

РЕЦЕНЗИЯ

от проф. д-р инж. Борислав Димитров Маринов
на дисертационния труд на гл. ас. инж. Христо Стоянов Николов
на тема: **“Приложение на нелинейни методи от теорията за разпознаване на образи в
дистанционните изследвания на Земята”**

за присъждане на образователната и научна степен **“доктор”** -
област на висше образование **4. Природни науки, математика и информатика,**
професионално направление **4.4 Науки за Земята,**
научна специалност **“Дистанционни изследвания на Земята и планетите”**

Представената рецензия е изготвена в съответствие с изискванията на ЗРАСРБ и Правилника за неговото прилагане и чл. 20 (3) от Правилника за условията и реда за придобиване на научни степени на ИКИТ-БАН и на основание на заповед на директора на ИКИТ-БАН за назначаване на Научно жури (зап. № 115 / 09.10.2018 г.) и решенията на журито, взети на неговото първо заседание, проведено на 15.10.2018 г.

1. Общи сведения за процедурата

Инж. Христо Николов е зачислен за докторант на самостоятелна подготовка към секция “Системи за дистанционни изследвания” на ИКИТ-БАН, считано от 09.07.2002 г. Научни консултанти на докторанта са били доц. д-р инж. Дойно Петков и проф. д-р Бойчо Бойчев. Той е положил специализираните си изпити по Дистанционни изследвания през м. февруари 2003 г. с отличен успех (5,75) и през м. юли 2006 г. с отличен успех (5,75). Докторантът е отчислен с право на защита на 12.07.2016 г. Процедурата за защита е открита след окончателното разглеждане на работата пред разширен състав на секцията “Системи за дистанционни изследвания”, проведено на 25.07.2018 г. Решението за избор на жури и откриване на процедура за защита е взето на заседание на Научния съвет на ИКИТ-БАН от 04.10.2018 г. (Протокол № 33).

По процедурата за защита докторантът е представил следните материали:

- автобиография;
- заповед за журито № 115/09.10.2018 г.;
- копия от протоколите за изпитите - 2 броя за специализирани изпити по Дистанционни изследвания;
- дисертация;
- копия на публикациите по дисертацията;
- автореферат на дисертацията.

2. Сведения за кандидата

Христо Стоянов Николов е роден на 19 март 1966 г. Той придобива образователно квалификационна степен магистър с квалификация “Инженер по автоматизация на минното производство” в Минно-геоложкия университет “Св. Иван Рилски” (Минно-геоложки институт), където следва в периода 1984-1990 г. През периода от м. януари 1998 г. до м. декември 2002 г. работи като инженер по изграждане и поддръжка на локална мрежа, системна поддръжка на сървъри към Геофизичния институт на БАН, в периода от м. януари 1990 г. до м. юни 1997 г. работи като софтуерен инженер - специализирано програмно осигуряване към “Лабораторията по слънчево-земни въздействия” на БАН, а в периода от м. юли 1997 г. до м. юни 2010 г. е научен сътрудник от III степен до I степен по геоинформатика и моделиране с ГИС, обработка на данни и изображения към “Лабораторията по слънчево-земни въздействия” на БАН. От м. юли 2010 г. до момента е главен асистент по геоинформатика и моделиране с ГИС, обработка на данни и изображения в ИКИТ-БАН.

Като научен сътрудник към “Лабораторията по слънчево-земни въздействия” на БАН е водил упражненията по дисциплината “Дистанционни изследвания и обработка на изображения” към Минно-геоложкия университет “Св. Иван Рилски”.

Инж. Христо Николов е представил 5 публикации, които са включени в неговата дисертационна разработка. Своите работи той е публикувал на международни симпозиуми и конференции в България и на конгреси в чужбина.

3. Анализ на дисертационния труд

Дисертационният труд е оформен в общ обем от 114 стр., от които 110 стр. текст, списъци на фигурите, таблиците и използваните съкращения. Текстовата част включва съдържание и разработката на дисертационния труд в обем от 101 стр., с включени 68 фигури и 27 таблици, както и справка за приносите, списък на публикациите по дисертацията, списък с използвана литература.

Дисертационният труд е оформен в пет глави, като първата глава е уводна.

3.1. Цели и задачи на работата

Целите и задачите на разработката са формулирани ясно в уводната първа глава. Темата на дисертационния труд е актуална и има важно практическо значение за подобряване на точността и достоверността на резултатите от класификацията на многоканални спътникови данни, използвани за анализи в различни стопански отрасли и научни области. Кръгът на поставените задачи включва тематична класификация на обекти от земната повърхност, посредством използването на нелинейни модели. Формулирани са теоретичните аспекти, свързани с прилагането на адекватен математически апарат при класификацията и практичен аспект, който се състои в алгоритмизиране на прилаганите методи и разработване на програмни продукти на езици от високо ниво и тяхното интегриране в цялостна програмна система.

