

# Детекција поплављеног пољопривредног земљишта у мултиспектралним сателитским снимцима

Немања И. Петровић, Предраг Лугоња, Милош Стојменовић,  
Дубравко Ђулибрк, Владимир Црнојевић

*Садржај* — У овом раду дајемо метод за откривање делимично или потпуно поплављених површина у пољопривредном земљишту, због којих долази до оштећења усева на самим тим смањења укупног приноса. Предложена метода за детекцију заснована је на генетском програмирању и користи мултиспектралне сателитске слике високе резолуције. Добијени детектори имају велику прецизност захваљујући способности генетског програмирања да пронађе скривене зависности којима се на основу спектралних података може прецизно оценити да ли је посматрана област поплављена или не. Развијени класификатори тестирани су и добијени резултати потврђују предност предложеног решења, како по питању тачности детекције поплављеног пољопривредног земљишта тако и по питању једноставности имплементације.

*Кључне речи* — детекција поплављених подручја, сателитски снимци, генетско програмирање.

## I. UVOD

САВРЕМЕНЕ методе за детекцију поплављених подручја углавном су усредсређене на проблематику лоцирања и надгледања великих поплављених или мочварних подручја зарад праћења еколошких последица промена у њима. У овом раду бавимо се донекле другачијим проблем. Циљ нам је да пронађемо мале поплављене области у пољопривредном земљишту. Посебно, заинтересовани смо да детектујемо такве области у равницама северне српске покрајине, где се оне повремено јављају у обрадивом земљишту као последица подизања нивоа подземних вода. До овог феномена обично долази после дуготрајних кишних периода; он захвата релативно мале, узајамно изолова-

Немања И. Петровић, Факултет техничких наука, Универзитет у Новом Саду, Трг Доситеја Обрадовића 6, 21000 Нови Сад, Србија (e-mail: [nemanja77@uns.ac.rs](mailto:nemanja77@uns.ac.rs)).

Предраг Лугоња, Факултет техничких наука, Универзитет у Новом Саду, Трг Доситеја Обрадовића 6, 21000 Нови Сад, Србија (e-mail: [elpeceef@gmail.com](mailto:elpeceef@gmail.com)).

Милош Стојменовић, Факултет техничких наука, Универзитет у Новом Саду, Трг Доситеја Обрадовића 6, 21000 Нови Сад, Србија (e-mail: [milos22@gmail.com](mailto:milos22@gmail.com)).

Дубравко Ђулибрк, Факултет техничких наука, Универзитет у Новом Саду, Трг Доситеја Обрадовића 6, 21000 Нови Сад, Србија (e-mail: [alef.tau@gmail.com](mailto:alef.tau@gmail.com)).

Владимир Црнојевић, Факултет техничких наука, Универзитет у Новом Саду, Трг Доситеја Обрадовића 6, 21000 Нови Сад, Србија (e-mail: [crnojevic@uns.ac.rs](mailto:crnojevic@uns.ac.rs)).

не области и нестаје након неколико недеља. Иако је привремен, његов утицај на усеве је веома значајан јер смањује принос. Стога је неопходно детектовати поплављене области ако би се правилно могла проценити настала штета. По својој структури овакве области сличне су мочварама јер су усеви и даље присутни у поплављеним њивама и негде чак земљиште и није потпуно поплављено већ је презасисацићено влагом. Имајући то у виду јасно је да проблем није нимало једноставан, јер се не своди на пуку детекцију области под водом.

Као што смо већ споменули, велики мочварни екосистеми детектују се и надзиру анализом сателитских снимака ниске резолуције (Landsat TM and ETM+) data [2] или радарских снимака (synthetic aperture radar - SAR) [1],[6]. Подаци које користимо у овом раду су осмоканалне слике високе резолуције добијене са сателита WorldView-2. Поред стандардних спектралних опсега присутних у сателитским сликама: црвеног, зеленог, плавог и блиског инфрацрвеног они садрже и приморско плави, жути, ивично-црвени и додатни блиско инфрацрвени опсег [3]. Сви они су доступни у високој просторној резолуцији од 50 cm. Имајући такве податке у могућности смо да разликујемо релативно ситне структуре на површини.

