представена архитектурата и функционалностите на разработената от докторантката система за разпознаване на он-лайн подписи. В нея са реализирани описаните в предишните глави методи за събиране и обработка на подписи, създаването на база данни, обучението на моделите невронни мрежи и избор на най-добрия от тях. Графичният интерфейс е със съвременен дизайн и удобен за работа.

### *ХАРАКТЕР НА ПРИНОСИТЕ*

Получените в дисертацията резултати могат да бъдат класифицирани основно като научно-приложни и приложни. Към научно-приложните приноси се отнасят следните:

1. Предложен е последователен подход за минимизиране на признакови пространства.
2. Предложен е комбиниран подход за класификация на базата на многослоен перцептрон с използване на най-добрия за всеки участник класификатор.
3. Разработката на обобщено-мрежов модел на процеса за разпознаване на он-лайн подписи.
4. Оценка на точността на разпознаването при използване на общи и индивидуални подмножества от признаци при наличие на умели и неумели фалшификати.

Към приложните приноси отнасям:

1. Разработената софтуерна система, реализираща цялостния процес на получаване и обработка на данни и разпознаване на подпис.
2. Разработената и събрана за целта собствена база данни.

### *ЛИЧЕН ПРИНОС НА ДИСЕРТАНТА*

Освен определянето на основните насоки на работата и методичното ръководство от страна на научния ръководител извършената изследователска и експериментална работа са лично дело на докторанта. Това личи и от първото място, което заема като автор във всички публикации по дисертацията.

### *ПУБЛИКАЦИИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИЯТА*

По темата на дисертацията са направени 6 публикации: 5 отпечатани и една приета за печат. Една от публикациите е самостоятелна, останалите в съавторство с научния ръководител. Всички с изключение на една са на английски език. Две са на конференции в чужбина (Москва и Мадрид), три са на конференции в България и една - в списанието CIT (Cybernetics and Information Technologies). Представените в допълнителен списък две публикации намирам за несвързани с темата на дисертацията.

### *АВТОРЕФЕРАТ*

Авторефератът от 37 страници отразява правилно съдържанието на дисертацията, направените изследвания и получените резултати.

### *КРИТИЧНИ БЕЛЕЖКИ*

Към предложения дисертационен труд имам следните критични бележки и въпроси:

1. Не става ясно защо в труда са използвани само глобални признаци, но не и локални. Това естествено предполага използването и сравнението и с други методи за класификация като скрит Марковски модел, динамично времево изкривяване, друг

тип невронни мрежи и др. В такъв случай изследването щеше само да спечели и да бъде цялостно.

2. На някои места има неточни твърдения и объркани понятия и формули:
  - на стр. 68 – когато има случаи с признаци, които по стойност силно се различават от стойностите на този признак в група от няколко случая, то се премахва не признака, а случая (реализацията), за която има силно различна стойност на признака;
  - на стр. 70 – с нарастване на прага броят на лъжливо отрицателните и истински отрицателните не намалява, а се увеличава;
  - на стр. 73 – когато дисперсията е с висока стойност и стандартното отклонение е с висока стойност, а не както докторантката твърди, че са различни;
  - на стр. 85 – налице е матрица-ред и стойностите  $[0,1]$  са различни във всеки от стълбовете, а не в редовете на матрицата;
  - на стр. 121 – докторантката твърди, че е предложила „комбиниран“ подход за създаване на класификатор. Според мен, предвид същността на подхода, той по-скоро е „интегриран“ подход.

Към работата имам следните препоръки:

1. Да се разшири обема на собствената база от подписи и се подготви публикация за някои от известните списания по разпознаване на образи.
2. Да се потърси възможност за предоставяне на разработената система на заинтересовани потребители.
3. Да се експериментират многонивови схеми с комбиниране на различни класификатори, така наречените ансамбъл от класификатори.

Независимо от забележките, в мен не остава никакво съмнение в стойността на дисертационния труд.

### **ЗАКЛЮЧЕНИЕ**

В дисертационният труд на маг. Десислава Бояджиева е извършено сериозно по обем и качество изследване върху един актуален проблем с възможност за приложение на получените резултати в практиката. Впечатлението ми е, че докторантката е добре информирана в предметната област и умело прилага съвременните изследователски методи в обработката на изображения и разпознаването на образи. Направените публикации показват, че получените в дисертацията резултати са направени достояние на достатъчно широка аудитория от наши и чуждестранни учени. Всичко това ми дава основание да считам, че представеният дисертационен труд „Комбиниран подход за разпознаване на он-лайн подписи“ удовлетворява изискванията на Закона за академично израстване за получаване на образователната и научна степен „доктор“ и предлагам на уважаваните членове на Научното жури да гласуват за присъждането на тази степен на маг. Десислава Николова Бояджиева.

Дата: 15.04.2014 г.