Предмет на изследването са нелинейни методи и модели, прилагани за класификация на многоканални спектрални данни (МКСД) за изследване на антропогенното въздействие в Панагюрския минен район за годините 1999 и 2006.

Основните цели на дисертацията са формулирани като създаване на иновативна методика за подбор и предварителна обработка на МКСД, избор и тестване на използваните модели при обработката и разработване на система от критерии за оценка на резултатите, получени при използване на различни методи за класификация.

Поставените цели са конкретизирани в поставените за решаване на задачи по създаване на геобаза данни (ГБД) за Панагюрския минен регион, нейното обновяване и актуализиране във времето, времеви анализ на промените в класовете на ГБД, теоретично обосноваване на прилагането на нелинейни методи от теорията за разпознаване на образи и тяхното адаптиране за тематична класификация на обекти, включването на допълнителна информация, различна от спектралните данни, формулиране на паралелни процедури, ползващи статистически методи, съвместно с модели на невронни мрежи и метода на опорните вектори, тяхното експериментално прилагане и валидиране на получаваните резултати и модели, интегриране на получените модели на разпознаващи процедури като външни модули в съществуващи програмни продукти за обработка на данни от дистанционни изследвания.

3.2. Съдържание на дисертационния труд

Първа глава от работата е въвеждаща и включва въпросите, засягащи обекта и обхвата на дисертационния труд, актуалността на проблема и съдържа формулировките на целите и задачите на работата.

Втора глава е посветена на проблематиката на дистанционните изследвания. Разгледани са областите на приложение и решаваните задачи, същността на експеримента при дистанционните изследвания, отразената и собствена радиация на природните образувания, методите и апаратурата за регистриране на отразената радиация във видимата и близката инфрачервена зона. **Първият раздел** формулира приетата в дисертационния труд дефиниция относно физическата същност на ДИ съгласно (Мишев 1981) - "дистанционните изследвания са свързани с измерване на количествени параметри на електромагнитното поле, излъчвано или отразено от природните и антропогенни образувания". Дистанционните изследвания са отнесени към категорията на недеструктивните методи за получаване на данни за конкретен обект, при които информация за изследвания обект се получава чрез провеждане на измервания от определено разстояние без директно съприкосновение с обекта. Дефинирани са научните и приложни области при дистанционните изследвания (ДИ) – метеорология, геофизика, геология, хидрология, селско и горско стопанство.

Вторият раздел на втора глава представя същността и особеностите на експерименталното изучаване при дистанционните изследвания, което се основава на специфичното взаимодействие на различните природни и антропогенни обекти със слънчевата

радиация. Посочени са основните компоненти, които участват в процеса на получаване на данни от дистанционни изследвания: обект на изследване, източник на електромагнитна енергия, преносна среда и регистрираща апаратура.

В *третия раздел* на втората глава са анализирани физическите принципи на излъчвана електромагнитна енергия в зависимост от параметрите на излъчващия източник и спектралните отражателни характеристики на обектите. Разгледана е същността на закона на Планк, който характеризира спектралния състав на излъчваната от абсолютно черно тяло енергия в зависимост от температурата, както и получаването на закона на Стефан-Болцман чрез неговото интегриране. Представени са също примерни спектрални отражателни характеристики на основни типове естествени или изкуствени обекти.

В *четвъртия раздел* на втората глава е направен преглед на методите и апаратурата за измерване на отразената радиация във видимата и близката инфрачервена зони на електромагнитния спектър. Дефинирани са основните компоненти на апаратурата, предназначена за регистриране на отразената радиация, а в някои случаи на собствената радиация от обекта. Посочени са особеностите при регистриращи системи за многоканални данни и особеностите на формиране на хиперспектралните изображения.

Основните изводи, обобщени от изложението във втора глава, са следните:

- дистанционните изследвания (ДИ) на Земята са съвременна и бързоразвиваща се мултидисциплинарна област, обединяваща постижения с теоретичен и приложен аспект и позволяваща решаване на широк спектър приложни задачи;

- ДИ се основават на теоретични закони от различни дялове на физиката, което е предпоставка за създаване на иновативни инженерни и технологични решения за получаване на МКСД от ДИ;

- информацията, извлечена от МКСД позволява установяване на моментното състояние на обекти от повърхността на Земята и проследяване на тяхната времева динамика;

- информацията от ДИ е надежден източник за създаването на тематични карти и разработване на модели за вземането на обосновани решения при управление на територии;

- съотношението цена/разходи при крайни продукти, получените от ДИ е на порядък по-ниско в сравнение със сходни такива, получени по други методи.