У овом раду предлажемо коришћење генетског програмирања (genetic programming - GP) [4], еволутивне методе за учење под надзором, за конструкцију класификатора који су у стању да препознају поплављено пољопривредно земљиште у сателитским сликама. Класификатори имају облик програмског стабла које као улазне параметре узима вектор спектралних компонента који одговара свакој појединачној локацији у слици и на основу њега одређује да ли је та локација поплављена или не.

## II. ДЕТЕКЦИЈА ПОПЛАВЉЕНОГ ЗЕМЉИШТА

Подаци које користимо у овом раду добијени су са комерцијалног сателита WorldView-2. Снимци су направљени у јулу 2010. године и садрже област претежно пољопривредног земљишта из околине Бечеја. Овакви подаци изузетно су занимљиви за анализу зато што је WorldView-2 први комерцијални сателит који има осам спектралних сензора у области видљивог и блиског инфрацрвеног спектра. Сваки сензор је уско фокусиран на свој део електромагнетног спектра и

```

minus (plus (X1, times (X7, X4)), plus (minus (plus (X1, times (times (X2, X6), times (X3, myif (times (X5, minus (X8, X1)), X7, plus (myif (X3, X7, -0.82742, X5), times (plus (X7, times (plus (X7, X6), myif (X8, X6, plus (-0.96248, -0.98426), X5))), minus (X2, plus (X7, times (minus (plus (X1, X6), plus (0.94681, X1))), myif (X8, X6, plus (X2, X6), plus (times (X2, X2), X4))))))), plus (X1, times (X3, X6))))), X2), myif (minus (plus (X2, times (plus (times (plus (plus (times (X1, X2), X5), myif (minus (plus (X2, times (times (X1, '-0.18002), minus (X8, plus (plus (X1, times (X3, X3))), times (X3, X3))), X3), myif (minus (plus (X4, times (times (X1, X2), X5))), X5, plus (plus (times (X1, X2), myif (minus (plus (X4, times (times (X6, X6), minus (times (-0.22957, -0.18002), X3))), plus (minus (X8, plus (-0.22957, X4)), X5)), X5, X2, X4))), myif (minus (plus (X1, times (X6, X6)), X6), X4, plus (X5, -0.98426), times (X3, X3))), times (X3, X3))), X3), myif (minus (plus (X4, times (times (X6, X6), minus (times (X6, X6), minus (myif (X8, X6, X4, times (X8, minus (X8, plus (X2, times (X1, X2))))), X3))), plus (minus (X8, plus (-0.22957, X4)), X5)), X5, myif (X2, X8, X7, X2), minus (plus (X2, times (times (X1, -0.18002), minus (X7, plus (X1, times (X7, -0.96248)))))), plus (times (X2, X2), X5))), X4), plus (minus (plus (plus (X1, times (times (myif (X8, X6, plus (X2, -0.98426), X5), -0.18002), times (-0.41115, X4))), times (times (X3, myif (X4, myif (X3, X7, -0.82742, X5), X6, -0.82742))), minus (X8, plus (X1, times (times (X1, X8), times (minus (plus (X1, times (X3, X6)), plus (times (X1, X1), X3)), X7))))), X1), myif (minus (plus (X1, times (plus (minus (X1, X2), myif (minus (plus (X1, times (times (X5, X6), X5))), plus (minus (X8, X2), X5))), X5, minus (X1, X8), times (X5, X6))), myif (X2, minus (plus (X1, times (times (X4, -0.18002), times (X7, X4))), plus (times (X1, X2), X5))), X5, -0.82742))), plus (minus (X8, plus (-0.22957, X4)), X4)), X5, X4, times (myif (X8, X6, plus (X2, -0.98426), X5), X4))), X5, X1, times (X5, myif (X7, X4, X7, X5))))))