В *трета глава* на работата са разгледани принципите на получаване на многоканални спектрални изображения, предварителната обработка и подобряване на идентификацията и интерпретацията, като и същността на основните методи за класификация. *Първият раздел* на трета глава съдържа информация за видовете, принципа на действие и характерните особености на многоканалните системи за регистрация на изображения. Посочени са основните принципи на разделяне на спектралните канали. Получаваните изображения са класифицирани като многоканални или хиперспектрални според броя и ширината на използваните канали. Изяснени са методите за визуализация на изображенията, анализирани са радиометричната разделителна способност и процедурите за преобразуване на стойностите пикселите в отражателна способност на обекта (reflectance).

Във *втори раздел* на трета глава е формулирана същността и са описани особеностите при формиране на еталонни многоканални спектрални данни. Процесът е част от процедурата на създаване на геобаза данни. Съдържаната информация включва освен спектралните отражателни характеристики и допълнителни данни като влажност, температура и условия на заснемане. Прилаганият подход осигурява възможност за дешифриране на съответния природен или антропогенен обект в различни изображения, а така също и възможност и дешифриране на други обекти от същия клас в същото или други изображения.

Третият раздел на трета глава описва процедурата за предварителна обработка на данните. Разгледани са особеностите на трите основни типа корекции: спектрометрична, радиометрична и геометрична. Систематизирани са източниците на радиометрични грешки.

Разгледани са геометричните корекции, които коригират геометрични неточности, предизвикани от положението на регистриращата апаратура в пространството и изкривяванията, внасяни от оптичната система на апаратурата. Формулирани са двата основни метода за тяхното отстраняване, посредством контролни точки от земната повърхност (GCP) или чрез трансформация между две изображения, едното от които е прието за еталон.

Описани са особеностите на изображения със средна разделителна способност от сензорите TM, ETM+ и OLI както и прилаганите процедури за радиометрични корекции,

преобразуващи първичните данни до данни за енергийна яркост, регистрирана от сензора и до стойности на отражение. Разгледани са корекции, отчитащи промени в средната яркост поради разположението в осветената или сенчестата част на склона, влияние на склона, изложение, вторично излъчване и други.

В **четвъртия раздел** на трета глава се разглежда интерпретираща процедура на основата на класифицираща процедура с обучение. Формулиране е същността на процедурата за тематична класификация, при която за всеки конкретен обект трябва да осигури съответствие между предварително формирани спектрални класове и търсени информационни класове. Тя се основава на метода на групирането на отделните видове природни и антропогенни образувания в пространството на характеризиращите признаци (ХП). Систематизирани са основните задачи, които трябва да се решат в процеса на нейната реализация.

За създаване на процедура, извършваща идентификация на обекти от земната повърхност на базата на МКСД от ДИ, е необходимо да бъдат намерени подходящи дискриминантни функции, осигуряващи надеждно разделяне на обектите в пространството на ХП, получени на базата на теоретични предпоставки или изведени емпирично.

В процеса на класификация на данни от ДИ се прилагат два метода – класификация без обучение (кластеризация) и класификация с обучение, като разглеждани самостоятелно всеки един от споменатите методи има своите ограничения. Препоръчва се прилагането на комбинация от двата метода с цел максимална ефективност, при отсъствие на достатъчно допълнителна информация за конкретния район на изследване.

Методът за класификация без обучение предполага прилагане на алгоритъм за намиране на “естествено” групирането в пространството на ХП обекти от цялото многоканално спектрално изображение (МКСИ). След края на процедурата се допуска, че всяка формирана група (кластер – спектрален клас) представлява разпределението на вероятността за всеки един клас от информационните класове, представляващи интерес.

Прилагането на класификация с обучение на МКСД се основава на предположението, че е възможно на базата на данни от наземни изследвания, карти, снимки и др. да се локализируют малки хомогенни области в МКСИ, в които със сигурност са разположени обектите, представляващи интерес за конкретното изследване. От тях се формират извадките за обучение и валидация на модела, който ще се създава за извършване на тематичната класификация.