```

### Сл. 1. Детектор конструисан генетским програмирањем, приказан у облику функције

снима слике у високој резолуцији која износи 50 cm. Основни циљ прикупљања оваквих података је обезбеђивање више информација за класификацију земљишта и вода које се посматрају. Овај сателит детектује следеће спектралне компоненте: црвена (630-690 nm), зелена (510-580 nm), плава (450-510 nm), блиска инфрацрвена (NIR1 - 770-895 nm), приморско плава (400-450 nm), жута (585-625 nm), ивично-црвена (705-745 nm) и додатна блиско инфрацрвена (NIR2 - 860-1040 nm). Последње четири компоненте су нове, у односу на оне које остали комерцијалне сателити за прикупљање података обезбеђују и свака од њих има своје особености. Приморско плава се најмање апсорбује у води али је апсорбује хлорофил у здравим биљкама. Компонента NIR2 се донекле поклапа са NIR1 компонентом али је мање осетљива на атмосферске утицаје. Ивично-црвена је специјално дизајнирана тако да се максимално рефлектује од вегетације и намењена је мерењу и класификацији биљака [3].

Директну зависност између различитих вредности ових спектралних компоненти и чињенице да је одређена област у слици полављена није могуће уочити на први поглед. Због тога смо применили генетско програмирање, да бисмо пронашли функцију која што је могуће веродостојније описује ту зависност.

#### А. Генетско програмирање

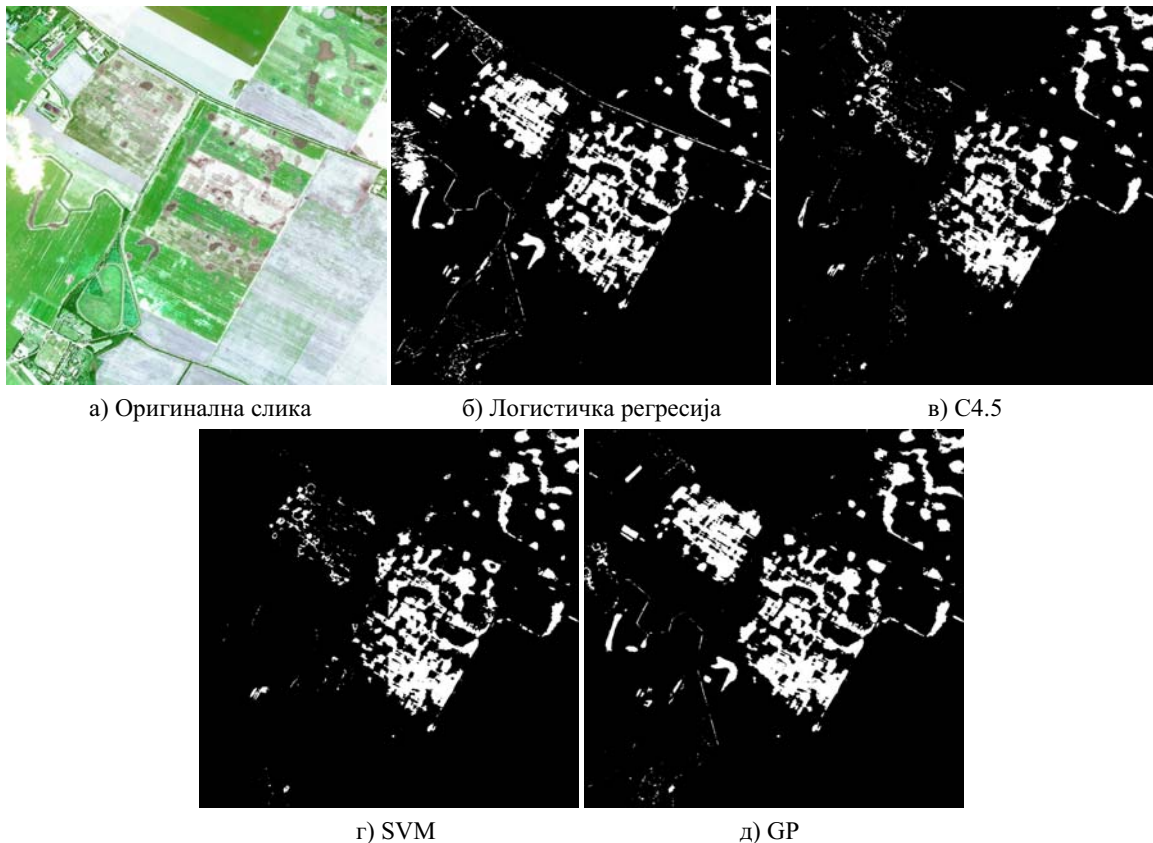
Генетско програмирање је релативно модерна техника машинског учења под надзором. Основна особина, због које је и занимљива за примену, је њена способност да научи изразито нелинеарне зависности између улазних атрибута и излазне одлуке класификатора. Оно користи еволутивни принцип природне селекције за претраживање простора свих могућих решења да би пронашао најбоље. Претрага започиње од почетне, насумично генерисане популације могућих решења од које генетско програмирање формира нове генерације решења уз помоћ генетских операција: мутације и укрштања. Принцип природне селекције користи се у процесу формирања нових генерација тако што боља решења, односно једике из популације имају, сразмерно квалитету, веће шансе да учествују у генетским операцијама, односно у стварању нове популације. Код класификационих проблема уобичајени начин мерења квалитета решења из популације је