В **петия раздел** на трета глава се разглежда класификацията в пространството на признаците чрез функция на разстоянието. Описани са методи от най-опростените към по-сложните. Формулирани са паралелепипедният метод и методът, използващ Махаланобиусово разстояние, а като най-често прилагани разстояния са посочени евклидовото разстояние и разстоянието, формирано като сума от абсолютните стойности на разликите. Посочено е, че често използван критерий е минимаксният, който минимизира разстоянията между точките в кластера и максимизира разстоянията между кластерите. Прави се извода, че най-често използваните алгоритми за кластеризация са на К-средните и Итеративния Само-Организиращ се Метод за Анализ на Данни- (ИСОМАД), известен още като ISODATA. Посочено е различието между тях по отношение на дефинирания брой кластери - променлив при метода на К-средните и предварително зададен при ИСОМАД.

В **шестия раздел** на трета глава се разглежда статистическият подход при класификация с обучение при многоканални данни, като се дефинират условията за неговото прилагане:

- минимизиране влиянието на случайните компоненти (например атмосфера, регистриращ елемент, промяна в обекта) върху данните, водещи до повишаване на грешката при класификация;

- възможност за получаване на оценки относно степента на принадлежност към даден клас в случай на припокриване на изследваните класове в пространството на признаците.

Систематизирани са основните изисквания и ограничения при прилагане на статистическия подход.

Формулирана е същността на оптималната стратегия за класификация по правилото за максимално правдоподобие, при която са дефинирани n класа, със зададени априорни вероятности и функция на плътност на вероятност. Основният критерий, по който се създава

модела при процедурата за тематична класификация, се състои в *минимизиране на грешката при класификация на всеки един от пикселите от изображението към определен, предварително дефиниран клас.*

На основата на изложеното в тази глава са формирани следните изводи:

- съвременните спектрометрични системи за получаване на МКСД, инсталирани на спътниково и аеро базирани комплекси, са надежден източник на данни за обекти от земната повърхност;

- основен източник на информация осигуряват актуални и архивни МКСД със средна ПРС поради възможността за безвъзмезден достъп и наличието на програмно осигуряване за тяхната предварителна обработка;

- предварителната обработка на МКСД от спътникови и аеро носители има съществено значение, тъй като позволява тяхното съвместно използване с данни от спектрометрични in-situ измервания, което осигурява създаване на еталонни обекти в пространството на ХП;

- коректната интерпретация на спектралните класове, налични в МКСД, като съответни информационни класове изисква дефиниране на пространство на ХП, извършване на кластеризация (по един от двата основни метода, базирани на статистическия подход), прилагане на дискриминантни функции, които дефинират подобластите за всеки един информационен клас.

Четвърта глава на работата съдържа детайлно описание на онези процедури от разпознаването на образи, които са обект на анализ в дисертационния труд. Те са формулирани като приложение на методологията на разпознаване на образи за класификация на данни от дистанционните изследвания.

Разгледани са теоретичните основи на два метода за създаване на модели за разпознаване на образи с предварително обучение на разпознаващата процедура от налични данни (learning machine approach), които се основават на различни подходи, но основна тяхна същност е възможността за групиране в по-големи или малки области от пространството на характеризиращите признаци. Функциите, които описват разделящите ги повърхнини, в повечето случаи са нелинейни, което е сериозна пречка при използването на статистическите методи, описани в предходната глава.

Специфична особеност при използване на МКСД за формиране на модел за класификация с обучение, дължащ се на крайния обем на данните, използвани за обучение на модела, се проявява в това, че при увеличаването на броя на използваните ХП с цел повишаване на точността, се достига до такива стойности на техния брой, след които точността спада - "феномен на Хю" Причина за това е затрудненото моделиране при големи размерности на пространството на признаците.

Анализът на методите, описани в предходните глави на дисертационния труд, позволява да се направи извода за предимството на приложението на нелинейни методи за разпознаване на образи, тъй като те позволяват използване на нелинейни дискриминантни функции за разделяне на спектралните класове в пространството на ХП, което обуславя мотивацията за последващото провеждане на експериментите с тях, в резултат на което са генерирани непараметрични модели, за които не е необходимо предварително да са известни функциите на вероятностите за разпределение за данните.

Формулираните в дисертационния труд предпоставки, обуславящи прилагането на нелинейните методи са:

- развитието на теоретични постижения, което води до предпочитането на този тип модели пред методите, ползващи статистическия подход, чрез което се осигурява подобряване на крайната точност;

- наличието на категорията от непараметрични методи, които не предполагат нормално разпределение, поради което могат да се прилагат и в пространство на ХП с висока размерност без да изпитват "ефекта на Хю" и съответно са по-неподатливи на преобучаване.

В работата обстойно са разгледани два основни нелинейни подхода, използвани при интерпретация на изображения: метод на невронните мрежи и метод на опорните вектори.