израчунавање тачности коју постиже на тренинг скупу. Класификатор конструисан генетским програмирањем има форму стабла које се састоји од терминала и примитивних функција. Терминали су заправо вредности атрибута података из тренинг скупа, а примитивне функције се користе за њихово комбиновање. Стабло формирано генетским програмирањем за сваки појединачни елемент тренинг скупа израчунава вредност у корену стабла користећи атрибуте тог елемента и класификује га поређењем добијене вредности са унапред дефинисаним прагом. Оваква структура омогућава прављење изузетно нелинеарних класификатора који са истовремено имају једноставну структуру. Детаљнији опис генетског програмирања излази ван оквира овог рада, а може се пронаћи у [4].

Изузетно важна особина генетског програмирања је њена отпорност на присуство бескорисних атрибута у тренинг скупу [11]. Ако поједини атрибути нису информативни за класификацију биће или у потпуности игнорисани или слабо заступљени у класификатору конструисаном помоћу генетског програмирања. Како не поседујемо знање о томе који су атрибути из наших осмоканалних мултиспектралних података са сателита WorldView-2 у корелацији са проблемом детекције поплавлених подручја, коришћење генетског програмирања за конструкције детектора је више него оправдано. Слично резоновање је већ коришћено за анализу хиперспектралних слика [5].

#### В. Конструисани детектор

Конструисали смо детектор који анализира сваки пиксел понаособ и доноси одлуку да ли је захваћен поплавом или не. Спектралне компоненте осмоканалне слике користили смо као атрибуте на основу којих се пиксели класификују. Према томе, скуп терминала коришћен за генетско програмирање састоји се од осам спектралних компоненти, а сваки од њих је посебно нормализован на опсег [-1,1], како би се појединачно процес учења [10]. Као терминали коришћене су и насумично генерисане константе у опсегу вредности [-1,1]. Оне се генеришу током GP еволуције у складу са униформном дистрибуцијом, што је стандардан метод који се код генетског програмирања користи за побољшање стохастичке претраге простора



Сл. 2. Детекција поплавлених подручја у пољоприврдном земљишту

решења [4]. Дакле, терминални скуп има укупно девет компонената: осам спектралних и једну која насумично генерише константу.

Пошто детектор анализира сваки пиксел понаособ, намера нам је била да направимо што је могуће једноставнији класификатор. Зато смо узели да се скуп функција састоји од једноставних двоулазних аритметичких функција: сабирања (*plus*), одузимања (*minus*) и множења (*times*) и посебно дефинисане четвороулазне *myif* функције, дефинисане на следећи начин:

$$myif(R1, R2, C1, C2) = \begin{cases} R1, & C1 \geq C2 \\ R2, & C1 < C2. \end{cases} \quad (1)$$

Захваљујући *myif* функцији постоји могућност да тренирани детектор моделује нелинеарне зависности између елемената скупа терминала и класификационог проблема, односно детекције пиксела који приказују поплављено земљиште.