Същността и процедурата на функциониране на *невронните мрежи* са разгледани в *раздел първи* на четвъртата глава. Описани са основните параметри и характеристики на невронните мрежи и начина на формиране на мрежовата топология. Детайлно е проследен процеса на обучение на невронната мрежа, изразяваща се в нейната пластичност, състояща се в

три процеса - поява на нова връзка, модификация на съществуваща или изчезване на съществуваща, както и в промяна на тежестните коефициенти между невронните елементи. Анализирани са процедурата за преодоляване на локалните минимума и достигане на глобален минимум на грешката.

Основните изводи, които се правят за метода на НМ са следните:

- за моделите с НМ най-важна е топологията на мрежата и стойностите на тежестните коефициенти, свързващи обработващите елементи в мрежата, от които зависи достигането до оптимален модел;

- на етап "създаване на модел" с НМ от особена важност е достигането до глобален минимум в повърхнината на грешката, което се постига с избора на подходящи начални параметри за този етап;

Същността на **метода на опорните вектори** (МОВ) и неговото прилагане при обработката на многоканални изображения са разгледани във **втория раздел** на четвъртата глава.

Теорията на опорните вектори (Support Vector Machines – SVM) се прилага при разработването на метода на опорните вектори (МОВ) за класификация на данни. Основна роля при него имат така наречените опорни вектори, които представляват онези точки от данните, които са разположени най-близко до линията или повърхнината, разделяща множествата и които като правило е най-трудно да бъдат коректно класифицирани. Моделите, изградени по МОВ, притежават по-добра способност за обобщаване за сметка на компромиса между точност и способност на модела за точно възпроизвеждане на конкретно обучаващо множество (капацитет на модела).

Основно предимство на МОВ е възможността за по-добра разделимост на класовете с предварително избрана функция, която най-често е нелинейна. Класификацията с нелинейни функции по МОВ теоретично е директно приложима и при линейно разделими класове, но ако данните в избрано пространство на ХП са линейно неразделими или грешката, получавана при линейно разделяне е неприемлива, то за тяхното разделяне се налага използване на нелинейни функции. Тяхното прилагане притежава недостатък, че моделът не достига до стабилно състояние, което налага да се извърши трансформация на данните от пространството на ХП в пространство с по-висока размерност, посредством трансформация с функционални ядра (позната в литературата като kernel trick).

Основните изводи, направени за МОВ, са следните:

- за моделите по МОВ разделянето на данните се реализира от линия или равнина независимо от трансформациите, на които се подлага пространството на ХП и от неговата размерност;

- независимо от приложената стратегия за достигане до оптимален модел за класификация с НМ и по МОВ основно предимство представлява способността на изградения модел за обобщаване, а не фокусирането изключително върху точното възпроизвеждане на данните от обучаващата извадка.

Експерименталното изследване на разгледаните нелинейни методи е представено в **глава пета** на работата, която е най-обширна.

В **раздел първи** на пета глава е описана методиката за извличане на единични спектри от МКСД. Резултатите от стъпките на първична обработка и получаваните векторни слоеве за региона на анализ са представени графично и таблично.

Във **втори раздел** на пета глава са описани програмните продукти, използвани в процеса на анализа. Това са програмата Neumapper, програмен компонент за невронна мрежа, която е част от геоинформационна система QGIS.

В **трети раздел** на пета глава са представени програмни продукти, които осъществяват класификация на МКСД по метода на опорните вектори. Разгледан е програмният компонент по МОВ като част от геоинформационната система QGIS, като има възможност за ползване на два варианта: SVM като част от по-ранна версия на ОТВ или libSVM като заимствана от библиотеката за компютърно зрение OpenCV. Друга възможност се предоставя от програмния компонент по МОВ като част от програмната система Opticks. В нея се ползва предложението от L. Bruzone алгоритъм за автоматичен анализ на разликово изображение, получено при класификация без обучение на мултitemпорални изображения.

В *четвърти раздел* на пета глава са сравнени осем варианта на експерименти с разглежданите програмни продукти.

Първият вариант анализира експерименти за класификация на МКСД по метода на класификация без обучение с учител по метода на К-средните и ИСОМАД. Сравнено е формирането на 22 класа, тъй като такъв е броят на наличните класове, получени от проекта Корине земно покритие за изследвания регион. Резултатите, получени по двата метода са сравнени таблично и графично като ползвани стълбовидни диаграми за процентното разпределение на формираните класове. Наличието на пиксели от тип “смесен клас” в обектите от обучаващата извадка е причина за получаване на невисоки стойности за точността при класификация по метод с предварително обучение – НМ и МОВ.

Вторият вариант анализира методика за съвместно прилагане на статистически подход и класифицираща процедура използваща невронни мрежи, което позволява да се повиши точността при класификация като се обединят предимствата на двата метода, въпреки нарасналата сложност при реализация на цялостната обработваща процедура.

Комбинацията включва метода на максималното правдоподобие, който е най-успешно прилаганият в дистанционните изследвания статистически метод за класификация и класифицираща процедура, използваща изкуствена невронна мрежа (НМ), чиято структура е от тип многослоен персептрон с входен слой, 1 скрит (обработващ) слой и един изходен слой. Схемата за комбинирана класификация на многоканални данни от дистанционни изследвания е изградена от три обработващи блока, при която първичните данни се подават на две паралелно работещи процедури за класификация, едната използваща статистически метод, а другата невронна мрежа, а “арбитражната” процедура се осъществява от НМ.

Реализацията на метода на опорните вектори чрез различни програмни продукти е анализирано във вариантите от трети до шести.

Третият вариант използва класификация по МОВ с програмния компонент ОТВ. Ползвани са обработващите модули SVM и libSVM. Като критерий за качество на изградения модел след етап “създаване на модел” се използва обобщената за всички класове оценка $kappa$, получена от матрицата на грешките, получена след класификация на данните, използвани на етап “валидация на модел”. В зависимост от стойността на флага за оптимизация анализът чрез валидация се извършва еднократно или се извършва многократна оптимизация, като за оптимални се приемат параметрите на модела, за които грешката получена при кросвалидация е минимална.

Четвъртият вариант анализира получените резултати при прилагане на класификация по МОВ с обработващ модул SVM - ОТВ.

Резултатите от етап “обучение на модел” с различни функционални ядра, които ще се използват за последваща тематична класификация, са представени таблично, за различни размери на обучаващата извадка и извадката за валидация, при различни стойности на основните параметри C и γ . Оценката на качеството се извършва чрез обобщената оценка $kappa$.

За варианта „обучение на модел“ по МОВ при прилагане на функции с радиален базис (ФРБ) получаваните стойности за приетия параметър за оценка на качеството на класификация варират в сравнително тесни граници. При този метод за създаване на модел промяната на основните параметри и размерите на извадките за обучение и валидация не оказват съществено значение за подобряване на точността на класификация. Основно негово предимство представлява краткото време за обучение на модела.

Петият вариант анализира получените резултати при прилагане на класификационна процедура по МОВ с обработващ модул libSVM - ОТВ.

Сравнени са резултати от модели изградени от обработващ модул libSVM и няколко вида функционални ядра – линейно, функция с радиален базис и полином. И в този случай основен параметър, оказващ влияние върху качеството на всеки един от създаваните модели в процеса на неговото обучение, е коефициент C на модела, а оценка на качеството на всеки един създаден модел бе извършено посредством обобщената оценка $kappa$.

Резултатите за различните ядра са сравнени таблично. Моделът с функционално ядро полином, създаван от този модул дава по-добри резултати от другите две ядра, като предоставя най-добри стойности за $kappa$, но за сметка на по-продължително време за обучение в

сравнение с другите два варианта. Резултатите за линейно и RBF ядро са представени и графично в две изображения.

За тази програмна реализация на МОВ са извършвани допълнителни експерименти за изграждане на модел посредством библиотека ОТВ – самостоятелен модул с графична среда и изпълнение на алгоритъма от команден ред. Очакванията за постигане на по-добро качество или намалено време на създаване на модела не се потвърждават и поради тази причина са провеждани експерименти в средата на QGIS, която позволява по-добър контрол върху изпълнението на процеса на създаване на модела.

Шестият вариант анализира получените резултати при прилагане на класификация с обработващ програмен компонент *SVM - Opticks*, но само за случая на линейно функционално ядро.

Прилагането на този програмен модул се реализира в няколко стъпки: предварително геореферирание на векторните файлове за районите към референтен елипсоид WGS84 и въвеждане в геобаза данни в програмата, преобразуване в райони за изследване (RoI – regions of interest), обучение и създаване на модел. Възможностите за задаване на броя параметри в процеса на обучение при този програмен модул са по-ограничени в сравнение с другата използвана програмна библиотека – ОТВ.

Седмият вариант анализира получените резултати при прилагане на класификация с обработващ модул за невронни мрежи – ОТВ.