Тренинг детектора је спроведен тако да се пиксели којима је вредност на излазу GP стабла већа или једнака нули класификују као захваћени поплавом, док се они са вредношћу мањом од нуле детектују као непоплављени. Тренинг скуп направљен је ручним означавањем области под усевима које су поплављене. Садржи ~200000 елемената и балансиран је, односно садржи подједнак број пиксела захваћених поплавом и оних који то нису.

Конструисани детектор приказан је на Сл. 1. У облику функције чија се вредност израчунава а затим одлука о класификацију доноси поређењем резултата са

нулум. Променљиве  $X_1, \dots, X_8$  означавају вредности нормализованих атрибута, односно спектралних компоненти: приморско плава (400-450 nm), плава (450-510 nm), зелена (510-580 nm), жута (585-625 nm), црвена (630-690 nm), ивично-црвена (705-745 nm) NIR1 (770-895 nm) и NIR2 (860-1040 nm), редом.

Детектор приказан на Сл.1. добијен је као најбољи из популације од 400 јединки, односно потенцијалних решења, након еволуције кроз 200 генерација. У еволутивном процесу 5% јединки нове генерације стварано је простим копрањем најбољих јединки из претходне генерације, 25% јединки добијено је мутацијом, преосталих 70% добијено је укрштањем јединки из претходне генерације.

### III. РЕЗУЛТАТИ

Класифајер направљен генеским програмирањем тестиран је поређењем са решењима базираним на логистичкој регресији [7], стаблима одлучивања и машинама са потпорним векторима (SVM) [8]. Сви класификатори су тренирани на истом тренинг скупу. Поред тренинг скупа, ручним означавањем поплављених њива формиран је и тест скуп који садржи нешто мање од 90000 пиксела. У табели 1. приказана је тачност поређених класификатора на тренинг и на тест скупу. Видимо да једино логистичка регресија постиже бољу тачност на тренинг скупу него на тест скупу, док је код осталих тестираних алгоритама обрнут случај. Ова нелогичност може се објаснити једино тако да

је тренинг скуп „компликованији“ за класификацију од тест скупа, односно да његови елементи заузимају део простора класификације који је већи од дела простора који заузимају елементи тест скупа и који га у потпуности покрива. У сваком случају, из табеле 1. видимо да сви алгоритми имају релативно добру генерализацију, а да се генетским програмирањем постижу најбољи резултати.

ТАБЕЛА 1: ТАЧНОСТ ДЕТЕКТОРА

Детектор	Тренинг скуп	Тест Скуп
Логистичка	97,58 %	95,21 %
C4.5	98,53 %	98,96 %
SVM	98,80 %	99,05 %
GP	98,87 %	99,62 %

Да бисмо илустровали резултате класификације посматрали смо део пољопривредног земљишта где поплављене области нису ручно означене. У Сл. 2. Приказан је тај комад земљишта, заједно са маскама које индикују детектоване области. Сл. 2(а) приказује само три спектрална канала из улазне осмоканалне слике која се користи у класификацији, док слике Сл. 2. (б)-(г) приказују поплављене области које су детектоване различитим класификаторима. На први поглед је јасно да најједноставнији класификатор, логистичка регресија има проблема са погрешним детекцијама. Он путеве и облаке брка са поплављеним подручјима. Стабло одлучивања направљено је уз помоћ C4.5 алгоритма [9]. Оно изоставља велике поплављене области, слично као и SVM детектор. Коначно, на Сл. 2. (д) можемо видети да резултат добијен помоћу класификатора направљеног генетским програмирањем даје боље резултате него остали детектори.

Уколико поредимо комплексност класификатора, логистичка регресија је свакако најмања јер садржи само један неурон са осам улаза, али јој је и квалитет најслабији. Стабло одлучивања направљено помоћу C4.5 алгоритма има дубину 23 и укупно 1217 чворова, док је стабло које одговара функцији са Сл. 1. једноставније структуре јер има дубину 19 и 343 чворова. Коначно, детектор конструисан помоћу SVM алгоритма је свакако најкомплекснији јер садржи преко 6000 потпорних вектора и користи Гаусове кернел функције што као резултат даје веома спору класификацију.