Реализираният в дисертационния труд модел на процедура за класификация на МКСД с обучение, използваща НМ, е от тип персептрон с един скрит слой, тъй като е доказано, че такъв модел е в състояние да апроксимира произволна функционална зависимост. Както и при МОВ моделът на класифицираща процедура по метод с НМ е изграждан и валидиран върху същите два набора МКСД – 1999 и 2006г. Зададените параметри на етапа “обучение на модел” са: брой на скритите слоеве на НМ; брой на обработващите елементи във всеки един от скритите слоеве; начални стойности на тежестните коефициенти; максимален брой итерации за обучение на модела; максимална стойност на грешката между две итерации; коефициента моментум (разликата в стойностите за един тежестен коефициент между две последователни итерации).

Матриците на грешките, получени след тематичната класификация на пълния обем данни за 1999 г. и за 2006 г. и от етап “обучение на модел” с НМ са показани таблично.

Фрагменти от тематичните изображения, получени в резултат на класификацията на пълния обем данни за от модели с НМ 1999г. и 2006г., включващи основно класове градска част, водна площ, горска растителност и други елементи, са представени графично и са графично сравнени с полигони, определени от проекта Корине земно покритие.

Осмият вариант съдържа обобщение на получените резултати от създадените модели за тематична класификация на МКСИ с НМ и по МОВ.

Таблично са обобщени получените стойности за средната точност и обобщената оценка *kappa* за създадените оптимални на модели по методи МОВ и НМ при класификация върху извадката за валидация и след класификация на пълния набор от МКСД за години 1999 и 2006. Прави се заключението, че моделите, създавани по МОВ, показват по-добри показатели върху извадката, използвана на етап “валидиране на модел”, но при класификация на пълния обем не се представят добре в сравнение с моделите, прилагащи НМ. Този факт прави моделите с НМ предпочитан избор за създаване на тематични изображения в случаите, когато пълният набор МКСД съдържа класове спектрални данни, за които облакът от стойности в пространството на ХП е доста по-голям от използвания на етап “създаване на модел”.

Изводите от проведените експерименти, описани в пета глава, обобщават съществените резултати от експерименталните изследвания:

- моделите, прилагащи нелинейни методи за класификация с обучение на МКСД предоставят достатъчно възможности за тяхното прецизиране чрез търсенето на оптимални параметри, което повишава крайната точност след тематична класификация;

- моделите по МОВ позволяват повишаване на точността при класификация, посредством трансформация в пространство на ХП с по-висока размерност;

- моделите с НМ предоставят по-добри възможности за причисляване към даден информационен клас за МКСД за данни извън обучаващата извадка;

- изяснени са двете основни причини, които понижават точността на крайните тематични изображения: първата – наличие на пиксели от тип “смесен клас” в извадката,

използвана на етап “обучение на модел” и втората причина - използването на референтни полигони от проект КЗП, съдържащи погрешни пиксели по периферията.

- усложняването на топологията на модела с НМ чрез използване на два обработващи слоя не води до съществено подобряване на точността на етап “валидация на модел”;

- установени са значителни вариации за точността за клас, който е от спектрално смесен тип, съгласно проекта Корине земно покритие (КЗП), а най-малък брой погрешно разпознати обекти са установени при два от класовете, притежаващи специфични спектрални характеристики, разграничаващи ги надеждно от останалите класове.

Изложението на проблемите в текста на дисертационния труд е детайлно и изчерпателно. Главите от втора до пета представляват систематизирано изложение от общите принципи на дистанционните изследвания до конкретното теоретично разглеждане и експериментално изследване на методите, анализирани в дисертационния труд.

Положително впечатление прави наличието на изводи към всяка от главите, което показва способността на автора за обобщение и задълбочен анализ.

Списъкът на представените фигури и таблици от пета глава позволява по-добро систематизиране на резултатите от работата, а списъкът на използваните съкращения улеснява ползването на текста на разработката.

3.3. Литературни източници

Литературната справка съдържа 101 заглавия, като от тях 81 са на латиница, а 10 са на кирилица и 10 са адреси в Интернет (също на латиница).

Обемът и обхвата на използваните литературни източници показва, че докторантът се е постарал да се запознае разностранно и в максимален обем с проблемите по дисертацията. Като слабост може да се отбележи, че литературните източници не са номерирани, което затруднява тяхното систематизиране, въпреки че това не е абсолютно необходимо, тъй като в текста на работата е възприето по-правилното реферирание на източниците по имената на авторите и годината на публикуване.

3.4. Познаване на проблема

Съдържанието на работата и задълбоченият анализ на съществуващите методи показва доброто познаване на разработваните проблеми от страна на докторанта и неговата способност за адекватен избор на подходяща методика в зависимост от особеностите на решаваната задача за класификация и характеристиките на използваните данни. Предлаганите в работата методика за извличане на спектрални данни, метод за съвместна обработка и формиране на оптимални модели, могат да служат като методическо ръководство при решаването на задачи за класификация и да намерят приложение при решаване на практически задачи за анализ на изследваните територии и използване на набори от данни от различни източници. Получените резултати се основават на задълбочените познания на докторанта в областта на изследване и показват неговата способност за творчески подход и оригинално решаване на проблемите.

4. Приноси и значимост на разработката

Основните претенции на автора за приноси са формулирани в 5 точки. Приносите имат основно научно-приложен характер. Те са формулирани в съответствие с изискванията на ЗРАСРБ.

Три от приносите са теоретико-приложни и имат методически характер и са свързани с разработване на методики и методи, прилагани при използваните нелинейни методи за класификация.

Два от приносите, които са с номера 4 и 5 имат приложен характер като представляват практическите резултати от прилаганите методи и изследвания.

Признавам претенциите на докторанта за приносите и считам, че те са добре формулирани.

5. Автореферат

Текстът на автореферата съответства на дисертационния труд и отразява основните положения от проблемите, разглеждани в разработката, като е разработен съгласно изискванията на ЗРАСРБ. Разработен е на 56 стр., от които 50 стр. представят основни моменти от петте глави на разработката. В автореферата са включени още и справка за приносите,

списък на публикациите, свързани с дисертацията, списък с цитираната литература и списък на използваните съкращения.

6. Оценка на публикациите по дисертационния труд

Докторантът е представил три публикации, свързани с дисертационния труд. Една от работите е докладвана на Международен конгрес в Австрия, две на национални конференции с международно участие в България и две на международни конференции в България. Три от публикациите са самостоятелни, а другите две са в съавторство, като едната е с един съавтор, а другата с трима съавтори. И петте работи са публикувани на английски език. Една от разработките е публикувана през 1993 г., а останалите са от периода 2008 г. - 2016 г. Три от публикациите са посветени на проблемите на невронните мрежи при дистанционните изследвания, а другите две третират проблемите на метода на опорните вектори. Разработки от публикациите са включени в представения дисертационен труд, но явно е цитирана само една от публикациите (от 1993 г.). Съдържанието им и компетентното описание на проблемите, разгледани в тях, дават основание да се заключи, че те са дело на автора.

7. Цитирания и използване на резултатите

Докторантът не е представил сведения за известни цитирания на представените публикации.

Предлаганите от автора методики и методи за обработка имат основно изследователски и методичен характер, което дава възможност за тяхното използване при други изследователски практически разработки.

8. Критични бележки и препоръки

Работата е добре оформена текстово и графично. Положително впечатление прави списъка на термините. Към работата и същността на разработката нямам забележки по същество.

Към работата могат да се направят основно бележки, които са основно от технически или методичен характер.

1. Представянето на знаменателя на зависимостта за M_v се отличава от класическата форма на представяне на закона на Планк (формула [3] на стр. 14).

2. Случаят на припокриване на класовете, посочен за наличен на Фигура 11, се представя по-ясно на Фигура 12 от същата глава (стр. 39, 2 ред отдолу).

3. По мнение на рецензента формулировката на четвърта глава от работата "Приложение на методологията за разпознаване на образи ..." е твърде обща и не съответства на съдържанието на главата и по - удачно би било да се уточни като "Приложение на нелинейни методи от методологията за разпознаване на образи ..." (стр. 42, ред 1 отгоре).

9. Лични впечатления за кандидата

Познавам инж. Христо Николов от неговото представяне на научни конференции и научно-образователни лекции. Той се проявява като ерудиран научен работник с подчертано критично отношение към разглежданите проблеми и способност за творческо мислене.

10. Заключение

Получените резултати в разработката на гл. ас. инж. Христо Николов, формулираните в нейната работа приноси, като и оформянето на нейната работа и публикации показват, че той притежава нужните качества за научно-изследователска работа и способности за формулировка на оригинални решения и тяхното практическо реализиране.

Имайки предвид гореизложеното, предлагам на уважаемото научно жури да присъди на инж. Христо Стоянов Николов образователната и научна степен "доктор" в област на висше образование **4. Природни науки, математика и информатика** по професионално направление **4.4 Науки за Земята, научна специалност "Дистанционни изследвания на Земята и планетите"**.

гр. София,

26.11.2018 г.

ВЯРНО С ОРИГИНАЛА



Рецензент:

BM

(проф.) д-р инж. Б. Маринов