#### IV. ЗАКЉУЧАК

Лоцирање поплављених подручја у пољопривредном земљишту битно због процене штете до које долази због уништавања усева. Детектор који смо развили у овом раду приказали у овом раду постиже заснован је на генетском програмирању и класификује пикселе мултиспектралних сателитских снимака високе резолуције. Притом постиже изванредне резултате: једноставнији је од других решења, а тачност му је већа. Наш будући рад биће усмерен на откривање детекцију појединачних биљних култура у пољопривредном земљишту и прецизније одређивање колико је која култура оштећена поплавом.

#### ЗАХВАЛНИЦА

Овај рад делом је спонзорисан „FP7 AGROSENSE“ пројектом.

#### ЛИТЕРАТУРА

- [1] E. S. Kasichke, K. B. Smith, L. L. Bourgeau-Chavez, E. A. Romanowicz, S. Brunzell and C. J. Richardson, “Effects of seasonal hydrologic patterns in south Florida wetlands on radar backscatter measured from ERS-2 SAR imagery,” *Remote Sensing of Environment*, vol. 88, no. 4, 2003, pp. 423–441.
- [2] E. R. De Roeck, N. E. Verhoest, M. H. Miya, H. Lievens, O. Batelaan, A. Thomas and L. Brendonck, “Remote Sensing and Wetland Ecology: a South African Case Study,” *Sensors*, vol. 8, no. 5, 2008, pp. 3542–3556.
- [3] Digital Globe, “The Benefits of the 8 Spectral Bands of WorldView-2,” *white paper*, 2010.
- [4] J. R. Koza, “Genetic Programming: On the Programming of Computers by Means of Natural Selection,” MIT Press, 1992.
- [5] B. J. Ross, A. G. Gualtieri, F. Fueten, P. Budkewitsch, “Hyperspectral image analysis using genetic programming,” *Applied Soft Computing*, vol. 5, no. 2, 2005, pp. 147–156.
- [6] S. Wdowski, S.-W. Kim, F. Amelung, T. H. Dixon, F. Miralles-Wilhelm and R. Sonenshein, “Space-based detection of wetlands' surface water level changes from L-band SAR interferometry,” *Remote Sensing of Environment*, vol. 112, no. 3, March 2008, pp. 681–696.
- [7] A. Agresti, “Building and applying logistic regression models,” in *An Introduction to Categorical Data Analysis*, Hoboken, New Jersey: Wiley, 2007, pp. 138.
- [8] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Machine Learning*, 20, 1995.
- [9] J. R. Quinlan, “Improved use of continuous attributes in c4.5,” *Journal of Artificial Intelligence Research*, vol. 4, 1996, pp. 77–90.
- [10] S.-S. Choi and B.-R. Moon, “Genetic and Evolutionary Computation,” *Lecture Notes in Computer Science*, vol. 2723, 2003, pp. 198.
- [11] W. A. Tackett, “Genetic programming for feature discovery and image discrimination,” in *Proc. 5th Int. Conf. Genetic Algorithms*, San Mateo, CA, 1993, pp. 303–309.

#### ABSTRACT

In this work, we develop a method for the detection and monitoring of small waterlogged areas in farmland which significantly deteriorate or completely damage the plants in the affected zone, thus reducing the average crop yield. We use high-resolution multispectral satellite images as input data and develop highly accurate wetland detection classifiers by utilizing genetic programming. We tested the developed techniques and the results confirm the advantage of the proposed method, both in the accuracy of the localization of wetlands and in implementational complexity.

#### DETECTION OF WETLAND AREAS IN MULTISPECTRAL IMAGES OF FARMLAND

Nemanja I. Petrović, Predrag Lugonja, Miloš Stojmenović, Dubravko Ćulibrk, Vladimir Crnojević